

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/78557> holds various files of this Leiden University dissertation.

Author: Dinis Fernandes, C.

Title: Multi-parametric MRI to guide salvage treatment of recurrent prostate cancer

Issue Date: 2019-09-25

MULTI-PARAMETRIC MRI TO GUIDE
SALVAGE TREATMENT OF
RECURRENT PROSTATE CANCER

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 25 september 2019
klokke 11.15 uur

door

Catarina Dinis Fernandes
geboren te Lissabon
in 1989

MULTI-PARAMETRIC MRI TO GUIDE
SALVAGE TREATMENT OF
RECURRENT PROSTATE CANCER

Catarina Dinis Fernandes

Multi-parametrische MRI voor de salvage behandeling van recidief prostaatkanker

Ressonância magnética multiparamétrica como meio de guiar o tratamento do cancro
da próstata recorrente

Promotor:

prof. dr. U.A. van der Heide

Co-promotor:

dr. M. Smolic (*NKI-AvL*)

Promotiecommissie:

prof. dr. C.A.M. Marijnen

prof. dr. R.C.M. Pelger

prof. dr. K. Haustermans (*KULeuven*)

prof. dr. A.G.J.M. van Leeuwen (*University of Amsterdam*)

ISBN: 978-94-6323-756-7

Printed by: Gildeprint - Enschede

An electronic version of this thesis can be found at

<https://openaccess.leidenuniv.nl>

Cover by Diana Costa

The work on this thesis was performed at the Netherlands Cancer Institute in Amsterdam, and was funded by the Dutch Cancer Society, grant NKI 2013 – 5937 and project 10088.

CONTENTS

| | | |
|---|--|-----|
| Chapter 1 | Introduction | 1 |
| Part 1 Magnetic resonance imaging in the treatment and outcome prediction of primary prostate cancer | | |
| Chapter 2 | Biochemical recurrence prediction after radiotherapy for prostate cancer with T2w magnetic resonance imaging radiomic features | 15 |
| Chapter 3 | Prostate fiducial marker detection with the use of multi-parametric magnetic resonance imaging | 37 |
| Part 2 Magnetic resonance imaging for the diagnosis and localisation of recurrent prostate cancer | | |
| Chapter 4 | Quantitative 3T multi-parametric MRI of benign and malignant prostatic tissue in patients with and without local recurrent prostate cancer after external-beam radiation therapy | 55 |
| Chapter 5 | Quantitative 3T multi-parametric MRI and step-section pathology of recurrent prostate cancer patients after radiation therapy | 81 |
| Chapter 6 | Multi-parametric MRI tumour probability model for the detection of locally recurrent prostate cancer after radiation therapy: pathologic validation and comparison with manual tumour delineations | 105 |
| Chapter 7 | General discussion | 131 |

| | |
|----------------------|-----|
| References | 143 |
| Summary | 161 |
| Samenvatting | 167 |
| Sumário | 173 |
| List of Publications | 179 |
| Acknowledgements | 183 |
| Curriculum Vitae | 185 |

LIST OF ABBREVIATIONS

| | |
|--------|---|
| ADC | Apparent diffusion coefficient |
| ADT | Androgen deprivation therapy |
| CBCT | Cone-beam computer tomography |
| CG | Central gland |
| CT | Computer tomography |
| CTV | Clinical tumour volume |
| CV | Cross-validation |
| DCE | Dynamic contrast enhanced |
| DWI | Diffusion weighted imaging |
| EBRT | External-beam radiotherapy |
| EPID | Electronic portal imaging device |
| GLMM | Generalized linear mixed-effect modelling |
| GTV | Gross tumour volume |
| HDR | High-dose rate |
| H&E | Hematoxylin and eosin |
| LDR | Low-dose rate |
| mp-MRI | Multi-parametric magnetic resonance imaging |
| MRI | Magnetic resonance imaging |
| PCa | Prostate cancer |
| PET | Positron emission tomography |

| | |
|------|------------------------------------|
| PSA | Prostate specific antigen |
| PSMA | Prostate-specific membrane antigen |
| PU | Periurethra |
| PZ | Peripheral zone |
| RT | Radiotherapy |
| SRP | Salvage radical prostatectomy |
| SV | Seminal vesicles |
| TP | Tumour probability |

INTRODUCTION

1.1 PROSTATE CANCER

The prostate gland is part of the male reproductive system and can be anatomically divided into three main zones: the transition and central zone, which together comprise the central gland, and the peripheral zone. The prostate can undergo vast changes during a man's lifetime. Along with prostatitis and benign prostatic hyperplasia, prostate cancer is a common pathological condition. Prostate cancer (PCa) affects more than 40% of men over the age of 60 [1]. Worldwide, it is the second most frequently diagnosed cancer in males next to lung cancer, and the most common non-cutaneous malignancy among men in the Western world [2, 3]. Incidence rates of PCa vary geographically but have seen an overall increase whereas mortality rates have declined in most countries [4]. Considering that the global population aged 60 or more is expected to nearly double by 2050 [5], the management of age related diseases such as PCa, while ensuring a good quality of life, presents a health challenge worldwide.

In clinical practice, increased prostate-specific antigen (PSA) levels are used as biomarkers for PCa. However, PSA screening is controversial because although there is evidence that it leads to a reduction in PCa-specific mortality, it does not seem to have an effect on overall mortality rates [6]. Additionally, PSA is not PCa specific, as a number of benign conditions can also result in a PSA rise. The ongoing debate is focused on how to prevent overdiagnosis and overtreatment of clinically insignificant cases. The main treatment options are radiotherapy (RT) and the surgical removal of the prostate gland (prostatectomy). The field of RT has seen great technological advances in the past decades, with improvements in patient position verification and delivery techniques allowing for more conformal dose distributions. Similarly, the in-

roduction of robot-assisted surgery allows surgeons to perform more precise dissections potentially leading to better preservation of functional structures [7]. Improvements in diagnosis and treatment options have led the 5-year age-adjusted survival rates to increase from 49% in 1970 to 89% in 2015 [8].

1.2 IMAGING FOR THE STAGING OF PRIMARY PROSTATE CANCER

1.2.1 Local staging

In the diagnosis of PCa, multi-parametric MRI (mp-MRI) refers to an MRI protocol typically consisting of high-resolution T2-weighted (T2w) anatomical scans, as well as at least two functional sequences [9]. When performed before biopsy, it is a sensitive tool to detect clinically relevant tumour and to localise sites for targeted biopsies [10]. This is the technique of choice for local tumour staging (T “staging”), as it allows the identification of extracapsular extension and seminal vesicle invasion.

The most relevant anatomical sequence for detection of primary PCa is a T2w scan. This sequence offers a good depiction of the gland anatomy, where a tumour region can be identified as hypointense due to its shorter T2 decay.

In PCa, the two main tissue properties probed with functional MRI are diffusion and perfusion. Diffusion-weighted imaging (DWI) probes the random translational motion of water molecules. As tumour tissue is characterised by increased cellular density, these densely packed regions restrict the extracellular movement of water molecules, resulting in hyperintense regions in the DWI sequence. An apparent diffusion coefficient (ADC) map can be derived from DWI images acquired at different gradient strengths. The ADC map provides a quantitative measure of restricted diffusion for each voxel and is usually low in tumour regions. Dynamic contrast-enhanced (DCE) MRI is used to quantify tissue perfusion by dynamically monitoring the distribution of injected contrast agent in the region of interest. The high energetic demand of growing tumour cells is accompanied by the formation of a new vascular bed which is often

disorganised, inefficient and leaky. These abnormal perfusion characteristics can be imaged and distinguished from those of healthy tissue, and a quantification of vessel permeability can be obtained by using tracer kinetic modelling. In prostate tissue, the Tofts model is commonly used, yielding values for the volume transfer constant of contrast agent molecules through the capillary endothelium (K^{trans}), the volume of the extravascular-extracellular space (v_e) and the rate constant between the extravascular-extracellular space and the plasma (k_{ep}). Tumour tissue is associated with elevated K^{trans} and k_{ep} values.

1.2.2 Nodal and metastatic staging

The detection of nodal spread (“N” staging) can be performed with abdominal CT, anatomical MRI or positron emission tomography (PET). Both CT and MRI methods indirectly assess nodal invasion by allowing the measurement of nodal diameter and morphology. However, the sensitivity of these methods is only around 40% on the basis of a 10 mm threshold for nodal involvement [11]. ^{68}Ga - or ^{18}F -labelled prostate-specific membrane antigen ligands (PSMA) PET is increasingly being used, since preliminary evaluations have shown promising sensitivity for the detection of lymph node involvement [12–14]. For the assessment of metastatic spread (“M” staging), which in PCa usually takes the form of bone metastasis, bone scintigraphy (also known as bone scan) has been the most widely used method [15], albeit with an increasing popularity of PSMA scans [16]. The tumour, nodal and metastatic staging results form together the TNM staging, a clinical parameter used to define tumour spread.

1.2.3 Prostate cancer risk groups

The Gleason grading system is used to evaluate the histopathology of the biopsy specimen by describing the glandular pattern and the degree of differentiation. In combination with the TNM staging and the measured PSA values, it is used to assess the tumour risk and to ground treatment decisions. Prostate cancer can thus be categorised in three different risk groups. According to the European Association of Urology (EAU) guidelines [15], low-risk PCa is defined by a tumour with a clinical T-stage of cT1-2a, with a Gleason grade <7, and PSA values <10 ng/ml. When the PSA values are between 10–20 ng/ml, or the Gleason grade = 7, or the clinical T-stage is cT2b, the tumour is categorised as

intermediate-risk PCa. If the PSA values are >20 ng/ml, or the Gleason grade is >7 , or the tumour extends to both lobes having thus a clinical T-stage of cT2c, the tumour is defined as high-risk PCa. While the previous descriptions relate to localised prostate cancer, high-risk tumours can also extend beyond the prostate in which case they are referred to as locally advanced. These are characterised as either a clinical T-stage \geq cT3 or positive regional lymph nodes (cN+) with any PSA and Gleason grade values.

1.3 RADIOTHERAPY AS A TREATMENT OPTION FOR PRIMARY PROSTATE CANCER

Alongside radical prostatectomy, RT is one of the main treatment options for primary PCa and can be administered in two main forms: interstitially in the form of brachytherapy or as external-beam radiotherapy (EBRT). Brachytherapy can be delivered by permanently implanting radioactive seeds (low dose-rate brachytherapy, LDR) or by using needles and an afterloading device to transport a radioactive source close to the target for a pre-determined period of time (high dose-rate brachytherapy, HDR). EBRT treatments deliver the dose by using a combination of radiation beams generated by a linear accelerator and adapted to maximise target coverage.

Disease extent, tumour differentiation grade and the patient's preferences influence the choice for a specific treatment strategy. Tumours of any differentiation grade that are locally confined without extracapsular extension can be treated with either radiotherapy or prostatectomy. According to the EAU guidelines, brachytherapy in the form of LDR should be restricted to prostate volumes of less than 50 cm^3 and is an appropriate treatment option for patients with locally confined and moderate or well-differentiated tumours, while HDR can be offered to men with poorly differentiated tumours and/or with higher tumour volumes [15]. EBRT (at times combined with HDR brachytherapy) is the treatment of choice for those with suspected extraprostatic extension. In patients with moderate and poorly differentiated tumours, EBRT can be combined with adjuvant or neo-adjuvant hormonal therapy.

During EBRT treatments, the dose is delivered in several fractions. Before each fraction, images are acquired for patient position verification. Visualisation of the prostate gland in the daily acquired cone-beam CT (CBCT) or electronic portal imaging device (EPID) images is challenging due to the poor soft-tissue contrast. Therefore, typically three to four gold fiducial markers are implanted in the prostate gland prior to treatment to allow daily verification of the patient position. These are visible in the planning CT and in the daily acquired CBCT or EPID images (obtained via devices integrated in the linear accelerator), and are used as a prostate surrogate. A transition towards an MR-only treatment planning workflow is currently under investigation. In the first stages, MRI is employed exclusively for treatment planning while the conventional Linac machines with CBCT position verification are used to deliver the treatment. In this setting, fiducial markers are still used for patient position verification, and one precondition to implement an MR-only treatment planning workflow is the ability to detect them on MRI. However, due to their paramagnetic properties they present as local signal voids on most MR images. Sequence specific parameters influence the appearance of the marker voids, increasing the complexity of reproducible and accurate manual marker localisation on MRI. While the implementation of MR-Linac systems will eventually render obsolete the use of fiducial markers for patient position verification, these systems have only recently been clinically implemented and are not widely available. Brachytherapy seeds are usually encapsulated in titanium, resulting in artefacts similar to those of the markers on the MR images. Since the markers and seeds are permanently implanted, these effects must be taken into consideration when assessing MRI scans of patients after primary RT treatment.

1.3.1 Treatment outcome prediction

Clinical nomograms, such as the Kattan nomogram [17], use the patient's clinical characteristics to provide an indication of the risk of biochemical recurrence. However, such clinical parameters fail to explain some of the variability seen in patient outcome. Improved stratification can potentially be obtained by including information from pre-treatment imaging.

MRI scans are commonly acquired for the purpose of tumour staging and treatment planning, and evidence is emerging that factors such as tumour ag-

gressiveness are associated with specific textural features from MR imaging [18, 19]. MRI-derived features could potentially provide insight into the recurrence risk. Visually scored, MRI-based, semantic attributes of PCa, such as seminal vesicle invasion or extracapsular extension, are predictors for 5-year biochemical recurrence free survival. These predictors improve staging, thereby augmenting the performance of predictive models which combine them together with clinical nomograms [20]. In heterogeneous cohorts, tumour-extracted imaging features from T2w MRI were associated with recurrent disease after RT [21]. This information might be used to determine in which patients conventional treatment is likely to succeed and those where intensification would be beneficial. Specifically for high-risk patients, who bear the highest recurrence rates [22], early identification of increased recurrence risk can potentially impact clinical management and subsequent follow-up.

1.4 RECURRENT PROSTATE CANCER AFTER RADIOTHERAPY: INCIDENCE AND TREATMENT OPTIONS

After RT, PSA values gradually decrease until they reach their lowest level (nadir). If after the initial treatment the PSA values rise above a predefined level there is a concern of a recurrent tumour, and this increase is defined as 'biochemical recurrence'. One of the most widely used definitions for biochemical recurrence is a PSA rise of 2 ng/ml above the nadir (Phoenix definition) [23]. A recent study analysing a series of 2694 patients treated with doses equal to or greater than 75.6 Gy reported 8-year biochemical recurrence risks of 10%, 23%, and 43% in low-, intermediate-, and high-risk PCa patients, respectively [24]. Thus, for a proportion of patients the primary RT will fail to eradicate the tumour.

For patients with recurrent disease after RT, the standard management is currently androgen deprivation therapy (ADT), adapted according to disease extent and progression rate. ADT is given with the intent of delaying disease progression, but it will not prevent metastatic spread and eventual death. It

is further associated with severe side-effects that can strongly impact a patient's lifestyle [25]. While most patients with recurrent disease have already developed metastases, Hruby *et al.* [26] and Bolla *et al.* [27] found that 17–20% of the patients will harbour locally recurrent disease only. For these, curative – commonly denominated “salvage” – treatment options can be offered, including radical prostatectomy, brachytherapy, cryotherapy, or high-intensity focused ultrasound (HIFU). The oncologic outcomes seem to be comparable between the available whole-gland salvage treatment options [28]. However, conventional whole-gland treatments can result in substantial morbidities, such as incontinence, rectal injury and bladder neck stricture [29, 30], and should only be offered to highly selected patients [31]. The toxicity associated with whole-gland approaches has made physicians reluctant to refer patients for salvage treatments. Reports indicate that only 2% of radiation failures are currently managed with local salvage treatments. The rest of the patients are either immediately treated with ADT, or managed with observation and, if deemed necessary, started on ADT [32].

To overcome the toxicity related to whole-gland strategies, while maintaining local control, focal treatments targeting the index lesion while sparing the surrounding uncompromised tissue can be offered to selected patients. The commonly offered whole-gland salvage approaches can then be designed with focal intent and with different ablation extents, ranging from ultrafocal, to hemi or quadrant-wise. Functional and oncological outcomes differ per modality and ablation extent, and no randomised trials have yet been performed to ascertain which modality and ablation scheme is optimal. Overall, the 5-year biochemical failure-free survival is between 46.5–54.5% [33]. Moreover, Duijzentkunst *et al.* [33] suggested that focal salvage treatments could decrease severe genitourinary (GU) and gastrointestinal (GI) toxicities, potentially preserving erectile function.

1.5 IMAGING IN RECURRENT PROSTATE CANCER

Radiographic detection of recurrent PCa in the context of biochemical recurrence is essential in evaluating disease spread, as well as to set expectations

and determine an appropriate treatment for the patient. The most accurate PET tracer for the detection of tumours and metastatic spread is currently ^{68}Ga PSMA-PET [13], with reported detection rates of 48 - 63.5% for locally recurrent PCa cancer after RT [26, 34]. On the other hand, mp-MRI, using both anatomical and functional sequences, is increasingly used to stage locally recurrent disease [35–37], and it has been reported to obtain good sensitivity in detecting radio-recurrent PCa [38–41]. The previous RT treatment has likely impact not only on the structure but also on the physiology of the prostatic tissue. So far, a quantitative characterisation of mp-MRI in the recurrent setting after RT has not been established and it remains to be defined which sequences are preferred for the detection of recurrent PCa. Currently, the sequences performed for the detection of recurrent disease are typically similar to those used in the primary setting.

Distinguishing between local recurrence and metastases is critical in determining whether treatment with local salvage therapy is sufficient, or if systemic treatment should be provided to control disease progression. To this end mp-MRI findings are used in combination with those from PET to reach a diagnosis.

1.5.1 Image assessment

^{68}Ga PSMA-PET is increasingly being used to assess metastatic spread, whereas mp-MRI can be used to evaluate intraprostatic disease. In the setting of primary PCa, the PI-RADS v2 consensus guidelines [42] were established to inform radiologists on how to interpret and rate the different mp-MRI sequences for the likelihood of tumour presence. However, these guidelines are not applicable in the recurrent setting.

After RT, the interpretation of mp-MRI is complicated due to the presence of RT-induced benign confounders [36, 37, 43]. The prostate tissue presents with an overall diffuse signal intensity reduction in the T2w sequence, which complicates tumour detection [37]. Furthermore, DCE-derived maps can present with a benign increased enhancement in the peri-urethral tissues [44]. Conflicting results have been published regarding the importance of the different MRI sequences in the detection of recurrent PCa. The use of 1.5T MR spectroscopy [45] as well as DWI- and DCE-MRI [39, 40] in addition to T2w were found to

improve local recurrence detection. On the other hand, Donati *et al.* [46] and Luzurier *et al.* [47] found no additional benefit of DCE when added to T2w and DWI. To date, there are no guidelines on how to score or delineate recurrent tumours on mp-MRI, and the relevance of each abovementioned imaging sequences has yet to be determined for recurrent PCa.

Little is known regarding the effect of RT on the mp-MRI characteristics of benign tissue. An accurate description of mp-MRI characteristics of benign tissue after RT is necessary to establish a baseline reference against which suspicious findings can be compared. This would facilitate the design of radiological reporting guidelines and result in easier image interpretation. Currently, and in the absence of further guidance, PI-RADS v2 is often still used as a starting point for image assessment.

1.6 FOCAL SALVAGE TREATMENTS: HOW TO LOCALISE THE AREA TO TREAT?

Understandably, delineation of the target volume is of great importance in RT. Nevertheless, it remains one of the weakest links in the treatment chain [48, 49]. While whole-gland treatments are standard for primary PCa, the number of studies reporting on focal approaches is increasing. Yet, the lack of contouring guidelines results in large reported inter-observer variability when delineating primary PCa on mp-MRI [49, 50]. It may be inferred that accurate delineations are just as important in the focal treatment of recurrent disease, potentially influencing local tumour control. Thus, methods to decrease variability in defining the region to treat are warranted.

Even though ^{68}Ga PSMA-PET offers high sensitivity in detecting recurrent PCa, the high-resolution anatomical and functional possibilities offered by MRI, together with a good accuracy in detecting recurrent disease extension [51], make this an attractive technique to use for the design of focal salvage treatments [52]. In studies reporting on focal salvage treatments, the target is commonly defined with the use of MRI, often combined with biopsy results [53, 54]. Given that there is yet to be a consensus on which imaging sequences

should be used to identify recurrent tumour, as well as the varying definitions of focal treatment (i.e. ultrafocal, hemi, etc.), the target volumes reported in literature are highly variable and not directly comparable.

To validate imaging findings the gold-standard is histopathology validation. Biopsies are prone to sampling errors and fail to capture the spatial heterogeneity of the underlying tissue. Thus, prostatectomy samples represent the most complete histopathology reference. However, salvage prostatectomy after RT is only rarely performed due to high treatment-related toxicity rates. Only a few whole-mount histopathology validated studies exist on the MRI properties of recurrent tumour after radiotherapy [45, 55, 56]. From the available studies with salvage prostatectomy samples, evidence is emerging that recurrent tumour is multifocal [57] but with intraprostatic failures occurring predominantly at the location of the primary index lesion [55]. There are no studies that have registered the histopathology to MRI, which would enable a more accurate characterisation of tumour and healthy tissue.

1.7 AIMS AND OUTLINE OF THIS THESIS

Despite the technical advancements of recent decades, around 25% of all PCa patients treated with RT will experience a biochemical recurrence and a significant minority will have an exclusively local failure. The general aim of this thesis is to improve the detection and localisation of locally recurrent PCa after RT using mp-MRI. The different sections of this thesis are focused on improving the stratification of recurrence risk, detection, and definition of the region to treat with focal salvage therapies. More specifically, the research questions to be addressed in the following chapters are:

Part 1 - Magnetic resonance imaging in the treatment and outcome prediction of primary prostate cancer

- Can pre-treatment MRI features be used to improve patient stratification for the risk of experiencing biochemical recurrence?

- Is it possible to automatically detect gold fiducial markers on MRI?

Part 2 - Magnetic resonance imaging for the diagnosis and localisation of recurrent prostate cancer

- What are the quantitative mp-MRI characteristics of benign tissue after radiotherapy?
- What are the quantitative mp-MRI characteristics of recurrent tumours after radiotherapy?
- Can mp-MRI be used to automatically define a region to treat with focal salvage treatments?

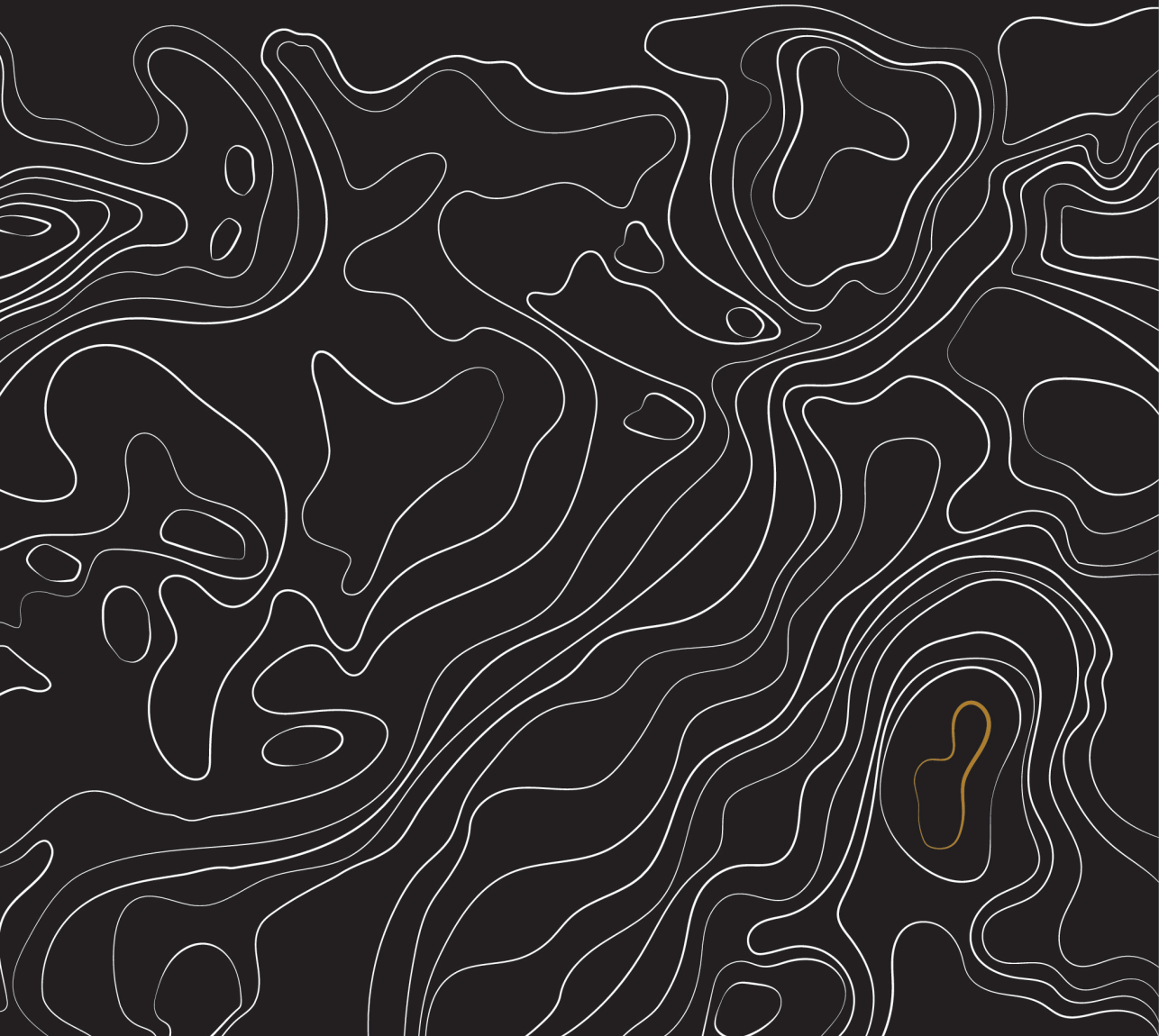
The studies described in the first part of this thesis are focused on the role of mp-MRI in the treatment of primary PCa. **Chapter 2** evaluates the potential of whole-gland T2w MR imaging features for 5-year biochemical recurrence risk prediction in high-risk patients treated with RT for local primary PCa. Recurrent disease is a significant problem for patients treated with primary RT for PCa. By identifying patients with high risk of recurrence prior to primary RT, treatment intensification could be offered, for instance, by providing adjuvant ADT or a dose boost to the index tumour lesion. Fiducial markers are implanted as part of the EBRT treatment for primary PCa and will therefore be present when imaging patients with a suspected recurrence after RT. **Chapter 3** describes how imaging features extracted from a combination of MRI sequences can be used with a machine learning classifier to automatically detect fiducial markers on MRI.

The second part of this thesis addresses the use of mp-MRI in the setting of recurrent disease after RT. The detection of recurrent tumours at an early stage, while disease is still confined to the gland, will potentially increase the success rate of salvage treatments. In order to allow early detection, we must improve the current diagnostic work-up for patients with suspected recurrence. After RT, MRI interpretation is complicated by benign RT-induced imaging confounders. To facilitate recurrent disease detection and distinction from benign changes, in **Chapter 4** a case-control study is used to establish a baseline reference by characterising benign tissue and comparing it to recurrent disease using mp-

MRI. By registering whole-mount slices of salvage prostatectomy specimens to the mp-MRI acquired prior to the surgery, the pattern of recurrence after RT as well as tumour imaging characteristics are presented in **Chapter 5**. Furthermore, manual MRI-based tumour delineations are compared to histopathology findings. **Chapter 6** describes how uncertainties in target volume delineation can be overcome by incorporating all mp-MRI sequences to obtain voxel-wise tumour probabilities that can be used to determine regions to target with focal salvage treatments. Finally, the impact of the findings reported in this thesis will be discussed in **Chapter 7**.

PART 1

Magnetic resonance imaging in the treatment and outcome prediction of primary prostate cancer



BIOCHEMICAL RECURRENCE PREDICTION
AFTER RADIOTHERAPY FOR PROSTATE
CANCER WITH T2W MAGNETIC RESONANCE
IMAGING RADIOMIC FEATURES

Catarina Dinis Fernandes¹
Cuong V Dinh¹
Iris Walraven¹
Stijn WTPJ Heijmink²
Milena Smolic¹
Joost JM van Griethuysen^{2,3}
Rita Simões¹
Are Losnegård^{4,5}
Henk G van der Poel⁶
Floris J Pos¹
Uulke A van der Heide¹

¹Department of Radiation Oncology, ²Radiology and ⁶Urology,
The Netherlands Cancer Institute, Amsterdam, The Netherlands

³GROW – School of Oncology and Developmental Biology
Maastricht University, Maastricht, The Netherlands

⁴University of Bergen, Bergen, Norway

⁵Haukeland University Hospital, Bergen, Norway

ABSTRACT

Background and purpose: High-risk prostate cancer patients are frequently treated with external-beam radiotherapy (EBRT). Of all patients receiving EBRT, 15-35% will experience biochemical recurrence (BCR) within five years. Magnetic resonance imaging (MRI) is commonly acquired as part of the diagnostic procedure and imaging-derived features have shown promise in tumour characterisation and biochemical recurrence prediction. We investigated the value of imaging features extracted from pre-treatment T2w anatomical MRI to predict five year biochemical recurrence in high-risk patients treated with EBRT.

Materials and methods: In a cohort of 120 high-risk patients, imaging features were extracted from the whole-prostate and a margin surrounding it. Intensity, shape and textural features were extracted from the original and filtered T2w-MRI scans. The minimum-redundancy maximum-relevance algorithm was used for feature selection. Random forest and logistic regression classifiers were used in our experiments. The performance of a logistic regression model using the patient's clinical features was also investigated. To assess the prediction accuracy we used stratified 10-fold cross validation and receiver operating characteristic analysis, quantified by the area under the curve (AUC).

Results: A logistic regression model built using whole-prostate imaging features obtained an AUC of 0.63 in the prediction of BCR, outperforming a model solely based on clinical variables (AUC = 0.51). Combining imaging and clinical features did not outperform the accuracy of imaging alone.

Conclusions: These results illustrate the potential of imaging features alone to distinguish patients with an increased risk of recurrence, even in a clinically homogeneous cohort.

2.1 INTRODUCTION

High-risk primary prostate cancer (PCa) patients are commonly treated with radiotherapy (RT). According to the Phoenix definition, biochemical recurrence (BCR) after RT occurs within five years in 15 - 35% of all cases [58–60]. Aiming at better patient stratification, clinical nomograms, such as the Kattan nomogram [17], have been developed to predict biochemical recurrence after RT. These incorporate factors such as the PSA level and biopsy Gleason score, known to be good predictors of biochemical recurrence, but are often limited by the accuracy of the measured variables. The prognosis of high-risk PCa patients is heterogeneous, however the available clinical nomograms are not tailored to distinguish patients within a single risk group.

Magnetic resonance imaging (MRI) is well established for PCa diagnosis and staging. T2-weighted (T2w) anatomical scans are used to assess the extent of the tumour. Visually scored semantic attributes of PCa visible on T2w-MRI, such as seminal vesicle invasion or extracapsular extension, are predictors for five year biochemical recurrence free survival. By improving staging, these predictors augment the performance of predictive models which combine them together with clinical nomograms [20].

Radiomics has emerged as a field in which a high number of quantitative imaging features are extracted from a region of interest (ROI) to quantitatively describe its phenotype. Texture analysis based on the grey level co-occurrence matrix (GLCM) [61], using second order statistics to characterise the spatial dependence of grey-levels in an image, has been applied extensively to evaluate both the location and aggressiveness of PCa [18, 19] based on T2w-MRI. By assessing tissue micro-architecture and tumour aggressiveness, being the latest by definition related with BCR, this modality could provide insight on recurrence risk prediction.

Local recurrence after RT is reported to occur predominantly at the site of the index lesion [55], and imaging features from the primary tumour were found to strongly associate with the probability of BCR following RT [21, 62]. However, these studies have relatively small and inhomogeneous patient cohorts, with the study by Gnep *et al.* [21] having a median follow-up time of only four years. An

early identification of increased recurrence risk can potentially impact clinical management and subsequent follow-up, particularly for high-risk disease as it is related with higher recurrence rates [22].

For many years, the standard for RT purposes was not multi-parametric MRI (mp-MRI) but anatomical imaging with the goal of prostate gland delineation. Thus, cohorts with longer follow-up are often restricted to T2w-MRI and the uncertainty in tumour localisation is high. Tumour delineations on mp-MRI are prone to inter-observer variations up to 2.3mm with smaller satellite lesions being often missed [49]. No guidelines are yet available for this task. Prostate delineations are also prone to inter-observer variability but radiological guidance on how to delineate in T2w-MRI is available [63]. Evaluation of whole-prostate imaging features avoids delineation uncertainty, is not restricted to the visible tumour area and might therefore be sensitive to both micro- and macroscopical features predictive of recurrence.

We here aimed to investigate the potential of whole-prostate imaging features for five year BCR prediction after RT of local PCa, in a clinically homogeneous cohort of high-risk biopsy-proven PCa patients. We further analysed the predictive value of features in the margin surrounding the prostate, as the presence of extracapsular extension and seminal vesicle invasion is related to the risk of relapse. The performance of models using imaging features was compared to one using solely clinical features.

2.2 MATERIALS AND METHODS

2.2.1 Dataset

In a single centre, patients with high-risk PCa were selected retrospectively from a consecutive cohort treated with external-beam radiotherapy (EBRT) between 2007 and 2011. Risk classification was performed according to the D'Amico definition [64]. Further inclusion criteria involved having received hormonal therapy (HT), a dose of 78Gy in 39 fractions and no other pelvic comor-

bidities before the treatment. Biochemical recurrence was diagnosed according to the Phoenix criteria [65] and all patients had five years of follow-up. A total of 120 patients satisfied the inclusion criteria.

Pre-treatment clinical predictors of biochemical recurrence after EBRT were chosen according to the input parameters of the Kattan nomogram [17].

2.2.2 MRI protocol

T2w anatomical MRI scans were acquired as part of the RT treatment planning procedure. Axial T2w turbo-spin echo (TSE) and T2w 3D VISTA (for nine patients) sequences were acquired on a 3T Philips Achieva MRI scanner (Philips Healthcare, Best, the Netherlands). For the TSE sequence the repetition times (TR) were longer than 3800 ms, the echo times (TE) between 120 –150 ms and the scans had an in-plane pixel pitch of 0.27 – 0.49 mm and slice thickness of 2.3 – 3.3 mm. For the VISTA sequence the TR = 2034 ms and TE = 120 ms, with isotropic voxels of 0.8 mm width. Functional sequences were not part of the RT clinical workup.

2.2.3 ROI segmentation

Prostate delineations were performed for all patients with an atlas-based approach using a research software version of ADMIRE 1.13.5 (Elekta AB, Stockholm, Sweden), with visual verification by the researcher (four years of experience in prostate delineation) and manual correction whenever necessary.

Based on the prostate delineation intra-observer variability reported by Nyholm *et al.* [66], prostate ROIs were created by expanding the delineation by 2mm to compensate for possible delineation uncertainties and to ensure whole-prostate coverage. The margin ROIs were defined as the region between an expansion and shrinkage of the prostate delineation by 2mm.

2.2.4 Image feature extraction

The pyradiomics 1.2.0 toolbox [67] was used for region-wise 3D feature extraction. For a consistent calculation of 3D features, all images were resampled to an isotropic grid of $2 \times 2 \times 2$ mm voxels using BSplines to avoid extreme image oversampling in the slice direction. Images were then normalised as described in Appendix A. The normalised images, as well as images obtained by further filtering with a Laplacian of Gaussian (LoG) with $\sigma = 1, 3,$ and 5 mm, were used as input for feature extraction. The extracted features were categorised in shape, intensity and texture. From the LoG filtered images only intensity and texture features were extracted.

Image discretization was performed by using a fixed bin width = 5. Rotational invariant textural features derived from the grey-level co-occurrence (GLCM) and run length (GLRLM) matrices were computed by averaging the values obtained over 13 angles ($0, 45, 90$ and 135° symmetrical angles in-plane and out-of-plane) using a displacement vector of one voxel. These features quantify regional heterogeneity.

A total of 254 region-level features were obtained per patient (Table 2.1 and Appendix B), all scaled as described in Appendix A.

2.2.5 Models and validation

Independent models were created using either clinical or imaging features. Separate imaging models were generated for each ROI. The clinical model was developed using PSA, Gleason and clinical stage variables.

All models were independently validated using stratified 10-fold cross-validation (CV), ensuring the folds preserve the percentage of samples for each class. Receiver operating curve (ROC) analysis with the use of the area under the curve (AUC) values per fold was applied to assess the prediction accuracy for the different ROIs.

Clinical and imaging models were created based on different features, potentially offering complementary information about the pattern to be classified.

Combining the results of the two to generate a consensus decision may improve efficiency and accuracy, as the sets of patterns misclassified by the different models would not necessarily overlap [68]. To evaluate this hypothesis, the posterior probabilities obtained for each patient for imaging and clinical models were averaged in one joint probability, and the performance of the combined models was assessed.

