OPTIMISING PROSTATE CANCER STAGING AND ACTIVE SURVEILLANCE

TIMO F.W. SOETERIK

OPTIMISING PROSTATE CANCER STAGING AND ACTIVE SURVEILLANCE

TIMO F.W. SOETERIK

Optimising Prostate Cancer Staging and Active Surveillance
PhD thesis, Utrecht University, The Netherlands

The research presented in this thesis was financially supported by means of independent grants from Astellas Pharma B.V., Amgen B.V. and St. Antonius Onderzoeksfonds.

Publication of this thesis was financially supported by St. Antonius Ziekenhuis Raad van Bestuur, de maatschap Urologen voor U, de maatschap urologie Canisius Wilhelmina Ziekenhuis, Hoogland Medical B.V., Pfizer B.V., Chipsoft B.V., Stöpler B.V., Coloplast B.V., Elekta B.V., IPSEN Farmaceutica B.V., Bayer B.V. and Ferring B.V.

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Cover design: Leon Postma, Visual Designer
Layout and design: Dagmar van Schaik, persoonlijkproefschrift.nl
ISBN: 978-94-6416-866-2

OPTIMISING PROSTATE CANCER STAGING AND ACTIVE SURVEILLANCE

Optimalisatie van Prostaatkanker Stadiëring en Actief Afwachtend Beleid
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 4 november 2021 des middags te 12.15 uur

door

Timotëus Franciscus Willibrordus Soeterik

geboren op 28 september 1991
te Weert

Promotoren: Prof. dr. D.H. Biesma
Prof. dr. J.A. Witjes

Copromotoren: Dr. L.M. Dijksman
Dr. J.P.A. van Basten

Voor mijn ouders

PREFACE

Prostate cancer is a disease with a tremendous individual, societal and economic impact. The overall burden of this entity and its associated treatments have been shown to be extremely high in both human and financial terms. Optimising prostate cancer staging can lead to higher treatment benefits for the patient with regard to oncological control, while retaining quality of life by minimizing intervention-related morbidity. Improved understanding of the disease stage also helps to successfully avoid or postpone interventions and associated side-effects in those with indolent tumours, while ensuring patients in whom active treatment is indicated are successfully identified.

In this thesis, we describe the main findings of our search towards optimised prostate cancer outcomes by means of improving initial tumour staging, combined with further exploring non-invasive prostate cancer management by expanding our knowledge on active surveillance.

CONTENTS

CHAPTER 1

General Introduction and Thesis Outline

PROSTATE CANCER

Prostate cancer: the importance of accurate staging

Patients with newly diagnosed prostate cancer are confronted with a disease that can present itself in a wide heterogeneity of forms. Initial presentation can vary from indolent cancer with very limited impact on life expectancy without need for immediate treatment, to subtypes that are characterized by rapid progression with potential lethal outcome, demanding a timely and invasive treatment strategy. The more accurately the aggressiveness and stage of disease are characterized at diagnosis, the higher the likelihood of selecting the treatment option providing the best value in terms of oncological and health-related quality of life outcomes.

Accurate staging of prostate cancer can be established by combining the most relevant baseline characteristics and by incorporating those into risk classification systems and prediction tools, enabling personalized treatment advice. Given the developments in the field of medical imaging (e.g. magnetic resonance imaging [MRI]), regular validation and updates of these tools and classification systems are essential to ensure prognostication and choice of treatment are done as accurate as possible given the current state of knowledge.

Prostate cancer epidemiology

Prostate cancer is a very common disease and is currently the second most frequently diagnosed cancer and third most common cause of cancer death among males worldwide. In 2019, a total of 13557 new diagnoses and 2896 prostate cancer-related deaths were reported in the Netherlands.[1] Taking into account expected population growth and aging, incidence is expected to increase between 40 and 60% in 2040.[2] Of all newly diagnosed Dutch patients, approximately 65% are affected with prostate cancer that is limited to the prostate.[1] Another 20% of the patients are diagnosed with locally advanced cancer (extension of the disease outside the prostate capsule and/or metastasis limited to the pelvic lymph nodes), whereas 15% constitutes of patients with distant metastasis (distant lymph node, bone and/or visceral metastasis).[1]

History of prognostic factors established at initial diagnosis

Several clinical, biochemical and histological characteristics found at initial diagnosis of prostate cancer have been evaluated as potential factors providing information regarding the expected disease course. Currently, the most established factors include clinical stage of the tumour (cT), serum prostate-specific antigen (PSA) level at diagnosis and highest Gleason score found on prostate biopsy.

Clinical T-stage

The cT describes the primary tumour's size and extent. In prostate cancer, the conventional modality to establish cT of the tumour is by digital rectal examination (DRE). Findings established by DRE, such as palpation of a nodule or extraprostatic extension, were considered valuable to establish disease stage.[3] Ever since the first prostate cancer staging system was established, information regarding primary tumour size, assessed by digital rectal examination, was included as an essential component.[3] Throughout the years, staging systems have evolved and the cT is now part of the widely accepted tumour, nodal and metastasis (TNM) classification of malignant tumours.[4] At present, the 8th Edition of the AJCC is used to define the cT (Table 1).[5]

Histological grading of prostate cancer

The first structured approach to grade prostate cancer based on the underlying histological architecture was developed in 1966 by Donald Gleason, who proposed a morphologic classification of prostate cancer.[6] The grading system showed to contribute significantly to the mortality rate prediction information, in addition to that provided by the clinical staging of the tumour.[7] To establish the Gleason score, the pathologist assigns a primary score and a secondary score. The primary score includes the most dominant pattern of the tumour, that has to be greater than 50% of the total pattern observed. The secondary score is based on the next-most frequent pattern, less than 50% of the total pattern, and at least 5%. Subsequently, the total score is summed resulting in the Gleason sum score, with a range of 2 to 10. Contemporarily referred to as the Gleason grading system, this scoring system has gained worldwide recognition allowing a more individualized approach to patients with prostate cancer. In 2005, the International Society of Urologic Pathology (ISUP) made the first revisions to the grading system. Modifications are based on novel insights regarding the prognostic implications of different histological patterns that can be found in prostate cancer specimens. Most recent updates were established during the ISUP consensus conference in 2014, wherein various grade patterns were defined and a new grading system of prostate cancer was proposed (Table 2).[8]

TABLE 1. The AJCC UICC 8th edition prostate cancer clinical T-stage classification

Clinical T (cT)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
T1	Clinically inapparent tumour that is not palpable
T1a	Tumour incidental histologic finding in 5% or less of tissue resected
T1b	Tumour incidental histologic finding in more than 5% of tissue resected
T1c	Tumour identified by needle biopsy found in one or both sides, but not palpable
T2	Tumour is palpable and confined within prostate
T2a	Tumour involves one-half of one side or less
T2b	Tumour involves more than one-half of one side but not both sides
T2c	Tumour involves both sides
T3	Extraprostatic tumour that is not fixed or does not invade adjacent structures
T3a	Extraprostatic extension (unilateral or bilateral)
T3b	Tumour invades seminal vesicle(s)
T4	Tumour is fixed or invades adjacent structures other than seminal vesicles such as external sphincter, rectum, bladder, levator muscles, and/or pelvic wall

TABLE 2. The ISUP 2014 New Grading System Morphologic Patterns and Grade Group Pattern Composition

Grade Group	Histopathological pattern
Grade Group 1 (Gleason score 3+3=6)	Only individual discrete well-formed glands
Grade Group 2 (Gleason score 3+4=7)	Predominantly well-formed glands with a lesser component of poorly formed/fused/cribriform glands
Grade Group 3 (Gleason score 4+3=7)	Predominantly poorly formed/fused/cribriform glands with a lesser component of well-formed glands
Grade Group 4 (Gleason sum score 8)	Only poorly formed/fused/cribriform glands, or Predominantly well-formed glands with a lesser component lacking glands, or Predominantly lacking glands with a lesser component of well-formed glands
Grade Group 5 (Gleason sum scores 9–10)	Lacks gland formation (or with necrosis) with or w/o poorly formed/fused/cribriform

Prostate-specific antigen

Serum blood markers have always played an important role in prostate cancer screening and diagnosis. The first important serum tumour marker was prostate acid phosphatase (PAP), discovered in 1938 by Gutman.[9] PAP is a glycoprotein synthesized in lysosomes of prostate epithelial cells, of which increased levels can be found in the circulation of patients with prostate cancer. Since its discovery, it has been used as a prostate cancer screening marker and as an important component of prostate cancer classification systems.[3] However, clinical use of PAP drastically decreased since Wang et al in 1979 managed to purify and characterize an antibody against a human antigen that was prostate specific for the first time, known as prostate-specific antigen (PSA).[10] In several comparative analyses, PSA was shown to be a more sensitive marker than PAP for prostate cancer screening.[11,12] PSA is also shown to be a superior follow-up tool, as PSA levels routinely fall to undetectable levels after a radical prostatectomy; whereas PAP always remains detectable.[13] Even though PSA was advised not to be used in a population-based screening programme, due to a low positive predictive value of only 47%, its use as a screening tool increased significantly over the years.[12] The following "PSA screening era" subsequently led to a drastic increase in number of prostate cancer diagnoses, along with migration toward higher detection of indolent disease at diagnosis.[14]

Because the balance between benefits and harms is still not well established, a PSA-based population screening programme is currently not recommended. The European Randomised Study of Screening for Prostate Cancer showed that although PSA testing resulted in a 21% relative reduction in prostate cancer mortality in favour of screening, overall mortality rates in the screening and non-screening study arms were comparable.[15] In addition, the main downside of screening is the high detection rate of clinically irrelevant cancer, which occurs in approximately 40% of the screen-detected cases.[16] The current EAU guidelines therefore recommend not to subject men to PSA testing without counselling them on the potential benefits and harms.[17] In addition to its role in screening and treatment follow-up, PSA was shown to be a valuable prognostic variable. For example, high preoperative PSA values are associated with increased odds of extracapsular extension, seminal vesicle invasion and risk of biochemical progression after radical prostatectomy.[18]