2.2.6 Feature ranking and selection

Due to the high dimensionality of the feature set, feature selection was implemented to address the curse of dimensionality [69]. Within the stratified 10-fold CV scheme we aimed at identifying a model hyperparameter - the number of features to select (nFeats). Firstly, feature ranking was performed in the training fold using the minimum-redundancy maximum-relevance (mRMR) algorithm [70]. This method maximises the dissimilarity and minimises the redundancy between features. To identify the hyperparameter nFeats, an inner stratified 10-fold CV scheme was implemented within the training set. Data was thus further split into training and test folds, and different sizes of features sets were tested: nFeats = 3,5,10,20,30,40,50,60,70,80,90,100,150,200,250,254. The AUC performance of each feature set when used with a logistic regression (LR) classifier was recorded for each fold. The LR classifier was chosen for its simplicity. The AUC was averaged for all folds of each feature set, and the nFeats with the highest value was chosen as being optimal. This optimal number of features was then selected from the outer training fold data, and given to train and test the model classifiers.

Pearson correlation between the top five features for the whole dataset was also calculated.

Table 2.1: Description of the features extracted for each region of interest (ROI). The index ¹ and ² in homogeneity and informal measure of correlation refer to the two used formulations used to calculate these measures. Further information about these features can be found in the Appendix B.

| Feature class | Description | Features extracted |
|---------------|---|--|
| Shape | 3D shape features | Sphericity, maximum 3D diameter, volume, spherical disproportion, surface area, surface volume ratio |
| Intensity | 1st order statistics (2D and 3D) | Root mean squared, maximum, median, standard deviation, variance, 90% percentile, minimum, mean absolute deviation, kurtosis, mean, energy, interquartile range, range, 10% percentile, skewness, total energy, robust mean absolute deviation, entropy, uniformity |
| Texture | Grey-level co-occurrence matrix, GLCM (2D and 3D) | Entropy, cluster tendency, inverse difference moment, inverse difference moment normalised, maximum probability, correlation, sum variance, homogeneity ₁ , homogeneity ₂ , energy, dissimilarity, informal measure of correlation ₁ , informal measure of correlation ₂ , inverse difference, inverse difference normalised, contrast, average intensity, difference average, sum squares, cluster shade, sum entropy, difference entropy, inverse variance, cluster prominence, auto correlation, sum average, difference variance |
| | Grey level run length matrix, GLRLM | Short run emphasis (SRE), long run emphasis (LRE), grey-level non-uniformity (GLN), grey-level non-uniformity normalised (GLNN), run length non-uniformity (RLN), run length non-uniformity normalised (RLNN), run percentage (RP), run entropy (RE), low grey-level run emphasis (LGLRE), high grey-level run emphasis (HGLRE), short run low grey-level emphasis (SRLGLE), short run high grey-level emphasis (SRHGLE), long run low grey-level emphasis (LRLGLE), long run high grey-level emphasis (LRHGLE), grey-level variance (GLV), run length variance (RLV) |
| Filtered | Laplacian of Gaussian filter | Order = 1, Sigma = 1,3,5 |

2.2.7 Classifiers

Parmar *et al.* [69] reported the Random Forest (RF) classifier to have obtained the highest prognostic performance with high stability against data perturbations. However, LR is by far one of the most widely used classification algorithms. Both algorithms are simple and computationally efficient, so the performance of RF and LR classifiers was evaluated here.

The LR model was built using a l2-penalty, tolerance=0.0001 and C=1. The RF model was built using 45 trees and a minimum of two samples per split. As ours is an imbalanced dataset (i.e. more non-recurrent than recurrent samples) for both classifiers the class weight was set to 'balanced' so that the weights are adjusted to be inversely proportional to the class frequencies in the input data. To investigate the stability of the RF model, the classification process was repeated 20 times, and the standard deviation (SD) of the obtained AUC values was calculated.

The stratified 10-fold CV was used to avoid overfitting by ensuring the test set was independent and not included in the feature selection process. A separate model was created for each of the ROIs and for both RF and LR classifiers, resulting in a total of 4 imaging models. The clinical model was developed using a LR classifier. A total of two combined models - using the predictions from RF and LR imaging classifiers combined with clinical predictions - were investigated per ROI.

Figure S2.1 (Supplementary Materials) illustrates the different steps in optimising feature selection, in classification and cross-validation.

2.2.8 Statistics

To test for significant differences between clinical features of recurrent and non-recurrent patients, a T-test (for continuous variables) and a chi-square test (for categorical variables) was used. The patient's clinical stage is decided in consensus between the urologist and radiologist, and is therefore potentially biased by the MRI findings. A bivariate Pearson correlation was calculated between the TNM stage and all MRI extracted features. The degree of correlation can

help reason as to whether the clinical TNM stage provides similar information as the imaging features; if so, the two models cannot be considered to provide independent predictions. The same correlation measure was performed between Gleason score and the MRI extracted features.

2.3 RESULTS

Patient characteristics are reported in Table 2.2. The follow-up time was five years or until biochemical recurrence. Median time to BCR was four years (range 1-5 years). Three recurrences were local [10%], six were locoregional [19%] (local with involved lymph nodes), two were regional [7%] (only lymph nodes), three were regional-distant [10%] and 11 were distant (typically bone metastasis) [35%]. For six [19%] recurrent patients the location was unreported. Table S2.1 (Supplementary Materials) reports the distribution of recurrence location according to T stage, Gleason score and PSA level. Recurrence location was determined by either Choline or ^{68}Ga -PSMA PET, SPECT/CT and at times MRI and/or biopsy.

Figure 2.1 illustrates the segmented ROIs from which imaging features were extracted. The original, as well as the resampled and filtered images provided for feature extraction are presented in Figure 2.2 for a representative patient.

No significant differences were found between the clinical variables of recurrent and non-recurrent patients (p-values were 0.40 for PSA, 0.14 for Gleason and 0.62 for TNM stage) highlighting the clinical homogeneity of the high-risk cohort. TNM stage and Gleason score were weakly correlated (maximum absolute value of 0.30 and 0.35 respectively) with the MRI extracted features of any of the ROIs. Thus, the clinical model was considered to provide an independent prediction from the imaging models.

Table 2.3 shows the AUC values for the different classifiers; for RF, the value is averaged over the folds and for 20 runs with the corresponding SD between different runs. For a single run, using whole-prostate imaging features, the average AUC (SD between folds) for 10 folds was for the RF classifier 0.56

Table 2.2: Patient characteristics.

*The numbers in brackets are the percentages rounded down to the nearest integer.
IQR – Interquartile range.*

| | Recurrent | Non-recurrent |
|--|------------------|----------------------|
| Number of patients (%) | 31 (26%) | 89 (74%) |
| Median pre-treatment PSA (ng/ml) [IQR] | 17 [25] | 15 [29] |
| PSA \leq 10 | 9 (29%) | 31 (35%) |
| 10 < PSA \leq 20 | 7 (23%) | 17 (19%) |
| PSA \geq 20 | 15 (48%) | 41 (46%) |
| Clinical tumour stage | | |
| T1 | 2 (6%) | 10 (11%) |
| T2 | 9 (29%) | 20 (22%) |
| T3 + T4 | 20 (65%) | 59 (66%) |
| Primary Gleason grade | | |
| Gleason 5 - 6 | 6 (19%) | 20 (22%) |
| Gleason 7 | 9 (29%) | 31 (35%) |
| Gleason 8 | 9 (29%) | 28 (31%) |
| Gleason 9 - 10 | 7 (23%) | 10 (11%) |

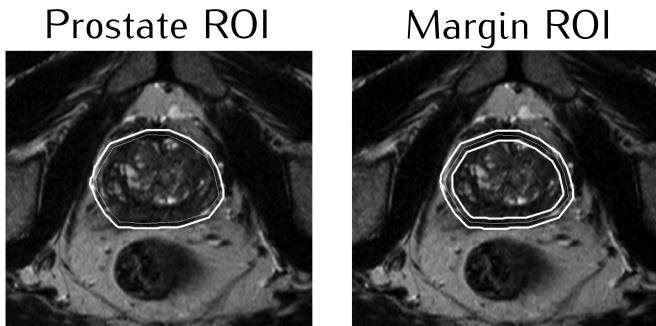


Figure 2.1: The different regions of interest (ROIs) used for feature extraction. The ROIs were created by expanding (and in the case of the margin also shrinking) the original delineations (thin lines) to obtain the final ROIs used for feature extraction (solid lines).

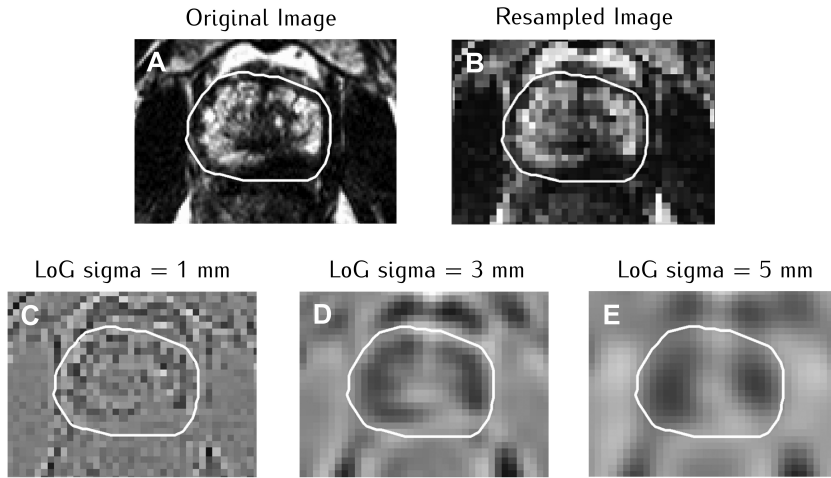


Figure 2.2: A. Original T2w image; B. Normalised and resampled image in the grid of $2 \times 2 \times 2 \text{ mm}^3$; C – E Normalised and resampled images filtered with a Laplacian of Gaussian (LoG) with sigmas = 1,3 and 5 mm. The resampled image as well as the filtered images were used as input for feature extraction. The white contour represents the prostate ROI.

(0.21) and 0.63 (0.18) for the LR. Margin imaging features obtained for a single run of 10 folds an average AUC (SD between folds) of 0.57 (0.20) for the RF classifier and 0.59 (0.21) for the LR.

The clinical model had a poor performance with an AUC (SD between folds) = 0.51 (0.18). The best performance with imaging features was obtained for the whole-prostate region with a LR classifier. The highest AUC for the combination of imaging and clinical features was of 0.58, using margin imaging features.

The most prevalent optimal number of features was three for the margin, and either three, ten or 20 for the prostate (all were chosen by three folds each). There was inter-fold variability regarding the optimal number of features. For the prostate, the optimal number of features varied between three [30%], 10 [30%], 20 [30%] and 90 [10%]. For the margin the values were of 3 [60%], 5 [20%], 30 [10%] and 60 [10%].

Table 2.3: AUC values obtained with the different feature selection methods and classifiers. Numbers in brackets show the standard deviation for average AUC for all folds between different rounds when using a random forest classifier.

AUC values (SD)

| | | Clinical | | | |
|----------|--|-----------------|---------------------------|-------------|-----------|
| | | 0.51 | | | |
| | | Imaging | Imaging + Clinical | | |
| ROI | | mRMR + RF | mRMR + LR | mRMR + RF | mRMR + LR |
| Prostate | | 0.55 (0.03) | 0.63 | 0.54 (0.02) | 0.56 |
| Margin | | 0.56 (0.02) | 0.59 | 0.58 (0.02) | 0.54 |

Table 2.4: Feature ranking for the different ROIs obtained using the mRMR method for the whole dataset.

| Prostate | | Margin |
|-----------------|--|---|
| #1 | LoG sigma 3 GLRLM LGLRE | No filter image first order 10th percentile |
| #2 | No filter image first order minimum | No filter image GLCM cluster shade |
| #3 | LoG sigma 5 GLCM inverse difference normalised | No filter image shape surface area |
| #4 | LoG sigma 5 GLCM cluster prominence | LoG sigma 5 first order maximum |
| #5 | LoG sigma 5 first order mean | LoG sigma 3 GLCM difference variance |

Names and description of the highest ranking features for the two ROIs, when ranked in the whole dataset, can be found in **Table 2.4**. The majority of the top 5 prostate features originated from the filtered images and for the margin from the no filter image. The LoG filter enhanced boundaries and overall changes in intensities. First order statistics were mostly associated with extreme values (e.g. minimum, maximum). Textural based features were related to homogeneity (e.g. GLCM inverse difference moment normalised), heterogeneity (e.g. GLCM difference variance and GLCM dissimilarity) as well as skewness, asymmetry and uniformity of the GLCM (e.g. GLCM cluster prominence and GLCM cluster shade). The highest Pearson correlation for the top five features was 0.20 for the prostate and 0.51 for the margin.

2.4 DISCUSSION

In this study we aimed at identifying high-risk PCa patients who are at a higher-risk of BCR up to five years after EBRT. The heterogeneous outcome reported in this homogeneously treated high-risk cohort raises the question of potential treatment intensification for a subgroup of very high-risk patients. A LR model with clinical features resulted in a poor performance predicting BCR. This is not surprising as no significant differences were found between clinical features of recurrent and non-recurrent patients, confirming the clinical homogeneity of the cohort and highlighting the difficulty in discriminating patients solely based on clinical information. Similar findings have been reported by Hegde *et al.* [71], stressing the importance of imaging in such a high-risk cohort. Literature reported values for the use of the Kattan nomogram are of AUC = 0.61 [72] and 0.58 [62], higher than our clinical model performance. The Kattan nomogram was originally developed using a combined population of intermediate and high-risk patients and the literature values reported above are obtained when applied in similarly mixed cohorts. We did not use the original Kattan nomogram but instead a model incorporating the same features and trained on our own clinically homogeneous cohort. Despite the differences in methodology and cohort characteristics, our findings are in line with the published literature.

Whole-prostate pre-treatment MRI radiomic features obtained an AUC of 0.63, outperforming standard clinical features in recurrence prediction. Several studies describe the association between tumour adjacent stroma and prostate microenvironment to relapse and disease progression [73, 74]. MRI is known to have limited accuracy in the detection of small tumour foci of less than 0.5 cm³ [49]. Thus it is impossible to rule out the presence of satellite lesions, not visible on MRI and missed by biopsy sampling, in the remaining prostatic region. Analysis of the prostate as a whole is less time consuming and takes into account all available information. The overall findings support the idea that relevant information can be found on a whole-prostate level as well as the potential of this region for BCR prediction. Due to the cohort's clinical homogeneity, combining clinical with imaging features did not improve performance and introduced noise.

Despite the AUC values being relatively low to extrapolate significant clinical decisions, these results offer a proof of concept of the potential of radiological images in the context of precision medicine. Published literature on the use of radiomics for outcome prediction reported similar AUC values [69, 75]. For our cohort with five years follow up, only T2w anatomical scans were available for all patients, whereas to date mp-MRI is considered standard of care for diagnostics and treatment. Functional parameters extracted from DCE-MRI and DWI have been found to be predictive of response in other cancer sites [76, 77] and in pre-clinical studies [78]. In particular, the association between DWI-derived ADC maps and Gleason score has been extensively reported [79, 80]. The inclusion of functional imaging can potentially enhance the performance of biochemical recurrence prediction models.

The MRI protocol underwent slight changes during the period in which this patient cohort was treated. The influence of different MRI scanners, parameters and setup on the extracted radiomic features is still under-investigated. Standardizing the extraction and use of radiomic features as well as evaluating the repeatability of MR-based radiomic features are important subjects. Various reviews [81, 82] highlight important topics to be addressed in designing future radiomics studies. To tackle the curse of dimensionality we use a feature selection method, as commonly done in the radiomics field. Optimising an RF classifier would be an alternative to regularise the model and address this issue.

Lastly, the results obtained in this study require external validation in similar cohorts, with our findings suggesting the use of a small number of features. Nonetheless these are encouraging findings as they provide pilot evidence of the relevance of imaging in outcome stratification of clinically homogeneous patients. The use of whole-prostate imaging characteristics to obtain information about five year biochemical recurrence risk can potentially be used to develop individualised treatment strategies.

APPENDICES

2.5 APPENDIX A

2.5.1 Image Normalisation

Images were normalised according to:

$$f(x) = \frac{s(x - \mu_x)}{\sigma_x}$$

where x and $f(x)$ are the original and normalised intensities, μ_x and σ_x are the ROI mean and standard deviation of the intensity values, and s is a scaling value here set to 100. Intensity values outside three standard deviations from the mean were considered outliers and set to $\mu + 3\sigma$ or $\mu - 3\sigma$ according to their location in the distribution.

2.5.2 Feature scaling

All imaging features were scaled by subtracting the median and scaling the data according to the interquartile range. This method provides increased robustness against outliers which can impact the estimation of the mean and variance used by typical scalers, and is implemented as a 'robust scaler' part of scikit-learn – a machine learning toolbox for Python.

2.6 APPENDIX B

The textural features homogeneity and informal measure of correlation were calculated using two different formulations:

$$\text{homogeneity}_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + |i - j|}$$

$$\text{homogeneity}_2 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + |i - j|^2}$$

$$\text{informal measure of correlation}_1 = \frac{HXY - HXY_1}{\max(HX, HY)}$$

$$\text{informal measure of correlation}_2 = \sqrt{1 - e^{-2(HXY_2 - HXY)}}$$

Where N_g is the number of discrete intensity levels in the image and therefore the size of the GLCM matrix; i and j are the elements from the matrix; p is the second-order joint probability function of an image region constrained by the mask; p_x is the marginal row probabilities; p_y is the marginal column probabilities; HX and HY are the entropy of p_x and p_y ; HXY is the entropy of $p(i,j)$ and

$$HXY_1 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log_2(p_x(i)p_y(j) + \varepsilon)$$

$$HXY_2 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log_2(p_x(i)p_y(j) + \varepsilon)$$

being ε an arbitrarily small non-negative number added to prevent $\log(0)$ which is undefined (-infinity).

2.7 APPENDIX C

2.7.1 Programs and settings

All programming and analysis were performed in Python 3.5 and MATLAB 2017a. The simple Insight Segmentation and Registration Toolkit was used to resample the images prior to feature extraction. An online available MATLAB implementation of the mRMR method by Peng *et al.* [83] was used. Feature scaling and the RF and LR classifiers were used as implemented in the scikit-learn 0.18.1 package for Python [84]. Statistics were performed using IBM SPSS Statistics 22.

SUPPLEMENTARY MATERIALS

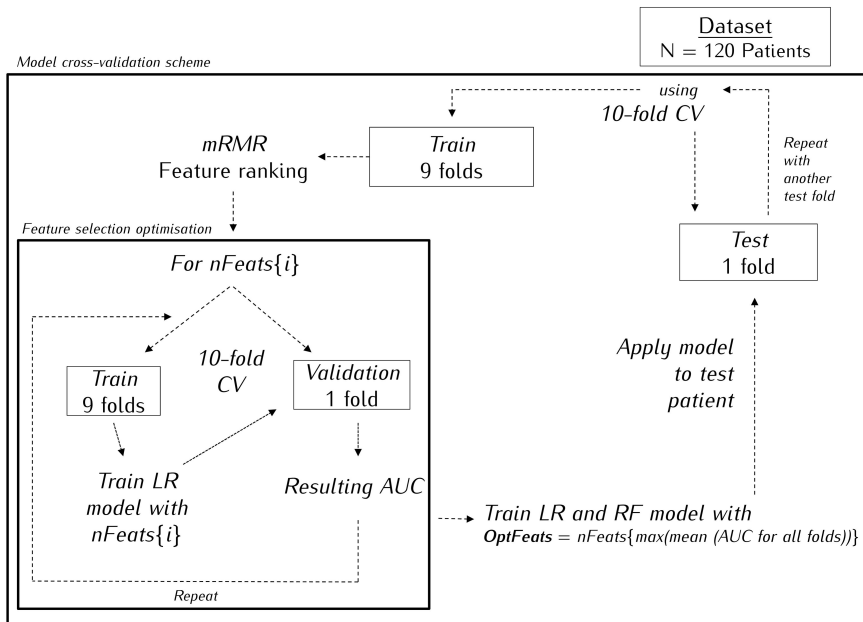


Figure S2.1: Feature selection, classification and cross-validation pipeline. The stratified 10-fold cross validation (CV) separates the patients in 10 folds, and then iterates the use of 9 folds for training and the remaining fold as test set. Feature selection is performed using stratified 10-fold CV on the training set where different subsets of features are tested. Variable $i = 1, 2, \dots, M$, where M equals the total number of $n\text{Feats}$ combinations to test. The number of features resulting in the maximum AUC value in the training set is used to train an LR and RF classifiers which are evaluated on the remaining independent test fold. This process is repeated 10 times.

Table S2.1: Recurrence location according to primary clinical tumour stage, Gleason grade and pre-treatment PSA values.

| | Recurrence location | | | | |
|----------------------------------|----------------------------|--------------|----------|------------------|---------|
| | Local | Locoregional | Regional | Regional-distant | Distant |
| Clinical T stage | | | | | |
| T1 | - | - | - | - | 1 |
| T2 | - | 3 | - | - | 4 |
| T3 + T4 | 3 | 3 | 2 | 3 | 6 |
| Gleason grade | | | | | |
| 6 | 1 | 1 | - | 1 | 2 |
| 7 | - | 1 | - | 2 | 4 |
| 8 | 1 | 3 | 2 | - | 1 |
| 9 + 10 | 1 | 1 | - | - | 4 |
| Pre-treatment PSA (ng/ml) | | | | | |
| PSA ≤ 10 | 1 | 2 | - | - | 5 |
| 10 < PSA ≤ 20 | 1 | - | - | 1 | 3 |
| PSA > 20 | 3 | 4 | 2 | 2 | 3 |

PROSTATE FIDUCIAL MARKER DETECTION
WITH THE USE OF MULTI-PARAMETRIC
MAGNETIC RESONANCE IMAGING

Catarina Dinis Fernandes¹
Cuong V Dinh¹
Marcel J Steggerda¹
Leon C ter Beek²
Milena Smolic¹
Laurens D van Buuren¹
Floris J Pos¹
Uulke A van der Heide¹

¹Department of Radiation Oncology and ²Radiology,
The Netherlands Cancer Institute, Amsterdam, The Netherlands

ABSTRACT

Background and purpose: The introduction of a magnetic resonance (MR)-only workflow in radiotherapy requires that fiducial markers, used for position verification, can be detected on MR images. Here we evaluate a model for marker detection in prostate cancer patients by combining information from our hospital standard multi-parametric (mp-) MRI protocol (T1-weighted - T1w, T2-weighted - T2w, B₀) with dedicated sequences (balanced steady-state free precession sequence - bTFE, susceptibility weighted imaging - SWI).

Materials and methods: Thirty two patients scheduled for external-beam radiotherapy received a mp-MRI and computed-tomography; the latter was used as ground truth location of the markers. A logistic regression model was implemented for marker detection by combining features from all imaging sequences. The performance of the individual and combined sequences was assessed by determining true and false positive detections.

Results: The combination of different sequences (mp-MRI) resulted in a better performance than the best imaging sequence alone (bTFE). Combining mp-MRI + bTFE resulted in good accuracy and a true positive detection rate of 0.94.

Conclusions: The standard mp-MRI provides valuable information to detect fiducial markers. The combination of different sequences outperforms the use of a single dedicated sequence. We recommend the addition of a bTFE to the standard mp-MRI protocol to improve fiducial marker detection.

3.1 INTRODUCTION

Prostate cancer patients scheduled to receive external-beam radiotherapy (EBRT) are commonly implanted with fiducial markers used for treatment position verification. Currently, the markers are identified using computed-tomography (CT) whereas structures are delineated on magnetic resonance imaging (MRI), requiring registration of MRI to CT. The registration procedure is associated with errors which will propagate throughout the whole treatment chain. The introduction of a magnetic resonance (MR)-only workflow is a possible way to avoid this. However, a precondition is the ability to identify the markers on MRI.

The most commonly used markers are made of gold, usually presenting as a local signal void on MR images. Several studies have investigated the accuracy of specifically optimised MRI sequences for the detection of implantable metallic structures such as fiducial markers. Some suggest the use of optimised spin-echo MRI sequences, improving void visibility [85], or gradient-echo sequences [86, 87] sensitive to T_2^* decay, using the susceptibility effects to get better marker depiction [88], as well as by combining it with contrast agents [89]. Others visually combine different MRI sequences [90, 91], while some groups have developed sequences that allow for positive contrast at the marker [92–95].

Typically, patients receive a multi-parametric (mp-) MRI used for radiotherapy (RT) target delineation as part of their standard clinical treatment. In our hospital this protocol includes anatomical (T1- and T2- weighted) and functional (diffusion-weighted (DWI) and dynamic contrast enhanced (DCE)) sequences as well as a B_0 map acquired for the post-processing of DWI-MRI. Sequence specific parameters result in distinctive marker voids between images, increasing the complexity of reproducible and accurate manual marker localisation. We propose the development of a marker detection model using a machine learning framework incorporating information from the clinical mp-MRI protocol. To investigate the performance of some of the most often used dedicated sequences we have incorporated a balanced steady-state free precession sequence (balanced Turbo-Field Echo or bTFE) [86, 87] and a susceptibility weighted imaging (SWI) sequence in the scanning protocol. Our goal is to in-

investigate the performance of a model based on the standard mp-MRI protocol and compare it to models based on a single, dedicated sequence.

3.2 MATERIALS AND METHODS

3.2.1 Dataset

Thirty two consecutive patients with biopsy proven prostate cancer, scheduled to receive external-beam radiotherapy (EBRT) were treated between October 2015 and January 2016. Up to three gold fiducial markers (RT Cast, 1.0×3 mm) were implanted per patient. Approximately 1 week after implantation, and on the same day, patients underwent a mp-MRI and a CT scan used for RT planning purposes.

The MRI data was acquired using a 3T Achieva dStream MRI scanner (Philips Healthcare, Best, The Netherlands) with the use of an anterior and posterior phased array coil. The mp-MRI protocol included transversal T1-weighted (T1w) T1 THRIVE and T2-weighted (T2w) T2 turbo spin-echo (TSE) images, B_0 map, and the dedicated sequences bTFE and SWI (details of the protocol can be found in [Table S3.1](#), Supplementary Materials). Coronal and sagittal T2w scans as well as functional sequences were also acquired as part of the protocol but were not used for further analysis. The CT scans were acquired using a 24-slice CT scanner (Somatom-Sensation-Open, Siemens), with image resolution $0.9 \times 0.9 \times 3$ mm³ (x, y, z).

A conservative threshold of 1900 Hounsfield units, resilient to possible streaking artefacts, was used to segment the markers on the CT images. The ground truth (GT) marker coordinates were automatically calculated as the centre of mass of the clusters resulting from the segmented CT image. All MR images were acquired within the same exam session. To account for possible within-session motion all images were rigidly registered to the T1w scan with the use of a clipbox around the prostate and periprostatic region [96]. Subsequently the T1w scan was registered to the CT using a similarly defined clipbox; the transformation was then applied to the other MRI sequences. This

was achieved by means of rigid-registration (mutual information). Registration accuracy was checked by visual inspection of the prostate boundaries and marker overlap and by assessing the registration error comparing the GT coordinates with a manually defined centre of mass of the marker artefact in the bTFE sequence (CM_{mbTFE}). The difference between GT and CM_{mbTFE} coordinates was used as a measure for the target registration error (TRE) between CT and MR.

The prostate gland was delineated using the T2w sequence. The ROI used for model analysis was defined as a 4 pixels (3.6 mm) expansion of the prostate delineation (**Figure 3.1A-F**). Within the ROI the signal intensities (SI) for each image (T1w, T2w, B_0 magnitude, bTFE and SWI images) were normalised to values between 0 and 1 using Min-Max scaling, ensuring the values were comparable between patients. All images were resampled to the grid of voxel size $1 \times 1 \times 1 \text{ mm}^3$.

3.2.2 MRI features and feature extraction

The markers present as signal voids in most MR images and their apparent position depends on their shape and orientation relative to the magnetic field. With exception of T2w, the artefact they induce is usually much larger than the actual physical size and often exhibits a blob-like shape in T1w, bTFE, B_0 magnitude and SWI images – **Figure 3.1A-F**. This same artefact also expands through slices in a cylindrical line-like pattern. In this study we have included multi-scale blobness [97] and line filtering [98, 99] of the prostate region as well as intensity based local statistical features such as signal intensity in the voxel i and the mean, median, minimum, maximum and standard deviation values from a local window ($3 \times 3 \times 9$ voxels in x, y, z, directions, where z is along the B_0) surrounding this voxel.

Each voxel was described by 42 features in total (6 features from T2w and 9 features from each T1w, B_0 magnitude, bTFE and SWI). **Table 3.1** summarises all extracted features with detailed description. Single sequence models include all features extracted for that specific imaging sequence.

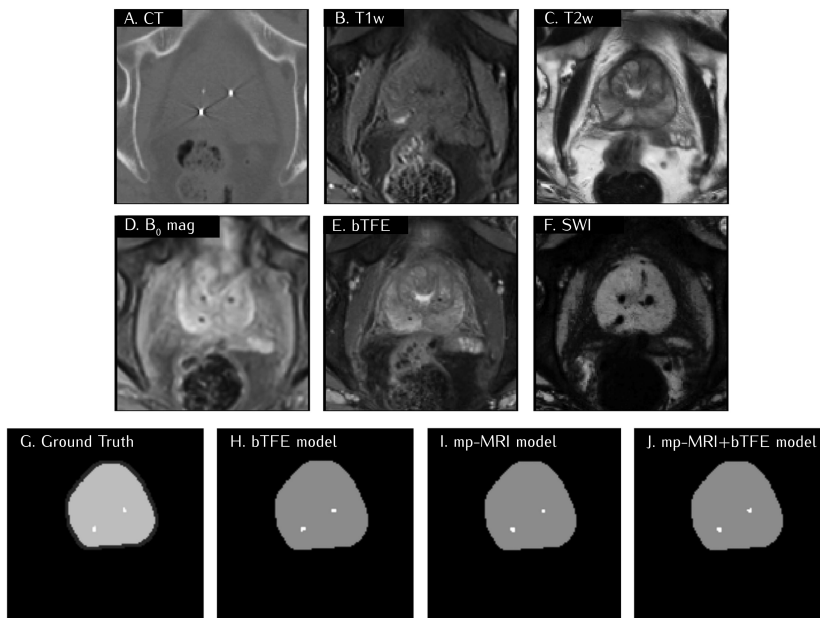


Figure 3.1: Model pipeline in one representative slice. (A-F) CT and all MRI sequences. The ground truth map (G) was constructed based on the information from the CT scan and delineations. (H-J) Different model outcomes; all models correctly identify the two markers present on this slice. Model results are prior to cluster selection.

Performance was investigated for models created using four sequence combination options: 1) mp-MRI only (T1w, T2w and B_0 magnitude); 2) mp-MRI + bTFE; 3) mp-MRI + SWI and 4) mp-MRI + bTFE + SWI. Models for each individual imaging sequence were also created in order to ascertain their individual value.

3.2.3 Model creation

Ground truth maps (GT_{maps}) were created based on the GT coordinates, incorporating the prostate and margin delineations; these were used to train and validate the model results.

Table 3.1: Features and what they capture.

| Category | Feature ID | Feature name | MRI sequence | Feature settings | Captures |
|----------------------------|-----------------------------------|--|--------------------|--|---------------------------------------|
| Intensity-based statistics | f ₁ – f ₆ | Signal intensity, mean, median, min, max and standard deviation values for each imaging sequence | T _{1w} | Kernel = 3 × 3 × 9 vox | Marker-induced susceptibility effects |
| | f ₁₀ – f ₁₅ | | T _{2w} | | |
| | f ₁₆ – f ₂₁ | | B ₀ mag | | |
| | f ₂₅ – f ₃₀ | | bTFE | | |
| | f ₃₄ – f ₃₉ | | SWI | | |
| Structure | f ₇ , f ₈ | Multi-scale blobness | T _{1w} | (in mm) | Focality and compactness |
| | f ₂₂ , f ₂₃ | | B ₀ mag | $\sigma = 0.9, 1.3, 1.7$ | |
| | f ₃₁ , f ₃₂ | | bTFE | $\sigma = 1, 1.5, 2$ | |
| | f ₄₀ , f ₄₁ | | SWI | $\sigma = 0.5, 1, 1.5$ $\sigma = 3, 3.5, 4$ | |
| | f ₉ | Multi-scale line | T _{1w} | (in mm) | Curvilinear structures |
| | f ₂₄ | | B ₀ mag | $\sigma = 0.9, 1.3, 1.7$ | |
| | f ₃₃ | | bTFE | $\sigma = 1, 1.5, 2$ | |
| | f ₄₂ | | SWI | $\sigma = 0.5, 1, 1.5$ $\sigma = 3, 3.5, 4$ | |

Data analysis was performed using MATLAB R2015a (MathWorks, Natick, MA, USA) and the PRTools toolbox [100]. PRTools is a pattern recognition toolbox that among others supports feature extraction, the use of classifiers and evaluation methods. First, we determined from the GT_{maps} the value of M , representing the average number of voxels labelled as markers. Then, for each of the individual sequences and for the four sequence combinations a model was trained by fitting a logistic regression to the labelled (marker or non-marker) training data. The logistic regression model, applied to an unseen dataset, results in a map of the probability for each voxel of being a marker. The M highest probability voxels were grouped into clusters. We defined the probability of each cluster candidate as the highest probability value of all voxels within it. Prior information regarding the number of markers implanted per patient was used to further post-process the results by selecting the $n(j)$

highest probability clusters, where $n(j)$ is the number of markers implanted in patient j . Other than the logistic regression, the process did not involve any optimization of parameters or feature selection.

Each model was evaluated using a leave-one-dataset-out cross-validation method to make sure that the training set – used to build the model – and test set are independent from each other. The different (individual and combination) models were trained and the rate of true positives (TP) and false positives (FP) was used to assess the performance of the model for each patient. Clusters were classified as a FP if the distance between the GT coordinates and the centre of mass of the clusters identified by the model (CM_{Model}) was larger than 3mm in plane or through slice. Model accuracy was assessed for each marker by calculating the distance between the GT coordinates and the CM_{Model} for all TP.

3.2.4 Statistics

A Shapiro-Wilk test was used to test the normality of the TRE between MRI and CT coordinates. For each model we determined the numbers of TP and FP per patient and combined them into distributions over the population. A non-parametric Wilcoxon signed-rank test was used to evaluate whether the difference in distributions between models was statistically significant ($p \leq 0.05$). Statistical analysis was performed using IBM SPSS Statistics 22.

3.3 RESULTS

For two patients air was present in the rectum which resulted in signal drop-off around this boundary in the bTFE sequence, complicating the detection of two markers (one in each patient) implanted in this region – **Figure 3.2**. Thirty patients were implanted with three markers; in the remaining two only two fiducial markers were implanted.

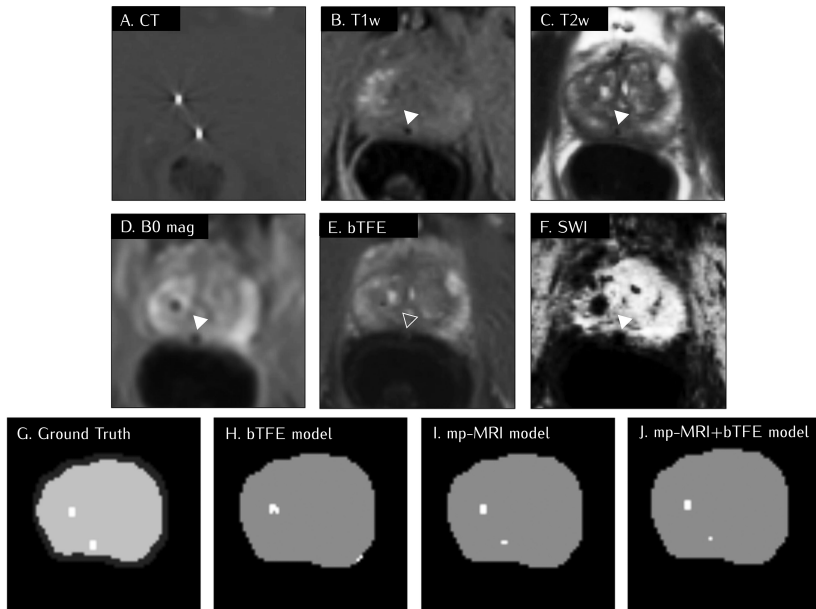


Figure 3.2: Model pipeline in one representative slice. (A-F) CT and all MRI sequences. The ground truth map (G) was constructed based on the information from the CT scan and delineations. (H-J) Different model outcomes. One of the markers is visible on T1w, T2w, B₀ mag and marginally visible on SWI (filled arrow head) but not on the bTFE sequence (open arrow head). This marker could not be identified on the bTFE model; the mp-MRI and mp-MRI + bTFE models successfully detect the two markers. Model results are prior to cluster selection.

Within-session motion between the different MRI sequences was also quantified: median of translations in all images relative to T1w, averaged for the population, were equal or less than 0.3 mm in all directions (interquartile range [IQR]: 0.9–1.4); average median rotation (in degrees) was 0 in all directions (IQR: 0.4–0.7).

Two markers were not visible on the bTFE sequence and were therefore excluded from the comparison between GT and CM_{mbTFE} coordinates. The TRE was not normally distributed, with median values of 0 mm for all directions and $IQR_{x,y} = 1\text{mm}$, and $IQR_z = 2\text{mm}$ – Figure 3.3A.