Risk classification systems

Prostate cancer risk classification systems provide a qualitative assessment of the likelihood of progression after initial therapy. The most commonly used risk classification system for prostate cancer is the D'Amico risk classification.[19] The D'Amico classification divides men into low-, intermediate- and high-risk categories of progression after initial radical treatment, based on cT, biopsy Gleason grade, and pre-treatment levels of PSA.[19] The D'Amico risk classification has been adopted by clinicians worldwide and forms the

basis of the classification system recommended by the European Association of Urology (Table 3).[17]

Information regarding the likelihood of prostate cancer progression after initial radical therapy can facilitate tailored treatment approaches, wherein treatment intensity can be either increased or decreased based on the expected disease course. For instance, in patients with high-risk prostate cancer selected for radical prostatectomy, a wide resection can be preferred over a nerve sparing approach, in order to minimize the risk of positive surgical margins and thus the risk of disease recurrence.[17]

Clinical nomograms

Besides risk classification systems, clinical nomograms are also frequently used for disease prognostication. A nomogram is based on a mathematical formula that can be used to calculate the probability of an event to occur. It is used as a graphical display of a prediction model or calculation tool in such way that multiple variables can be taken into account. In the traditional implementation of the nomogram, the user draws pencil lines between axes, counts up points, and then reads off a prediction.[20] An advantage of nomograms over risk classification systems is the ability of establishing an individualized risk prediction, wherein the probability for a specific outcome is based on the characteristics of the individual patient.

An example of an established nomogram, based on PSA, clinical T-stage and Gleason score is the preoperative nomogram for prediction of disease recurrence after radical prostatectomy, developed by Kattan et al.[21] Another well-established nomogram based on the same parameters includes a nomogram for prediction of pelvic lymph node invasion.[22] Over the years, several novel nomograms have emerged such as the Memorial Sloan Kettering Cancer Center (MSKCC) pre-radical prostatectomy nomogram and the Briganti nomogram, which can be used to predict presence of pelvic lymph node metastasis.[23,24] The nomograms include additional prognostic factors such as relative number of positive biopsy cores taken on systematic biopsy at diagnosis.[23,24] For prostate cancer staging and treatment selection, several nomograms are recommended by the current EAU guidelines to support clinical decision-making.[17]

TABLE 3. EAU risk groups for biochemical recurrence of localised and locally advanced prostate cancer

Definition			
Low-risk	**Intermediate-risk**	**High-risk**	
PSA < 10 ng/mL	PSA 10–20 ng/mL	PSA > 20 ng/mL	any PSA
and GS < 7 (ISUP grade 1)	or GS 7 (ISUP grade 2/3)	or GS > 7 (ISUP grade 4/5)	any GS (any ISUP grade)
and cT1–2a	or cT2b	or cT2c	cT3–4 or cN+
Localised			**Locally advanced**

MAGNETIC RESONANCE IMAGING

Introduction of magnetic resonance imaging in risk prediction

The first generation of prediction models and risk classification systems for prostate cancer, based on baseline PSA, biopsy Gleason score and cT have substantially improved risk stratification and disease prognostication.[25,26] However, the utility of these tools can be further improved by incorporating additional specific tumour characteristics, such as tumour size, location and EPE. Novel prognostic indicators can be collected using advanced imaging modalities such as magnetic resonance imaging (MRI). Moreover, improved visualization of the prostate cancer lesion at initial diagnosis can also lead to more accurate estimation of the predictor variables. For example, targeting MRI-identified suspicious lesions using MRI-guided target biopsy can reduce the risk of tumour sampling error, resulting in a more accurate estimation of the prostate cancer Gleason grade.

Prostate cancer detection

Since its first application for prostate cancer imaging in the mid-1980s, MRI has evolved from a promising technique into a mature prostate imaging modality.[27] MRI can provide functional tissue information along with anatomic information. To increase the accuracy, anatomic T2-weighted MRI and imaging techniques such as dynamic contrast agent-enhanced imaging and diffusion-weighted imaging can be combined in an integrated multiparametric magnetic resonance (mpMRI) examination.[28]

Use of mpMRI has increased substantially in the last decade, fulfilling the unmet need of a non-invasive accurate screening tool for prostate cancer in men with elevated PSA.[29] MpMRI-visualized lesions can be targeted using mpMRI-guided biopsy, decreasing the risk of sampling error compared with traditional transrectal ultrasonography guided systematic biopsy. Studies on this subject, including the PROMIS, PRECISION, and 4M study, have shown that use of prebiopsy MRI and subsequent target biopsy leads to higher detection rates of clinically significant prostate cancer, and lower detection rates of clinically insignificant prostate cancer.[30–32] The current EAU guidelines therefore recommend that mpMRI should be performed before prostate biopsy.[17] In case of a positive mpMRI (PI-RADS ≥3), MRI target and systematic biopsies should be combined. In case of a negative mpMRI, combined with a low clinical suspicion for prostate cancer, biopsy may be omitted based on shared decision-making with the patient.[17]

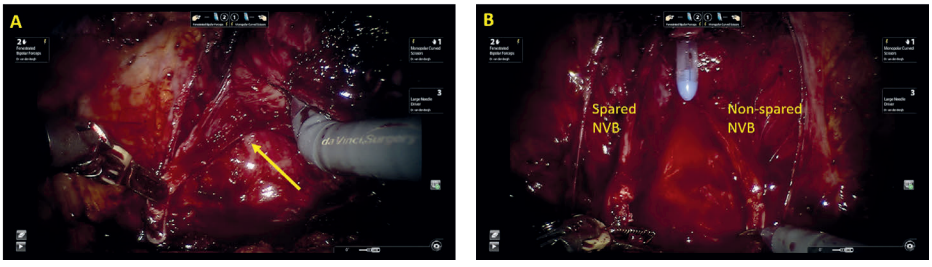
Local tumour staging

Besides prostate cancer detection, MRI can be used for local staging of the tumour. Information regarding the localization, diameter and presence of extraprostatic extension (EPE) is crucial for determining the optimal treatment strategy. For instance,

MRI can help localize the side of present EPE so that only the ipsilateral neurovascular bundle has to be resected in order to achieve negative surgical margins following robot-assisted radical prostatectomy (RARP).[33] Although a recent meta-analysis of prior studies on this subject did not show nerve sparing to be associated with increased risk of positive surgical margins, it is generally assumed that preservation of the neurovascular bundle increases the risk of positive margins.[34] This assumption is best supported by reviewing the anatomy during the surgical procedure (Figure 1). As shown in Figure 1, preservation of the neurovascular bundle requires close dissection to the prostate. Therefore, when nerve sparing is preferred by the patient, accurate risk estimation of EPE should be performed to ensure the safety of the procedure. Since there are two neurovascular bundles crossing both posterolateral sides of the prostate, EPE risk assessment should be performed in a side-specific manner.

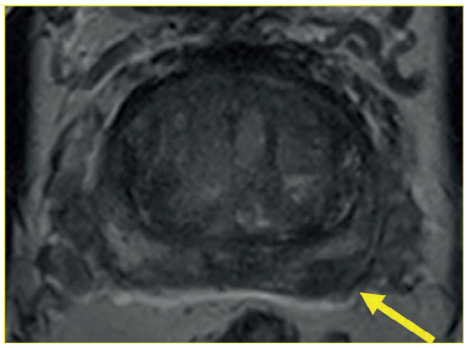
Although mpMRI has the potential to overcome the limitations of DRE with regard to local tumour staging, the technique still has its limitations. This is reflected by a recent meta-analysis, showing that the pooled sensitivity of mpMRI for the detection of EPE is 57%.[35] The low sensitivity is mainly attributable to the very limited ability of mpMRI to detect microscopic EPE. An example derived from routine clinical care is presented in Figure 2A & 2B. In Figure 2A, the index tumour lesion is visualized via mpMRI, interpreted by the radiologist as "organ-confined". However, on histopathological evaluation, microscopic EPE was present (Figure 2B).

FIGURE 1. Nerve sparing robot-assisted radical prostatectomy

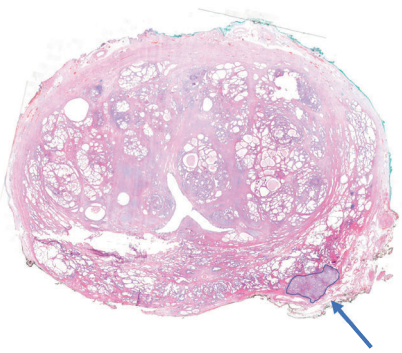


A: Procedure of preserving the left neurovascular bundle (NVB) (arrow) during prostatectomy.
B: After removal of the prostate; resulting in preserving neurovascular bundle tissue or greater thickness (spared NVB), compared to the residual neurovascular tissue after a wide excision (non-spared NVB).

FIGURE 2. Extraprostatic extension on magnetic resonance imaging and histopathological whole mount section



A. Magnetic resonance imaging (T2), tumour stage classified as "organ confined".



B. Whole mount section after radical prostatectomy revealing extraprostatic extension (arrow).

Value of MRI in risk classifications and nomograms

The current EAU guidelines recommend the use of a risk classification system, based on the D'Amico classification, using DRE to determine cT.[17] The outcomes of DRE are crucial, and subtle differences in DRE findings determine the risk group a patient is assigned to (e.g. cT2a [low-risk] vs. cT2b [intermediate-risk]). Because MRI enables visualization of the prostate gland in total, whereas with DRE the prostate can only be evaluated dorsally, definition of the cT may be more accurate when using MRI information. For example, risk classification by means of cT assessed by DRE is associated with the risk of understaging and subsequent undertreatment of prostate cancer.[36] Although MRI may eventually improve local tumour staging, there is a lack of suitable studies providing information on the impact of using the cT assessed by prostate MRI. Therefore, the current standpoint includes that although MRI provides significant additional information, it cannot replace DRE as the clinical staging standard.[37]

The most established clinical nomograms used for the prediction of unfavourable histopathological outcomes after surgery, including prediction of EPE and pelvic lymph node involvement (LNI), also include cT assessed by DRE.[38] MpMRI information could also improve clinical risk prediction models; as mpMRI local staging information is potentially more robust compared to local staging information assessed by DRE. Thus, the combined value of mpMRI information and the traditional clinical parameters in nomograms may enhance risk prediction of unfavourable histopathological outcomes. Preliminary findings on this topic are promising, as two separate study groups observed significant increases in model discrimination in terms of area under the curve after incorporation of MRI information, for commonly used tools including the MSKCC pre-radical prostatectomy nomogram.[39,40] The potential benefit of using mpMRI clinical staging information,

compared with DRE, on existing and novel prediction models includes an important research area that deserves further exploration.