3.3.1 Evaluation of model performance

Figure 3.1 (A-F) displays the acquired CT and MRI sequences for an example patient and slice, as well as the respective GT_{maps} – Figure 3.1G – and model results – Figure 3.1 (H-J).

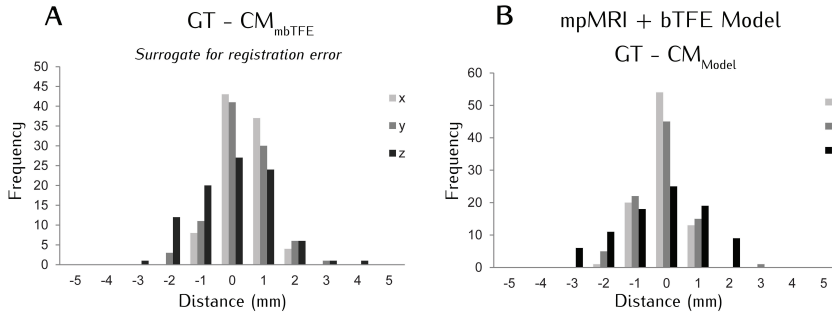


Figure 3.3: A) Histogram of the registration errors (mm) between CT and MRI in x , y and z directions; B) Accuracy of the mp-MRI + bTFE model in three spatial directions (x , y , z). GT – ground truth marker location based on segmented CT image; CM – center of mass of the cluster identified by the model. The plots represent the differences between the two coordinates for each separate marker.

Based on the GT_{maps} , each patient had on average $M = 103$ ($SD = 11$) voxels labeled as markers. Therefore clusters were formed by selecting the $M = 100$ highest probability voxels. In our dataset there were in total 94 implanted markers. Table 3.2 lists the number of TP and FP. Cluster selection based on the number of implanted markers per patient successfully reduced the number of FP for all models. The number of TP was also reduced but by a lesser extent than the FP. The number of false negatives (FN) can be inferred from the TP and FP: when $TP + FP = 94$, FN is equal to FP; if $TP + FP < 94$, $FN = 94 - TP$.

The mp-MRI model correctly detected a total of 83 markers ($FP = 10$) but with fewer positives than markers. The addition of the bTFE or SWI to the standard mp-MRI model improved performance, resulting in a total of 88 and 87 TP, respectively. The single imaging modality model with highest true positive rate (TP/P) was the bTFE ($TPR = 0.77$) – Table 3.2.

Table 3.2: Model results. In white the model results using individual sequences and in grey the results from the combined models.

† In total there are fewer positives (TP + FP) than markers

* Significantly different ($p \leq 0.05$) from the bTFE model using a Wilcoxon signed-rank test

| | Total TP | Total FP | Total TP selected | Total FP selected | True positive rate (TPR) | Model accuracy (mm) GT- CM_{Model} |
|---------------------|----------|----------|-------------------|-------------------|--------------------------|--------------------------------------|
| T1- THRIVE | 74 | 649 | 36 | 58 | 0.38 | 1.8 |
| T2 TSE | 41 | 1163 | 7 | 87 | 0.07 | 1.7 |
| B ₀ mag | 76 | 33 | 70† | 20 | 0.74 | 1.9 |
| bTFE | 88 | 340 | 72 | 22 | 0.77 | 1.7 |
| SWI | 51 | 51 | 48† | 25 | 0.51 | 2.0 |
| mp-MRI | 89 | 39 | 83† | 10 | 0.88 | 1.6 |
| mp-MRI + bTFE | 93 | 32 | 88* | 6* | 0.94 | 1.6 |
| mp-MRI + SWI | 88 | 14 | 87* | 6* | 0.93 | 1.7 |
| mp-MRI + bTFE + SWI | 92 | 8 | 89* | 4* | 0.95 | 1.6 |

The combination of mp-MRI with a dedicated sequence improved marker detection. The combination mp-MRI+bTFE resulted in a significantly better performance ($p = 0.005$) than using the bTFE alone. The mp-MRI + bTFE model accurately detected all markers for 26 out of 32 patients (81%), the mp-MRI in 75% and the bTFE in only 47% of the patients.

Model accuracy was described in terms of the Euclidean distance between the GT coordinates and the centre of mass of correctly identified markers CM_{Model} – **Table 3.2**. **Figure 3.3B** illustrates the distance distributions for the markers detected by the mp-MRI+bTFE model, discriminated for the three directions, where $IQR_x = 0$, $IQR_y = 1$ and $IQR_z = 2$. Distances became smaller by taking the registration error into account - removing this error from the mp-MRI + bTFE model results decreased the standard deviation of the model accuracy distribution. The manually selected CM_{mbTFE} and the correctly identified marker location in the mp-MRI+bTFE model had mean deviations (mm) of -0.49 ± 0.56 (x), 0.46 ± 0.58 (y), -0.24 ± 1.10 (z), median of 0 and $IQR_{x,y} = 1$, $IQR_z = 2$.

3.4 DISCUSSION

Moving towards an MR-only environment will mitigate registration errors as well as contouring inconsistencies [101–105]. A requisite for the implementation of an MR-only environment in RT is the ability to identify prostate fiducial markers on MRI. Manual marker detection represents a significant challenge for accurate and reproducible marker centroid detection, is time consuming, affected by pixel-size, noise, and quality of software used for detection [106]. We investigated the performance of MRI-based fiducial marker detection models, using CT as the ground truth. Provided that a delineation of the prostate is available - this is routinely performed and increasingly automated - the models presented here are fully automatic.

In our study, the mp-MRI model performs significantly better than the best imaging sequence alone (bTFE). The bTFE sequence has characteristics which optimise its potential for manual marker detection such as high SNR and good contrast, the markers presenting as small signal voids. However these same characteristics also increase the complexity for automated methods, due to the high similarity between a marker void and noise. In our detection setting the bTFE sequence benefits from being combined with the standard mp-MRI.

Schieda *et al.* [88] described how multi-echo gradient recall sequences might be of particular importance in the detection of markers located outside of the prostate gland. We implemented a multi-echo SWI sequence which has shown limited performance when used individually for automatic marker detection. The combination mp-MRI+SWI had good performance but a decreased model accuracy when compared to other combined models. The large extent of the artefact and its possible asymmetries when compared to the marker physical structure possibly compromised accurate localisation of the marker centroid. Moreover the sequence's long scanning times render it unpractical for clinical implementation.

Although no statistically significant differences were found between the mp-MRI and mp-MRI + bTFE models, the latter improves marker detection in cases where the mp-MRI model has failed, decreasing also the total number of FP. The mp-MRI + bTFE model has originally identified 93 markers; however

some of them have been rejected in favor of confounders upon cluster selection. Structural changes such as calcifications – found in over 40% of prostate cancer patients [107] – can mimic the appearance of markers on MRI, resulting in misclassifications. Fiducial markers implanted close to the rectum boundary can be difficult to detect, specifically in the bTFE sequence in the presence of air pockets in the rectum; in these cases combining different sequences can increase the likelihood of detection. The inclusion of the bTFE sequence in the clinical mp-MRI protocol has a minimal impact on scanning time (roughly two extra minutes), enabling an accurate localisation of the marker centroid, which is important for patient positioning and accurate dose delivery [106]. The use of a high bandwidth (BW) sequence (Table S3.1) reduces possible geometric distortions.

Validation of the results is non-trivial. Our framework is initially based on a voxel-wise logistic regression model but the final results are obtained after cluster selection. The region effectively occupied by the markers is significantly smaller than the whole prostate. This disproportion between TP and FP versus the total number of non-marker voxels, makes calculating sensitivity, specificity and accuracy for the voxel-wise model uninformative in this setting. Thus, results such as TP, FP and TPR are reported on a cluster level.

Distances between manual MRI-based coordinates (CM_{mbTFE}) and mp-MRI + bTFE model results are very small and comparable to Ghose et al. [106]. The apparent and actual marker positions on MRI have been reported to deviate < 1mm at 1.5T [108]. MRI-CT registration errors are small (median = 0 mm) with IQR up to 2 mm, well within reported literature values in this setting [96, 109–111]. Although larger deviations from the GT can be explained by rigid registration errors, these do not fully describe the overall model inaccuracies. Asymmetries (in shape and extent) of the artefact on MRI may explain some of the discrepancies between CT and MRI. The MRI artefact is dependent on marker size and orientation relative to the magnetic field [112], as well as on sequence BW, with possible asymmetries arising in the frequency encoding direction [108] and phase encoding direction for echo-planar imaging sequences [113]. The 3 mm limit for TP acceptability was chosen to remove outliers but not to be so restrictive given the size of the artefact. Reported IMRT systematic positioning errors, obtained using 2D electronic portal imaging, are approximately 1 mm with random variations from 0.9 to 4 mm in the three

orthogonal directions [114–116]. Overall geometric accuracy of the system is approximately 2mm [117]. Average accuracy of the mp-MRI+bTFE model is 1.6 mm.

Tokuda *et al.* [118] described fully automatic segmentation of fiducial frames for needle placement in MRI-guided prostate biopsies, by using a 3D multiscale line filter, but the method was not applied for gold fiducial marker detection. Ghose *et al.* [106] have also used machine learning methods for automatic fiducial marker detection using MRI. The authors use two dedicated gradient echo sequences (T1 and T2*w) and a dataset of 15 patients to develop a template-matching approach for marker detection. The described framework correctly identified all markers in 67% of the patients, compared to 73% by manual detection. Our mp-MRI+bTFE model has detected all markers for 81% of the patients. Validation of the results in a totally independent cohort, possibly even on a different scanner and field strength, would be useful to understand to what extent the model is transferable.

Fully unsupervised automated patient set-up requires 100% marker detection efficiency. The results indicate that none of the models is yet capable of doing so. A semi-automatic approach might be a valid solution, as the mp-MRI + bTFE model, prior to cluster selection, identifies all but one of the markers. Manual intervention can be used to ensure that all three markers are correctly identified in the case of missed or incorrect detection.

We have here evaluated an automatic fiducial marker detection framework with the use of MRI images. The core images used in the build of this model are part of the clinical RT planning protocol; the bTFE sequence is available in the clinical setting for the majority of scanners. Further optimization of image pre-processing (e.g. interpolation), extracted features and selection methods, enhancing the differentiation between markers and biological phenomena, can improve model performance. The overall findings show promise in the application of automatic fiducial marker detection methods in an MR-only workflow.

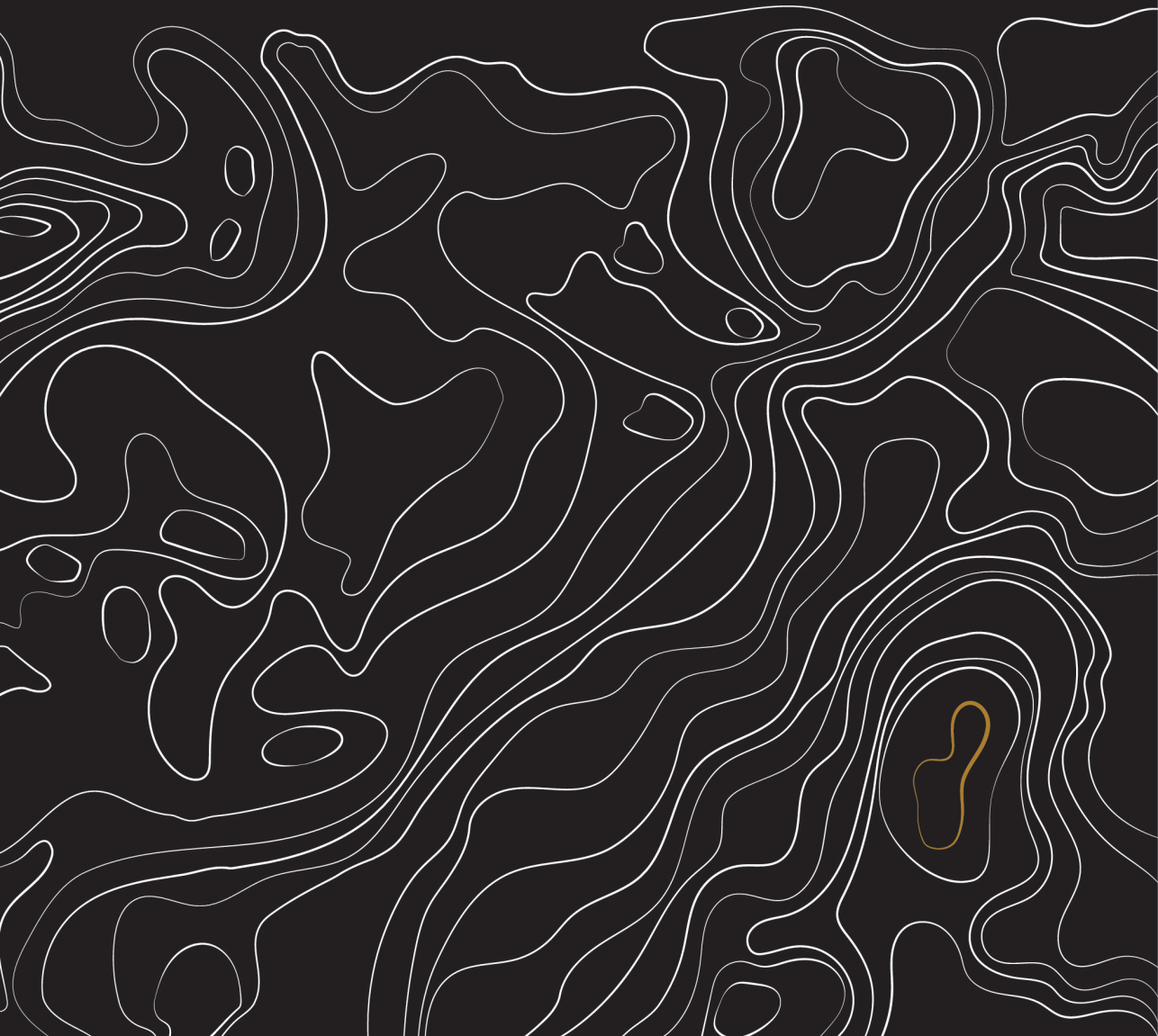
SUPPLEMENTARY MATERIALS

Table S3.1: Sequence parameters for the multi-parametric MRI prostate protocol performed with the anterior and posterior phased array coil at 3T.

| | Imaging plane | Scan mode | Field of view (mm) | Matrix size | Slice thickness / gap (mm) | Reconstructed voxel size (mm) | TR/TE (ms) | Echo train length | Flip angle | Acceleration factor | Receiver bandwidth (Hz/voxel) | Phase encoding direction | Acquisition time | Number of signal averages |
|-----------------------|---------------|-----------|--------------------|-------------|----------------------------|-------------------------------|------------------------|-------------------|------------|-----------------------------------|-------------------------------|--------------------------|------------------|---------------------------|
| T ₁ THRIVE | Axial | 3D | 280 x 400 | 280 x 276 | 1.0 / 0 | 0.85 | 3.6 / 1.81 | 88 | 10 | halfscan = 0.625 & SENSE factor 1 | 718.4 | RL | 1min 47sec | 3 |
| T ₂ TSE | Axial | 2D | 140 x 140 | 304 x 242 | 3.0 / 0 | 0.40 | 31.43 / 120 | 26 | 90 | N/A | 236.9 | RL | 3min 02sec | 3 |
| B ₀ | Axial | 3D | 360 x 316 | 164 x 144 | 3.0 / 0 | 1.09 | 5.9 / 2.3 | 40 | 21 | N/A | 2540.7 | RL | 46 sec | 2 |
| bTTE | Axial | 3D | 250 x 250 | 252 x 234 | 1.0 / 0 | 0.49 | 3.7 / 1.85 | 36 | 30 | N/A | 1240.1 | RL | 2min 23sec | 1 |
| SWI | Axial | 3D | 230 x 189 | 384 x 314 | 1.0 / 0 | 0.30 | 31 / 7.2; ΔTE = 6.2 | N/A | 17 | SENSE factor 2 | 255.3 | RL | 4min 29sec | 1 |

PART 2

Magnetic resonance imaging for the diagnosis and localisation of recurrent prostate cancer



QUANTITATIVE 3T MULTI-PARAMETRIC MRI
OF BENIGN AND MALIGNANT PROSTATIC
TISSUE IN PATIENTS WITH AND WITHOUT
LOCAL RECURRENT PROSTATE CANCER
AFTER EXTERNAL-BEAM RADIATION THERAPY

Catarina Dinis Fernandes¹

Petra J van Houdt¹

Stijn WTPJ Heijmink²

Iris Walraven¹

Rick Keesman¹

Milena Smolic¹

Ghazaleh Ghobadi¹

Henk G van der Poel³

Ivo G Schoots²

Floris J Pos¹

Uulke A van der Heide¹

¹Department of Radiation Oncology, ²Radiology and ³Urology
The Netherlands Cancer Institute, Amsterdam, The Netherlands

ABSTRACT

Background: Post-radiotherapy locally recurrent prostate cancer (PCa) patients are candidates for focal salvage treatment. Multi-parametric MRI (mp-MRI) is attractive for tumour localisation. However, radiotherapy-induced tissue changes complicate image interpretation. To develop focal salvage strategies, accurate tumour localisation and distinction from benign tissue is necessary.

Purpose: To quantitatively characterise radio-recurrent tumour and benign radiation-induced changes using mp-MRI, and investigate which sequences optimise the distinction between tumour and benign surroundings.

Study type: Prospective case-control.

Subjects: Thirty-three patients with biochemical failure after external-beam radiotherapy (cases), 35 patients without post-radiotherapy recurrent disease (controls), and 13 patients with primary PCa (untreated).

Field Strength/ Sequences: 3T; quantitative mp-MRI: T₂-mapping, ADC, and K^{trans} and k_{ep} maps.

Assessment: Quantitative image-analysis of prostatic regions, within and between cases, controls, and untreated patients.

Statistical Tests: Within-groups: non-parametric Friedman analysis of variance with post-hoc Wilcoxon signed-rank tests; between-groups: Mann-Whitney tests. All with Bonferroni corrections. Generalised linear mixed modelling to ascertain the contribution of each map and location to tumour likelihood.

Results: Benign imaging values were comparable between cases and controls ($P = 0.15$ for ADC in the central gland up to 0.91 for k_{ep} in the peripheral zone), both with similarly high peri-urethral K^{trans} and k_{ep} values (min⁻¹) (median [range]: K^{trans} = 0.22 [0.14–0.43] and 0.22 [0.14–0.36], $P = 0.60$, k_{ep} = 0.43 [0.24–0.57] and 0.48 [0.32–0.67], $P = 0.05$). After radiotherapy, benign central gland values were significantly decreased for all maps ($P \leq 0.001$) as

well as T_2 , K^{trans} , and k_{ep} of benign peripheral zone (all with $P \leq 0.002$). All imaging maps distinguished recurrent tumour from benign peripheral zone, but only ADC, K^{trans} , and k_{ep} were able to distinguish it from benign central gland. Recurrent tumour and peri-urethral K^{trans} values were not significantly different ($P = 0.81$), but k_{ep} values were ($P \leq 0.001$). Combining all quantitative maps and voxel location resulted in an optimal distinction between tumour and benign voxels.

Data Conclusion: Mp-MRI can distinguish recurrent tumour from benign tissue.

4.1 INTRODUCTION

One of the main treatment options for localised prostate cancer (PCa) is external-beam radiotherapy (EBRT). Depending on the risk-group, 5-year disease free survival varies from 67 to 80% for PCa patients who are treated with whole-gland radiotherapy (RT) to 78 Gy [58]. A small but significant proportion of patients failing the primary treatment, will harbour locally recurrent disease only [26, 27], for which focal treatment strategies, targeting the tumour region while sparing the surrounding uncompromised tissue, might offer a curative treatment option. To this end, multi-parametric magnetic resonance imaging (mp-MRI) with both anatomical and functional properties is attractive for loco-regional evaluation of recurrent PCa and for tumour boundary definition. However, RT-induced tissue changes pose a challenge to MRI interpretation. T2-weighted (T2w) MRI is the reference anatomical sequence, however it has limited performance in this setting as tumour conspicuity is decreased and zonal anatomy is often lost [37]. The use of functional sequences such as 1.5T MR spectroscopy [45] as well as diffusion weighted imaging (DWI) and dynamic contrast enhanced (DCE)-MRI [38–40] were found to surpass T2w-MRI in the detection of recurrent disease.

To successfully develop focal salvage strategies for recurrent PCa, an accurate detection and localisation of the tumour as well as distinction from radiation affected benign tissue is necessary.

The objectives of this study were to: 1) investigate RT-induced tissue changes and how these impacted the distinction of recurrent tumour from surrounding benign tissue; and, 2) use quantitative sequences to characterise both benign tissue and recurrent PCa after RT, and investigate which sequences resulted in optimal tumour localisation.

4.2 MATERIALS AND METHODS

4.2.1 Patient Recruitment and Matching

This prospective single-institution study was approved by the Institutional Review Board and took place between December 2015 and July 2018. All patients signed informed consent prior to participation in the study. An overview of the study design can be seen in **Figure 4.1**.

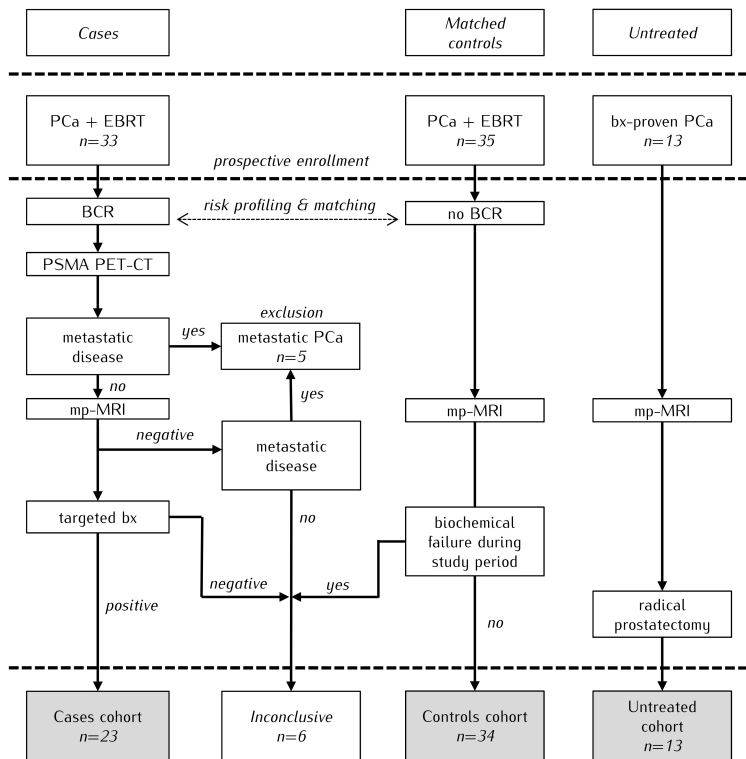


Figure 4.1: Study design with an overview of the three analysed cohorts. BCR, biochemical recurrence; bx, biopsy; EBRT, external-beam radiation therapy

Cases (patients with recurrent PCa)

Patients with a biochemical recurrence according to the Phoenix definition [65], were enrolled in the study. Patients could only be included more than 24 months after completion of EBRT and if they qualified for salvage treatment. A gallium-68 prostate-specific membrane antigen positron emission tomography (^{68}Ga -PSMA-PET) scan was acquired to exclude metastatic disease. Exclusion criteria were hormonal therapy in the past year, the use of anticoagulants that could not be stopped temporarily to acquire biopsies, contra-indications for an MRI exam and other treatments for cancer in the pelvis. Patients received a mp-MRI. When a tumour suspected region was identified, an MR-fused ultrasound-guided biopsy was performed targeting this region. If no tumour suspected region could be identified, either only prostate-specific antigen (PSA) follow-up would be continued or a standard systematic biopsy could be performed. This decision was made by the treating physician. Patients with positive imaging and biopsy findings for locally recurrent PCa constituted the cohort of case patients.

An inconclusive diagnosis was given to patients for whom a local or metastatic recurrence could not be detected on MRI, or the biopsies were negative or both. These patients were analysed further but separately from the cohort of cases.

Controls (patients without recurrent PCa)

Matched controls (patients without evidence of recurrent disease) were recruited and received a mp-MRI scan. The matching criteria were: time since primary RT (years), use of hormonal therapy and the risk group of the primary tumour. Risk stratification was based on the definition proposed by the European Association for Medical Oncology (ESMO) in 2010 [119], incorporating changes suggested in the review by Rodrigues *et al.* [120] on the importance of the amount of high-grade cancer, differentiating between Gleason 3+4 and 4+3. Risk stratification was defined as: low-risk - T1-T2a, Gleason ≤ 6 and PSA ≤ 10 ng/ml; intermediate-risk - T2b-T2c, Gleason = 7 (3+4) and PSA ≤ 20 ng/ml, not otherwise low- or high-risk; high-risk - T3-T4 or Gleason ≥ 7 (4+3) or PSA > 20 ng/ml. The exclusion criteria were identical to those applied for the cases.

If a patient developed a biochemical recurrence during the study period, he was moved to the inconclusive cohort.

Untreated (patients at primary diagnosis)

A previously reported cohort of 13 primary PCa patients [121, 122], scanned prior to treatment with a mp-MRI protocol, was used as reference. As part of the study, these patients underwent two mp-MRI scanning sessions. Following the second MRI examination, the patients were treated with a radical prostatectomy. This cohort is referred to as the untreated cohort.

4.2.2 MRI Protocol

All patients were scanned in a 3T Achieva dStream (A) or Ingenia (B) scanner (Philips Healthcare, Best, The Netherlands). Twenty patients were scanned in system A with a 6-channel cardiac coil and 48 were scanned in system B with a 16-channel anterior and 12-channel posterior coil. With the exception of 4 patients (3 scanned in system A), all were scanned with the use of an endorectal coil.

The mp-MRI protocol consisted of an axial, sagittal and coronal T2w turbo spin echo; a 3D gradient echo T1-weighted (T1w); a balanced steady-state free precession (bSSFP); an axial multi-echo spin echo k-t-T₂ sequence [121], with twelve echoes acquired at a spacing of 16ms, starting at 32ms; a transversal diffusion weighted sequence (DWI) and a dynamic contrast-enhanced sequence (DCE). DWI was acquired using a single-shot spin-echo echo-planar imaging sequence with diffusion encoding b-values ranging from 0 to 800 sec/mm². DCE was acquired with a 3D T1w spoiled gradient echo sequence at a temporal resolution of 2.6 seconds over 5 minutes, before, during and after intravenous administration of 15 ml of Dotarem (0.5 mM Dotarem, Guerbet, Roissy CDG, France). The contrast agent was administered using a power injector (MedRad, Warrendale, PA, USA), followed by a 30 ml saline flush at a flow rate of 3 ml/sec. Further sequence specific details can be found in **Table S4.1** (Supplementary Materials).

Apparent diffusion coefficient (ADC) maps were derived from the DWI sequence using $b = 200$ and 800 sec/mm^2 . The standard Tofts model [123] was used to calculate the pharmacokinetic maps K^{trans} and k_{ep} . This was done using a T_1 map generated based on the variable flip angle method [124], using flip angle = 3° , 6° , 10° , 20° , and 30° , to convert signal intensity into concentration, and an arterial input function (AIF) with parameters derived from an in-house study population. T_2 maps were created by converting the signal to the logarithmic scale and performing a weighted linear fit.

All sequence-derived functional maps were generated using MATLAB R 2017b (Mathworks, Natick, MA, USA).

Visual inspection was used to assess possible displacements between the functional sequences (and respective quantitative maps) and the T2w-MRI scans. Rigid registration based on mutual information was used whenever necessary to correct for these displacements. All images were resampled to the T2w grid.

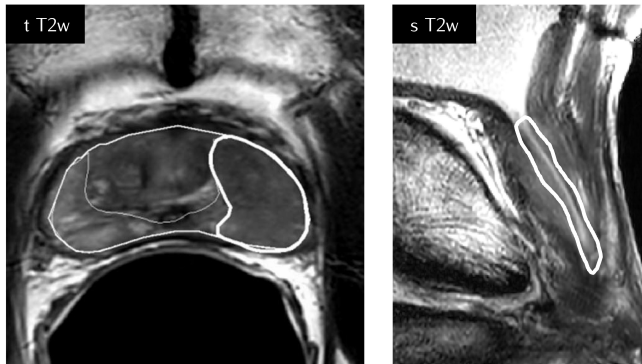


Figure 4.2: Illustration of the delineated ROIs. In the left image the eroded prostate contour, the peripheral zone, and the tumour can be seen delineated on the transversal T2w image. On the right, the peri-urethral contour is presented overlaid on the sagittal T2w scan.

4.2.3 Region of Interest (ROI) segmentation

The prostate, the peripheral zone (PZ), the central gland (CG) and the region surrounding the urethra – peri-urethral tissue (PU) – were delineated. The PU region was delineated using both the sagittal and transversal T2w images. The tumour suspected regions were delineated by a radiologist (14 years of experience) based on the MRI, PET and biopsy report information. As the Prostate Imaging Reporting and Data System 2 (PI-RADS v2) [42] is not applicable to recurrent prostate cancer, tumour was defined as a region with low signal-intensity (SI) on T2w-MRI, high SI in the $b=800$ DWI scan, low SI on the ADC map, and increased enhancement in the K^{trans} and k_{ep} maps. For the untreated cohort tumour delineations were originally performed on the histopathology and propagated to imaging. **Figure 4.2** illustrates the ROIs as delineated for a case patient.

The bSSFP sequence was used to identify and exclude from further analysis the implanted fiducial markers. Prostate delineations were eroded in all directions by 1mm to ensure that only prostate tissue was analysed.

Region of interest (ROI) analysis was performed using Python 3 (Python Software Foundation, Delaware, United States).

4.2.4 Statistics

A t-test for continuous variables (PSA) and a chi-square test for categorical variables (Gleason and TNM stage) were used to compare the clinical characteristics of the cohorts.

Normality was checked using the Shapiro Wilk test. In a case of non-normality a non-parametric Friedman's ANOVA for dependent groups was used to compare the different regions within the groups (untreated, cases and controls). If significant, the differences were further evaluated using a post-hoc Wilcoxon signed-rank test. To compare the same region between groups, a Mann-Whitney test was used. To correct for multiple testing, a Bonferroni correction was applied to the significance level of $\alpha = 0.05$.

Univariate and multivariate generalised linear mixed-effect modelling (GLMM) was applied to assess the predictive value of imaging on the voxel-wise likelihood of tumour. Only case patients were used in this analysis. Voxels were grouped into benign (resulting from a combination of unaffected PZ and CG) and tumour (based on the radiologist delineations). To obtain the likelihood of tumour on a voxel level, fixed and random effects were included. The quantitative imaging maps and voxel anatomical location (PZ or CG) were accounted as predictive variables and therefore included as fixed effects. Random effects accounted for spatial clustering by incorporating voxel location within the prostate (the relative distance in x , y and z from the prostate centre of mass) and patient identifiers. When the association between model parameters and tumour probability was non-linear, the parameters were grouped in quartiles and regression coefficients were estimated for each group considering the first quartile as reference. The model fit was assessed using the Bayesian Information Criterion (BIC) (a decrease of 10 points reflects an improved fit) and by evaluating the residual random error. Statistics were performed with the Statistical Package for Social Sciences, version 22.0 (SPSS, Chicago, IL, USA). Analysis was performed in R [125] using the lme4 package [126].

4.3 RESULTS

4.3.1 Patient Characteristics

In total, 33 case patients and 35 controls were prospectively included in the study. Adjuvant hormonal therapy was given for a maximum period of 3 years to 82% of the cases and to 94% of the controls upon the treatment of primary PCa. However, none of the patients received hormonal therapy within 1 year prior to the MRI exam.

Cases (patients with recurrent PCa)

Ten of the 33 included patients had either regional metastases without evidence of intraprostatic recurrence (n=5) or an inconclusive diagnosis (n=5). Twenty-three patients remained that fitted all the inclusion criteria for cases. For 3 of the 23 case patients, the first biopsy was negative and a repeated navigated biopsy confirmed the MRI suspected tumour location.

In 2 out of 23 patients the k-t-T₂ sequence was not acquired and for another the ADC map was not used due to poor image quality.

For the 23 cases a total of 27 suspected tumour foci were delineated, with an average tumour volume of 1.37 cm³. From these, 4 were located in CG. The median time to recurrence for the 23 cases was 7 years since primary EBRT treatment.

Controls (patients without recurrent PCa)

One patient was initially recruited as a control but had a biochemical failure during the study period. This patient was moved to the inconclusive cohort. The remaining 34 matched controls (patients without evidence of locally recurrent PCa) had a median time of 6 years since the EBRT treatment.

Untreated (patients at primary diagnosis)

The 13 untreated patients had a median time of 20 (range, 5-65) days between the MRI and prostatectomy. None of these patients received hormonal therapy before imaging. For one patient, the small dimensions of the histopathological tumour resulted in a sub-voxel tumour ROI in MRI. For this patient only benign tissue was analysed.

Table 4.1 describes the patient characteristics for cases, controls and untreated patients. For the untreated patients, the characteristics reported in the table are based on the biopsy report and clinical TNM stage.

The clinical characteristics of the cohorts were compared. TNM stage was grouped in two (T1+T2 and T3+T4) and the Gleason score in three groups (Gleason ≤ 6 , 7, ≥ 8). Cases and controls had comparable clinical characteristics (PSA: $p=0.789$; TNM stage: $p=0.393$; Gleason score: $p=0.271$). Both case and control patients had significantly higher-risk primary PCa when compared with the untreated patients (PSA: $p=0.008$ and $p=0.003$; TNM stage: $p=0.004$ and $p<0.001$; Gleason score: $p=0.03$ and $p=0.004$ for cases and controls, respectively).

Figure 4.3 illustrates the values of T_2 , ADC, K^{trans} and k_{ep} for untreated, controls, and case patients. Median imaging values for the 3 cohorts are presented in Table 4.2.

Table 4.1: Patient characteristics. Recurrent tumours were not assigned a Gleason score as radiation induced atypia can be a confounder for pathological interpretation.

* In six case patients TNM stage of recurrent tumour was not reported

† For one case patient only the Anderson score (=2) was available

‡ In one case and one control the PSA of the primary tumour could not be retrieved

| | Cases | | Controls | Untreated |
|---------------------------|----------------|------------|----------------|----------------|
| | Primary tumour | Recurrence | Primary tumour | Primary tumour |
| TNM stage * | | | | |
| T1 | - | - | 5 | 5 |
| T2 | 11 | 13 | 6 | 7 |
| T3 | 11 | 4 | 23 | 1 |
| T4 | 1 | - | - | - |
| Gleason score † | | | | |
| ≤ 6 | 4 | - | 6 | 7 |
| 7 | 12 | - | 12 | 6 |
| ≥ 8 | 6 | - | 16 | - |
| PSA (ng/ml) ‡ | | | | |
| ≤ 10 | 10 | 22 | 11 | 12 |
| $10 < \text{PSA} \leq 20$ | 6 | 1 | 14 | 1 |
| ≥ 20 | 6 | - | 8 | - |

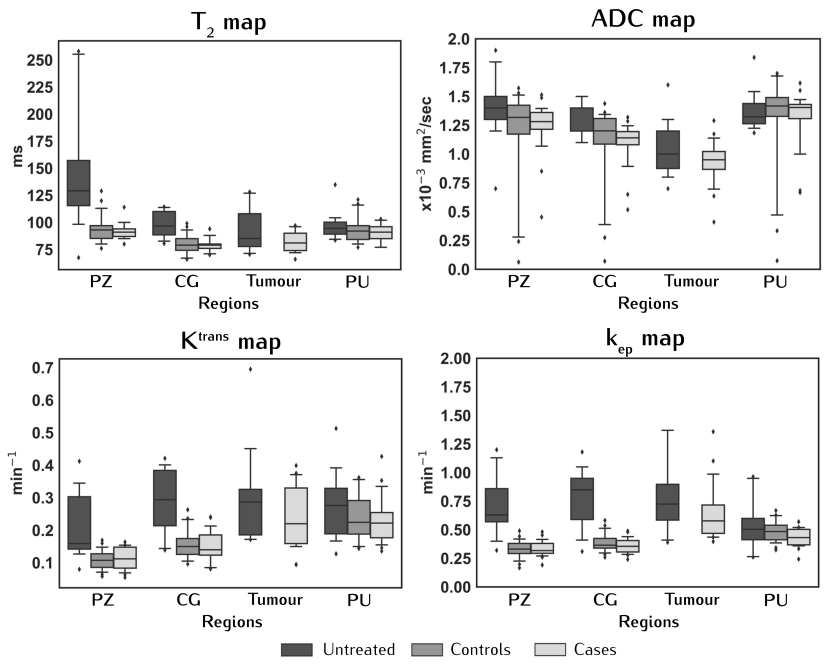


Figure 4.3: Boxplots with median values for all regions, imaging modalities, and the three cohorts of patients: untreated, controls (non-recurrent), and cases (recurrent). Edges of the boxes are the first (25th) and third (75th) quartiles, and whiskers represent the 5th and 95th percentile. PZ, peripheral zone; CG, central gland; PU, periurethral tissue.