ACTIVE SURVEILLANCE

Contemporary state of knowledge on active surveillance

In the mid-twentieth century, prostate cancer was mostly detected when it had reached an advanced and incurable stage.[3,41] Nowadays, the most common presentation includes screen-detected asymptomatic localised prostate cancer.[42] Although detection of the disease in an earlier stage by screening leads to an overall higher likelihood of curation, it also increases diagnosis of indolent cancers which may not have caused any harm.[43] To prevent unnecessary treatment-related morbidity in newly diagnosed patients, expectant management strategies can be initiated. Expectant management can grossly be divided into two strategies: watchful waiting (WW) or active surveillance (AS). WW refers to conservative management for patients deemed unsuitable for treatment with a curative intent because of limited life expectancy. Within this context, patients are "watched" for the development of local or systemic progression. Palliative treatment can eventually be initiated according to their symptoms. AS aims to avoid unnecessary treatment in men who have prostate cancer who do not need immediate treatment, but at the same time aims to achieve the correct timing for curative treatment in those who eventually do.[17]

Since the introduction of AS, significant progression has been made in reducing unnecessary treatment in patients who harbor prostate cancer showing to follow an indolent course. Based on the current ongoing AS registries, using AS can lead up to 58% of patients that remain untreated at 15 years of follow-up.[44,45] Several prospective cohort studies have been initiated showing favourable results.[46]

However, AS outcomes have been described for a specific selected subgroup of patients with very low-risk disease characteristics. Although there is some variation regarding eligibility criteria used in AS studies, the selection criteria generally include: clinical stage T1c, Gleason sum score ≤6 (Gleason grade group 1), PSA <10 ng/ml, maximum of two biopsy cores with cancer and ≤50% involvement in any core.[47] Strictly adhering to these selection criteria leads to, potentially unjustified, exclusion of a substantial number of low-risk patients from AS. For instance, patients are advised against AS if in more than two biopsy cores prostate cancer is found, even if this includes ISUP 1 grade disease; the least aggressive form of prostate cancer of which it is known to only rarely metastasize.[48] It is attainable that AS can also be a safe strategy in patients with higher-risk tumour characteristics, not fulfilling all very low-risk criteria. Given the potential benefit of AS in terms in quality of life outcomes, applicability of AS in patients that harbor higher-risk tumour characteristics should be explored. Unfortunately, due to the strict inclusion criteria used in the ongoing AS studies, the number of patients with higher risk characteristics on AS in a trial setting is limited. Only few studies have described

the outcomes of AS in patients not fulfilling one or more of the traditional AS inclusion criteria.[44,49] More data are urgently needed to assess whether the inclusion criteria can be expanded.

To enable timely detection of disease progression, intensive follow-up schemes are used to ensure the safety of AS. Although different AS follow-up protocols have been described, follow-up is generally characterized by periodical repeat prostate biopsies (every 1-3 years) and PSA monitoring (every 3-6 months) to assess if the disease progresses to a stage for which active treatment is indicated.[46] Strict adherence to AS follow-up protocols requires commitment of the urologist and the patient, especially since prostate biopsies are associated with pain, hematuria and an increased risk of urinary tract infections.[50] When applied in clinical practice, AS follow-up may not be compliant with current AS protocols. For instance, Loeb and colleagues evaluated the frequency of PSA testing and repeat biopsies in 5,192 patients on AS, and concluded that although over 80% of patients had 1 or more prostate specific antigen tests per year, fewer than 13% underwent prostate biopsy beyond the first 2 years follow-up.[51] Question remains whether this lack of compliance results in unfavourable outcomes for the patient, or if lower-intensive follow-up is safe in specific patient subgroups.

To conclude, although prior studies have demonstrated that AS is an effective and safe strategy to reduce overtreatment of prostate cancer, it remains to be determined if AS is also a safe when applied outside the strict trial setting. Areas that remain to be explored include the compliance with AS protocols regarding patient selection and follow-up intensity in real-world practice. In addition, it should be further evaluated in which patients, and to what extent, follow-up intensity and invasiveness can be reduced to further minimize treatment-related burden. Lastly, it should be established if selection for AS can be widened, by evaluating the outcomes of AS in patients not fulfilling all standard eligibility criteria.

SANTEON

The Santeon Prostate Cancer Value Based Healthcare programme

The Santeon consortium consists of seven hospitals including OLVG Amsterdam, Maasstad Hospital Rotterdam, St. Antonius Hospital Utrecht/Nieuwegein, Canisius Wilhelmina Hospital Nijmegen, Catharina Hospital Eindhoven, Martini Hospital Groningen and the Medisch Spectrum Twente Enschede (in close collaboration with the Hospital Group Twente Almelo/Hengelo).

Within the context of the Santeon Prostate Cancer Value Based Healthcare (VBHC) project, the hospitals have worked together intensively for over 10 years, comparing clinical and patient-reported outcomes. Since the initiation of the VBHC project, a data-infrastructure has been organized. The designed data-infrastructure facilitates short-cyclic feedback loops to improve clinical outcomes. In addition, the collected data

enabled the initiation of several scientific projects and quality improvement initiatives, including "Care for Outcome".[52] Collaborations of the Santeon hospitals are growing and are extending beyond the Dutch border. For instance, Santeon contributes to the True NTH Prostate Cancer global registry, an initiative funded by the Movember foundation.[53] The present thesis, describing the outcomes of Santeon prostate cancer research, further builds on the prior work of the Santeon VBHC prostate cancer project.

THESIS OBJECTIVES AND STRUCTURE

Thesis objectives

The objectives of this thesis are to evaluate the outcomes of a large real-world cohort of prostate cancer patients with regard to: (1) the outcomes of active surveillance, focusing on patient selection criteria and follow-up intensity, (2) the impact of MRI on staging as well as consequences of staging such as extent of surgery, (3) implementation of MRI information into available and novel prediction models.

This thesis is structured into five parts, including 10 chapters.

Chapter 1 includes a general introduction and summary of the current state of knowledge.

Part I

Optimising Active Surveillance

In **Part I** of this thesis, we report the outcomes of AS in a real-world multi-centre cohort. In **Chapter 2**, we will describe patient selection patterns for AS in daily practice. We test the hypothesis that selection for AS in real-world practice may be less stringent compared with AS performed in a strict trial setting, by determining the percentage of patients not meeting one or more AS eligibility criteria used in the largest ongoing AS study (PRIAS). In addition, we will assess the outcomes of patients meeting all criteria (PRIAS-eligible), and those not meeting one or more criteria (PRIAS-ineligible); assuming that PRIAS-ineligible patients have an increased risk of disease progression as well as development of metastasized disease.

The second chapter of **Part I (Chapter 3)** regards an analysis of the same cohort, focusing on the compliance with the follow-up protocol of the PRIAS study. Besides less stringent patient selection for AS, we assumed that follow-up intensity may also be less strict in daily clinical practice, compared to the trial setting. We will determine the compliance with the PRIAS follow-up protocol, with regard to repeat biopsy testing and PSA monitoring. In addition, we will test the hypothesis that non-compliant monitoring increases the risk of unfavourable outcomes, by comparing rates of metastasis of patients with PRIAS non-compliant and PRIAS compliant monitoring.

Part II

Impact of MRI on Prostate Cancer Risk Classification

Performing MRI to assist prostate cancer detection and staging is becoming standard of care. However, to what extent MRI local staging information can be incorporated into existing risk classification systems remains unclear. Due to the potentially improved local staging of the tumour by MRI, compared with DRE, cT of the tumour could be established more accurately. It remains unclear how this impacts the currently used D'Amico risk classification system. To evaluate whether MRI can replace DRE for determining cT, the accuracy of MRI as a local staging tool, as well as its impact on prostate cancer risk classification in terms of stage migration, needs further exploration. We assumed that use of MRI would lead to upstaging when used for determining cT, due to higher detections rates of EPE and seminal vesicle invasion, compared with DRE. To determine the correctness of the established cT by MRI and DRE, we will compare cT assessed by both modalities with definite surgical pathology. Comparison of the cT assessed by both strategies will provide insight into which entity should be used to determine the cT and prostate cancer risk classification **(Chapter 4)**.

Part III

Association Between Nerve Sparing and Positive Surgical Margins

Preservation of the neurovascular bundles during radical prostatectomy can be performed to optimise the probability of retaining postoperative erectile function and urinary continence. It is generally assumed that nerve sparing increases the risk of positive surgical margins. However, the majority of prior studies have resulted in contra-intuitive results, showing that nerve sparing is not associated with an increased risk of positive margins compared to non-nerve sparing surgery. These findings could be consequential to methodological limitations of prior studies, not exempt of selection bias. We assumed that by performing a study including an analysis with a side-specific approach, adjusting for a large number of covariates (including MRI staging information), would lead to more reliable results. Our findings regarding the association between side-specific nerve sparing and ipsilateral positive surgical margins are presented in **Chapter 5**.

Part IV

Incorporation of MRI into Clinical Prediction Models

Accurate determination of the local extent of the prostate cancer tumour is regarded to be crucial for determining the eligibility for preservation of the neurovascular bundles during radical prostatectomy. However, available tools that can be used for side-specific prediction of EPE are limited. The majority of these tools also do not include MRI local staging information. As MRI may improve local staging, we hypothesized that a nomogram including side-specific MRI staging information can lead to accurate

prediction of side-specific EPE. At the time of our study, a nomogram including side-specific MRI information was developed by Martini and colleagues, showing good model performance when applied in the development cohort.[54] To determine if this nomogram can be used safely in external populations, we will perform an external validation study **(Chapter 6)**.