Table 4.2: Quantitative imaging values for the three cohorts in the different prostatic regions. Median of all individual patients' median values, with the minimum and maximum presented between brackets.

| | | PZ | CG | Tumour | PU |
|---|-----------|------------------|------------------|------------------|--------------------|
| T_2 (ms) | Untreated | 129 (67-258) | 97 (81-114) | 85 (70-128) | 94 (84 - 135) |
| | Controls | 93 (76-129) | 79 (66-99) | - | 92 (77-121) |
| | Cases | 91 (80-114) | 79 (70-94) | 81 (66-97) | 91 (77-103) |
| ADC ($\times 10^{-3}$ mm ² /s) | Untreated | 1.40 (0.70-1.90) | 1.20 (1.10-1.50) | 1.00 (0.70-1.60) | 1.32 (1.20 - 1.80) |
| | Controls | 1.32 (0.06-1.57) | 1.20 (0.07-1.44) | - | 1.41 (0.07-1.70) |
| | Cases | 1.28 (0.45-1.51) | 1.14 (0.52-1.32) | 0.95 (0.41-1.29) | 1.40 (0.67 - 1.62) |
| K^{trans} (min ⁻¹) | Untreated | 0.16 (0.08-0.41) | 0.29 (0.14-0.42) | 0.29 (0.17-0.70) | 0.28 (0.13-0.51) |
| | Controls | 0.11 (0.06-0.17) | 0.15 (0.10-0.26) | - | 0.22 (0.14-0.36) |
| | Cases | 0.11 (0.05-0.16) | 0.14 (0.08-0.24) | 0.22 (0.09-0.40) | 0.22 (0.14-0.43) |
| k_{ep} (min ⁻¹) | Untreated | 0.63 (0.32-1.20) | 0.85 (0.31-1.18) | 0.73 (0.39-3.89) | 0.50 (0.26-0.96) |
| | Controls | 0.33 (0.17-0.49) | 0.37 (0.26-0.58) | - | 0.48 (0.32-0.67) |
| | Cases | 0.32 (0.19-0.48) | 0.36 (0.24-0.49) | 0.58 (0.40-1.36) | 0.43 (0.24-0.57) |

4.3.2 Characteristics of irradiated benign tissue

Figure 4.4 illustrates the quantitative maps and suspected tumour regions for a representative case (**Figure 4.4A**) and control (**Figure 4.4B**) patients.

The values for benign PZ and CG were comparable between cases and controls ($P = 0.15$ for ADC in the CG up to 0.91 for the k_{ep} in the PZ). Values in the PU were also similar between the groups (with the lowest $P = 0.05$ for $k_{\text{ep}} > 0.05/3$). Both cases and controls presented with PU enhancement.

For controls, significant differences were found between PZ and CG and between CG and PU ($P < 0.001$ for all comparisons between imaging values and regions).

The values obtained for the PZ, CG and PU in the inconclusive cohort were not significantly different from those of the cases (with the lowest $P = 0.03$ for T_2 in the PZ $> 0.05/3$) or controls (with the lowest $P = 0.04$ for T_2 in the CG

> 0.05/3). The quantitative imaging values for this specific cohort are present in the Table S4.2 (Supplementary Materials).

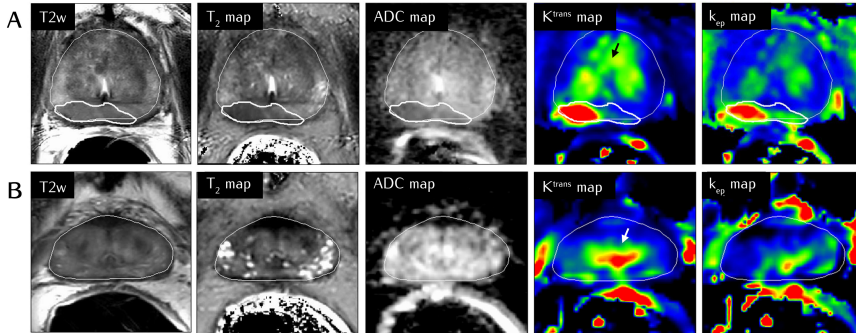


Figure 4.4: Representative example of a case (A) patient with recurrence and a control (B) without recurrent disease, with all anatomical and functional maps. Thin and thick white lines represent prostate and tumour delineations, respectively. Black and white arrows in the K^{trans} maps point to the increased enhancement in the peri-urethral tissue, present in both cases and controls.

4.3.3 Radiation-induced changes

To investigate radiation-induced tissue changes, the values from the untreated patients were used as a population reference against which the values of the cases were compared.

The different prostate regions were more homogeneous after RT and the range of values was in general smaller than before treatment. With the exception of ADC (for which $P = 0.03 > 0.05/4$), T_2 , K^{trans} and k_{ep} values in the PZ were significantly lower (T_2 , k_{ep} $P < 0.001$ and K^{trans} $P = 0.002$) after RT. The CG imaging values were for all maps significantly lower after RT (all with $P \leq 0.001$). In the PU, despite the higher K^{trans} values seen before treatment, the imaging values for all maps remained comparable before and after treatment ($P = 0.19$ for T_2 up to $P = 0.56$ for ADC).

Similar characteristics were seen for all maps for recurrent and primary tumours ($P = 0.17$ for k_{ep} up to $P = 0.22$ for ADC).

4.3.4 Tumour localisation

To compare recurrent tumour with the remaining irradiated benign gland, the cohort of case patients was assessed. With all imaging maps, median tumour values were significantly different from benign tissue within the PZ ($P < 0.001$ for all imaging modalities). No differences were seen between the T_2 values of suspected tumour and benign CG tissue ($P = 0.08$). However, values of ADC, K^{trans} and k_{ep} from suspected tumour regions were significantly different (all with $P < 0.001$) from those of the CG. When compared with the surrounding benign tissue, heightened K^{trans} and k_{ep} values were seen in the PU. The PU K^{trans} values were comparable with those of suspected tumour ($P = 0.81$), but with k_{ep} a significant distinction ($P < 0.001$) could be made between the two regions. The PZ and CG ROIs were significantly different for all maps except k_{ep} ($P = 0.05 > 0.05/6$).

Table 4.3: Model parameters obtained when combining T_2 map, ADC, K^{trans} and k_{ep} imaging values to predict tumour at the voxel-wise level. Imaging values and location are included as fixed effects; patient and spatial coordinates as random effects.

| Fixed effects (MRI) | Regression coefficients (β) | Std. Error | P |
|--|-------------------------------------|----------------|----------------|
| T_2 map reference group (24 – 73) | 0 | - | - |
| T_2 map 2 nd quartile (74 – 83) | 0.08 | 0.01 | <0.001 |
| T_2 map 3 rd quartile (84 – 97) | -0.40 | 0.01 | <0.001 |
| T_2 map 4 th quartile (98 – 1188) | -1.50 | 0.02 | <0.001 |
| ADC | -2.86 | 0.02 | <0.001 |
| K^{trans} | 6.52 | 0.09 | <0.001 |
| k_{ep} | 4.85 | 0.04 | <0.001 |
| Location | -0.98 | 0.01 | <0.001 |
| Intercept (β_0) | -5.55 | 0.46 | <0.001 |
| Random effects | Variance (σ^2) | | |
| Patients | 1.41 | | |
| Distance CM_x | 0.67 | | |
| Distance CM_y | 8.50 | | |
| Distance CM_z | 13.30 | | |
| Residuals (ϵ_0) | Median | First quartile | Third quartile |
| | -0.07 | -0.13 | -0.03 |

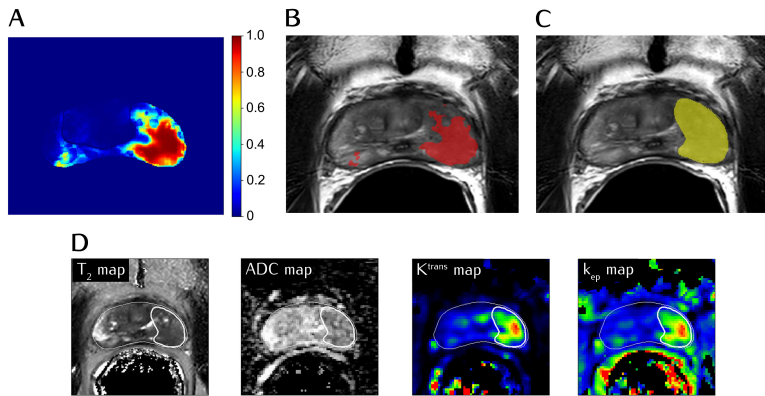


Figure 4.5: A reconstructed probability map based on the generalised linear mixed model (A). In red, overlaid on the T2w scan (B), is the same probability map thresholded at 0.5 and in (C), in yellow, the original tumour delineation. The anatomical and functional maps are presented in (D), together with the prostate and tumour delineations in white.

Using the GLMM, we observed that univariately all MRI parameters were significantly associated with tumour likelihood ($P < 0.001$). Location within the gland was also significantly associated with tumour likelihood. The association between T_2 and tumour probability was non-linear, and therefore this parameter was grouped in quartiles with univariate regression coefficients $\beta = 0$ (reference group), 0.19, -0.59 and -1.83. The regression coefficients were for ADC: $\beta = -5.06$, K^{trans} : $\beta = 17.27$, k_{ep} : $\beta = 8.42$ and location: $\beta = -0.64$. The multivariate model combining all mp-MRI parameters and location obtained the best fit and was the most predictive for tumour likelihood (BIC of 392212, $P < 0.001$). The values for the regression coefficients and standard errors are presented in [Table 4.3](#). A reconstructed probability map for a case patient is presented in [Figure 4.5](#).

4.4 DISCUSSION

In this study mp-MRI was used to quantitatively describe and distinguish benign prostate tissue from recurrent PCa after RT. The use of two patient cohorts,

with and without recurrent PCa, allowed us to show that tumour unaffected areas were similar in both groups. Compared with the pre-treatment setting, and with exception of ADC in the PZ, all other imaging values (T_2 , K^{trans} and k_{ep}) were lower, but also more homogeneous between the PZ and CG. This is in agreement with literature describing the loss of zonal distinction in the post-RT setting [37]. Still, significant differences between the imaging values of PZ and CG were seen in both recurrent and non-recurrent patients. As in primary disease, recurrent tumour occurred predominantly in the PZ. This is not surprising as recurrence usually happens at the site of primary PCa [55]. For all maps, values of recurrent tumour were not significantly different from those of primary PCa.

With all imaging maps, we found that the PZ was significantly different from recurrent tumour, illustrating that all quantitative maps seem to have potential to distinguish tumour in this region. However, tumour detection in the CG is challenging using solely the T_2 map, as the values for benign CG and tumour are not significantly different. This distinction is possible when evaluating ADC, K^{trans} and k_{ep} . The PI-RADS v2 [42] is the standard radiological guideline for the diagnose of primary PCa. These guidelines do not apply to the detection of suspected recurrent PCa following treatment. Yet, in the absence of further guidance, PI-RADS v2 is often used as a starting point for image assessment. The guideline describes T2w as the dominant mp-MRI sequence for the diagnosis of primary PCa in the transition-zone. Our results suggest this sequence is of less relevance compared to the other mp-MRI sequences when diagnosing recurrent PCa in this region. The results also suggest that DCE-MRI has an important role in the diagnosis of recurrent disease. In non-irradiated prostate tissue, the presence of benign prostatic hyperplasia (BPH) in the central gland is an important imaging confounder for the use of DCE-MRI in this region. In a previous study, pathology slides from salvage prostatectomy specimens of recurrent PCa patients after radiotherapy were registered to mp-MRI acquired prior to the surgery. In the analysed pathology, no BPH could be found [127]. This suggests that BPH vanishes after radiotherapy. The enhancement in the peri-urethral area remains and needs to be considered. It has been hypothesised that the fibrotic changes observed in irradiated prostates [128] and the decreased microvasculature associated with atrophic tissue, enhance the neovascularity of recurrent tumour relative to the remainder benign atrophic tissue [43].

With overall median values below 100ms, T_2 is decreased for the prostate as a whole after RT. Measured T_2 values (ms) were in agreement with a previous study observing the effects of RT in prostate tissue [129] (at week 8 of the RT treatment, mean \pm SD: PZ = 89 ± 13 , CG = 76 ± 5 , Tumour = 75 ± 9) and they provide evidence that the PZ and CG values do not increase after these 8 weeks.

The ADC values (in $\times 10^{-3}$ mm²/sec) obtained for recurrent tumour were similar to those reported in other studies (mean \pm SD: 1.0 ± 0.1 and 0.98 ± 0.23) [130, 131]. Tumour values in the untreated cohort were also comparable to those previously described (tumour PZ, mean \pm SD: 1.08 ± 0.39) [132]. Values for benign prostate tissue in our case and control cohorts were slightly lower than those reported by others (mean \pm SD: 1.6 ± 0.2 and 1.60 ± 0.21 ; $\times 10^{-3}$ mm²/sec) [130, 131]. We note that these studies did not distinguish PZ from CG and reported values for a composite of both regions. In our untreated cohort, benign prostate tissue showed slightly lower values than those seen by Sato *et al.* (mean \pm SD: 1.80 ± 0.41 for PZ and 1.58 ± 0.37 for the transition-zone) [132].

In controls the PU region exhibited high K^{trans} and k_{ep} values without signs of malignancy. Similar high values were found in cases and this enhancement strongly overlapped with tumour values, which might represent a confounder for tumour detection in the CG. Yet, with k_{ep} significant differences were seen between tumour and PU, suggesting this could be a relevant parameter when assessing the CG. This effect has been previously described in the recurrent setting [38, 44], potentially as a result of remaining BPH or peri-urethral vasculature. Donati *et al.* [46] found no additional benefit of DCE when complementing T2w and DWI, which could have been the result of benign CG enhancement.

The GLMM analysis revealed tumour distinction was optimised with the use of mp-MRI, highlighting the importance of using a combination of anatomical and functional sequences when identifying and localizing tumour for salvage treatment strategies. Our study did not include spectroscopic data. Nonetheless, spectroscopic data has been suggested as a valuable tool for recurrent tumour distinction [130].

Quantitative tissue characterization enables a more straightforward comparison between subjects, studies and centres. It also conveys important information to adapt the scanning protocols to better image this study population. As an example, for patients with suspected recurrence, the sampling echo times for T_2 mapping should be shortened as T_2 decay is expected to be shorter compared to untreated patients.

We aimed to circumvent the limitations of a 2D standard transrectal ultrasound (TRUS)-guided systematic biopsy by using a 3D MRI-TRUS fusion approach to target the tumour suspected region. With this technique's improved spatial information, it has been described that fewer cores were necessary to detect more clinically significant cancer [133], often leading to reclassification of primary tumours in higher risk categories [134]. Nonetheless, for a subset of our patients, cancer was only confirmed with repeated biopsy. Even when navigated, the biopsy procedure is limited by possible co-registration errors and in delivering the biopsy needle to the intended point [135]. Tumour delineations performed using 1.5T T2w, were shown to underestimate the true tumour volume when compared with prostatectomy samples [55]. As a result, the detection rates of MRI-TRUS fusion biopsies might also be affected by errors in defining the lesion to target.

Our study has several limitations. Firstly, it was a single hospital study with a limited number of patients. The controls were recruited to match the characteristics of cases as best as possible, and no significant differences were seen between their clinical parameters. However, our cohort of cases and controls had significantly higher-risk primary PCa when compared with the untreated patients. Thus, the reported imaging differences for benign tissue could theoretically arise from baseline imaging dissimilarities. No significant imaging differences were seen between the tumours in the untreated and recurrent cohorts.

Ideally, using salvage prostatectomy specimens as the standard of reference would be preferred to accurately identify cancer location, but this procedure is only sporadically performed. Thus, patients with an inconclusive diagnosis were analysed separately to avoid a potential bias. The diagnosis of recurrent disease is in practice made with a combination of mp-MRI and PSMA-PET findings. In this study we opted to individually investigate the properties of

MRI which, due to its higher spatial resolution, is better suited for the purpose of tumour localisation. A combination of both modalities can certainly help better select patients suitable for salvage treatment. Lastly, our GLMM was not tested in an independent cohort. Thus further validation is required to establish applicability to other populations. The model is also potentially reflecting delineation practices.

In conclusion, tumour unaffected areas were similar between recurrent and non-recurrent patients, and generally with lower and more homogeneous values than before RT. The PU region presents with increased enhancement in the pharmacokinetic maps for both cases and controls. Analysis of mp-MRI, with the quantitative maps T_2 , ADC, K^{trans} and k_{ep} together with location information resulted in optimal distinction between tumour and benign voxels. The ability to accurately localise tumour and distinguish it from benign tissue with the use of mp-MRI will help in the design of focal salvage treatment strategies for locally recurrent PCa.

SUPPLEMENTARY MATERIALS

Table S4.1: Sequence specific parameters.

Two case patients were scanned using a different protocol. T2w had a TR (ms) = 3952 and 3969 and TE (ms) = 120ms, with a reconstructed voxel size of $0.4 \times 0.4 \times 3 \text{ mm}^3$. T1w had TR (ms) = 3.61 and 3.67 and TE (ms) = 1.80 and 1.85, and a reconstructed voxel size (mm) of 0.85 and 0.84 with a 2mm slice thickness. The DWI sequence had a TR (ms) = 5355 and 4931 and a TE (ms) = 61.8 and 54.0, with a reconstructed voxel size of $1.06 \times 1.06 \times 3 \text{ mm}^3$. For these patients a balanced TFE and k-t-T₂ sequences were not acquired. The DCE sequence was the same as described in the table above.

| | Imaging plane | Scan mode | Field of view (mm) | Matrix size | Slice thickness (mm) | Reconstructed voxel size (mm) | TR/TE (ms) | Echo train length | Flip angle (degrees) | Acceleration factor | Receiver bandwidth (Hz/voxel) | Phase encoding direction | Acquisition time | Number of signal averages |
|--------------------|---------------|-----------|--------------------|-------------|----------------------|-------------------------------|-----------------------------|-------------------|----------------------|----------------------|-------------------------------|--------------------------|------------------|---------------------------|
| T1 THRIVE | Axial | 3D | 280 x 448 | 448 x 378 | 2.0 | 0.88 | 3.7 / 1.86 | 88 | 10 | 0.6 & SENSE factor 1 | 720 | RL | 2min 42sec | 3 |
| T2 TSE | Axial | 2D | 140 x 140 | 304 x 242 | 3.0 | 0.27 | 3143 / 120 | 26 | 90 | N/A | 236 | RL | 3min 25sec | 1 |
| k-t-T ₂ | Axial | 2D | 170 x 170 | 212 x 212 | 3.0 | 0.39 | 2469 / 32; $\Delta TE = 16$ | 12 | 90 | SENSE factor 2 | 169 | RL | 5min 6sec | 1 |
| DWI | Axial | 2D | 160 x 180 | 144 x 144 | 2.7 | 1.03 | 3131 / 55 | 49 | 90 | SENSE factor 2 | 17.2 | RL | 2min 23sec | 1, 4, 8 for b=0, 200, 800 |
| DCE | Axial | 3D | 260 x 260 | 144 x 144 | 6.0 | 1.16 | 4.0 / 1.9 | 1 | 20 | SENSE factor 4 | 2117 | RL | 4min 56sec | 1 |
| bTFE | Axial | 3D | 250 x 250 | 252 x 234 | 2.0 | 0.49 | 3.70 / 1.85 | 36 | 30 | N/A | 1240 | RL | 2min 23sec | 1 |

Table S4.2: Median of all patient's median values for each region and imaging sequence for the inconclusive cohort. Minimum and maximum values are presented between brackets.

| Inconclusive | | | |
|--|--------------------|--------------------|--------------------|
| | PZ | CG | PU |
| T_2 <i>ms</i> | 84 (76 - 100) | 70 (62 - 82) | 83 (76 - 111) |
| ADC $\times 10^{-3} \text{ mm}^2/\text{s}$ | 1.31 (1.20 - 1.48) | 1.13 (1.03 - 1.36) | 1.39 (1.16 - 1.68) |
| K^{trans} <i>min</i> ⁻¹ | 0.09 (0.04 - 0.12) | 0.14 (0.08 - 0.24) | 0.19 (0.12 - 0.25) |
| k_{ep} <i>min</i> ⁻¹ | 0.29 (0.12 - 0.33) | 0.33 (0.21 - 0.40) | 0.43 (0.28 - 0.50) |

QUANTITATIVE 3T MULTI-PARAMETRIC MRI
AND STEP-SECTION PATHOLOGY OF
RECURRENT PROSTATE CANCER PATIENTS
AFTER RADIATION THERAPY

Catarina Dinis Fernandes¹
Ghazaleh Ghobadi¹
Henk G van der Poel²
Jeroen de Jong³
Stijn WTPJ Heijmink⁴
Ivo G Schoots⁴
Iris Walraven¹
Petra J van Houdt¹
Milena Smolic¹
Floris J Pos¹
Uulke A van der Heide¹

¹Department of Radiation Oncology, ²Urology and ⁴Radiology
The Netherlands Cancer Institute, Amsterdam, The Netherlands

³Department of Pathology, Reinier Haga Medisch Diagnostisch Centrum,
Delft, The Netherlands

European Radiology 2018; in press

DOI: [10.1007/s00330-018-5819-y](https://doi.org/10.1007/s00330-018-5819-y)

ABSTRACT

Objectives: Diagnosis of radio-recurrent prostate cancer using multi-parametric MRI (mp-MRI) can be challenging due to the presence of radiation effects. We aim to characterise imaging of prostate tissue after radiation therapy (RT), using histopathology as ground truth, and to investigate the visibility of tumour lesions on mp-MRI.

Methods: Tumour delineated histopathology slides from salvage radical prostatectomy patients, primarily treated with RT, were registered to MRI. Median T2-weighted, ADC, K^{trans} and k_{ep} values in tumour and other regions were calculated. Two radiologists independently performed mp-MRI-based tumour delineations which were compared with the true pathological extent. General linear mixed-effect modelling was used to establish the contribution of each imaging modality and combinations thereof in distinguishing tumour and benign voxels.

Results: Nineteen of the twenty-one included patients had tumour in the available histopathology slides. Recurrence was predominantly multifocal with large tumour foci seen after external-beam radiotherapy, whereas these were small and sparse after low-dose-rate brachytherapy. MRI-based delineations missed small foci and slightly underestimated tumour extent. The combination of T2-weighted, ADC, K^{trans} and k_{ep} had the best performance in distinguishing tumour and benign voxels.

Conclusions: Using high resolution histopathology delineations, the real tumour extent and size was found to be underestimated on MRI. Mp-MRI obtained the best performance in identifying tumour voxels. Appropriate margins around the visible tumour suspected region should be included when designing focal salvage strategies. Recurrent tumour delineation guidelines are warranted.

5.1 INTRODUCTION

Depending on the risk group, 5-year disease free survival varies from 67 – 80% for prostate cancer patients who are treated with whole-gland radiation therapy (RT) to 78 Gy [58]. Most patients with recurrent disease have developed metastasis, yet a small but significant proportion will harbour locally recurrent disease only [26, 27]. For these patients, salvage treatment options include radical prostatectomy, brachytherapy, cryotherapy or high-intensity focused ultrasound [33]. Whole-gland approaches, such as salvage radical prostatectomy (SRP), are reported to obtain good disease control, however with a high chance of severe gastrointestinal and genitourinary complications when compared with radical prostatectomy in primary disease [136].

Focal salvage strategies aim at reducing these comorbidities by sparing the uninvolved tissue while specifically targeting recurrent cancer. When evaluated in radio-recurrent PCa, the complications of focal approaches are comparable or lower than in whole-gland strategies, with a 5-year biochemical disease-free survival of 46.5 – 54.5% [33].

Successful development of focal salvage strategies for local recurrent PCa requires an accurate detection and localisation of the tumour. Positron emission tomography (PET) with ^{68}Ga PSMA-ligands, offers high sensitivity to detect recurrent PCa [13], yet the poor spatial resolution limits its use for focal salvage strategies. Magnetic resonance imaging (MRI) is alongside PET the most used modality for recurrence diagnosis. High-resolution anatomical and functional imaging make MRI attractive for the preparation of focal salvage treatments [52].

Post-RT benign confounders pose a challenge to MRI interpretation. Prostate tissue shows diffuse signal intensity (SI) reduction on T2-weighted (T2w) MRI, complicating tumour detection [37]. The use of 1.5T MR spectroscopy [45] as well as DWI- and DCE-MRI [40] have been found to surpass T2w in detecting local recurrence. Thus multi-parametric MRI (mp-MRI) with diffusion-weighted imaging (DWI) and dynamic contrast enhanced (DCE) MRI is often preferred. Zattoni *et al.* [51] reported 3T mp-MRI to have good accuracy in detecting recurrent disease extension. Conversely, Donati *et al.* [46] found no additional

benefit of DCE when added to T2w and DWI. These studies used either biopsy or SRP samples for validation, but none attempted accurate registration of imaging and histopathology. To date, no guidelines exist on how to score or delineate recurrent tumour suspected regions.

In this study 3T mp-MRI is used to characterise irradiated prostate tissue, using SRP specimens registered to mp-MRI as ground truth. Using high resolution tumour delineations in histopathology, we describe radio-recurrent PCa and use mixed modelling to establish which imaging sequences result in an optimal distinction between tumour and benign voxels. We further investigate the accuracy of tumour detection in mp-MRI, a crucial element for the development of focal salvage strategies.

5.2 MATERIALS AND METHODS

5.2.1 Patients

Twenty one patients with radio-recurrent PCa, who underwent mp-MRI for local staging prior to SRP between 2011 and 2017, were retrospectively included. Biochemical recurrence was established according to the Phoenix criteria [65]. All patients had biopsy-proved recurrence and absence of distant metastasis on Choline PET scans. External-beam radiotherapy (EBRT) and low-dose rate (LDR) brachytherapy treated patients were included.

5.2.2 MRI acquisition

Patients were scanned on a 3T Achieva (16), Achieva dStream (3) or Ingenia (2) MRI scanner (Philips Healthcare) between June 2011 and March 2017. Thirteen patients were scanned using an endorectal coil in addition to a torso or cardiac phased array coil. The mp-MRI protocol included triplanar T2w turbo spin-echo (TE 90-130 ms, TR > 2600 ms) and axial T1-weighted (T1w) gradient echo (TE < 2.3 ms, TR < 5.3 ms) sequences. DWI was acquired using

a single-shot spin-echo echo-planar imaging sequence (b-values between 100 and 1000 s/mm²), from which apparent diffusion coefficient (ADC) maps were generated. DCE was acquired with a 3D T1w spoiled gradient echo sequence (TE/TR = 2/4 ms), dynamic interval 2.3-2.9 seconds over 5-6 minutes, with intravenous administration of 7.5 mmol gadoteric acid (Dotarem). For seven patients the DCE sequence was not acquired. The pharmacokinetic parameter maps K^{trans} and k_{ep} were derived from the DCE scans. A balanced Steady-State Free Precession (bSSFP) sequence was acquired for fiducial and seed visualization in seven patients. MRI acquisition and post-processing details can be found in the Supplementary Materials.

Visual inspection was used to assess displacements between the functional sequences and the T2w scans, with rigid registration performed when necessary. The T2w voxel size in plane resolution was between 0.3-0.8 mm, with a slice thickness of 3 mm. All maps were resampled to the T2w grid.

5.2.3 Pathology

Whole-mount axial slides, stained with hematoxylin and eosin (H&E), were used for histopathological validation. Slides at the apex and base of the specimen were sliced parasagittally for extra-prostatic extension evaluation and were excluded from analysis. The slides were digitalised using an Aperio ScanScope XT (Aperio Technologies). Supervised by an uro-pathologist (6 years of experience), tumour delineations were made (resolution 0.5 $\mu\text{m}/\text{pixel}$) on the digitised slides using ImageScope.

5.2.4 T2w-MRI and H&E registration

The registration between H&E slides and the transversal T2w MRI was performed by two observers in consensus. A T2w slice was visually assigned to each H&E slide. Slide matching was performed considering the order of the slides, the location of the apex and base of the prostate, visible anatomical landmarks, and the relative size and shape of subsequent H&E and T2w slices. Each H&E slide was then registered to its matched T2w slice using deformable

registration based on landmark points (Coherent Point Drift) [137] implemented in MATLAB R2015a (The MathWorks). Landmark points included the urethra, nodules, the prostate boundary, and the tumour.

The registration error was estimated by selecting one landmark per pathology slide and measuring the Euclidean distance between the point in the T2w and in the registered pathology.

After registration, tumour delineations on the H&E slides were propagated to MRI. To ensure that only tumour voxels were used to characterise imaging and train the mixed model, tumour delineations were eroded in all directions by 1 mm.

All further analyses were restricted to the MRI slices for which a matching H&E slide existed.

5.2.5 Region of interest segmentation

A ROI was delineated in the levator ani muscle and the median SI in this region was used to normalise the T2w images. For MRI slices with H&E match the entire prostate and peripheral zone (PZ) were delineated using the transversal T2w MRI. The central gland (CG) was defined as the remaining non-PZ region, comprising the central and transition zone. The peri-urethral (PU) region was delineated with both sagittal and transversal T2w MRI. All voxels within these regions were included in the analysis. The location of implanted seeds and fiducial markers, as visible in the T1w or bSSFP sequences, was removed from the analyses.

Two uro-radiologists (14 and 7 years of experience) independently delineated suspected tumour regions, having access to mp-MRI and the radiological report, the PET Choline scans and the biopsy results. Since PI-RADS v2 [42] is not applicable to recurrent prostate cancer, tumour was defined as a region with low signal-intensity (SI) on T2w-MRI, high SI in the $b = 800$ DWI scan, low SI on the ADC map, and increased enhancement in the K^{trans} and k_{ep} maps. Tumour regions delineated on mp-MRI were compared with the delineations propagated from histopathology, and the distance and overlap were described.

When overlapping, the distance and spatial overlap were quantified using the 95% Hausdorff distance between contours and the Dice coefficient. Delineated and pathological volumes were estimated using the T2w grid.

For LDR brachytherapy patients with seminal vesicle (SV) invasion, the visible tumour areas as well as a benign region were delineated but no distances with histopathology contours were determined.

5.2.6 Statistics

Significant differences ($p < 0.05$) of median imaging values between ROIs were tested with a non-parametric Friedman's ANOVA. To further examine differences a post-hoc Wilcoxon signed rank test was used with a Bonferroni correction.

Using the EBRT patients, univariate and multivariate generalised linear mixed-effect modelling was applied to assess the predictive value of imaging on the voxel-wise likelihood of tumour. Voxels were grouped as: benign, resulting from a combination of unaffected PZ and CG for which H&E was available; and tumour, based on pathology delineations. To obtain the likelihood of tumour on a voxel level, fixed and random effects were included.

The T2w, ADC, K^{trans} , k_{ep} maps and voxel anatomical location (PZ or CG) were included as fixed effects. Random effects accounted for spatial clustering, by incorporating the voxel relative distance from the prostate centre of mass and patient identifiers. When the association between model parameters and tumour probability was nonlinear, the parameters were grouped in quartiles and regression coefficients were estimated for each group considering the first quartile as reference. The model fit was assessed using the Bayesian Information Criterion (BIC) (a decrease of 10 points reflects an improved fit) and by evaluating the residual random error. The contribution of predictors such as time since RT, use of hormonal therapy (HT) and treatment modality, included as fixed effects, was also investigated. A calibration curve was used to assess the quality of the best model. Analyses were performed in R (RStudio) using the lme4 package.

5.3 RESULTS

Thirteen patients had been primarily treated with EBRT and eight with LDR brachytherapy between 1999 and 2013. Median time between treatment and the MRI for recurrence diagnosis was 84 months for EBRT and 60 months for LDR brachytherapy. Following the D'Amico definition [64], 1, 8 and 12 patients had low, intermediate and high-risk primary PCa. Eleven patients received HT as part of their primary treatment, but none received HT at the time of imaging. Patient characteristics can be found in **Table 5.1**. Gleason score was not assigned to the recurrent tumour as radiation-induced atypia can be a confounder for pathological interpretation [138].

5.3.1 Pathological findings

EBRT patients had an average tumour volume of 1.2 cm^3 , LDR patients 0.51 cm^3 . For all patients except one, recurrent tumour after LDR brachytherapy extended to the SV. In this subset, tumour within the gland was mainly small and sparse. For 2 patients no tumour was visible in the available H&E slides, with their pathology reports describing tumours of small dimensions in the prostate apex, bladder neck and/or extension to the SV. These patients were excluded from further analysis.

The average registration error for all slides was $1.1 \pm 0.9 \text{ mm}$. Quantitative image analysis based on the matched H&E slides was performed for 19 patients and 56 slides - **Figure 5.1**. Multifocal recurrent disease was present for all but one patient. Sixty-three percent of the tumours were located in the PZ and 11% in the CG. The remaining 26% covered both PZ and CG. Benign prostatic hyperplasia (BPH) could not be seen in the pathology slides matched to MRI. The pre-SRP biopsy report contained information about the location of positive cores for 17/21 patients. When compared to the SRP histopathology, biopsy correctly detected 42% of SV-located tumours, as well as 33% of left and 75% of right prostate gland tumours.

Table 5.1: Patient characteristics.

IQR – Interquartile range

| | |
|---|-----------|
| Nadir PSA (ng/ml) [IQR] | 0.7 [0.9] |
| Median PSA level at MR imaging (ng/ml) [IQR] | 5.4 [3.4] |
| Median time from MRI to SRP (months) [IQR] | 3 [3] |
| Median iPSA level (ng/ml) [IQR] | 15 [22] |
| Primary clinical tumour stage | |
| T1c | 4 |
| T2a | 4 |
| T2b | 2 |
| T2c | 1 |
| T3a | 7 |
| T3b | 3 |
| Primary Gleason grade | |
| Gleason 6 | 6 |
| Gleason 7 | |
| 7 (3+4) | 10 |
| 7 (4+3) | 3 |
| Gleason 8 (5+3) | 1 |
| Gleason 9 (5+4) | 1 |
| Recurrent pathological tumour stage | |
| pT2b | 1 |
| pT2c | 4 |
| pT3a | 3 |
| pT3b | 9 |
| pT3c | 3 |
| pT4a | 1 |

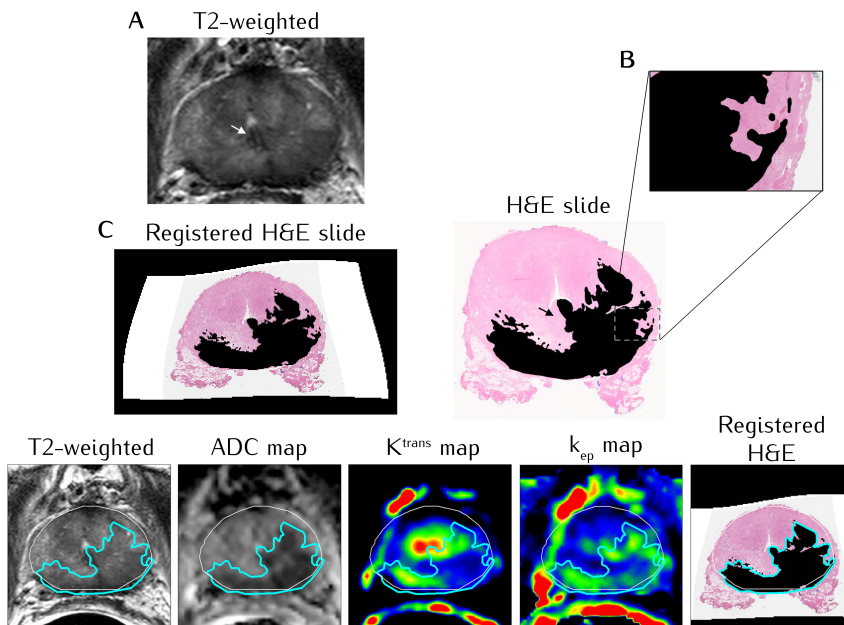


Figure 5.1: Example of an EBRT patient. The arrow points at the urethra, location used for the match of the H&E slide (B) with T2w MRI (A). The registered H&E slide (C) is placed side by side with the mp-MRI images where the prostate is delineated in white and the tumour in blue.

5.3.2 Radiologists' delineations compared to histopathology

Median delineated volumes (cm^3) were 0.34 (range, 0.04–2.14) for radiologist 1 and 0.56 (range, 0.06–3.50) for radiologist 2, and missed tumour foci were smaller than 1 and 0.5 cm^3 , respectively.

Radiologists' contours did not visually overlap with pathology for 5 patients, with 3 of them missed by both. When overlapping with pathology, the median delineated volumes (cm^3) – 0.61 (range: 0.05–1.76) and 0.68 (range: 0.06–2.82) – slightly underestimated the corresponding median pathological tumour volumes (cm^3) – 0.74 (range: 0.01–2.82) and 0.74 (range: 0.02–2.82) – **Figure 5.2A-C**. Contours overlapping with pathology had a median Dice coefficient of 0.64 (range: 0.1–0.8) and 0.58 (range: 0.01–0.83), and a median 95% Hausdorff distance (mm) of 3.6 (range: 0.4–20.9) and 6.2 (range: 0.9–17.1) for radiologist 1 and 2 respectively.

The pattern of recurrence after LDR brachytherapy, showing small and sparse foci (**Figure 5.2D**), together with the artefacts caused by the seeds, made detection of recurrent tumour within the gland challenging. Disease extension to the SV was more readily visible as it presented characteristic tumour features on MRI.