To further investigate our assumption that inclusion of MRI may improve the accuracy of nomograms for the prediction of EPE, we aimed to develop and externally validate novel nomograms for the prediction of side-specific EPE. The nomograms will include MRI information, biochemical parameters and histopathological biopsy characteristics. The performance of four nomograms, each including different combinations of parameters, will be established and compared **(Chapter 7)**.

In **Chapter 8**, we will evaluate if cT assessed by MRI can replace cT assessed by DRE as an impute parameter of two established nomograms used for LNI risk prediction (MSKCC 2018 preoperative nomogram and the Briganti 2012 nomogram). We assumed that cT assessed by MRI could potentially be more accurate compared with DRE. However, question remains if this would also lead to improved LNI risk prediction; since both nomograms are originally developed using cT assessed by DRE. Therefore, we will assess how the incorporation of cT assessed by MRI impacts the accuracy of nomogram-based risk prediction of prostate cancer LNI.

Part V

General Summary and Discussion

Part V includes the General Summary **(Chapter 9)**, the Discussion of the main findings presented in chapter 2-8 and Future Perspectives **(Chapter 10)**.

REFERENCES

1. IKNL. Incidentie prostaatkanker. https://www.iknl.nl/kankersoorten/prostaatkanker/registratie/ incidentie. Published 2019. Accessed December 21, 2020.
2. RIVM. Demografische prognose ziekten 2015-2040. https://www.volksgezondheidenzorg.info/ onderwerp/demografische-prognose-ziekten-en-aandoeningen/toelichting-demografische-prognose. Published 2017. Accessed December 21, 2020.
3. Mellinger GT. Carcinoma of the Prostate. *Lancet*. 1963;83:94-105.
4. Hermanek P, Sobin L. International Union Against Cancer. TNM Classification of Malignant Tumours, ed 4. 1987.
5. Buyyounouski MK, Choyke PL, McKenney JK, et al. Prostate cancer - major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017;67:245-253.
6. Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother reports*. 1966;50:125-128.
7. Gleason D, Mellinger G, The Veterans Administratoin Cooperative Urological Research Group. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J Urol*. 1974;111:58-64.
8. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol*. 2016;40:244-252.
9. Gutman AB, Gutman EB. An "acid" phosphatase occurring in the serum of patients with metastasizing carcinoma of the prostate gland. *J Clin Investig*. 1938;17:473-478.
10. Wang MC, Valenzuela LA, Murphy GP, Chu TM. Purification of a human prostate specific antigen. *Invest Urol*. 1979;17:159-163.
11. Delaere KPJ, Van Dieijen-Visser MP, Gijzen AHJ, Brombacher PJ. Is prostate-specific antigen the most useful marker for screening in prostate cancer? *Am J Clin Oncol*. 1988; Suppl 2:S65-7.
12. Powell CS, Fielding AM, Rosser K, Ames AC, Vaughton KC. Prostate Specific Antigen—a Screening Test for Prostatic Cancer? *Br J Urol*. 1989;64:504-506.
13. Stamey TA, Yang N, Hay AR, Mcneal J, Freiha FS, Redwine E. Prostate-Specific Antigen as a Serum Marker for Adenocarcinoma of the Prostate. *N Engl J Med*. 1987;317:909-916.
14. Potosky AI, Miller BA, Albertsen PC, Kramer BS. The role of increasing detection in the rising incidence of prostate cancer. *JAMA J Am Med Assoc*. 1995;273:548-552.
15. Schroder FH, Hugosson J, Roobol MJ, et al. The European Randomized Study of Screening for Prostate Cancer – Prostate Cancer Mortality at 13 Years of Follow-up. *Lancet*. 2014;384:2027-2035.
16. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: Estimates from the European randomized study of screening for prostate cancer. *J Natl Cancer Inst*. 2003;95:868-878.
17. Mottet N, Bellmunt J, Briers E, et al. EAU Guidelines. Edn. Presented at the EAU Annual Congress Amsterdam; 2020.
18. Freedland SJ, Hotaling JM, Fitzsimons NJ, et al. PSA in the New Millennium: A Powerful Predictor of Prostate Cancer Prognosis and Radical Prostatectomy Outcomes - Results from the SEARCH Database. *Eur Urol*. 2008;53:758-766.
19. D'Amico A V., Whittington R, Bruce Malkowicz S, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *J Am Med Assoc*. 1998;280:969-974.
20. Vickers AJ, Fearn P, Scardino PT, Kattan MW. Why Can't Nomograms Be More Like Netflix? *Urology*. 2010;75:511-513.