5.3.3 MR Imaging parameters

Diffuse SI reduction throughout the prostate and a decreased conspicuity of zonal anatomy was observed on T2w. For most patients (10/17 with DCE), the PU region showed K^{trans} enhancement without signs of malignancy – **Figure 5.1**. Quantitative imaging values are reported for all ROIs in **Figure 5.3** for EBRT and **Figure S5.1** (Supplementary Materials) for LDR brachytherapy patients. Further information about the values per ROI can be found in **Table S5.1** (Supplementary Materials).

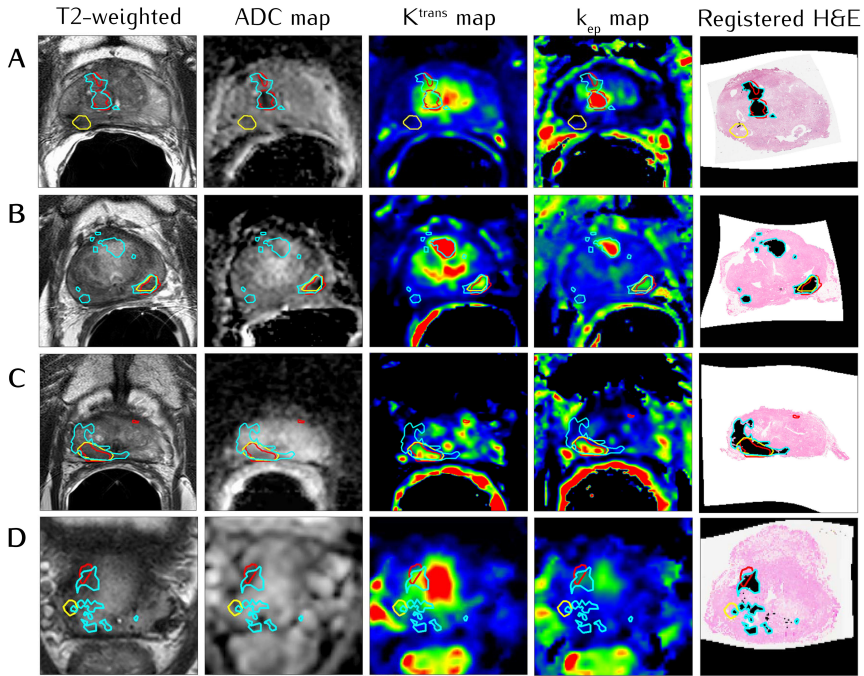


Figure 5.2: Example patients treated with primarily EBRT (A, B and C) and LDR brachytherapy (D) with histopathology delineations propagated to MRI (in blue) and tumour-suspected regions delineated by the experienced uro-radiologists (in yellow and red).

EBRT patients

The Friedman's ANOVA test revealed significant differences between regions. A Wilcoxon signed rank test between tumour and PZ, with a corrected $\alpha = 0.05/6 = 0.008$, found median ADC (PZ = $1.3 \times 10^{-3} \text{ mm}^2/\text{s}$, Tumour = $0.9 \times 10^{-3} \text{ mm}^2/\text{s}$), K^{trans} (PZ = 0.13 min^{-1} , Tumour = 0.23 min^{-1}) and k_{ep} (PZ = 0.32 min^{-1} , Tumour = 0.45 min^{-1}) to be significantly different. No imaging modality was significantly different between tumour and the CG. The PU region had the highest values for all but k_{ep} , with K^{trans} (PU = 0.35 min^{-1}) and ADC (PU = $1.4 \times 10^{-3} \text{ mm}^2/\text{s}$) significantly higher than tumour. For T2w and K^{trans} there was a considerable overlap between tumour and CG values – Figure 5.3. The tumour had the lowest T2w and ADC and the highest k_{ep} values.

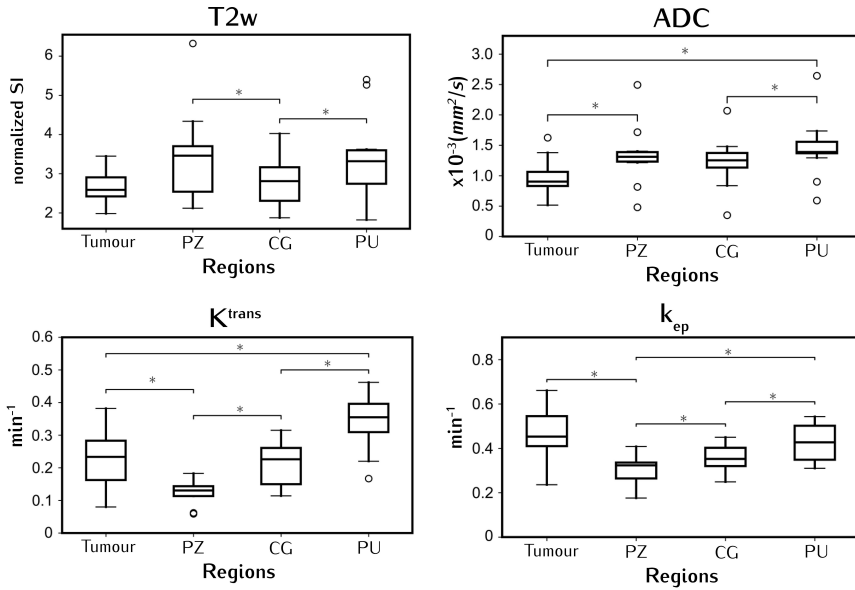


Figure 5.3: Median imaging values for all ROIs of the 13 EBRT patients. The boxes represent the first (25th) and third (75th) quartile; the horizontal line indicates the median and the whiskers the limit $Q_1 - 1.5 \times Q_1$ and $Q_3 + 1.5 \times Q_3$; dots represent outliers. The asterisk denotes significant differences ($p < 0.008$).

LDR brachytherapy patients

For LDR brachytherapy patients, no significant differences were found between ROIs for all imaging modalities – Figure S5.1 (Supplementary Materials). The majority of the foci in the gland were small and did not have typical tumour characteristics on mp-MRI. The DWI was degraded due to the presence of the seeds– Figure 5.4. Similarly to EBRT patients, the PU had the highest values for all but k_{ep} maps. For the seven patients with SV invasion – Table 5.2 – tumour in the SV had the highest values for K^{trans} and k_{ep} and lowest ADC and T2w values.

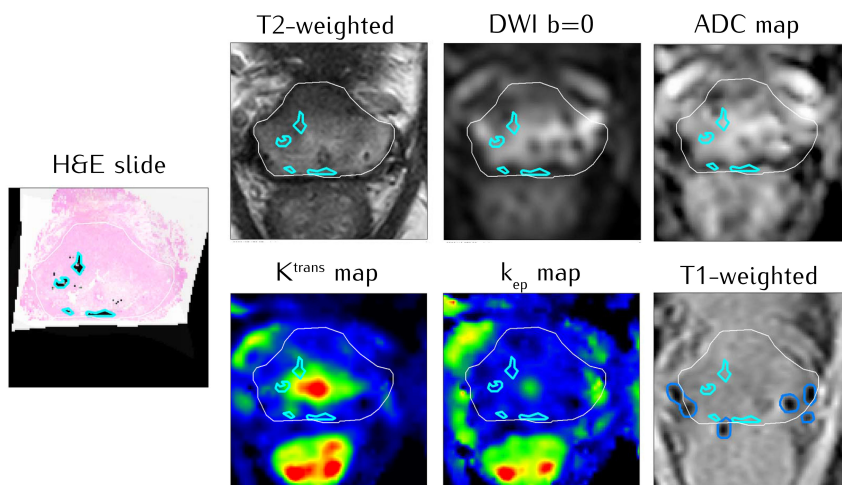


Figure 5.4: Example of a LDR-brachytherapy patient in which the DWI sequence is distorted by the presence of the seeds. The region masked out of the analysis due to the presence of the seeds (in dark blue) was delineated using the T1w gradient echo sequence.

Table 5.2: Median imaging values for tumour and benign ROIs in the seminal vesicles for the LDR-brachytherapy patients presented as median (Q₁₀ – Q₉₀).

| Imaging Parameter | Seminal Vesicles | |
|--|--------------------|--------------------|
| | Tumour | Benign |
| T2w | | |
| Normalised T ₂ values (SI) | | |
| Median | 2.9 (2.4 – 4.1) | 6.6 (4.2 – 8.4) |
| DWI | | |
| ADC (×10 ⁻³ mm ² /s) | | |
| Median | 1.0 (0.9 – 1.9) | 1.7 (1.3 – 3.2) |
| DCE | | |
| K ^{trans} (min ⁻¹) | | |
| Median | 0.20 (0.17 – 0.33) | 0.13 (0.06 – 0.22) |
| k _{ep} (min ⁻¹) | | |
| Median | 0.59 (0.48 – 0.79) | 0.47 (0.29 – 0.55) |

5.3.4 Mixed-modelling

Univariately, all MRI parameters were significantly associated with tumour likelihood, k_{ep} having the best statistical performance (BIC of 239552, compared to 286265, 247260 and 254273 for T2, ADC and K^{trans}). Location (PZ = 0 or CG = 1) was the only non-imaging parameter that improved discrimination, with a voxel location in the CG decreasing tumour likelihood. A multivariate model combining all mp-MRI parameters and location obtained the best fit and was the most predictive for tumour (BIC of 198942, $p < 0.001$) – **Table S5.2** (Supplementary Materials). A calibration curve for this model can be found in the Supplementary Materials (**Figure S5.2**), showing that model predicted probabilities correspond well to the actual fractions of voxels with tumour in histopathology. **Figure 5.5** shows an example of a reconstructed tumour probability map originating from the use of the mp-MRI model.

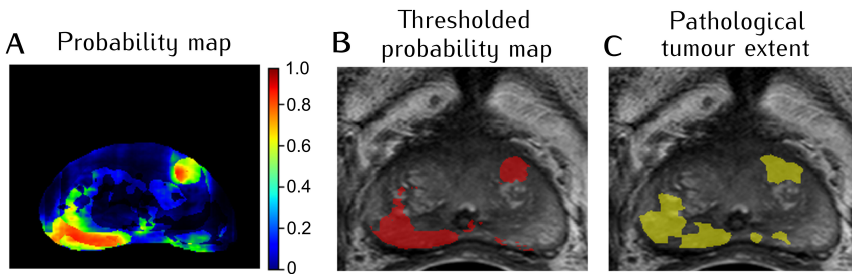


Figure 5.5: (A) Tumour likelihood (with probabilities between 0 and 1) for every prostate voxel, obtained with the use of the multivariate mp-MRI model. (B) Region with probability values above 0.5 (red overlay on the T2w). (C) True pathological extent based on the matched H&E slide (yellow overlay on the T2w).

5.4 DISCUSSION

In selected patients, focal salvage can achieve local control while reducing the risk of adverse events. An accurate representation of tumour location is necessary and mp-MRI has been used for this purpose [139–142]. In a quadrant analysis, T2w alone was reported to result in a similar tumour evaluation accuracy

as in the pre-treatment setting, but with substantial interobserver variability [56]. Benign confounders in previously irradiated prostates can complicate tumour detection. No consensus exists yet on the optimal MRI sequences for detecting local recurrence. Since patients only sporadically receive SRP after RT, and with limited availability of mp-MRI prior to surgery, studies reporting on these patients are usually based on small numbers [45, 55, 56]. Our study is unique in registering pathology to MRI to propagate high resolution tumour delineations.

Consistent with literature, we found radio-recurrence disease to be often multifocal [57] yet with a different recurrence pattern in EBRT and LDR brachytherapy treated patients. No BPH was found in the available pathological slides, suggesting that it is not an imaging confounder as in the untreated prostate.

Our findings have implications for the design of focal salvage treatments. Smaller foci were often missed, and when overlapping with pathology, radiologists' delineations often underestimated the real tumour size. Both radiologists performed quite similarly regarding the evaluated metrics. Our results are in line with findings in the de-novo setting, comparing mp-MRI tumour detection to prostatectomy specimens: Borofsky *et al.* [143] reported that in 26% of the patients clinically important lesions were missed and in 8% there was substantial tumour size underestimation; according to Bratan *et al.* [144] two observers underestimated tumour volume with every pulse sequence and Steenbergen *et al.* [49] described how teams of observers missed all small satellites and parts of the 18/22 correctly detected dominant lesions. In the setting of recurrent PCa, tumour volume underestimation in 1.5T T2w MRI has been described using SRP samples [55]. The 95% Hausdorff distances here reported were above the 2.3mm error observed in primary tumour delineations on mp-MRI [49]. Our results suggest a margin could improve index lesion coverage in focal salvage strategies. The development of MRI scoring and recurrent tumour delineation guidelines would likely improve consistency between observers and studies.

Analysis of the quantitative imaging maps showed that for EBRT patients K^{trans} , k_{ep} and ADC could successfully distinguish tumour in the PZ. Tumour detection in the CG was hampered by the presence of PU enhancement of K^{trans} . Recurrent PCa after LDR brachytherapy was challenging to image due to the

seeds, causing severe distortions in the DWI, and potentially unreliable ADC values. For these patients, no significant differences were found between ROIs, possibly due to the small sample size.

Using mixed-modelling, the mp-MRI combination obtained the best performance in distinguishing tumour and benign tissue on a voxel-level. Location significantly contributed to tumour likelihood, mimicking primary tumour prevalence maps in which the PZ has higher tumour incidence. The addition of DCE parameters improved tumour distinction. The decreased accuracy reported by Donati *et al.* [46] may have been caused by the PU enhancement confounder.

The limited prostate coverage by the H&E slides made it impossible to reconstruct the overall shape of the gland, limiting registration accuracy between the slides and MRI. The main study limitation is the small sample size.

In conclusion, when using MRI to guide focal salvage treatments, all imaging modalities should be used for delineation, and the recommended treatment plan should encompass adequate margins beyond the visible tumour to accommodate for size underestimation. Tumour multifocality should also be accounted for. De-escalated whole gland treatment with a focal tumour boost could potentially fulfil these requisites.

SUPPLEMENTARY MATERIALS

5.5.1 MRI acquisition and post-processing

The T2w scans were acquired with an echo time (TE) between 90 – 130 ms, and a repetition time (TR) between 2600 – 6000 ms. The T1w gradient echo scans were acquired with a TE between 1.8 – 2.3 ms and a TR between 3.6 – 5.3 ms. For one patient the T1w sequence was a turbo spin-echo with TE/TR = 8/572 ms. The DCE sequence was acquired with a TE/TR = 2/4 ms. Contrast injection was followed by a 30 ml saline flush at a flow rate of 3 ml/s. The apparent diffusion coefficient (ADC) maps were derived from the DWI data excluding b-value = 0. In 12 patients the b-values (s/mm^2) = 200 and 800 were used; for 2 the used values were b (s/mm^2) = 500 and 1000 and in 1 patient the values of b (s/mm^2) = 100 and 1000. For 6 patients, up to 4 b-values between 100 – 800 were used. Pharmacokinetic analysis was performed with the standard Tofts model [145], using an average T1 value of 1597 ms [146] and an arterial input function with parameters derived from an in-house study population. DCE and DWI- derived maps were generated using MATLAB R2015a (The MathWorks). Further region of interest (ROI) analysis was performed using Python 3 (Python Software Foundation).

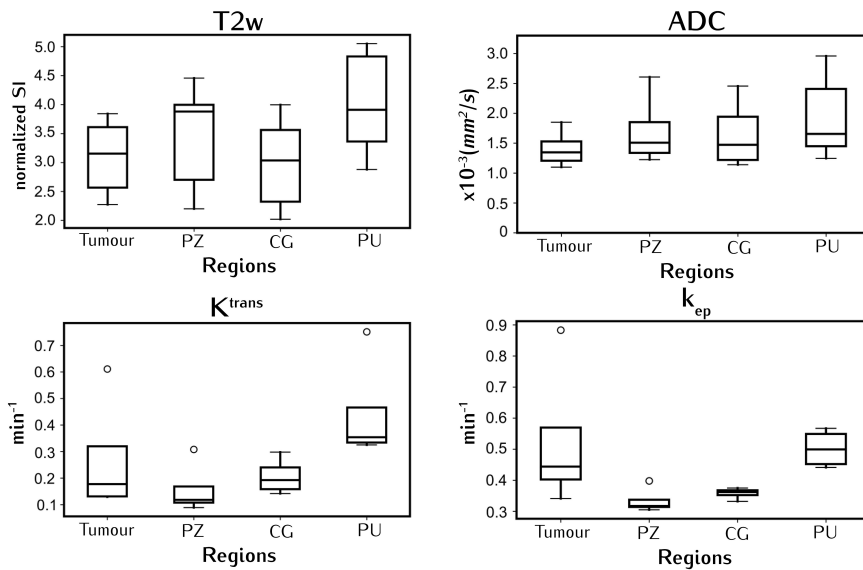


Figure S5.1: Median imaging values for all ROIs of the six LDR brachytherapy patients. The boxes represent the first (25th) and third (75th) quartile; the horizontal line indicates the median and the whiskers the limit $Q_1 - 1.5 \times IQR$ and $Q_3 + 1.5 \times IQR$; dots represent outliers.

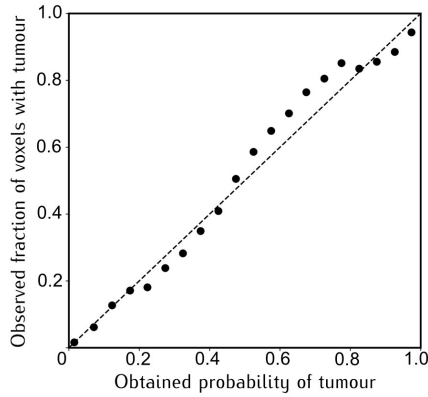


Figure S5.2: Relation between the fraction of voxels with tumour as observed in histopathology and the probability of tumour as obtained from the model. For voxels within a probability interval of 0.05, the fraction of voxels identified as tumour in the histopathology was calculated. For each interval, this fraction was plotted against the averaged predicted probability by the model. The dashed line represents perfect correspondence between observed fraction and predicted probability.

Table S5.1: Median values for both subsets of patients and all ROIs. Median ($Q_{10} - Q_{90}$).

| EBRT | | | | |
|--|--------------------|--------------------|--------------------|--------------------|
| | Tumour | PZ | CG | PU |
| T2w (normalised SI) | 2.6 (2.3 - 3.3) | 3.5 (2.3 - 4.2) | 2.8 (2.0 - 3.7) | 3.3 (2.3 - 4.9) |
| ADC ($\times 10^{-3}$ mm ² /s) | 0.9 (0.7 - 1.4) | 1.3 (0.9 - 1.7) | 1.3 (0.9 - 1.5) | 1.4 (1.0 - 1.7) |
| K ^{trans} (min ⁻¹) | 0.23 (0.15 - 0.33) | 0.13 (0.07 - 0.17) | 0.23 (0.14 - 0.31) | 0.35 (0.23 - 0.45) |
| k _{ep} (min ⁻¹) | 0.45 (0.37 - 0.61) | 0.32 (0.24 - 0.37) | 0.35 (0.27 - 0.44) | 0.43 (0.32 - 0.52) |
| LDR brachytherapy | | | | |
| | Tumour | PZ | CG | PU |
| T2w (normalised SI) | 3.2 (2.4 - 3.7) | 3.9 (2.3 - 4.2) | 3.0 (2.1 - 3.9) | 3.9 (3.1 - 5.0) |
| ADC ($\times 10^{-3}$ mm ² /s) | 1.3 (1.1 - 1.7) | 1.5 (1.3 - 2.3) | 1.5 (1.2 - 2.3) | 1.7 (1.3 - 2.8) |
| K ^{trans} (min ⁻¹) | 0.18 (0.13 - 0.49) | 0.12 (0.10 - 0.25) | 0.19 (0.15 - 0.27) | 0.35 (0.33 - 0.64) |
| k _{ep} (min ⁻¹) | 0.44 (0.37 - 0.76) | 0.32 (0.31 - 0.37) | 0.36 (0.34 - 0.37) | 0.50 (0.45 - 0.56) |

Table S5.2: Model parameters obtained when combining T2w, ADC, K^{trans} and k_{ep} imaging values to predict tumour at the voxel-wise level. Imaging values and location are included as fixed effects; patient and spatial coordinates as random effects.

| Fixed effects (MRI) | Regression coefficients (β) | Std. Error | P |
|--|-------------------------------------|----------------|----------------|
| T2w | -0.29 | 0.01 | <0.001 |
| ADC reference group ($0 - 0.924 \times 10^{-3}$) | 0 | – | – |
| ADC 2 nd quartile ($0.925 \times 10^{-3} - 1.222 \times 10^{-3}$) | -1.70 | 0.02 | < 0.001 |
| ADC 3 rd quartile ($1.223 \times 10^{-3} - 1.438 \times 10^{-3}$) | -2.58 | 0.02 | < 0.001 |
| ADC 4 th quartile ($1.439 \times 10^{-3} - 3.825 \times 10^{-3}$) | -3.35 | 0.03 | < 0.001 |
| K^{trans} reference group ($0 - 0.120$) | 0 | – | – |
| K^{trans} 2 nd quartile ($0.121 - 0.160$) | 0.30 | 0.02 | 0.594 |
| K^{trans} 3 rd quartile ($0.161 - 0.248$) | 1.70 | 0.02 | <0.001 |
| K^{trans} 4 th quartile ($0.249 - 2.494$) | 2.26 | 0.03 | <0.001 |
| k_{ep} | 4.95 | 0.05 | <0.001 |
| Location | -1.44 | 0.02 | <0.001 |
| Intercept (β_0) | -2.35 | 0.41 | <0.001 |
| Random effects | Variance (σ^2) | | |
| Patients | 2.43 | | |
| Distance CM_x | 0.71 | | |
| Distance CM_y | 0.38 | | |
| Distance CM_z | 0.35 | | |
| Residuals (ϵ_0) | Median | First quartile | Third quartile |
| | -0.128 | -0.240 | -0.06 |

MULTI-PARAMETRIC MRI TUMOUR
PROBABILITY MODEL FOR THE DETECTION
OF LOCALLY RECURRENT PROSTATE CANCER
AFTER RADIATION THERAPY: PATHOLOGIC
VALIDATION AND COMPARISON WITH MANUAL
TUMOUR DELINEATIONS

Catarina Dinis Fernandes¹
Rita Simões¹
Ghazaleh Ghobadi¹
Stijn WTPJ Heijmink²
Ivo G Schoots²
Jeroen de Jong³
Iris Walraven¹
Henk G van der Poel⁴
Petra J van Houdt¹
Milena Smolic¹
Floris J Pos¹
Uulke A van der Heide¹

¹Department of Radiation Oncology, ²Radiology and ⁴Urology,
The Netherlands Cancer Institute, Amsterdam, The Netherlands

³Department of Pathology, Reinier Haga Medisch Diagnostisch Centrum,
Delft, The Netherlands

ABSTRACT

Purpose: Focal salvage treatments of recurrent prostate cancer (PCa) after radiation therapy require accurate delineation of the target volume. Magnetic resonance imaging (MRI) is used for this purpose; however radiation therapy-induced changes complicate image interpretation, and guidelines are lacking on the assessment and delineation of recurrent PCa. A tumour probability (TP) model was trained and independently tested using multi-parametric MRI (mp-MRI) of patients with radio-recurrent PCa. The resulting probability maps were used to derive target regions for radiotherapy treatment planning.

Methods and Materials: Two cohorts of patients with radio-recurrent PCa were used in this study. All patients underwent mp-MRI (T2 weighted, diffusion-weighted imaging, and dynamic contrast-enhanced). A logistic regression model was trained using imaging features from 21 patients with biopsy-proven recurrence who qualified for salvage treatment. The test cohort consisted of 17 patients treated with salvage prostatectomy. The model was tested against histopathology-derived tumour delineations. The voxel-wise TP maps were clustered using k-means to generate a gross tumour volume (GTV) contour for a voxel-level comparisons with manual tumour delineations performed by 2 radiologists and with histopathology validated contours. Later, k-means was used with 3 clusters to define a clinical target volume (CTV), high-risk CTV and GTV, with increasing tumour risk.

Results: In the test cohort, the model obtained a median (range) area under the curve of 0.77 (0.41–0.99) for the whole prostate. The GTV delineation resulted in a median sensitivity of 0.31 (0–0.87) and specificity of 0.97 (0.84–1.0) with no significant differences between model and manual delineations. The 3-level clustering GTV and high-risk CTV delineations had median sensitivities of 0.17 (0–0.59) and 0.49 (0–0.97) and specificities of 0.98 (0.84–1.00) and 0.94 (0.84–0.99), respectively.

Conclusions: The TP model obtained a good performance in predicting voxel-wise presence of recurrent tumour. Model-derived tumour risk-levels achieved sensitivity and specificity similar to manual delineations in localising recurrent

tumour. Voxel-wise TP derived from mp-MRI can in this way be incorporated for target definition in focal salvage of radio-recurrent PCa.

6.1 INTRODUCTION

Radiation therapy is an important treatment modality for primary prostate cancer (PCa). The 5-year disease-free survival of patients with PCa treated with external-beam radiation therapy (EBRT) to 78 Gy varies from 67 to 80% depending on the risk group [58]. Although most of the patients with recurrent disease will bear metastases, a significant proportion will harbour locally recurrent disease only [26]. Because of the toxicity rates associated with whole-gland salvage treatments, and given the pathological evidence that recurrence occurs predominantly at the site of the primary index lesion [55], focal salvage strategies are emerging for locally recurrent PCa without metastatic spread. This could limit toxicity. The extent of ablation varies from focal with a specified margin around the lesion, to hemi or zonal [147]. A variety of techniques, such as cryotherapy [148], HIFU [139], Iodine-125 (low dose rate [LDR]) [53] and high dose-rate brachytherapy [54] obtained control rates in-line with biochemical disease-free survival outcomes from whole-gland approaches.

Accurate detection and delineation of tumours is important for the design of focal treatments. Positron emission tomography (PET) with ^{68}Ga prostate-specific membrane antigen (PSMA) ligands is used to identify metastatic disease and offers a high sensitivity to detect recurrent PCa [13]. Studies of focal salvage treatments frequently use multi-parametric MRI (mp-MRI) together with biopsy confirmation to define the region to treat [33]. High-resolution anatomical and functional imaging makes MRI attractive for guiding focal salvage treatments. In the setting of recurrent PCa after radiation therapy, benign MRI confounders are present and can complicate image interpretation and tumour detection [37]. The Prostate Imaging Reporting and Data System (PI-RADS) version 2 guidelines do not apply in the recurrent setting [42], and to date, no consensus guidelines exist on how to delineate recurrent PCa.

Automated methods could aid radiation oncologists in delineating the recurrent tumour. As an alternative to manual contouring, automated methods were developed for primary PCa in which voxel-wise tumour probability (TP) maps, representing the probability of tumour presence in each voxel, were derived from MR imaging information [62, 149–152].

The purpose of this study was to train and test a TP model for the detection and delineation of recurrent PCa. We used the model structure described by Dinh *et al.* [149], which we trained and validated on 2 independent cohorts of patients with recurrent PCa after radiation therapy. The performance of the model in detecting and delineating recurrent PCa was compared to histopathology and with manual tumour delineations.

6.2 METHODS AND MATERIALS

6.2.1 Patients

Two independent cohorts were used in this study. The training cohort consisted of 21 patients with biopsy-proven recurrent PCa after EBRT treatment for primary PCa who qualified for salvage treatment. All patients were imaged more than 24 months after completion of EBRT and none received hormonal therapy (HT) in the year before imaging. This was a subgroup of a prospectively collected cohort [153] that had given informed consent prior to participation in the study. The test cohort consisted of retrospectively collected data from 17 patients treated with salvage radical prostatectomy for recurrent PCa after radiotherapy [127]. Written informed consent was waived by the institutional review board. None of the patients received HT at the time of imaging. All training and test patients received mp-MRI for the diagnosis of recurrent disease.

6.2.2 Image acquisition

The training cohort patients were scanned on a 3T Achieva dStream (n=3) or Ingenia (n=18) MRI scanner (Philips Healthcare) between September 2016 and June 2018. With the exception of 4 patients (3 scanned in the Achieva dStream), all were scanned with the use of an endorectal coil. The test cohort patients were scanned on a 3T Achieva (n=12), Achieva dStream (n=3)

Table 6.1: Sequence specific parameters for the two cohorts.

| | Training cohort | Test cohort |
|---------------------------------|---------------------------------------|------------------|
| tT2w | | |
| Sequence | turbo spin-echo | |
| Pixel size/slice thickness (mm) | 0.27/3.0 | 0.3- 0.8/3.0 |
| TR/TE (ms) | 3143/120 | 2600-6000/90-130 |
| DWI | | |
| Sequence | single-shot spin-echo echo-planar | |
| Pixel size/slice thickness (mm) | 1.03/2.7 | 2.73/4.0 |
| TR/TE (ms) | 3131/55 | 7110-9900/67 |
| b-Values (s/mm^2) | 0, 200, 800 | 0, 100 - 2000 |
| DCE | | |
| Sequence | 3D T1w spoiled gradient echo sequence | |
| Pixel size/slice thickness (mm) | 1.16/6.0 | 1.37-1.68/4.0 |
| TR/TE (ms) | 4/1.9 | 2/4 |
| Dynamic scan time (s) | 2.6 | 2.3-2.5 |

or Ingenia (n=2) MRI scanner between June 2011 and March 2017. Eleven patients were scanned with an endorectal coil in addition to a torso or cardiac phased array coil. In both cohorts, the mp-MRI protocol consisted of transverse, coronal and sagittal T2-weighted (T2w) turbo spin-echo, axial T1-weighted (T1w) gradient echo, diffusion-weighted imaging (DWI), and dynamic contrast enhanced (DCE) MRI. **Table 6.1** describes sequence specific parameters. Apparent diffusion coefficient (ADC) maps were derived from the DWI data using a mono-exponential fit, excluding b-value = 0. In the training cohort the ADC map was calculated using the values $b = 200$ and $800 s/mm^2$, whereas in the test cohort the values used varied on a patient level between $100 - 1000 s/mm^2$. The pharmacokinetic maps K^{trans} and k_{ep} were derived from the DCE scans with the Tofts model using a population-based arterial input function [145, 154].

To minimise the impact of different acquisition protocols, the T2w, K^{trans} , and k_{ep} values were normalised to the median value in the peripheral-zone. The T2w scans do not provide quantitative information and, although K^{trans} and k_{ep} are in theory quantitative, the repeatability of the exams is challenging. Normalisation was not necessary for the ADC maps.

6.2.3 Ground-truth tumour delineations

The ground-truth tumour delineations in the training cohort were defined on mp-MRI based on the radiologic report, biopsy and PSMA-PET findings [153]. In the test cohort, the ground-truth tumour delineations were made on the digitalised hematoxylin and eosin (H&E) slides at a resolution of 0.5 $\mu\text{m}/\text{pixel}$. The first step of the registration between H&E slides and the transversal T2w MRI was performed visually, assigning a T2w slice to each H&E slide. Each H&E slide was then registered to its matched T2w slice using deformable registration based on landmark points (Coherent Point Drift) [137] implemented in MATLAB R2015a (The MathWorks). The ground-truth tumour delineations were defined based on the annotated pathology slides and tumour contours propagated to MRI [127]. All images were registered to the T2w scan, and all maps were resampled to the grid of the T2w scans in the training cohort (voxel size = 0.27 x 0.27 x 3mm³). Per patient, the dataset consisted of ADC, normalised T2w, K^{trans} and k_{ep} images, plus tumour and prostate delineations.

6.2.4 Tumour probability model

In total, 31 intensity and textural-based image features were extracted from the transversal T2w, ADC, K^{trans} and k_{ep} as described in [149]. For each voxel, the intensity values of the ADC and normalised T2w, K^{trans} and k_{ep} were included. From the T2w images, the extracted textural features consisted of Gaussian derivatives up to second order at 4 exponentially increasing scales ($\sigma = 1.5, 2.4, 3.8, 6$ mm), summing up to 6 features per scale. Normalised multiscale blobness textural features were extracted from the ADC, K^{trans} , and k_{ep} images over the same 4 scales, resulting in 1 feature per image. It is known that the use of HT causes a homogenisation of the signal across the prostate. To investigate if the use of HT had an effect on the normalised T2w signal intensity (SI), the features from patients with and without previous HT treatment were compared in both the training and test set.

As described in the same publication, a logistic regression model was fitted to the data to derive a per-voxel TP for the presence of tumour. This same methodology was used in the current study to derive TP maps for the presence

of recurrent PCa. Prior knowledge, such as biopsy or prevalence maps, was not used. Time since radiation therapy was also not included because in previous research this parameter did not significantly improve voxel-wise recurrent tumour discrimination [127].

The importance of every feature was assessed by ranking it in the training set using the Mahalanobis distance [155] between the 2 classes (benign vs tumour) as a ranking criterion. The model coefficients were retrieved. MATLAB R2015a (The MathWorks) was used for feature extraction and to develop the tumour probability model.

6.2.5 From probabilities to risk-levels

The model output was a voxel-wise TP map, permitting the use of a binary threshold to define the tumour region (gross tumour volume [GTV]). However, instead of defining a binary threshold for tumour presence, all voxel-wise probability information can be incorporated [122]. A more moderate solution is to stratify the probability maps in 3 levels of low-, intermediate- and high-risk for the presence of tumour. To make the connection with radiation therapy and to define understandable treating volumes, the 3 risk categories can be labeled clinical target volume (CTV), high-risk CTV and GTV [156, 157]. In this study we analysed both approaches by applying a k-means clustering algorithm to partition the data. The k-means algorithm was used with 2 clusters to define a binary GTV delineation and surrounding CTV, and with 3 clusters to identify the CTV, high-risk CTV and GTV. The squared arithmetic distance was the metric used for minimisation. The clustering procedure was repeated 3 times with new randomly selected initial cluster centroid positions, and the solution with the lowest within-cluster sums of point-to-centroid distances was selected. The 3 clusters were sorted in ascending order according to their mean TP values, defining the CTV, high-risk CTV and GTV.

6.2.6 Comparison between manual- and model-derived tumour delineations

On the test cohort, manual delineations of tumour suspected regions were independently performed by 2 uro-radiologists (14 and 7 years of experience) based on the mp-MRI and radiologic report, choline PET scans, and biopsy information [127]. Because PI-RADS v2 [42] is not applicable to recurrent PCa, tumour was defined as a region with low SI on T2w MRI, high SI on the $b = 800$ DWI scan or otherwise comparable b -value DWI scan, low SI on the ADC map, and increased enhancement in the K^{trans} and k_{ep} maps.

The model-defined binary GTV delineations and the delineations by the 2 radiologists were compared to the histopathology-defined ground-truth tumour delineations. A voxel-wise assessment of true positives and true negatives obtained by each delineated region resulted in per-patient specificity and sensitivity values.

6.2.7 Statistics

Clinical characteristics of primary and recurrent tumours were compared between cohorts. For the statistical analysis, and due to the small number of patients in some of the highest and lowest categories, the TNM stage was grouped in 2 (T1+T2 and T3+T4) and the Gleason score in 3 groups (Gleason score ≤ 6 , 7, and ≥ 8). Continuous variables (prostate-specific antigen [PSA]) were compared using a t -test and categorical variables (Gleason score and TNM stage) with a chi-square test. Significance was set at $p < 0.05$. In the test cohort, a Wilcoxon signed rank test was used to compare the sensitivity and specificity values between radiologists and against model-derived tumour delineations. The significance value was Bonferroni corrected by the number of tests. When performing comparisons using the binary GTV delineation the number of tests was 6 with $p = 0.0083$, while for the 3-level risk $n=10$, resulting in a $p=0.005$. Statistical tests were performed with the Statistical Package for Social Sciences, version 22.0 (SPSS, Chicago, IL, USA).

The area under the curve (AUC) values were determined on a voxel-by-voxel basis. For the training cohort, individual AUC values were obtained by using leave-one-patient-out cross validation. For the test cohort, the model

was individually tested and an AUC value was obtained per patient. Population median and ranges are reported.

6.3 RESULTS

6.3.1 Patient characteristics

Patient characteristics from both cohorts are reported in **Table 6.2** for primary PCa and **Table 6.3** for recurrent disease. The test cohort consisted of both EBRT (n=13) and LDR brachytherapy-treated (n=4) patients. Characteristics of primary PCA were not significantly different between cohorts (TNM stage, $p=0.75$; Gleason score, $p=0.319$; PSA, $p=0.095$; and use of HT, $p=0.293$). However, recurrent tumour PSA levels and TNM stage were significantly higher in the test cohort (TNM stage, $p=0.003$; PSA, $p=0.014$). Gleason score was not assigned to recurrent tumour [138].

Summary statistics for the values of features in both training and test cohorts can be found in the **Table S6.1** in the Supplementary Materials. The median (5th – 95th percentile) T2w SI in the training cohort was 0.89 (0.49 – 1.35) for patients who had previously received HT and 0.88 (0.49-1.31) for those without HT. In the test cohort, the median T2w SI was 0.92 (0.49 – 1.46) for patients who had previously had HT treatment and 0.95 (0.47 – 1.45) for those without HT.

Table 6.2: Patient characteristics for primary prostate cancer.

** For one patient in training cohort no information was available about Gleason score or PSA levels of primary tumour.*

| Primary PCa | Training cohort | Test cohort |
|-----------------------------|------------------------|--------------------|
| T stage | | |
| T1 | - | 2 |
| T2 | 10 | 5 |
| T3 | 10 | 10 |
| T4 | 1 | - |
| Gleason score * | | |
| 5 | 1 | - |
| 6 | 3 | 3 |
| 3+4 | 7 | 9 |
| 4+3 | 3 | 3 |
| 8 | 4 | 1 |
| 9 | 2 | 1 |
| PSA levels (ng/ml) * | | |
| <10 | 8 | 4 |
| 10 ≤ PSA <20 | 6 | 6 |
| 20 ≤ | 6 | 7 |
| Hormonal therapy | | |
| Yes | 17 | 11 |
| No | 4 | 6 |

Table 6.3: Patient characteristics for recurrent prostate cancer.

[†] For two patients in the training cohort no information about TNM stage of recurrent tumour was available.

| Recurrent PCa | Training cohort | Test cohort |
|-----------------------------|------------------------|--------------------|
| T stage [†] | | |
| T1 | - | - |
| T2 | 15 | 5 |
| T3 | 4 | 11 |
| T4 | - | 1 |
| Gleason score | | |
| 5 | - | - |
| 6 | - | - |
| 3+4 | - | - |
| 4+3 | - | - |
| 8 | - | - |
| 9 | - | - |
| PSA levels (ng/ml) | | |
| <10 | 20 | 14 |
| 10 ≤ PSA <20 | 1 | 3 |
| 20 ≤ | - | - |
| Hormonal therapy | | |
| Yes | - | - |
| No | - | - |

Table 6.4: Feature ranking. $G_i(s)$ stands for the Gaussian derivative in the direction i , with standard deviation s (in mm).

| Rank | Feature name |
|------|------------------------------|
| 1 | k_{ep} - Intensity |
| 2 | k_{ep} - Blobness |
| 3 | K^{trans} - Blobness |
| 4 | K^{trans} - Intensity |
| 5 | ADC - Blobness |
| 6 | ADC - Intensity |
| 7 | T2w - Texture, $G(3.8)$ |
| 8 | T2w - Texture, $G(6)$ |
| 9 | T2w - Texture, $G(2.4)$ |
| 10 | T2w - Texture, $G_{xx}(1.5)$ |

6.3.2 TP model performance

Training cohort

Based on the MRI-based delineations used as ground-truth, the tumour prevalence in the training cohort was 6%. The median (range) tumour volume was 0.70 (0.11-4.15) cm^3 . **Table 6.4** lists the top 10 ranked features. The 4 highest ranking features were perfusion based, with k_{ep} ranking first. The model coefficients can be found in **Table S6.2** in the Supplementary Materials.

The TP model achieved a median AUC = 0.94 (range, 0.62-0.99). The binary GTV delineation obtained a median (range) sensitivity of 0.50 (0.17-0.90) and specificity of 0.98 (0.89-1.00). The GTV delineation had a median (range) volume of 0.83 (0.33-3.87) cm^3 .

Test cohort

Based on the histopathology ground-truth, the tumour prevalence in the test cohort was 9%. The median (range) histopathology tumour volume was 0.67 (0.07-2.93) cm^3 .

The TP model obtained a median AUC = 0.77 (range, 0.41–0.99). The binary GTV delineation resulted in a median (range) sensitivity of 0.31 (0–0.87) and specificity of 0.97 (0.84–1.00). An example of the input mp-MRI and the resulting TP and risk-level maps is illustrated in **Figure 6.1** for a patient in the test cohort. In 1 patient, the tumour was entirely missed by the model-defined GTV region. **Table S6.3** in the Supplementary Materials reports the sensitivity and specificity values for all individual patients. The model-defined GTV had a median (range) volume of 0.65 (0.07–2.62) cm³. The calibration (**Figure S6.1**) and ROC curves (**Figure S6.2**) for the training and test cohorts can be found in the Supplementary Materials. The calibration curve of the test set revealed that for low probabilities the model underestimated the fraction of tumour voxels and for high probabilities it overestimated the prevalence of tumour.

The regions defined by the 2 radiologists obtained a median (range) sensitivity of 0.33 (0–0.76) and 0.34 (0–0.91) and specificity of 0.99 (0.93–1.00) and 0.99 (0.72–1.00). The sensitivity and specificity values were not significantly different between the 2 radiologists. In total, the radiologists missed the tumour in 4 patient, and 2 of these were missed by both experts. One of these tumours was also missed by the model-defined GTV region. No significant differences were found between the sensitivity and specificity values of model and manual delineations. The delineations performed by the 2 radiologists had a median (range) volume of 0.17 (0–2.22) and 0.44 (0–3.52) cm³.

Three-level risk map

The 3-level risk map defined a CTV, high-risk CTV and GTV. In the training cohort, the GTV resulted in a median (range) sensitivity of 0.26 (0.06–0.69) and specificity of 0.99 (0.96–1.00), whereas the region encompassed by the high-risk CTV resulted in a median sensitivity of 0.71 (0.28–0.98) and specificity of 0.95 (0.78–0.99). The delineated GTV had a median (range) volume of 0.54 (0.09–2.23) cm³, and the high-risk CTV of 1.78 (0.57–7.71) cm³.

In the test cohort, and compared to the histopathology-derived ground-truth delineations, the GTV resulted in a median (range) sensitivity of 0.17 (0–0.59) and specificity of 0.98 (0.84–1.00), whereas the high-risk CTV delineation obtained a median sensitivity of 0.49 (0.00–0.97) and specificity of 0.94

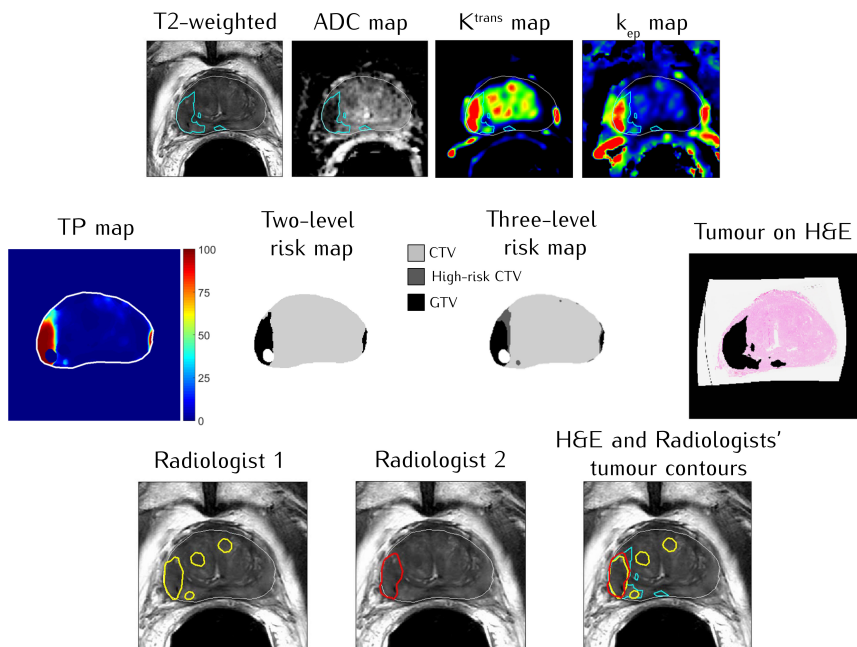


Figure 6.1: Example patient from the test cohort with input MR images, tumour probability map, and defined risk-regions. In the first row, are the the mp-MRI scans with the prostate delineation in white and the ground-truth tumour delineation in cyan. In the second row, from left to right, a TP map, the two- and three-level risk maps and the histopathology section. In the third row, the individual delineations from the first (yellow) and second (red) radiologists are displayed, as well as an overlap of both delineations with the ground-truth tumour delineation (bottom right corner). The fiducial markers were removed from the analysis, explaining the hole in the tumour probability map.

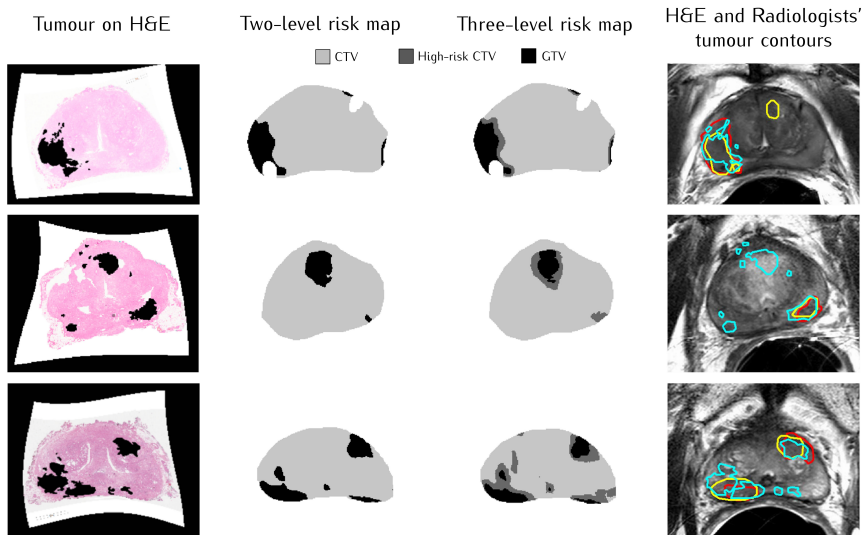


Figure 6.2: Visual representation of the results obtained for three patients in the test cohort. From left to right: the hematoxylin and eosin (H&E) slides, the model-derived two- and three-level risk maps, and the corresponding radiologist's delineations for 3 patients in the test cohort. The yellow and red contours represent the delineations by the 2 radiologists, and the cyan is the ground-truth tumour delineation.

(0.84-0.99). In 1 patient the tumour was entirely missed by the model-defined GTV and by the high-risk CTV. The delineated GTV had a median (range) volume of 0.27 (0.02-2.62) cm³ and the high-risk CTV of 1.20 (0.07-2.67) cm³. With the exception of the specificity values between the high-risk CTV and the first radiologist ($p < 0.001$), no significant differences were found between the sensitivity and specificity of model-derived and manual delineations.

Figure 6.2 illustrates the H&E slides next to the model-derived 2- and 3-level risk maps as well as the corresponding radiologist's delineations for 3 patients in the test cohort.

6.4 DISCUSSION

Accurate detection and localisation of the recurrent tumour is essential for the design of focal salvage treatments. However, to date no guidelines exist on how to interpret mp-MRI of recurrent PCa after radiation therapy. To deal with this lack of radiological guidelines, we propose the use of a TP model with different risk levels. Radiation therapy using dose painting by numbers, in which higher doses are given to voxels with a higher TP, is a possible strategy for focal salvage treatments. Our model fits this idea as it provides a probabilistic estimate of tumour presence. To translate the model outcome to a practical input for radiation therapy, the probability values were clustered to define 3 risk levels for tumour presence: CTV, high-risk CTV and GTV. Although variable treatment schemes are reported in the literature, brachytherapy offers the possibility to provide an ultra-conformal treatment while permitting dose escalation. Doses for LDR brachytherapy range from 120 to 145 Gy for ¹²⁵I, and 90 to 113 Gy for ¹⁰³Pd [158]. The 3-level risk map derived from our MRI-based voxel-wise TP model could be used to target the different regions with varying doses. Based on salvage LDR brachytherapy studies [53, 142, 159] the GTV could be prescribed to a dose of 144 Gy, the high-risk CTV could function as a margin around the GTV for decreasing treatment dose, while ensuring the CTV would receive a dose of 108 Gy.

The TP model developed here was tested on an independent cohort with accurate histopathology-based ground-truth and obtained an AUC = 0.77 for the voxel-wise prediction of recurrent tumour presence. According to [160], the AUC value of our model can be interpreted as good diagnostic accuracy ($0.7 < \text{AUC} < 0.8$). To our knowledge no studies report voxel-wise sensitivity and specificity values for recurrent PCa detection. Using T2w + DWI + DCE and biopsy as ground-truth, Luzurier *et al.* [47] used 2 teams of radiologists to assign a Likert score to each prostatic sextant for the likelihood of recurrent tumour after radiotherapy. The AUC varied between 0.898 and 0.946 on a lobe-level and 0.853 and 0.905 on a sextant-level. Many studies have derived models for the detection of primary PCa, with reported AUC values ranging from 0.61 to 0.96 [62, 149, 150, 152, 161]. The AUC value obtained by our model is well within the interval of values reported for primary PCa detection.

In our model, the top ranking features were perfusion based. This suggests an increased importance of DCE-derived parameters in the recurrent setting when compared with primary tumour detection. When assessing primary PCa according to the PI-RADS v2 guidelines, DCE-MRI has a limited role compared to T2w and DWI [42] since the presence of benign prostatic hyperplasia (BPH) in the central-gland is a major imaging confounder. On the other hand, identifying recurrent disease with DCE-MRI is easier than primary PCa detection because of the different enhancing patterns between tumour and benign tissue [37]. In addition, BPH was not found in the analysed histopathology of irradiated prostates [127]. Nonetheless, attention should be given to the increased enhancement around the peri-urethral tissues in K^{trans} and k_{ep} maps without evident signs of malignancy [153]. T2w-derived features were found to be considerably less important than in the pre-treatment setting, potentially as a result of the overall decrease in T2 values after radiation therapy [37]. The previous use of HT up to 1 year prior to the MRI examination did not seem to have an impact on the T2w signal intensity in patients after radiation therapy.

The voxel-level sensitivity and specificity in detecting recurrent tumour was compared between model-defined binary GTV contours and the delineations by 2 expert uro-radiologists. The sensitivity and specificity of the high-risk CTV and GTV delineations obtained when clustering the data in 3 risk-levels for tumour presence were also assessed. All of the methods obtained comparable and very high specificity and low sensitivity. We observed a wide range of

sensitivity values for the model in both training and test cohorts, as well as with radiologist-made delineations. These results suggest that some tumours are intrinsically hard to detect. As reported by Greer *et al.* [162] in primary PCa, the combination of a computer-aided diagnosis (CAD) system with expert knowledge might result in improved sensitivity also in recurrent PCa. Nevertheless, this sensitivity improvement was traded against a decrease in specificity due to CAD-induced false positives. Because recurrent PCa tends to occur at the location of the primary tumour [55] another way to improve the model sensitivity and remove false positives could be to include the pre-treatment MRI scans or the location of the primary tumour. In addition, the high sensitivity and specificity of PSMA-PET in detecting recurrent PCa (9) suggests that adding this imaging modality would improve model performance. To suppress central-gland confounders, a prevalence map was included in the model by Dinh *et al.* [149]. After radiation therapy, confounders such as BPH are expected to have vanished, and the benefit of including a prevalence map for recurrent PCa detection will likely not outweigh the decrease in sensitivity in the central-gland. Since the training cohort patients received MRI-transrectal ultrasound fusion biopsies, the information provided by a biopsy map was not independent from the MRI; thus these were not included as features.

Despite having similar feature characteristics, the 2 cohorts had different definitions of ground-truth, as seen in the calibration curve of the test data set, where for low values of TP the actual fraction of tumour was higher than the model prediction. This is consistent with a training based on manual delineations, where poorly visible cancers are often missed entirely. For high values of TP, we noticed that the actual fraction of tumour was lower than predicted. This is also consistent with a manual delineation that is typically generous and overestimates the size of a visible lesion. Using a salvage prostatectomy cohort for training of the model would likely overcome the bias introduced by manual delineations. However, because of high treatment-related toxicity, salvage prostatectomy is not commonly performed and retrieving such a data set is practically very difficult. We have chosen to use the available salvage prostatectomy cohort to test the model because in this way it was possible to compare the model results with histopathology ground-truth, resulting in more solid conclusions about the model performance. Other study limitations include the relatively small number of patients and that the training and test cohorts had different recurrent tumour characteristics. Although power calculations for

binary classifications are possible [163], our problem is not a classification on a patient level, since all patients have a recurrence. Power calculations on voxel level would disregard the fact that the many voxels are not necessarily independent. Our validation in an independent test cohort of patients did show a decrease in AUC from 0.94 to 0.77. Although this can, to some extent, be explained by overfitting, it must also be attributed to the use of manual delineations as ground-truth in the training cohort, compared to whole-mount section histopathology in the test cohort. In the test cohort, tumour prevalence was biased because only slices with histopathology validation were used. Lastly, despite striving for the best possible match between histopathology and imaging, registration errors arising from this process result in some uncertainty in the exact boundary between tumour and benign tissue. The registration can be complicated when prostate deformations caused by imaging with an endorectal coil are present. However, we observed that by using deformable image registration, the matches between H&E and T2w slices were overall successful. As reported in the original publication describing this salvage prostatectomy cohort [127], the in-plane average registration error was for all slides of 1.1 ± 0.9 mm. In the present study, registration inaccuracies were not taken into account. This could mean that we might be underestimating the AUC values compared to when registrations errors are compensated (eg, by accounting neighbouring voxels). When designing radiation therapy treatments this problem is mitigated because the limited steepness of the dose fall-off means that a voxel next to a suspected tumour voxel would still receive a considerable dose.

6.5 CONCLUSION

This study demonstrates that it is feasible to use mp-MRI information to derive target volumes for focal salvage treatments in radio-recurrent PCa. An independently validated TP model based on mp-MRI obtained sensitivity and specificity values comparable to those of manually-defined regions in contouring recurrent tumour. Thus, TP maps could potentially be used as an input for the planning of focal salvage treatments.

SUPPLEMENTARY MATERIALS

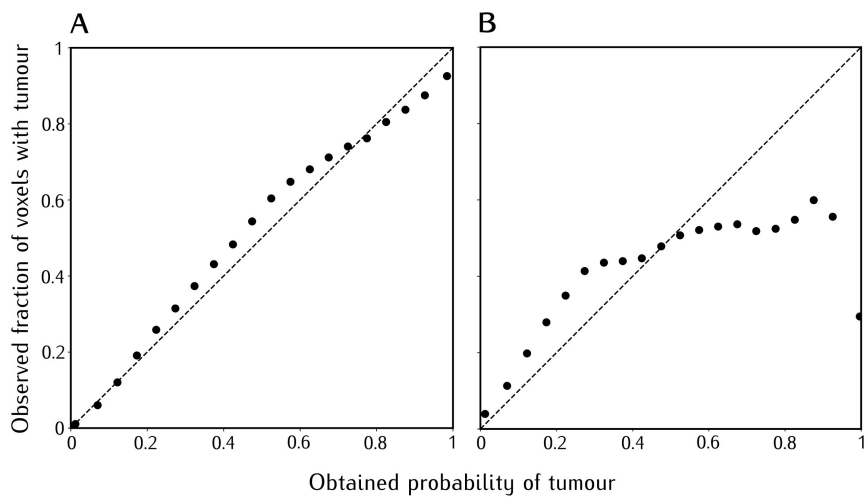


Figure S6.1: Calibration plots in the train (A) and test (B) cohorts.

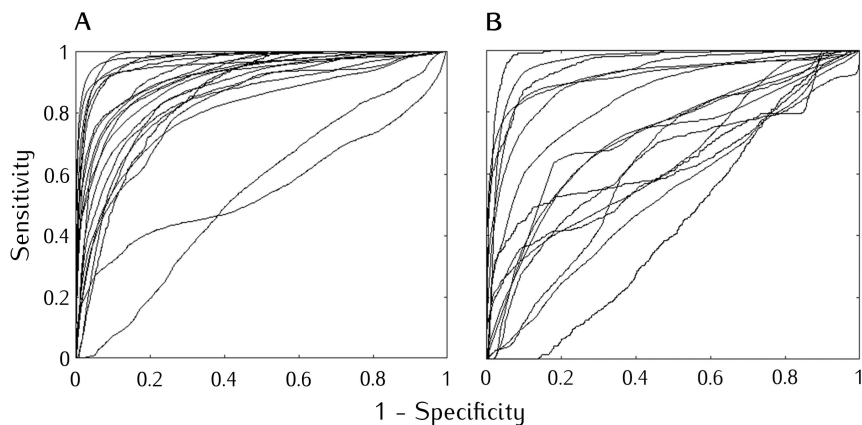


Figure S6.2: Receiver operating characteristic (ROC) curves for each individual patient in the training (A) and test (B) cohorts.

Table S6.1: For every feature the median value are reported for the training and test cohorts. In brackets are the 5th and 95th percentiles. With exception of the 95th percentile of the 31st feature in the test cohort, which is inflated by a single patient with abnormally high k_{ep} blobness values, the values are comparable between training and test cohorts.

| Features | Train | Test |
|--------------------------------------|--------------------------------------|-------------------------------------|
| T2w | 0.89 (0.49-1.35) | 0.93 (0.49 – 1.46) |
| T2w - Texture, G(1.5) | 0.89 (0.56 – 1.24) | 0.93 (0.59 – 1.34) |
| T2w - Texture, G _x (1.5) | 0.00 (-0.11 – 0.10) | 0.01 (-0.12 – 0.13) |
| T2w - Texture, G _y (1.5) | 0.00 (-0.09 – 0.09) | 0.00 (-0.09 – 0.11) |
| T2w - Texture, G _{xx} (1.5) | 0.02 (-0.10 – 0.13) | 0.01 (-0.14 – 0.14) |
| T2w - Texture, G _{yy} (1.5) | 0.00 (-0.05 – 0.05) | 0.00 (-0.06 – 0.06) |
| T2w - Texture, G _{xy} (1.5) | 0.00 (-0.10 – 0.10) | 0.00 (-0.11 – 0.14) |
| T2w - Texture, G(2.4) | 0.90 (0.60 – 1.20) | 0.95 (0.64 – 1.31) |
| T2w - Texture, G _x (2.4) | -0.01 (-0.13 – 0.11) | 0.01 (-0.14 – 0.15) |
| T2w - Texture, G _y (2.4) | 0.01 (-0.10 – 0.11) | 0.01 (-0.10 – 0.14) |
| T2w - Texture, G _{xx} (2.4) | 0.03 (-0.10 – 0.15) | 0.02 (-0.15 – 0.16) |
| T2w - Texture, G _{yy} (2.4) | 0.00 (-0.06 – 0.06) | 0.00 (-0.08 – 0.08) |
| T2w - Texture, G _{xy} (2.4) | 0.01 (-0.10 – 0.14) | 0.01 (-0.12 – 0.18) |
| T2w - Texture, G(3.8) | 0.91 (0.65 – 1.19) | 0.96 (0.70 – 1.30) |
| T2w - Texture, G _x (3.8) | -0.01 (-0.15 – 0.12) | 0.01 (-0.18 – 0.19) |
| T2w - Texture, G _y (3.8) | 0.01 (-0.12 – 0.17) | 0.02 (-0.12 – 0.19) |
| T2w - Texture, G _{xx} (3.8) | 0.04 (-0.11 – 0.18) | 0.01 (-0.16 – 0.17) |
| T2w - Texture, G _{yy} (3.8) | 0.00 (-0.09 – 0.09) | 0.00 (-0.10 – 0.11) |
| T2w - Texture, G _{xy} (3.8) | 0.02 (-0.11 – 0.20) | 0.02 (-0.13 – 0.23) |
| T2w - Texture, G(6.0) | 0.93 (0.71 – 1.20) | 0.99 (0.76 – 1.32) |
| T2w - Texture, G _x (6.0) | -0.01 (-0.21 – 0.14) | 0.01 (-0.21 – 0.22) |
| T2w - Texture, G _y (6.0) | 0.03 (-0.16 – 0.21) | 0.04 (-0.18 – 0.22) |
| T2w - Texture, G _{xx} (6.0) | 0.04 (-0.15 – 0.19) | -0.01 (-0.22 – 0.18) |
| T2w - Texture, G _{yy} (6.0) | 0.00 (-0.12 – 0.11) | 0.00 (-0.13 – 0.12) |
| T2w - Texture, G _{xy} (6.0) | 0.04 (-0.17 – 0.25) | 0.04 (-0.18 – 0.26) |
| ADC | 1.1 (0.4 – 1.6) × 10 ⁻³ | 1.3 (0.6 – 2.5) × 10 ⁻³ |
| ADC blobness | -0.2 (-0.4 – 0.2) × 10 ⁻³ | 0.2 (-0.4 – 0.3) × 10 ⁻³ |
| K ^{trans} | 1.15 (0.48 – 2.70) | 1.19 (0.54 – 2.83) |
| K ^{trans} blobness | -0.17 (-0.88 – 0.42) | -0.18 (-0.94 – 0.40) |
| k _{ep} | 1.03 (0.52 – 1.83) | 1.05 (0.44 – 2.62) |
| k _{ep} blobness | 0.11 (-0.43 – 0.48) | 0.14 (-0.47 – 798.32) |

Table S6.2: Feature ranking and coefficients. $G_i(s)$ stands for the Gaussian derivative in the direction i , with standard deviation s (in mm).

| Rank | Feature name | β | Rank | Feature name | β |
|------|-------------------------------|---------|------|-------------------------------|---------|
| 1 | k_{ep} - Intensity | 2.7 | 17 | T2w - Texture, G_{xy} (2.4) | -134.0 |
| 2 | k_{ep} - Blobness | -0.9 | 18 | T2w - Texture, G_{yy} (6.0) | 1.5 |
| 3 | K^{trans} - Blobness | -2.7 | 19 | T2w - Texture, G_{yy} (3.8) | -0.7 |
| 4 | K^{trans} - Intensity | -1.5 | 20 | T2w - Texture, G_{xx} (6.0) | 1.6 |
| 5 | ADC - Blobness | 28206.0 | 21 | T2w - Texture, G_x (6.0) | -0.6 |
| 6 | ADC - Intensity | -965.0 | 22 | T2w - Texture, G_{yy} (2.4) | -7.2 |
| 7 | T2w - Texture, G (3.8) | 86.6 | 23 | T2w - Texture, G_x (3.8) | -1.2 |
| 8 | T2w - Texture, G (6.0) | 329.1 | 24 | T2w - Texture, G_{xx} (3.8) | -5.9 |
| 9 | T2w - Texture, G (2.4) | -281.9 | 25 | T2w - Texture, G_y (1.5) | 2.5 |
| 10 | T2w - Texture, G_{xx} (1.5) | -3.8 | 26 | T2w - Texture, G_y (2.4) | -118.6 |
| 11 | T2w - Texture, G_{xy} (6.0) | -70.9 | 27 | T2w - Texture, G_y (3.8) | -203.0 |
| 12 | T2w - Texture, G (1.5) | -133.7 | 28 | T2w - Texture, G_{yy} (1.5) | 5.5 |
| 13 | T2w - Texture, G_y (6.0) | -59.4 | 29 | T2w - Texture, G_x (2.4) | 5.0 |
| 14 | T2w - Intensity | -0.2 | 30 | T2w - Texture, G_x (1.5) | 1.9 |
| 15 | T2w - Texture, G_{xy} (3.8) | -183.3 | 31 | T2w - Texture, G_{xy} (1.5) | -18.4 |
| 16 | T2w - Texture, G_{xx} (2.4) | 5.1 | | Offset | -3.75 |

Table S6.3: Sensitivity and specificity values for the model-derived binary GTV delineation compared to the radiologist’s manual delineations. The high-risk CTV and GTV defined by clustering the results in 3-levels are also presented in the table. The reported values are from every individual patient in the test set.

| Patient ID | Sensitivity | | | | | Specificity | | | | |
|------------|-------------|---------------|---------------|------|---------------|-------------|---------------|---------------|------|---------------|
| | Binary GTV | Radiologist 1 | Radiologist 2 | GTV | High-risk CTV | Binary GTV | Radiologist 1 | Radiologist 2 | GTV | High-risk CTV |
| 1 | 0.40 | 0.00 | 0.00 | 0.31 | 0.47 | 0.97 | 0.99 | 1.00 | 0.99 | 0.89 |
| 2 | 0.24 | 0.51 | 0.34 | 0.17 | 0.46 | 0.94 | 0.93 | 0.95 | 0.97 | 0.85 |
| 3 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.98 | 0.99 | 1.00 | 0.99 | 0.94 |
| 4 | 0.18 | 0.43 | 0.75 | 0.09 | 0.26 | 0.96 | 0.93 | 0.72 | 0.98 | 0.94 |
| 5 | 0.87 | 0.33 | 0.68 | 0.59 | 0.97 | 0.93 | 1.00 | 0.99 | 0.98 | 0.88 |
| 6 | 0.51 | 0.61 | 0.71 | 0.36 | 0.75 | 1.00 | 1.00 | 1.00 | 1.00 | 0.97 |
| 7 | 0.22 | 0.18 | 0.15 | 0.17 | 0.28 | 0.98 | 1.00 | 1.00 | 0.99 | 0.97 |
| 8 | 0.30 | 0.76 | 0.91 | 0.12 | 0.54 | 0.99 | 0.95 | 0.89 | 0.99 | 0.97 |
| 9 | 0.47 | 0.00 | 0.66 | 0.33 | 0.60 | 0.97 | 0.98 | 0.97 | 0.99 | 0.95 |
| 10 | 0.67 | 0.44 | 0.00 | 0.22 | 0.71 | 0.99 | 0.99 | 1.00 | 1.00 | 0.99 |
| 11 | 0.09 | 0.00 | 0.57 | 0.02 | 0.30 | 0.96 | 1.00 | 0.99 | 0.98 | 0.91 |
| 12 | 0.31 | 0.56 | 0.00 | 0.15 | 0.49 | 0.95 | 1.00 | 1.00 | 0.97 | 0.89 |
| 13 | 0.56 | 0.63 | 0.66 | 0.49 | 0.71 | 0.97 | 0.95 | 0.97 | 0.98 | 0.95 |
| 14 | 0.18 | 0.02 | 0.11 | 0.08 | 0.27 | 0.96 | 1.00 | 1.00 | 0.98 | 0.93 |
| 15 | 0.59 | 0.28 | 0.21 | 0.59 | 0.59 | 0.84 | 0.98 | 0.98 | 0.84 | 0.84 |
| 16 | 0.57 | 0.65 | 0.91 | 0.23 | 0.71 | 0.97 | 0.99 | 0.96 | 0.98 | 0.95 |
| 17 | 0.03 | 0.18 | 0.12 | 0.01 | 0.09 | 0.96 | 1.00 | 0.97 | 0.98 | 0.87 |

GENERAL DISCUSSION

This thesis describes studies on the use of multi-parametric MRI for the detection and localisation of locally recurrent PCa after RT. We have found evidence that pre-treatment T2w imaging features could potentially improve recurrence-risk prediction models for high-risk PCa patients treated with RT. For the purpose of RT treatment planning we have shown that, as an alternative to CT, a combination of multiple MRI sequences can be used to detect fiducial markers automatically. Some patients will experience biochemical recurrence after primary RT. Despite some studies describing the qualitative properties of multi-parametric MRI after RT, no guidelines exist on how to evaluate and assess recurrent PCa on mp-MRI. By characterising mp-MRI properties of irradiated benign tissue and recurrent tumour, this thesis contributes towards a better understanding of imaging in this challenging setting and can be used as a foundation to establish guidelines on how to assess and score MRI scans. Improving the general radiological knowledge in the radio-recurrent setting will aid in the interpretation of images for the purpose of staging and patient selection for salvage treatments. By providing evidence on how to define the region to treat with local salvage strategies this thesis also contributes to the refinement of localised salvage treatments.

PART 1: MAGNETIC RESONANCE IMAGING IN THE TREATMENT AND OUTCOME PREDICTION OF PRIMARY PROSTATE CANCER

Can pre-treatment MRI features be used to improve prediction of treatment outcome?

Chapter 2 reports a study investigating the value of pre-treatment T2w MRI features in predicting the risk of biochemical recurrence in high-risk PCa patients treated with EBRT. The standard clinical nomograms have limited ability to distinguish different recurrence risk within a specific risk-group. Predicting which men will develop biochemical recurrence could help identify patients who would benefit from additional or more aggressive therapies but also targeted follow-up for early detection of recurrent cancer. Multi-parametric MRI offers the possibility to investigate prostate architecture and tumour heterogeneity non-invasively and with this, devise imaging biomarkers. These biomarkers can potentially describe some of the variability not explained by typically acquired clinical parameters and allow for individualised patient risk-stratification. The field of radiomics investigates the macroscopic heterogeneity of quantitative features computed from medical images with the goal to probe the characteristics of the underlying tissue [164]. In this study, we have found that combining high-dimensional whole-prostate radiomic features, extracted from pre-treatment T2w MRI scans, with a logistic regression classifier outperformed a model based solely on clinical variables (AUC = 0.63 versus 0.51). Considering that this was a clinically homogeneous cohort, the poor performance of a clinically-based model was expected. According to PI-RADS v2, T2w alone is insufficient for tumour detection [42] and thus whole-prostate features were used. By doing so, no assumptions were made about tumour location and all available information was taken into account, resulting in a less time-consuming procedure by circumventing the need to delineate the tumour region. Our results support previous findings describing an association between tumour adjacent stroma and prostate micro-environment to relapse and disease progression [73, 74]. As recurrence occurs mainly at the site of the primary largest and/or highest-grade index lesion [55, 57, 165, 166], it is reasonable to assume that the most crucial information would lie within the pre-treatment visible tumour region. In this case, the use of features from the whole prostatic area might dilute the information

provided by tumour-derived features and could explain the higher performance obtained in the study by Gnep et al. [21] when using solely tumour-derived features. The addition of functional MR imaging could potentially improve the model predictions by providing physiological information and by enabling better tumour localisation. In fact, there is new evidence that bi-parametric MRI, using T2w and ADC-derived features, results in improved outcome predictions when compared to the use of a single modality [167]. In addition to mp-MRI, ^{18}F Choline- or ^{68}Ga PSMA-PET might provide additional insight into the risk of biochemical recurrence. Further research is needed to ascertain which imaging sequences, features and region of interest would be optimal. The results presented in this study offer a proof of concept of the potential of pre-treatment radiological information in the context of precision medicine. Similar models could help discriminate for which patients treatment intensification would be beneficial and to design tailored follow-up schemes, with more frequent visits for patients with higher risk of recurrence (*Figure 7.1, Pre-treatment imaging*).

Is it possible to automatically detect gold fiducial markers on MRI?

Fiducial markers (FM) are implanted in the prostate with the purpose of daily position verification during the course of EBRT treatments. Moving towards an MR-only environment will mitigate systematic spatial uncertainties present when registering computed tomography (CT) and MRI [168]. However, a requisite for the implementation of an MR-only workflow in RT is the ability to identify the gold FM on MRI. In **Chapter 3** we describe how our standard mp-MRI sequences (T1-weighted, T2-weighted and a B_0 map) acquired for the purpose of treatment planning can be combined with dedicated sequences (balanced steady-state free precession – bTFE, susceptibility weighted imaging – SWI) to derive the FM position. Using CT-derived marker coordinates as the ground-truth, we found that the combination of different sequences resulted in a better performance than the best single imaging sequence (bTFE). By combining mp-MRI with bTFE the highest accuracy was obtained, with a true positive rate of 0.94. Similarly, a study evaluating manual MR-based FM detection found that multiple sequences led to better localisation performance compared with the use of a single sequence [169].

Accurate and reproducible manual FM localisation using MRI can be challenging. The procedure is time consuming, affected by pixel-size and noise as well as the quality of the software used for detection [106]. In addition, marker appearance on imaging is impacted by sequence parameters and marker orientation with respect to the magnetic field. The use of automatic detection algorithms aims at facilitating MR-based FM localisation. A recent study comparing automatic with manual MR-based FM detection found that the spatial accuracy and precision were comparable for both methods [170]. An automated FM detection model could then be implemented to identify and superimpose the FM in the derived pseudo-CT (*Figure 7.1, Treatment planning*). Brachytherapy seeds and FM will be present on the images acquired for the detection of recurrent disease in EBRT and LDR-brachytherapy treated patients. Thus, similar automated detection algorithms could be of use to identify them in this setting (*Figure 7.1, Diagnosis of recurrent PCa*).

This is only one of the alternatives proposed to address this issue. Other options, mostly confined to the research setting currently, are the use of markers that provide positive contrast on the MR images [171] and MR sequences adapted to localise paramagnetic objects with high positive contrast [94]. Instead of gold FM, the use of electromagnetic beacons that allow for real-time prostate localisation has also been proposed [172]. The implementation of an MR-only workflow mitigates the errors arising from multimodality registration. Currently, automatic FM detection methods cannot yet fully eliminate manual interaction but they can be used for an initial automatic localisation with possible manual intervention in more complicated cases. The clinical implementation of MR-Linac systems will allow for real-time MR-guided RT and will ultimately eliminate the need to use FM.

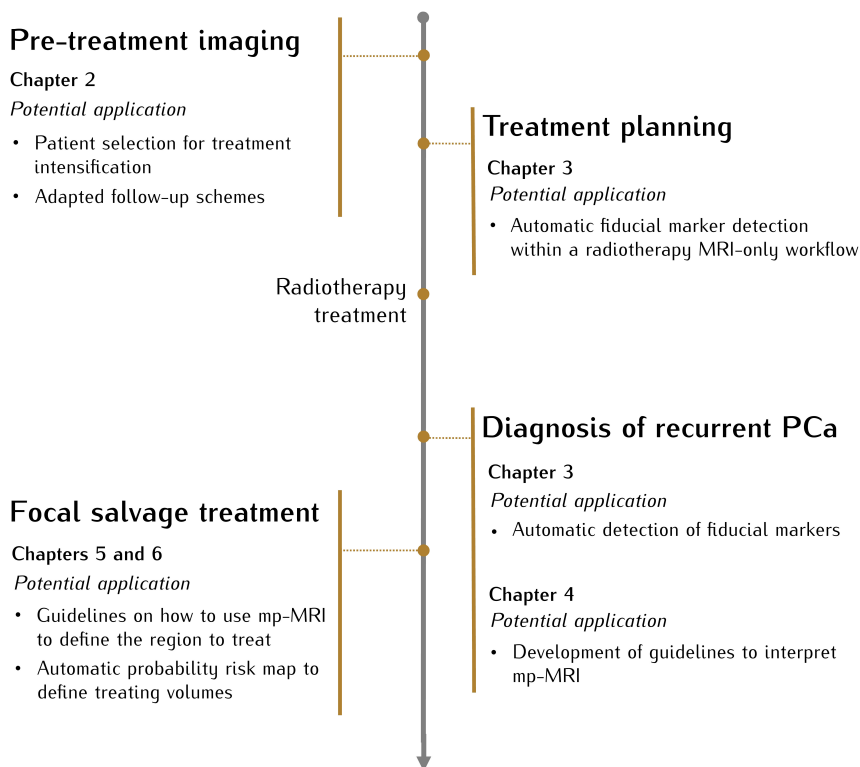


Figure 7.1: Schematic representation of a patient treatment chain and potential applications of the findings from each of the chapters at the different time points.

PART 2: MAGNETIC RESONANCE IMAGING FOR THE DETECTION AND LOCALISATION OF RECURRENT PROSTATE CANCER

Focal salvage treatments can be devised to overcome the toxicity associated with whole-gland approaches. Having an accurate detection and a clear definition of local or distant disease at the time of biochemical failure is a crucial aspect for patient selection, as the therapeutic ratio (treatment efficacy versus treatment-related toxicity) is narrow in the recurrent setting. Multi-parametric MRI is a non-invasive imaging tool that can be used to localise recurrent PCa. However, MRI information is often underutilised as guidelines are lacking on image interpretation in this setting. In the second part of this thesis we char-

acterised and assessed mp-MRI for detection and localisation of recurrent PCa after RT.

Radiation effects versus recurrent disease

What are the quantitative mp-MRI characteristics of benign tissue after radiotherapy?

What are the quantitative mp-MRI characteristics of recurrent tumours after radiotherapy?

These two questions have been addressed in **Chapter 4** and **5**, respectively.

A case-control set-up was used to describe and compare the imaging characteristics of irradiated benign prostate tissue and radio-recurrent disease (**Chapter 4**). Imaging maps for benign central-gland and peripheral-zone were found to be similar in both recurrent and non-recurrent patients. The peri-urethral tissue had heightened enhancement in the K^{trans} and k_{ep} maps without apparent signs of malignancy for both recurrent and non-recurrent patients. This enhancement could be considered a confounder for tumour detection in the central-gland. While K^{trans} values for recurrent tumour and peri-urethral tissue were not significantly different, k_{ep} values were. This suggests that k_{ep} maps can be used to discriminate benign enhancement from recurrent tumour in the central-gland.

While all imaging maps distinguished recurrent tumour from benign peripheral-zone, the distinction between tumour and benign tissue in the central-gland was only possible with ADC, K^{trans} and k_{ep} . While T2w is the dominant sequence in PI-RADS v2 for the detection of primary PCa in the transition zone, its use is limited in the recurrent setting. Using mixed modelling we found that a combination of the quantitative maps T_2 , ADC, K^{trans} and k_{ep} together with voxel location resulted in optimal distinction between tumour and benign voxels.

When comparing the imaging characteristics of patients with primary PCa before RT and those with recurrent disease after RT, we observed that the T_2 , K^{trans} and k_{ep} values for both peripheral-zone and central-gland were significantly lower after the treatment (with the exception of the ADC values in

the peripheral-zone). Nonetheless, no significant differences were seen for all imaging values between recurrent and primary tumours.

While in **Chapter 4** recurrent disease was defined on imaging and proved with targeted biopsies, in **Chapter 5** we aimed to characterise recurrent disease more accurately. To do so, whole-mount histopathology slices of tissue from patients treated with salvage prostatectomy for recurrent PCa after initial RT were registered to MRI. Tumour delineations performed on pathology could in this way be transferred to the MRI grid. We have found radio-recurrent disease to be often multifocal and with a different recurrence pattern in EBRT and LDR-brachytherapy-treated patients. While recurrent disease after EBRT consisted of large tumour foci, these were mainly small and sparse after LDR-brachytherapy. In addition, recurrent PCa after LDR-brachytherapy was challenging to image due to the presence of the seeds, which caused severe distortions in the DWI sequence potentially affecting the reliability of ADC values. The clinical relevance and implications of tumour multifocality for the design of focal salvage treatments are yet to be addressed. The contrasting patterns of recurrence in EBRT and LDR-brachytherapy treated patients suggests that different treatment strategies might need to be devised for each group.

The study presented in **Chapter 5** further compared histopathology derived delineations with those made manually by two radiologists, and found that delineations based on mp-MRI tend to miss smaller foci and slightly underestimate the tumour volume of local recurrent PCa. A mixed modelling analysis performed best in distinguishing tumour from benign tissue by using a combination of all imaging maps together with location information. The addition of DCE parameters improved tumour distinction. The location significantly contributed to tumour likelihood because it acts as a surrogate to tumour prevalence, which is typically higher in the peripheral-zone.

The results from these two studies suggest that DCE-MRI has an important role in diagnosing locally recurrent disease. We did not find benign prostatic hyperplasia (BPH) in the analysed pathology (**Chapter 5**), suggesting that BPH vanishes after RT and is no longer a benign central-gland confounder. However, increased enhancement of K^{trans} and k_{ep} in the peri-urethral region without signs of malignancy is present and needs to be considered when designing guidelines. The value of T2w imaging is diminished in the detection and

localisation of intraprostatic recurrent disease, and we suggest that its main role should be confined to identifying extracapsular extension and seminal vesicle invasion. The DWI sequence is limited by the presence of brachytherapy seeds as these create severe susceptibility artefacts. A study exclusively investigating recurrence after LDR-brachytherapy treatment found that DCE-MRI seems to be superior to T2w and DWI in detecting recurrent disease [18]. The implementation of a DWI sequence with an alternative sampling scheme [173, 174] (other than echo-planar imaging) has been proposed to deal with this limitation. The information presented in these studies can be used to devise guidelines to interpret mp-MRI in the setting of recurrent PCa after RT (*Figure 7.1, Diagnosis of recurrent PCa*).

Defining the volume to target with focal salvage treatments

Can mp-MRI be used to automatically define a region to treat with focal salvage treatments?

In order to provide focal salvage treatments of recurrent PCa a delineation of the target volume is necessary. MRI is commonly used for this task but, as discussed in the previous chapters, no guidelines exist on how to assess and delineate recurrent PCa on mp-MRI. In addition, in **Chapter 5** we observed that mp-MRI derived contours of recurrent tumour tend to underestimate the true tumour volume. Following these findings, in **Chapter 6** we propose a model to derive regions to treat with focal salvage RT. A logistic regression model, also referred to as tumour probability (TP) model, was trained on mp-MRI imaging features. When tested on an independent cohort, the TP model obtained a good diagnostic accuracy (AUC = 0.77). To deal with intrinsic image confounders and the uncertainties in defining a binary delineation we proposed the use of a TP model with three levels of risk stratification (CTV, high-risk CTV and GTV). In order to define these regions, the voxel-wise tumour probability maps were further clustered on increasing probability of tumour presence. This approach treats not only the GTV but it aims at mapping voxel-wise tumour probabilities to proportionally increasing doses, and with this aims at compensating possible tumour volume underestimation and covering microscopic spread.

Model-derived tumour risk-levels achieved similar sensitivity and specificity as manual delineations in localising the recurrent tumour.

Our previous studies have led to the conclusion that DCE-MRI is important for the detection of recurrent disease after RT. In fact, in **Chapter 6** the pharmacokinetic maps k_{ep} and K^{trans} seem to be the imaging parameters with the greatest relevance in defining a voxel's tumour risk. Finally, our results suggest that to optimise tumour coverage, margins should be included around the visible tumour. In this study, the high-risk clinical target volume surrounding the gross tumour volume was patient specific and defined based on the mp-MRI characteristics. In primary PCa, Hollmann *et al.* [175] reported the tumours to be predominantly multifocal, and found that a fixed CTV margin would not effectively cover all satellite tumours. We have also observed recurrent PCa to be predominantly multifocal (**Chapter 5**). However, the clinical relevance of tumour satellites is as yet unknown. If these can be considered indolent, then focal salvage treatments covering only the index tumour lesion might be sufficient. But it is also possible that these need to be treated to ensure local control. In that case, and following Hollmann's reasoning, it might be hard to identify a fixed margin for the treatment of recurrent PCa. If considering re-irradiation, a dose de-escalation regimen from the index tumour lesion could be a suitable compromise between tumour coverage and normal tissue toxicity.

The overall findings demonstrate that it is possible to use mp-MRI information to automatically derive volumes to target that could be used as a starting point for treatment planning (*Figure 7.1, Focal salvage treatment*) for focal salvage approaches in radio-recurrent PCa. Additional research should be performed to validate the devised image-based models and to test the feasibility of including such approaches within the patient treatment workflow.

IMPLICATIONS FOR CLINICAL PRACTICE

Currently, patients with high-risk characteristics are typically given more aggressive treatment regimens, but their oncologic outcome is still heterogeneous. Incorporating pre-treatment high-dimensional imaging-features can potentially improve patient stratification. In addition, accurate delivery of RT treatment is crucial in ensuring local tumour control while minimising normal tissue toxicity. The implementation of an MR-only treatment planning workflow would not only be logistically beneficial but it would also eliminate inter-modality registration errors. A commercial MR-only simulation for RT planning is currently being tested for the treatment of primary PCa at the Netherlands Cancer Institute.

A proportion of patients treated with RT will experience recurrent disease, and those with locally recurrent tumours are candidates for focal salvage treatments. Despite the fact that MRI is commonly used to diagnose and evaluate disease extent, there are presently no guidelines on how to assess disease recurrence from mp-MRI after RT. In this thesis, DCE-MRI was found to be very important in detecting and localising recurrent PCa. The data collected over the course of these studies could ultimately be used in teaching sessions, and as a foundation for establishing mp-MRI interpretation guidelines in the setting of radio-recurrent PCa. Moreover, the different patterns of recurrent disease after EBRT and LDR-brachytherapy treatments suggest that different salvage strategies might need to be devised for the two groups. The design, and potentially the effectiveness, of focal treatments are highly dependent on accurate tumour localisation.

In this thesis we propose a model that uses mp-MRI information to derive target volumes that could be used to initialise treatment planning. Focal salvage interventions described in literature are typically MR-guided [54, 142, 176, 177], and could therefore benefit from the methodology presented here. Besides mp-MRI, metrics from Choline and PSMA-PET should also be incorporated into the design of focal salvage treatments.

In this thesis we provide data that enables a better understanding of mp-MRI in the context of recurrent PCa after RT. This body of work provides an overview of mp-MRI and its relationship with histopathology, which will be useful to

establish radiological guidelines and help define better clinical practices in the detection and localisation of radio-recurrent PCa. These can be used to optimise the design of focal salvage treatments.

REFERENCES

1. Zlotta, A. R. *et al.* Prevalence of prostate cancer on autopsy: cross-sectional study on unscreened Caucasian and Asian men. *Journal of the National Cancer Institute* **105**, 1050–1058 (2013).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **68**, 394–424 (2018).
3. Wibmer, A. G. *et al.* Molecular Imaging of Prostate Cancer. *Radiographics* **36**, 142–159 (2016).
4. Wong, M. C. S. *et al.* Global Incidence and Mortality for Prostate Cancer: Analysis of Temporal Patterns and Trends in 36 Countries. *European Urology* **70**, 862–874 (2016).
5. Department of Economic and Social Affairs, U. N. *World Population Ageing Report 2017* <https://www.un.org/en/development/desa/population/>.
6. Ilic, D. *et al.* Prostate cancer screening with prostate-specific antigen (PSA) test: a systematic review and meta-analysis. *BMJ* **362**. doi:10.1136/bmj.k3519 (2018).
7. De Carvalho, P. A. *et al.* Retrograde Release of the Neurovascular Bundle with Preservation of Dorsal Venous Complex During Robot-assisted Radical Prostatectomy: Optimizing Functional Outcomes. *European Urology*. doi:10.1016/j.eururo.2018.07.003 (2019).
8. *Integraal Kankercentrum Nederland, Cijfers over kanker.* (visited on 2019-02-15). <https://www.cijfersoverkanker.nl>.
9. Barentsz, J. O. *et al.* ESUR prostate MR guidelines 2012. *European Radiology* **22**, 746–757 (2012).

10. Ahmed, H. U. *et al.* Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet* **389**, 815–822 (2017).
11. Hövels, A. M. *et al.* The diagnostic accuracy of CT and MRI in the staging of pelvic lymph nodes in patients with prostate cancer: a meta-analysis. *Clinical Radiology* **63**, 387–395 (2008).
12. Obek, C. *et al.* The accuracy of ^{68}Ga -PSMA PET/CT in primary lymph node staging in high-risk prostate cancer. *European Journal of Nuclear Medicine and Molecular Imaging* **44**, 1806–1812 (2017).
13. Perera, M. *et al.* Sensitivity, Specificity, and Predictors of Positive ^{68}Ga -Prostate-specific Membrane Antigen Positron Emission Tomography in Advanced Prostate Cancer: A Systematic Review and Meta-analysis. *European Urology* **70**, 926–937 (2017).
14. Van Leeuwen, P. J. *et al.* Prospective evaluation of ^{68}Ga -prostate-specific membrane antigen positron emission tomography/computed tomography for preoperative lymph node staging in prostate cancer. *BJU International* **119**, 209–215 (2017).
15. Mottet, N. *et al.* *EAU prostate cancer guidelines 2019* (visited on 2019-05-23). <https://uroweb.org/guideline/prostate-cancer/>.
16. Lengana, T. *et al.* ^{68}Ga -PSMA PET/CT Replacing Bone Scan in the Initial Staging of Skeletal Metastasis in Prostate Cancer: A Fait Accompli? *Clinical Genitourinary Cancer* **16**, 392–401 (2018).
17. Zelefsky, M. J. *et al.* Pretreatment nomogram predicting ten-year biochemical outcome of three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for prostate cancer. *Urology* **70**, 283–287 (2007).
18. Vignati, A. *et al.* Texture features on T2-weighted magnetic resonance imaging: new potential biomarkers for prostate cancer aggressiveness. *Physics in Medicine & Biology* **60**, 2685–2701 (2015).
19. Nketiah, G. *et al.* T2-weighted MRI-derived textural features reflect prostate cancer aggressiveness: preliminary results. *European Radiology* **27**, 3050–3059 (2017).

20. Fuchsjäger, M. H. *et al.* Predicting Post-External-beam Radiation Therapy PSA Relapse of Prostate Cancer Using Pre-treatment MRI. *International Journal of Radiation Oncology*Biological*Physics* **78**, 743–750 (2010).
21. Gnep, K. *et al.* Haralick textural features on T2 -weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer. *Journal of Magnetic Resonance Imaging* **45**, 103–117 (2017).
22. Vora, S. A. *et al.* Outcome and toxicity for patients treated with intensity modulated radiation therapy for localized prostate cancer. *The Journal of Urology* **190**, 521–526 (2013).
23. Roach, M. *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. *International Journal of Radiation Oncology*Biological*Physics* **65**, 965–974 (2006).
24. Zumsteg, Z. S. *et al.* The natural history and predictors of outcome following biochemical relapse in the dose escalation era for prostate cancer patients undergoing definitive external beam radiotherapy. *European Urology* **67**, 1009–1016 (2015).
25. Nguyen, P. L. *et al.* Adverse effects of androgen deprivation therapy and strategies to mitigate them. *European Urology* **67**, 825–836 (2015).
26. Hruby, G. *et al.* Delineating biochemical failure with ⁶⁸Ga-PSMA-PET following definitive external beam radiation treatment for prostate cancer. *Radiotherapy and Oncology* **122**, 99–102 (2017).
27. Bolla, M. *et al.* External irradiation with or without long-term androgen suppression for prostate cancer with high metastatic risk: 10-year results of an EORTC randomised study. *Lancet Oncology* **11**, 1066–1073 (2010).
28. Philippou, Y., Parker, R. A., Volanis, D. & Gnanapragasam, V. J. Comparative Oncologic and Toxicity Outcomes of Salvage Radical Prostatectomy Versus Nonsurgical Therapies for Radiorecurrent Prostate Cancer: A Meta-Regression Analysis. *European Urology focus* **2**, 158–171 (2016).

29. Bianco, F. J. J. *et al.* Long-term oncologic results of salvage radical prostatectomy for locally recurrent prostate cancer after radiotherapy. *International Journal of Radiation Oncology*Biophysics* **62**, 448–453 (2005).
30. Nguyen, P. L., D'Amico, A. V., Lee, A. K. & Suh, W. W. Patient selection, cancer control, and complications after salvage local therapy for postradiation prostate-specific antigen failure: a systematic review of the literature. *Cancer* **110**, 1417–1428 (2007).
31. Cornford, P. *et al.* EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part II: Treatment of Relapsing, Metastatic, and Castration-Resistant Prostate Cancer. *European Urology* **71**, 630–642 (2017).
32. Tran, H., Kwok, J., Pickles, T., Tyllesley, S. & Black, P. C. Underutilization of local salvage therapy after radiation therapy for prostate cancer. *Urologic oncology* **32**, 701–706 (2014).
33. Duijzentkunst, D. A. S., Peters, M., van der Voort van Zyp, J. R. N., Moerland, M. A. & van Vulpen, M. Focal salvage therapy for local prostate cancer recurrences after primary radiotherapy: a comprehensive review. *World Journal of Urology* **34**, 1521–1531 (2016).
34. Einspieler, I. *et al.* Detection Efficacy of Hybrid (68)Ga-PSMA Ligand PET/CT in Prostate Cancer Patients with Biochemical Recurrence After Primary Radiation Therapy Defined by Phoenix Criteria. *Journal of Nuclear Medicine* **58**, 1081–1087 (2017).
35. Maurer, T., Eiber, M., Fanti, S., Budäus, L. & Panebianco, V. Imaging for Prostate Cancer Recurrence. *European Urology Focus* **2**, 139–150 (2016).
36. Barchetti, F. & Panebianco, V. Multiparametric MRI for Recurrent Prostate Cancer Post Radical Prostatectomy and Postradiation Therapy. *BioMed Research International* **2014**, 316272 (2014).
37. Vargas, H. A., Wassberg, C., Akin, O. & Hricak, H. MR imaging of treated prostate cancer. *Radiology* **262**, 26–42 (2012).
38. Rouviere, O. *et al.* Recurrent prostate cancer after external beam radiotherapy: value of contrast-enhanced dynamic MRI in localizing intraprostatic tumor—correlation with biopsy findings. *Urology* **63**, 922–927 (2004).

39. Haider, M. a. *et al.* Dynamic contrast-enhanced magnetic resonance imaging for localization of recurrent prostate cancer after external beam radiotherapy. *International Journal of Radiation Oncology*Biography*Physics* **70**, 425–30 (2008).
40. Akin, O. *et al.* Incremental value of diffusion weighted and dynamic contrast enhanced MRI in the detection of locally recurrent prostate cancer after radiation treatment: preliminary results. *European Radiology* **21**, 1970–8 (2011).
41. Arumainayagam, N. *et al.* Accuracy of multiparametric magnetic resonance imaging in detecting recurrent prostate cancer after radiotherapy. *BJU International* **106**, 991–7 (2010).
42. Weinreb, J. C. *et al.* PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2. *European Urology* **69**, 16–40 (2016).
43. Oppenheimer, D. C., Weinberg, E. P., Hollenberg, G. M. & Meyers, S. P. Multiparametric Magnetic Resonance Imaging of Recurrent Prostate Cancer. *Journal of Clinical Imaging Science* **6**, 18 (2016).
44. Moman, M. R. *et al.* Focal salvage guided by T2-weighted and dynamic contrast-enhanced magnetic resonance imaging for prostate cancer recurrences. *International Journal of Radiation Oncology*Biography*Physics* **76**, 741–6 (2010).
45. Pucar, D. *et al.* Prostate cancer: correlation of MR imaging and MR spectroscopy with pathologic findings after radiation therapy-initial experience. *Radiology* **236**, 545–553 (2005).
46. Donati, O. F. *et al.* Multiparametric prostate MR imaging with T2-weighted, diffusion-weighted, and dynamic contrast-enhanced sequences: are all pulse sequences necessary to detect locally recurrent prostate cancer after radiation therapy? *Radiology* **268**, 440–450 (2013).
47. Luzurier, A. *et al.* Dynamic contrast-enhanced imaging in localizing local recurrence of prostate cancer after radiotherapy: Limited added value for readers of varying level of experience. *Journal of Magnetic Resonance Imaging* **48**, 1012–1023 (2018).
48. Salembier, C. *et al.* ESTRO ACROP consensus guideline on CT- and MRI-based target volume delineation for primary radiation therapy of localized prostate cancer. *Radiotherapy and Oncology* **127**, 49–61 (2018).

49. Steenbergen, P. *et al.* Prostate tumor delineation using multiparametric magnetic resonance imaging: Inter-observer variability and pathology validation. *Radiotherapy and Oncology* **115**, 186–190 (2015).
50. Van Schie, M. A. *et al.* Contouring of prostate tumors on multiparametric MRI: Evaluation of clinical delineations in a multicenter radiotherapy trial. *Radiotherapy and Oncology* **128**, 321–326 (2018).
51. Zattoni, F. *et al.* Detection of recurrent prostate cancer after primary radiation therapy: An evaluation of the role of multiparametric 3T magnetic resonance imaging with endorectal coil. *Practical Radiation Oncology* **7**, 42–49 (2017).
52. Kollmeier, M. A., Zelefsky, M. & McBride, S. Magnetic resonance imaging-based salvage brachytherapy: Moving toward a focal paradigm. *Brachytherapy* **16**, 770–777 (2017).
53. Feasibility of MR imaging/MR spectroscopy-planned focal partial salvage permanent prostate implant (PPI) for localized recurrence after initial PPI for prostate cancer. *International Journal of Radiation Oncology*Biology*Physics* **85**, 370–7 (2013).
54. Murgic, J. *et al.* Focal Salvage High Dose-Rate Brachytherapy for Locally Recurrent Prostate Cancer After Primary Radiation Therapy Failure: Results From a Prospective Clinical Trial. *International Journal of Radiation Oncology*Biology*Physics* **102**, 561–567 (2018).
55. Pucar, D. *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. *International Journal of Radiation Oncology*Biology*Physics* **69**, 62–69 (2007).
56. Sala, E. *et al.* Endorectal MR imaging before salvage prostatectomy: tumor localization and staging. *Radiology* **238**, 176–183 (2006).
57. Jalloh, M. *et al.* Patterns of Local Failure following Radiation Therapy for Prostate Cancer. *The Journal of Urology* **194**, 977–982 (2015).
58. Michalski, J. *et al.* Clinical Outcome of Patients Treated With 3D Conformal Radiation Therapy (3D-CRT) for Prostate Cancer on RTOG 9406. *International Journal of Radiation Oncology*Biology*Physics* **83**, e363–e370 (2012).

59. Nomiya, T. *et al.* Management of high-risk prostate cancer: radiation therapy and hormonal therapy. *Cancer Treatment Reviews* **39**, 872–878 (2013).
60. Cahlon, O. *et al.* Ultra-High Dose (86.4 Gy) IMRT for Localized Prostate Cancer: Toxicity and Biochemical Outcomes. *International Journal of Radiation Oncology*Biography*Physics* **71**, 330–337 (2008).
61. Haralick, R., Shanmugan, K. & Dinstein, I. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 610–621 (1973).
62. Ginsburg, S. B., Rusu, M., Kurhanewicz, J. & Madabhushi, A. *Computer extracted texture features on T2w MRI to predict biochemical recurrence following radiation therapy for prostate cancer in SPIE Medical Imaging* **9035** (2014), 903509–903513. doi:[10.1117/12.2043937](https://doi.org/10.1117/12.2043937).
63. Villeirs, G. M., L Verstraete, K., De Neve, W. J. & De Meerleer, G. O. Magnetic resonance imaging anatomy of the prostate and periprostatic area: a guide for radiotherapists. *Radiotherapy and Oncology* **76**, 99–106 (2005).
64. D'Amico, A., Whittington, R., Malkowicz, S. & Al, E. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *The Journal of the American Medical Association* **280**, 969–974 (1998).
65. Abramowitz, M. C. *et al.* The Phoenix definition of biochemical failure predicts for overall survival in patients with prostate cancer. *Cancer* **112**, 55–60 (2008).
66. Nyholm, T. *et al.* Variability in prostate and seminal vesicle delineations defined on magnetic resonance images, a multi-observer, -center and -sequence study. *Radiation Oncology* **8**, 126 (2013).
67. Van Griethuysen, J. J. *et al.* Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research* (2017).
68. Kittler, J., Hatef, M., Duin, R. P. & Matas, J. On Combining Classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 226–239 (1998).
69. Parmar, C., Grossmann, P., Bussink, J., Lambin, P. & Aerts, H. J. W. L. Machine Learning methods for Quantitative Radiomic Biomarkers. *Scientific Reports* **5**, 13087 (2015).

70. Ding, C. & Peng, H. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology* **3**, 185–205 (2005).
71. Hegde, J. V. *et al.* Pre-Treatment MRI Staging Predicts for Biochemical Failure in High-Risk Prostate Cancer Treated with Combination High-Dose-Rate Brachytherapy and External Beam Radiotherapy. *Brachytherapy* **16**, S18–S19 (2017).
72. Westphalen, A. C. *et al.* Prostate cancer: prediction of biochemical failure after external-beam radiation therapy—Kattan nomogram and endorectal MR imaging estimation of tumor volume. *Radiology* **261**, 477–486 (2011).
73. Wikstrom, P., Marusic, J., Stattin, P. & Bergh, A. Low stroma androgen receptor level in normal and tumor prostate tissue is related to poor outcome in prostate cancer patients. *The Prostate* **69**, 799–809 (2009).
74. Leach, D. A. *et al.* Stromal androgen receptor regulates the composition of the microenvironment to influence prostate cancer outcome. *Oncotarget* **6**, 16135–16150 (2015).
75. Oakden-Rayner, L. *et al.* Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Scientific Reports* **7**, 1648 (2017).
76. Teruel, J. R. *et al.* Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. *NMR in Biomedicine* **27**, 887–896 (2014).
77. Braman, N. M. *et al.* Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast cancer research* **19**, 57 (2017).
78. Roe, K., Kakar, M., Seierstad, T., Ree, A. H. & Olsen, D. R. Early prediction of response to radiotherapy and androgen-deprivation therapy in prostate cancer by repeated functional MRI: a preclinical study. *Radiation Oncology* **6**, 65 (2011).

79. Boesen, L., Chabanova, E., Logager, V., Balslev, I. & Thomsen, H. S. Apparent diffusion coefficient ratio correlates significantly with prostate cancer gleason score at final pathology. *Journal of Magnetic Resonance Imaging* **42**, 446–453 (2015).
80. Fehr, D. *et al.* Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 6265–6273 (2015).
81. Yip, S. S. F. & Aerts, H. J. W. L. Applications and limitations of radiomics. *Physics in Medicine & Biology* **61**, R150–66 (2016).
82. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563–577 (2016).
83. Peng, H., Long, F. & Ding, C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE transactions on pattern analysis and machine intelligence* **27**, 1226–1238 (2005).
84. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2012).
85. Bowes, D., Crook, J. M., Rajapakshe, R., Araujo, C. & Parker, B. Defining a magnetic resonance scan sequence for permanent seed prostate brachytherapy postimplant assessment. *Brachytherapy* **12**, 25–9 (2013).
86. Vidakovic, S., Jans, H. S., Alexander, A. & Sloboda, R. S. Post-implant computed tomography-magnetic resonance prostate image registration using feature line parallelization and normalized mutual information. *Journal of Applied Clinical Medical Physics* **8**, 21–32 (2007).
87. Kapur, T., Egger, J., Damato, A., Schmidt, E. J. & Viswanathan, A. N. 3-T MR-guided brachytherapy for gynecologic malignancies. *Magnetic Resonance Imaging* **30**, 1279–1290 (2012).
88. Schieda, N., Avruch, L., Shabana, W. M. & Malone, S. C. Multi-echo Gradient Recalled Echo Imaging of the Pelvis for Improved Depiction of Brachytherapy Seeds and Fiducial Markers Facilitating Radiotherapy Planning and Treatment of Prostatic Carcinoma. *Journal of Magnetic Resonance Imaging* **720**, 715–720 (2015).

89. Bloch, B. N. *et al.* Prostate postbrachytherapy seed distribution: comparison of high-resolution, contrast-enhanced, T1- and T2-weighted endorectal magnetic resonance imaging versus computed tomography: initial experience. *International Journal of Radiation Oncology*Biolog*Physics* **69**, 70–78 (2007).
90. Katayama, N. *et al.* T2*-weighted Image/T2-weighted Image Fusion in Postimplant Dosimetry of Prostate Brachytherapy. *Journal of Radiation Research* **52**, 680–684 (2011).
91. Ohashi, T. *et al.* Impact of MRI-based postimplant dosimetric assessment in prostate brachytherapy using contrast-enhanced T1-weighted images. *Brachytherapy* **11**, 468–75 (2012).
92. Kuo, N., Lee, J., Tempany, C. & Stuber, M. *MRI-based prostate brachytherapy seed localization* in *Proceedings IEEE International Symposium on Biomedical Imaging* (2010), 1397–1400. doi:[10 . 1109 / ISBI . 2010 . 5490259](https://doi.org/10.1109/ISBI.2010.5490259).
93. De Leeuw, H., Seevinck, P. R. & Bakker, C. J. G. Center-out radial sampling with off-resonant reconstruction for efficient and accurate localization of punctate and elongated paramagnetic structures. *Magnetic Resonance in Medicine* **69**, 1611–1622 (2013).
94. Seevinck, P. R., de Leeuw, H., Bos, C. & Bakker, C. J. G. Highly localized positive contrast of small paramagnetic objects using 3D center-out radial sampling with off-resonance reception. *Magnetic Resonance in Medicine* **65**, 146–156 (2011).
95. Dong, Y., Chang, Z., Xie, G., Whitehead, G. & Ji, J. X. Susceptibility-based positive contrast MRI of brachytherapy seeds. *Magnetic Resonance in Medicine* **00**, 1–11 (2014).
96. McLaughlin, P. W. *et al.* The use of mutual information in registration of CT and MRI datasets post permanent implant. *Brachytherapy* **3**, 61–70 (2004).
97. Lindeberg, T. Feature detection with automatic scale selection. *International Journal of Computer Vision* **30**, 79–116 (1998).
98. Frangi, A. F., Niessen, W. J., Vincken, K. L. & Viergever, M. A. *Multiscale vessel enhancement filtering BT - Medical Image Computing and Computer-Assisted Intervention — MICCAI* in (1998), 130–137.

99. Sato, Y. *et al.* Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Medical Image Analysis* **2**, 143–168 (1998).
100. Duin, RPW and Juszczak, P and Ridder, DD and Paclik, P and Pekalska, E and Tax, DMJ. *PRTools version: A Matlab Toolbox for Pattern Recognition* version 5.2.3. 21st Dec. 2014. <http://37steps.com/prtools/updates/>.
101. Villeirs, G. M. *et al.* Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlenther Onkol* **181**, 424–430 (2005).
102. Rasch, C. *et al.* Definition of the prostate in CT and MRI: a multi-observer study. *International Journal of Radiation Oncology*Biography*Physics* **43**, 57–66 (1999).
103. Gao, Z. *et al.* A study of prostate delineation referenced against a gold standard created from the visible human data. *Radiotherapy and Oncology* **85**, 239–246 (2007).
104. Roach, M. 3., Faillace-Akazawa, P., Malfatti, C., Holland, J. & Hricak, H. Prostate volumes defined by magnetic resonance imaging and computerized tomographic scans for three-dimensional conformal radiotherapy. *International Journal of Radiation Oncology*Biography*Physics* **35**, 1011–1018 (1996).
105. Debois, M. *et al.* The contribution of magnetic resonance imaging to the three-dimensional treatment planning of localized prostate cancer. *International Journal of Radiation Oncology*Biography*Physics* **45**, 857–865 (1999).
106. Ghose, S. *et al.* MRI-alone radiation therapy planning for prostate cancer: Automatic fiducial marker detection. *Medical physics* **43**, 2218 (2016).
107. Hong, C. G. *et al.* The Prevalence and Characteristic Differences in Prostatic Calcification between Health Promotion Center and Urology Department Outpatients. *Korean Journal of Urology* **53**, 330–334 (2012).
108. Jonsson, J. H., Garpebring, A., Karlsson, M. G. & Nyholm, T. Internal Fiducial Markers and Susceptibility Effects in MRI-Simulation and Measurement of Spatial Accuracy. *International Journal of Radiation Oncology*Biography*Physics* **82**, 1612–1618 (2012).

109. Parker, C. C. *et al.* Magnetic resonance imaging in the radiation treatment planning of localized prostate cancer using intra-prostatic fiducial markers for computed tomography co-registration. *Radiotherapy and Oncology* **66**, 217–224 (2003).
110. Roberson, P. L. *et al.* Use and uncertainties of mutual information for computed tomography/magnetic resonance (CT/MR) registration post permanent implant of the prostate. *Medical Physics* **32**, 473 (2005).
111. Van Dalen, J. A., Huisman, H. J., Welmers, A. & Barentsz, J. O. *Semi-automatic Image Registration of MRI to CT Data of the Prostate Using Gold Markers as Fiducials BT - Biomedical Image Registration in Biomedical Image Registration* (2003), 311–320.
112. Schenck, J. F. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Medical Physics* **23**, 815–850 (1996).
113. Rylander, S., Thörnqvist, S., Haack, S., Pedersen, E. M. & Muren, L. P. Intensity profile based measurement of prostate gold markers influence on 1.5 and 3T diffusion-weighted MR images. *Acta Oncologica* **50**, 866–872 (2011).
114. Van der Heide, U. A. *et al.* Analysis of fiducial marker-based position verification in the external beam radiotherapy of patients with prostate cancer. *Radiotherapy and Oncology* **82**, 38–45 (2007).
115. Beaulieu, L. *et al.* Performing daily prostate targeting with a standard V-EPID and an automated radio-opaque marker detection algorithm. *Radiotherapy and Oncology* **73**, 61–64 (2004).
116. McNair, H. A. *et al.* A Comparison of the Use of Bony Anatomy and Internal Markers for Offline Verification and an Evaluation of the Potential Benefit of Online and Offline Verification Protocols for Prostate Radiotherapy. *International Journal of Radiation Oncology*Biological*Physics* **71**, 41–50 (2008).
117. Partridge, M., Symonds-Tayler, J. R. N. & Evans, P. M. IMRT verification with a camera-based electronic portal imaging system. *Physics in Medicine & Biology* **45**, N183–N196 (2000).

118. Tokuda, J., Song, S.-E., Tuncali, K., Tempany, C. & Hata, N. *Configurable Automatic Detection and Registration of Fiducial Frames for Device-to-Image Registration in MRI-Guided Prostate Interventions BT - Medical Image Computing and Computer-Assisted Intervention – MICCAI* in (2013), 355–362.
119. Horwich, A., Parker, C., Bangma, C. & Kataja, V. Prostate cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **21 Suppl 5**, v129–33 (2010).
120. Rodrigues, G. *et al.* Pre-treatment risk stratification of prostate cancer patients: A critical review. *Canadian Urological Association Journal* **6**, 121–127 (2012).
121. Van Houdt, P. J. *et al.* Performance of a fast and high-resolution multi-echo spin-echo sequence for prostate T2 mapping across multiple systems. *Magnetic Resonance in Medicine* **79**, 1586–1594 (2018).
122. Van Schie, M. A. *et al.* Repeatability of dose painting by numbers treatment planning in prostate cancer radiotherapy based on multiparametric magnetic resonance imaging. *Physics in Medicine & Biology* **62**, 5575–5588 (2017).
123. Murase, K., Yamazaki, Y. & Miyazaki, S. Deconvolution analysis of dynamic contrast-enhanced data based on singular value decomposition optimized by generalized cross validation. *Magnetic Resonance in Medical Sciences* **3**, 165–175 (2004).
124. Schabel, M. C. & Morrell, G. R. Uncertainty in T1 mapping using the variable flip angle method with two flip angles. *Physics in Medicine & Biology* **54**, N1–8 (2009).
125. Team, R. D. C. *R: A Language and Environment for Statistical Computing* Vienna, Austria, 2014. <http://www.r-project.org/>.
126. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software; Vol 1, Issue 1 (2015)* (2015).
127. Dinis Fernandes, C. *et al.* Quantitative 3-T multi-parametric MRI and step-section pathology of recurrent prostate cancer patients after radiation therapy. *European Radiology*, 1–9 (2018).

128. Sheaff, M. T. & Baithun, S. I. Effects of radiation on the normal prostate gland. *Histopathology* **30**, 341–348 (1997).
129. Foltz, W. D. *et al.* Changes in apparent diffusion coefficient and T2 relaxation during radiotherapy for prostate cancer. *Journal of Magnetic Resonance Imaging* **37**, 909–916 (2013).
130. Westphalen, A. C. *et al.* Multiparametric 3T endorectal mri after external beam radiation therapy for prostate cancer. *Journal of Magnetic Resonance Imaging* **36**, 430–437 (2012).
131. Kim, C. K., Park, B. K. & Lee, H. M. Prediction of locally recurrent prostate cancer after radiation therapy: incremental value of 3T diffusion-weighted MRI. *Journal of Magnetic Resonance Imaging* **29**, 391–397 (2009).
132. Sato, C. *et al.* Differentiation of noncancerous tissue and cancer lesions by apparent diffusion coefficient values in transition and peripheral zones of the prostate. *Journal of Magnetic Resonance Imaging* **21**, 258–262 (2005).
133. Valerio, M. *et al.* Detection of Clinically Significant Prostate Cancer Using Magnetic Resonance Imaging-Ultrasound Fusion Targeted Biopsy: A Systematic Review. *European Urology* **68**, 8–19 (2015).
134. Dix, D. B. *et al.* How Would MRI-Targeted Prostate Biopsy Alter Radiation Therapy Approaches in Treating Prostate Cancer? *Urology* **122**, 139–146 (2018).
135. Martin, P. R., Cool, D. W., Fenster, A. & Ward, A. D. A comparison of prostate tumor targeting strategies using magnetic resonance imaging-targeted, transrectal ultrasound-guided fusion biopsy. *Medical physics* **45**, 1018–1028 (2018).
136. Rosoff, J. S., Savage, S. J. & Prasad, S. M. Salvage radical prostatectomy as management of locally recurrent prostate cancer: outcomes and complications. *World Journal of Urology* **31**, 1347–1352 (2013).
137. Myronenko, A. & Song, X. Point set registration: coherent point drift. *IEEE transactions on pattern analysis and machine intelligence* **32**, 2262–2275 (2010).

138. Crook, J. *et al.* Postradiotherapy prostate biopsies: what do they really mean? Results for 498 patients. *International Journal of Radiation Oncology*Biography*Physics* **48**, 355–367 (2000).
139. Ahmed, H. U. *et al.* Focal salvage therapy for localized prostate cancer recurrence after external beam radiotherapy: a pilot study. *Cancer* **118**, 4148–4155 (2012).
140. Baco, E. *et al.* Hemi salvage high-intensity focused ultrasound (HIFU) in unilateral radiorecurrent prostate cancer: a prospective two-centre study. *BJU international* **114**, 532–540 (2014).
141. Nguyen, P. L. *et al.* Magnetic resonance image-guided salvage brachytherapy after radiation in select men who initially presented with favorable-risk prostate cancer: a prospective phase 2 study. *Cancer* **110**, 1485–1492 (2007).
142. Peters, M. *et al.* Focal salvage iodine-125 brachytherapy for prostate cancer recurrences after primary radiotherapy: a retrospective study regarding toxicity, biochemical outcome and quality of life. *Radiotherapy and Oncology* **112**, 77–82 (2014).
143. Borofsky, S. *et al.* What Are We Missing? False-Negative Cancers at Multiparametric MR Imaging of the Prostate. *Radiology*, 152877 (2017).
144. Bratan, F. *et al.* How accurate is multiparametric MR imaging in evaluation of prostate cancer volume? *Radiology* **275**, 144–154 (2015).
145. Tofts, P. S. *et al.* Estimating Kinetic Parameters From Dynamic Contrast-Enhanced T1-Weighted MRI of a Diffusible Tracer: Standardized Quantities and Symbols. *Journal of Magnetic Resonance Imaging* **10**, 223–232 (1999).
146. De Bazelaire, C. M. J., Duhamel, G. D., Rofsky, N. M. & Alsop, D. C. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. *Radiology* **230**, 652–659 (2004).
147. Van den Bos, W. *et al.* Focal therapy in prostate cancer: international multidisciplinary consensus on trial design. *European Urology* **65**, 1078–1083 (2014).
148. Katz, A., Kosinski, K. E. & Berg, C. Is salvage focal cryotherapy curative treatment for patients with localized recurrent prostate cancer? *Journal of Clinical Oncology* **34**, 26 (2016).

149. Dinh, C. V. *et al.* Multicenter validation of prostate tumor localization using multiparametric MRI and prior knowledge. *Medical Physics* **44**, 949–961 (2017).
150. Gaur, S. *et al.* Can computer-aided diagnosis assist in the identification of prostate cancer on prostate MRI? a multi-center, multi-reader investigation. *Oncotarget* **9**, 33804–33817 (2018).
151. Viswanath, S. E. *et al.* Central Gland and Peripheral Zone Prostate Tumors have Significantly Different Quantitative Imaging Signatures on 3 Tesla Endorectal, In Vivo T2-Weighted Magnetic Resonance Imagery. *Journal of Magnetic Resonance Imaging* **36**, 213–224 (2012).
152. Hambrock, T., Vos, P. C., Hulsbergen–van de Kaa, C. A., Barentsz, J. O. & Huisman, H. J. Prostate Cancer: Computer-aided Diagnosis with Multiparametric 3-T MR Imaging—Effect on Observer Performance. *Radiology* **266**, 521–530 (2013).
153. Dinis Fernandes, C. *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. *Journal of Magnetic Resonance Imaging* **50**, 269–278 (2019).
154. Murase, K. Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging. *Magnetic resonance in medicine* **51**, 858–862 (2004).
155. De Maesschalck, R., Jouan-Rimbaud, D. & Massart, D. L. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems* **50**, 1–18 (2000).
156. Groenendaal, G. *et al.* Pathologic Validation of a Model Based on Diffusion-Weighted Imaging and Dynamic Contrast-Enhanced Magnetic Resonance Imaging for Tumor Delineation in the Prostate Peripheral Zone. *International Journal of Radiation Oncology*Biological*Physics* **82**, e537–e544 (2012).
157. Korporaal, J. G. *et al.* The use of probability maps to deal with the uncertainties in prostate cancer delineation. *Radiotherapy and Oncology* **94**, 168–172 (2010).

158. Tisseverasinghe, S. A. & Crook, J. M. The role of salvage brachytherapy for local relapse after external beam radiotherapy for prostate cancer. *Translational Andrology and Urology* **7**, 414–435 (2018).
159. Aaronson, D. S. *et al.* Salvage permanent perineal radioactive–seed implantation for treating recurrence of localized prostate adenocarcinoma after external beam radiotherapy. *BJU International* **104**, 600–604 (2009).
160. Šimundić, A.-M. Measures of Diagnostic Accuracy: Basic Definitions. *EJL-FCC* **19**, 203–211 (2009).
161. Peng, Y. *et al.* Validation of quantitative analysis of multiparametric prostate MR images for prostate cancer detection and aggressiveness assessment: a cross-imager study. *Radiology* **271**, 461–471 (2014).
162. Greer, M. D. *et al.* Computer-aided diagnosis prior to conventional interpretation of prostate mpMRI: an international multi-reader study. *European Radiology* **28**, 4407–4417 (2018).
163. Riley, R. D. *et al.* Minimum sample size for developing a multivariable prediction model: PART II – binary and time-to-event outcomes. *Statistics in medicine* **38**, 1276–1296 (2019).
164. Morin, O. *et al.* A Deep Look Into the Future of Quantitative Imaging in Oncology: A Statement of Working Principles and Proposal for Change. *International Journal of Radiation Oncology*Biography*Physics* **102**, 1074–1082 (2018).
165. Arrayeh, E. *et al.* Does local recurrence of prostate cancer after radiation therapy occur at the site of primary tumor? Results of a longitudinal MRI and MRSI study. *International Journal of Radiation Oncology*Biography*Physics* **82**, 787–793 (2012).
166. Chopra, S. *et al.* Pathological predictors for site of local recurrence after radiotherapy for prostate cancer. *International Journal of Radiation Oncology*Biography*Physics* **82**, 441–8 (2012).
167. Shiradkar, R. *et al.* Radiomic features from pretreatment biparametric MRI predict prostate cancer biochemical recurrence: Preliminary findings. *Journal of Magnetic Resonance Imaging* **48**, 1626–1636 (2018).

168. Nyholm, T., Nyberg, M., Karlsson, M. G. & Karlsson, M. Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments. *Radiation Oncology* **54**. doi:10.1186/1748-717X-4-54 (2009).
169. Maspero, M. *et al.* Evaluation of gold fiducial marker manual localisation for magnetic resonance-only prostate radiotherapy. *Radiation Oncology* **13**, 105 (2018).
170. Maspero, M. *et al.* Evaluation of an automatic MR-based gold fiducial marker localisation method for MR-only prostate radiotherapy. *Physics in Medicine & Biology* **62**, 7981–8002 (2017).
171. De Roover, R. *et al.* Characterization of a novel liquid fiducial marker for multimodal image guidance in stereotactic body radiotherapy of prostate cancer. *Medical Physics* **45**, 2205–2217 (2018).
172. Bittner, N. *et al.* Electromagnetic tracking of intrafraction prostate displacement in patients externally immobilized in the prone position. *International Journal of Radiation Oncology*Biological*Physics* **77**, 490–495 (2010).
173. Koch, K. M. *et al.* Multispectral diffusion-weighted imaging near metal implants. *Magnetic Resonance in Medicine* **79**, 987–993 (2018).
174. Czarniecki, M. *et al.* Role of PROPELLER-DWI of the prostate in reducing distortion and artefact from total hip replacement metalwork. *European Journal of Radiology* **102**, 213–219 (2018).
175. Hollmann, B. G. *et al.* Gross tumor volume and clinical target volume in prostate cancer: How do satellites relate to the index lesion. *Radiotherapy and Oncology* **115**, 96–100 (2015).
176. Maenhout, M. *et al.* Focal MRI-Guided Salvage High-Dose-Rate Brachytherapy in Patients With Radiorecurrent Prostate Cancer. *Technology in cancer research & treatment* **16**, 1194–1201 (2017).
177. Kanthabalan, A. *et al.* Focal salvage high-intensity focused ultrasound in radiorecurrent prostate cancer. *BJU International* **120**, 246–256 (2017).

SUMMARY

MULTI-PARAMETRIC MRI TO GUIDE SALVAGE TREATMENT OF RECURRENT PROSTATE CANCER

Depending on the tumour characteristics, up to 50% of patients treated with radiotherapy for primary prostate cancer will experience recurrent disease within 10 years of the primary treatment.

Chapter 2 shows evidence that pre-treatment T2w MRI features could potentially improve the biochemical recurrence risk prediction in a clinically homogeneous patient cohort. The use of whole-prostate imaging features resulted in an AUC = 0.63, outperforming a model solely based on clinical variables (AUC = 0.51). With the goal of removing inter-modality registration errors and improving the logistics of radiotherapy treatments, a new MR-only radiotherapy treatment planning workflow is currently under investigation. Despite being clearly visible on CT, the fiducial markers used for daily position verification can be hard to detect on MRI. In *Chapter 3* of this thesis we described how features from T1w, T2w, B₀ and a balanced TFE sequence can be combined to automatically detect fiducial markers with a good accuracy and a true positive detection rate of 0.94.

No clinical consensus exists on how to treat locally recurrent prostate cancer. Up to now, due to difficulties in assessing localised recurrent disease, salvage treatment modalities were typically targeted at the entire gland. However, whole-gland salvage treatments resulted in high failure rates and severe toxicity. Due to its high spatial resolution and functional possibilities, multi-parametric MRI (mp-MRI) (comprising T2-weighted, diffusion-weighted imaging (DWI) and dynamic contrast enhanced (DCE) MRI together with their derived maps) has, in recent years, become the modality of choice to detect and isolate locally recurrent cancer. Improved tumour visualisation and localisation has fuelled the development of focal salvage treatment approaches, which aim at targeting the tumour area while sparing the surrounding uncompromised tissue. Nevertheless, information is lacking on the characteristics of prostate tissue after radiotherapy on mp-MRI, and no guidelines or consensus exists on how to detect and localise the region to target with focal salvage treatments. The main purpose of the last three studies presented in this thesis was to improve the de-

tection and localisation of recurrent prostate cancer after primary radiotherapy treatment.

Chapter 4 uses mp-MRI to characterise benign irradiated tissue and recurrent prostate cancer. Imaging values for benign irradiated tissue were comparable between patients with and without recurrent prostate cancer, both presenting with increased peri-urethral enhancement in the pharmacokinetic K^{trans} and k_{ep} maps without evident signs of malignancy. Benign central-gland values were for all maps significantly decreased after radiotherapy, as well as T_2 , K^{trans} , and k_{ep} values of benign peripheral zone. Distinguishing recurrent tumour from benign central-gland was not possible with T_2 map. For all maps, recurrent and primary prostate cancer tumours had similar imaging characteristics.

In *Chapter 5*, the use of high-resolution histopathology delineations has enabled an accurate study of recurrent tumour characteristics and a comparison between manually-defined tumour regions and the ground-truth pathological extent on mp-MRI. A different recurrence pattern was observed in external-beam radiotherapy and low-dose brachytherapy treated patients. Delineations based on mp-MRI tended to underestimate the real tumour extent and size, suggesting that focal salvage strategies should include appropriate margins around the visible tumour. In both studies (*Chapter 4 and 5*) the use of all mp-MRI sequences (T_2w / T_2 map, ADC, K^{trans} and k_{ep} maps) together with the voxel anatomical location (peripheral zone or central gland), resulted in an optimal distinction between tumour and benign voxels.

Chapter 6 uses mp-MRI information to create a tumour probability model from which voxel-wise tumour probabilities can be obtained. The model's highest ranking features were derived from the DCE pharmacokinetic maps. Clustering of the voxel-wise tumour probabilities was used to define a clinical target volume (CTV), a high-risk CTV and a gross tumour volume (GTV). The sensitivity and specificity of the high-risk CTV and GTV model-derived contours in localising recurrent tumour were similar to those of manual tumour delineations based on mp-MRI. This suggests that voxel-wise probabilities could in this way be incorporated for target definition in focal salvage treatments of recurrent prostate cancer after radiotherapy.

Chapters 4, 5 and 6 of this thesis highlight the importance of DCE-MRI derived parameters for the detection and localisation of recurrent prostate cancer after radiotherapy. The findings presented in this thesis can be used as a foundation to establish guidelines for tumour detection and definition of the region to target with focal salvage treatments.

SAMENVATTING

MULTI-PARAMETRISCHE MRI VOOR DE SALVAGE BEHANDELING VAN RECIDIEF PROSTAATKANKER

Afhankelijk van de tumorkarakteristieken, zal tot 50% van de patiënten behandeld met radiotherapie voor primaire prostaatkanker binnen 10 jaar na de behandeling een recidief ontwikkelen. In *Hoofdstuk 2* tonen we in een klinisch homogene groep patiënten aan dat kenmerken in T2-gewogen MRI beelden, gemaakt vóór de behandeling, een betere voorspelling mogelijk maken van het risico op een biochemisch recidief. Het gebruik van karakteristieken in de beeldvorming van de gehele prostaat, resulteerde in een $AUC = 0.63$, wat een verbetering is ten op zichte van een model dat uitsluitend gebaseerd is op klinische variabelen ($AUC = 0.51$).

Om registratiefouten tussen MRI en CT te vermijden en de logistiek van radiotherapiebehandelingen te verbeteren, wordt momenteel een nieuwe werkwijze onderzocht voor een behandeling met radiotherapie, die voor beeldvorming alleen gebaseerd is op MRI. Gouden markers, geïmplanteerd in de prostaat, zijn op CT beelden goed zichtbaar en worden routinematig gebruikt voor dagelijkse positieverificatie. Echter, op MRI beelden zijn deze markers moeilijk te detecteren. In *Hoofdstuk 3* van dit proefschrift hebben we beschreven hoe met een combinatie van T1w, T2w, B₀ en balanced-TFE-sequenties deze markers automatisch kunnen worden gedetecteerd.

Tot op heden bestaat er geen klinische consensus over de behandeling van een lokaal recidief van prostaatkanker. Tot nog toe waren behandelingen hiervan doorgaans gericht op de gehele prostaat. Dit resulteerde echter in veel bijwerkingen. Met een hoge spatiële resolutie en variatie in contrast, is multi-parametrische MRI (mp-MRI), bestaande uit T2-gewogen, diffusie-gewogen (DWI) en contrast-versterkte (DCE) MRI, de aangewezen manier om een lokaal recidief van prostaatkanker te detecteren en af te lijnen. Dit biedt de mogelijkheid om focale behandelingen uit te voeren, die zich enkel richten op de recidief tumor, waarbij omliggend weefsel wordt gespaard. Echter, kennis over de beeldvorming met mp-MRI van de bestraalde prostaat is beperkt. Richtlijnen ontbreken over hoe de beelden moeten worden geïnterpreteerd en hoe de tumor moet worden afgelijnd. De laatste drie studies van dit proefschrift

zijn daarom gericht op het verbeteren van de detectie en aflijning van recidief prostaatcancer na eerdere behandeling met radiotherapie.

Hoofdstuk 4 beschrijft het gebruik van mp-MRI om de effecten van bestraling op prostaatweefsel te karakteriseren. Kwantitatieve parameters, afgeleid van mp-MRI waren bij patiënten zonder recidief vergelijkbaar met de waarden die bij patiënten met recidief werden gevonden in de delen van de prostaat waarin geen tumor voorkwam. In beide groepen was een verhoging te zien rondom de urethra van de farmacokinetische parameters K^{trans} en k_{ep} zonder duidelijke tekenen van maligniteit. Ten opzichte van onbehandelde patiënten waren in de rest van de prostaat T_2 , K^{trans} en k_{ep} verlaagd. Dat verklaart waarom het moeilijk is een recidief tumor te onderscheiden van omliggend weefsel. De recidief tumoren hadden vergelijkbare karakteristieken op mp-MRI als onbehandelde tumoren.

In *Hoofdstuk 5* hebben we beschreven hoe karakteristieken op mp-MRI beelden samenhangen met histopathologie. Met gedetailleerde intekeningen van recidief prostaat tumoren op de histopathologische coupes werden de intekeningen van een radioloog op mp-MRI beelden gevalideerd. Bij patiënten die behandeld waren met uitwendige radiotherapie, werd een andere verdeling van recidief prostaatcancer waargenomen dan bij patiënten behandeld met lage dosis brachytherapie. Intekeningen op mp-MRI bleken de omvang van de tumor te onderschatten, zoals waargenomen in histopathologie. Dit betekent dat bij een focale behandeling een veilige marge rondom de zichtbare tumor in acht moet worden genomen. In de studies beschreven in *Hoofdstuk 4* en *5* bleek dat met een combinatie van alle mp-MRI sequenties (T_2w / T_2 map, ADC, K^{trans} en k_{ep} mappen) het beste onderscheid gemaakt kon worden tussen tumor- en goedaardige voxels.

Hoofdstuk 6 werd op basis van mp-MRI een model gemaakt waarbij op voxelniveau de kans op aanwezigheid van tumor wordt weergegeven. De belangrijkste MRI karakteristieken voor het model waren gebaseerd op de farmacokinetische parameters k_{ep} en K^{trans} die zijn afgeleid van DCE-MRI. Op basis van de kans op tumor werd een klinisch doelvolumen (CTV) gedefinieerd waarin zich microscopische tumor kan bevinden, een volume waarvoor een ernstige verdenking bestaat van aanwezigheid van tumor (high-risk CTV) en een macroscopisch tumorvolume (GTV). De sensitiviteit en specificiteit van het high-

risk CTV en GTV zoals afgeleid van het model, waren vergelijkbaar met die van handmatige intekeningen. Dit betekent dat een model dat de kans op tumor op voxelniveau weergeeft bruikbaar is voor het bepalen van een doelgebied voor focale behandeling van recidief prostaatkanker na radiotherapie.

Hoofdstuk 4, 5 en 6 van dit proefschrift onderstrepen het belang van DCE-MRI voor de detectie en lokalisatie van recidief prostaatkanker na radiotherapie. De resultaten die in dit proefschrift zijn gepresenteerd, kunnen gebruikt worden als basis voor het opstellen van richtlijnen voor detectie van recidief tumoren in de prostaat en het aflijnen van het doelgebied voor een focale behandeling.

SUMÁRIO

RESSONÂNCIA MAGNÉTICA MULTIPARAMÉTRICA COMO MEIO DE GUIAR O TRATAMENTO DO CANCRO DA PRÓSTATA RECORRENTE

O cancro da próstata é uma condição patológica que afecta mais de 40% dos homens com idades acima de 60 anos. Um dos tratamentos mais comuns para o cancro da próstata é a radioterapia. Contudo, mesmo depois do tratamento de radioterapia, existe a possibilidade de que o tumor volte a crescer e, caso tal aconteça, falamos de uma recidiva ou um tumor recorrente. Vários factores podem influenciar o risco de recidiva, como por exemplo a extensão ou a agressividade do tumor inicial. O *Capítulo 2* descreve os resultados obtidos quando uma sequência anatómica (T2w) de ressonância magnética (RM) adquirida antes do tratamento inicial de radioterapia é usada para prever o risco que um paciente tem em sofrer uma recidiva. Parâmetros clínicos, tal como a extensão do tumor ou o grau de diferenciação, são usados como meio de prever o risco de recorrência. O cancro da próstata pode ser assim categorizado em três grupos de risco (baixo, médio e alto), e a escolha do tratamento é feita com base nestes parâmetros. Em concreto, este estudo focou-se num grupo clinicamente homogêneo de pacientes com tumores de alto risco, em que em alguns dos casos o paciente sofreu uma recidiva enquanto em outros não. Neste estudo demonstramos que um classificador que utiliza as características extraídas da imagem na região da próstata é capaz de superar um classificador que usa apenas parâmetros clínicos no que toca a prever o risco de recidiva. Havendo a possibilidade de ter acesso *a priori* à informação de que um paciente tem um alto risco de recidiva, é possível oferecer um regime de tratamento intensificado ou um maior acompanhamento.

Durante o planeamento do tratamento de radioterapia, a RM é utilizada para a delimitação das estruturas enquanto que as doses são calculadas usando a tomografia axial computadorizada (TAC). Recentes avanços no planeamento do tratamento de radioterapia focam-se na completa remoção do TAC, sendo então as suas tarefas substituídas pela RM. Marcadores fiduciais são implantados no paciente para o posicionamento diário durante as diferentes frações do tratamento. Apesar de visíveis no TAC, estes são difíceis de detectar na RM. No *Capítulo 3* demonstramos que uma combinação de várias sequências de RM

podem ser utilizadas para automaticamente identificar os marcadores fiduciais com uma boa precisão.

Para os pacientes com uma recidiva localizada sem a presença de metástases distantes, há uma última opção possível de tratamentos curativos, aos quais se dá o nome de tratamentos de resgate. Existem vários tipos de tratamentos de resgate, mas estes não são frequentemente oferecidos pois as chances de sucesso são limitadas e é comum resultarem em efeitos colaterais sérios e persistentes. Com o objectivo de diminuir os efeitos tóxicos dos tratamentos convencionais de resgate que tratam a próstata por completo, está a ser investigada a possibilidade de tratar apenas o tumor. A perspectiva é de que ao tratar apenas a zona onde se encontra o tumor, a toxicidade para os tecidos e órgãos na periferia da próstata seja diminuída. Para poder oferecer esta opção de tratamento é necessária uma identificação atempada de recorrência enquanto esta ainda é localizada. Métodos de diagnóstico, tal como a ressonância magnética multiparamétrica (mp-RM) que possibilita a obtenção de imagens não só da anatomia mas também da fisiologia dos tecidos, oferecem uma oportunidade para a detecção atempada do tumor recorrente. Este é também o método mais comum de localizar o tumor que deve ser visado com os métodos focais de tratamento de resgate. Contudo, pouca informação existe sobre as características do tecido da próstata depois do tratamento inicial com radioterapia. O objectivo principal dos capítulos apresentados na segunda parte da tese é o de melhorar a detecção e a localização do tumor recorrente após o tratamento inicial com radioterapia.

Para tal, o *Capítulo 4* utiliza mp-RM para a caracterização tanto do tumor recorrente como dos tecidos benignos. A estrutura do estudo foi caso-controle, e portanto o grupo analisado era constituído tanto por pacientes com tumor recorrente como por pacientes sem suspeita de recorrência. Foram ainda analisadas imagens de pacientes antes de receberem o tratamento de radioterapia. Como tal, foi assim possível fazer uma caracterização quantitativa de mudanças benignas provavelmente induzidas pelo tratamento de radioterapia. Tanto imagens anatómicas como funcionais foram analisadas nas diferentes regiões da próstata e no tumor recorrente. A informação descrita neste estudo ajuda na interpretação das imagens e para que o tumor consiga ser distinguido de mudanças benignas originadas pela radioterapia.

O *Capítulo 5* foca-se no estudo do tumor recorrente, utilizando não só imagens adquiridas aquando do diagnóstico da recorrência, mas também a amostras patológicas obtidas quando estes pacientes foram sujeitos a uma cirurgia de remoção da próstata como tratamento de resgate. Foi assim possível utilizar delineações de alta resolução, obtidas com base na patologia e transferidas para as imagens de RM, para um estudo preciso das características do tumor recorrente. Delineações foram também feitas por dois radiologistas experientes com base na mp-RM. Quando estas últimas delineações foram comparadas com as que foram obtidas através da patologia foi possível observar que a verdadeira extensão e tamanho do tumor tendem a ser subestimados, sugerindo que estratégias focais de tratamento de resgate devem incluir margens à volta do tumor visível na mp-RM. A combinação das imagens anatómicas T2w /T₂ map e funcionais ADC, K^{trans} and k_{ep} em conjunto com a localização dentro da próstata resultou em ambos os estudos (*Capítulos 4 e 5*) na melhor distinção entre tumor e tecido benigno.

Por fim, o *Capítulo 6* demonstra como as sequências de mp-RM podem ser usadas num modelo para criar um mapa que indica para toda a próstata qual é a probabilidade de conter tumor. As imagens que mais contribuíram para esta distinção foram o k_{ep} and K^{trans}, derivadas com base nas propriedades de perfusão dos tecidos. As probabilidades foram posteriormente aglomeradas para criar 3 zonas com um risco incremental de conterem tumor. Estas regiões foram comparadas com as delineações obtidas a partir da patologia. Foram observadas sensibilidades relativamente baixas mas alta especificidade em detectar o tumor recorrente. Ao comparar as regiões definidas pelo modelo com as delineações feitas por dois radiologistas experientes foi possível observar que estas eram semelhantes em termos de sensibilidade e especificidade. Estes resultados sugerem que estes mapas de probabilidades podem ser incorporados no planeamento de tratamentos focais de resgate.

Em conclusão, os resultados descritos nos capítulos desta tese podem ser usados como base para estabelecer directrizes em como detectar e localizar a região a tratar com os tratamentos focais de resgate. Os últimos três capítulos realçam ainda a importância dos parâmetros de perfusão na identificação do tumor recorrente. Por fim, modelos que utilizem a mp-RM podem ser úteis para uma definição automática do tumor a tratar com opções focais de resgate.

LIST OF PUBLICATIONS

This thesis is based on the following publications:

Prostate fiducial marker detection with the use of multi-parametric magnetic resonance imaging

Dinis Fernandes, C., Dinh, C.V., Steggerda, M.J., ter Beek, L.C., Smolic, M., van Buuren, L.D., Pos, F.J. and van der Heide, U.A.

Physics and Imaging in Radiation Oncology **1**, 14–20 (2017).

DOI: [10.1016/j.phro.2017.02.001](https://doi.org/10.1016/j.phro.2017.02.001)

Biochemical recurrence prediction after radiotherapy for prostate cancer with T2w magnetic resonance imaging radiomic features

Dinis Fernandes, C., Dinh, C.V., Walraven, I., Heijmink, S.W.T.P.J., Smolic, M., van Griethuysen, J.J.M., Simões, R., Losnegård, A., van der Poel, H.G., Pos, F.J. and van der Heide, U.A.

Physics and Imaging in Radiation Oncology **7**, 9–15 (2018).

DOI: [10.1016/j.phro.2018.06.005](https://doi.org/10.1016/j.phro.2018.06.005)

Quantitative 3T multiparametric MRI of benign and malignant prostatic tissue in patients with and without local recurrent prostate cancer after external-beam radiation therapy

Dinis Fernandes, C., van Houdt, P.J., Heijmink, S.W.T.P.J., Walraven, I., Keesman, R., Smolic, M., Ghobadi, G., van der Poel, H.G., Schoots, I.G., Pos, F.J. and van der Heide, U.A.

Journal of Magnetic Resonance Imaging **50**(1), 269–278 (2019).

DOI: [10.1002/jmri.26581](https://doi.org/10.1002/jmri.26581)

Quantitative 3-T multi-parametric MRI and step-section pathology of recurrent prostate cancer patients after radiation therapy

Dinis Fernandes, C., Ghobadi, G., van der Poel, H.G., de Jong, J., Heijmink, S.W.T.P.J., Schoots, I.G., Walraven, I., van Houdt, P.J., Smolic, M., Pos, F.J. and van der Heide, U.A.

European Radiology (2018), *Epub ahead of print*

DOI:[10.1007/s00330-018-5819-y](https://doi.org/10.1007/s00330-018-5819-y)

Multiparametric MRI tumor probability model for the detection of locally recurrent prostate cancer after radiation therapy: pathologic validation and comparison with manual tumor delineations

Dinis Fernandes, C., Simões, R., Ghobadi, G., Heijmink, S.W.T.P.J., Schoots, I.G., de Jong, J., Walraven, I., van der Poel, H.G., van Houdt, P.J., Smolic, M., Pos, F.J. and van der Heide, U.A.

*International Journal of Radiation Oncology*Biophysics* (2019), *Epub ahead of print*.

DOI: [10.1016/j.ijrobp.2019.05.003](https://doi.org/10.1016/j.ijrobp.2019.05.003)

Other publications:

Stroke recovery and lesion reduction following acute isolated bilateral ischaemic pontine infarction: A case report

Varsou, O., Stringer, M.S., **Dinis Fernandes, C.**, Schwarzbauer, C., MacLeod, M.J.

BMC research notes **7**(1), 728 (2014)

DOI:[10.1186/1756-0500-7-728](https://doi.org/10.1186/1756-0500-7-728)

Use of wavelet transforms and detrended fluctuation analysis in the evaluation of human skin microcirculation

Ferreira, H.A., **Dinis Fernandes, C.**, Pinto, P.C., Rodrigues, L.M.

Biomedical and Biopharmaceutical Research **9**(1), 79-86 (2011)

CONFERENCE ABSTRACTS

- 2015 Characterising irradiated prostate tissue in radio-recurrent PCa patients using mp-MRI and histopathology
Dinis Fernandes, C*, Ghobadi, G., de Jong, J., Dinh, C.V., van Houdt, P.J., Smolic, M., van der Poel, H.G., van der Heide, U.A.
Biology-Guided Adaptive Radiotherapy, Aarhus, Denmark
*Poster presentation
- 2016 Prostate fiducial markers detection with the use of multiparametric-MRI
Dinis Fernandes, C*, Dinh, C.V., Steggerda, M.J., ter Beek, L.C., Smolic, M., van Buuren, L.D., van Houdt, P.J., van der Heide, U.A.
ESTRO 35, Turin, Italy
*Oral presentation
- 2017 Outcome prediction after local prostate cancer treatment with the use of T2w magnetic resonance imaging (MRI): A radiomics approach
Dinis Fernandes, C*, Dinh, C.V., Walraven, I., Smolic, M., Losnegård, A., Heijmink, S.W.T.P.J., van der Poel, H.G., Muren, L.P., Pos, F.J., van der Heide, U.A.
5th MR in RT Symposium, Sydney, Australia
*Oral presentation
- 2017 Quantitative mp-MRI of recurrent prostate cancer post-radiotherapy: a matched case-control study
Dinis Fernandes, C*, van Houdt, P.J., Heijmink, S.W.T.P.J., de Jong, J., van der Poel, H.G., Smolic, M., Pos, F.J., van der Heide, U.A.
ESMRMB 2017, Barcelona, Spain
*Oral presentation
- 2018 Multi-parametric imaging for dose targeting in salvage radiotherapy of prostate cancer
van der Heide, U.A.*, **Dinis Fernandes, C**
ESTRO 37, Barcelona, Spain
*Oral presentation

2018 Tumour probability model for focal salvage strategies in recurrent prostate cancer patients after radiotherapy

Dinis Fernandes, C*, Ghobadi, G., van Houdt, P.J., van der Poel, H.G., de Jong, J., Heijmink, S.W.T.P.J., Smolic, M., Pos, F.J., van der Heide, U.A.

6th MR in RT Symposium, Utrecht, The Netherlands

*Oral presentation

ACKNOWLEDGEMENTS

Uulke, first of all I would like to thank you for giving me this opportunity and for the countless hours dedicated to this thesis. I really appreciate that you have always supported and encouraged me to pursue my own ideas. Your enthusiasm, trust and motivation have kept me going when I was in doubt; and your guidance, around-the-clock availability and critical eye made sure this thesis came to be. It has been a great pleasure learning from and working with you.

Milena, despite our very different fields, it was your mentorship, guidance and encouragement that made it possible for me to reach this point. Thank you for your unconditional support and availability.

A special thank you goes to my paranympths **Edzo** and **Yenny**, whom, in very different ways, played such a huge role during my PhD. I'm proud to say we are not just colleagues but also good friends. Edzo, it has been 5 remarkable years sharing an office with you and drinking more coffee than it should be humanly possible to. You were always generous, supportive and positive which helped us both navigate the ups and downs of a PhD. Yenny, we have totally different personalities but that hasn't stopped us from becoming good friends! You were always caring and honest, and showed me so many different new ways to look at and live life outside of work. I had so much fun with both of you and have lost count of the occasions in which we laughed uncontrollably. You are both such an inspiration, and know that I am truly grateful for your unconditional friendship and encouragement.

During my PhD I had the privilege to work with so many spectacular people from this and other departments. I would like to thank all of my fellow **PhD students** and **all other colleagues**, both present and past, for making these the most amazing 5 years! **Diedie** and **Patricia**, thank you for always being available to solve any problem, you are the moving force of this department. In

particular, I would like to acknowledged those I had the chance to work with more closely – **Ahmed, Anke, Cuong, Ernst, Ghazaleh, Iris, Laurens, Marcel, Peter, Petra, Rick, Rita, Roelant, Roque** – it has been fun and a great pleasure working with you all! **Kleopatra** and **Natasja**, it was really special having you around, thank you for the many incredible and memorable times in and (mostly) outside of the office. **Roel**, it wouldn't have been the same without you fuelling my surfing dreams.

To all friends that live near and far, and whom in one way or another were present and supportive – thank you all. A special mention goes to: **Nicole** for your nonstop positivity and encouragement; **Lian** and **Jeroen** for being always available, supportive and with a word of advice; **Joep, Nico, Tamara, Wieke**, (Riouwstraat!), your constant cheering, interest in my work and celebration of every small victory had such a huge impact on how I experienced this PhD. Thank you, these have been two unforgettable years.

Obrigada a todos os amigos portugueses que, e apesar da distância, estão de um modo ou outro *sempre* por perto. Em especial, gostaria de agradecer à **Diana** e à **Lotte** que, e independentemente das voltas que a vida dê, são uma fonte de inspiração e positivismo.

I would also like to express my gratitude to the whole **Kuilboer** family, for always welcoming me as one of their own.

Dear **David**, thank you for being unconditionally loving, supportive, caring, understanding, patient and motivational. You always made sure there was a freshly cooked meal and a comforting word after a long working day, and you stood by my side even when I was far from being the best company. I am incredibly blessed to have you in my life and I most certainly could not have made it without you. Now it's time to pick our van, load the surfboards and go on our long awaited trip!

Um obrigado muito especial vai para os meus queridos **pais** e para os meus **avós**: não existem palavras suficientes para expressar o quanto eu estou grata pelo vosso constante apoio e carinho, sem os quais esta tese não seria possível.

CURRICULUM VITAE

My name is Catarina Dinis Fernandes and I was born in Lisbon on the 28th of January 1989. I grew up in Sintra, where I also attended high school. In 2007, I started an integrated bachelor's and master's program in Biophysics and Biomedical Engineering at the University of Lisbon. In the summer of 2011, and as part of my degree, I did a three-month research project at the Centre for Advanced Biomedical Imaging of the University College in London. During this period, I focused on developing stimulation paradigms and improving functional MRI sequences on a 9.4T pre-clinical MRI scanner. This work ultimately culminated in my bachelor's thesis. It was also during this time that I became interested in medical imaging and specifically MRI. In 2012, I joined the Aberdeen Biomedical Imaging Centre (ABIC), Scotland, for an internship that formed the foundation of my master's thesis. There, I spent one year working on resting-state functional MRI in health and disease states. At the end of 2013, I graduated as a Master of Science with a specialisation in 'Radiation for Diagnosis and Therapy' from the University of Lisbon. I spent the following months working at the ABIC as a research assistant, but I was curious to learn how MRI could be used for purposes other than neuroimaging. This led me to Amsterdam where, in May 2014, I started a PhD at the Radiation Oncology Department of the Netherlands Cancer Institute. This thesis describes the research conducted during this PhD on the use of multi-parametric MRI in recurrent prostate cancer after initial radiotherapy, and how it can be used for the planning of focal salvage treatments.