21. Kattan MW, Eastham JA, Stapleton AMF, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst*. 1998;90:766-771.
22. Cagiannos I, Karakiewicz P, Eastham JA, et al. A preoperative nomogram identifying decreased risk of positive pelvic lymph nodes in patients with prostate cancer. *J Urol*. 2003;170:1798-1803.
23. Briganti A, Larcher A, Abdollah F, et al. Updated nomogram predicting lymph node invasion in patients with prostate cancer undergoing extended pelvic lymph node dissection: The essential importance of percentage of positive cores. *Eur Urol*. 2012;61:480-487.
24. Memorial Sloan Kettering Cancer Center. Pre-radical prostatectomy tool to predict probability of lymph node involvement in prostate cancer patients. www.mskcc.org/nomograms/prostate/pre_op. Published 2018. Accessed December 21, 2020.
25. Shariat SF, Karakiewicz PI, Suardi N, Kattan MW. Comparison of nomograms with other methods for predicting outcomes in prostate cancer: A critical analysis of the literature. *Clin Cancer Res*. 2008;14:4400-4407.
26. Capitanio U, Briganti A, Gallina A, et al. Predictive models before and after radical prostatectomy. *Prostate*. 2010;70:1371-1378.
27. Hricak H, Williams RD, Spring DB, et al. Anatomy and pathology of the male pelvis by magnetic resonance imaging. *Am J Roentgenol*. 1983;141:1101-1110.
28. Hoeks CMA, Barentsz JO, Hambrock T, et al. Prostate cancer: Multiparametric MR imaging for detection, localization, and staging. *Radiology*. 2011;261:46-66.
29. Oberlin DT, Casalino DD, Miller FH, Meeks JJ. Dramatic increase in the utilization of multiparametric magnetic resonance imaging for detection and management of prostate cancer. *Abdom Radiol*. 2017;42:1255-1258.
30. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet*. 2017;389:815-822.
31. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *N Engl J Med*. 2018:1767-1777.
32. van der Leest M, Cornel E, Israël B, et al. Head-to-head Comparison of Transrectal Ultrasound-guided Prostate Biopsy Versus Multiparametric Prostate Resonance Imaging with Subsequent Magnetic Resonance-guided Biopsy in Biopsy-naïve Men with Elevated Prostate-specific Antigen: A Large Prospective Multicenter Clinical Study. *Eur Urol*. 2019;75:570-578.
33. McClure TD, Margolis DJA, Reiter RE, et al. Use of MR imaging to determine preservation of the neurovascular bundles at robotic-assisted laparoscopic prostatectomy. *Radiology*. 2012;262:874-883.
34. Nguyen LN, Head L, Witiuk K, et al. The Risks and Benefits of Cavernous Neurovascular Bundle Sparing during Radical Prostatectomy: A Systematic Review and Meta-Analysis. *J Urol*. 2017;198:760-769.
35. de Rooij M, Hamoen EHJ, Witjes JA, Barentsz JO, Rovers MM. Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis. *Eur Urol*. 2016;70:233-245.
36. Grossfeld GD, Chang JJ, Broering JM, et al. Under staging and under grading in a contemporary series of patients undergoing radical prostatectomy: results from the cancer of the prostate strategic urologic research endeavor database. *J Urol*. 2001;165:851-856.
37. Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW, Amin MB. Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers. *Eur Urol*. 2018;73:560-569.
38. Campbell JM, Raymond E, O'Callaghan ME, et al. Optimum Tools for Predicting Clinical Outcomes in Prostate Cancer Patients Undergoing Radical Prostatectomy: A Systematic Review of Prognostic Accuracy and Validity. *Clin Genitourin Cancer*. 2017;15:e827-e834.

39. Feng TS, Sharif-Afshar AR, Wu J, et al. Multiparametric MRI Improves Accuracy of Clinical Nomograms for Predicting Extracapsular Extension of Prostate Cancer. *Urology*. 2015;86:332-337.
40. Rayn KN, Bioom JB, Gold SA, et al. Added Value of Multiparametric Magnetic Resonance lmaging to Clinical Nomograms tor Predicting Adverse Pathology in Prostate Cancer sr. *J Urol*. 2019;200:1041-1047.
41. McClelland JC. The early diagnosis of cancer in the bladder, prostate and kidney. *Can Med Assoc J*. 1934;31:165-167.
42. Butler SS, Muralidhar V, Zhao SG, et al. Prostate cancer incidence across stage, NCCN risk groups, and age before and after USPSTF Grade D recommendations against prostate-specific antigen screening in 2012. *Cancer*. 2020;126:717-724.
43. Loeb S, Bjurlin MA, Nicholson J, et al. Overdiagnosis and overtreatment of prostate cancer. *Eur Urol*. 2014;65:1046-1055.
44. Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. *J Clin Oncol*. 2015;33:272-277.
45. Carlsson S, Benfante N, Alvim R, et al. Long-Term Outcomes of Active Surveillance for Prostate Cancer: The Memorial Sloan Kettering Cancer Center Experience. *J Urol*. 2020;203:1122-1127.
46. Jeffrey J. Tosoian, H. Ballentine Carter, Abbey Lepor and SL. Active Surveillance for Prostate Cancer: Contemporary State of Practice. *Nat Rev Urol*. 2016;116:1477-1490.
47. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. *J Am Med Assoc*. 1994;271:368-374.
48. Donin NM, Laze J, Zhou M et al. Gleason 6 prostate tumors diagnosed in the PSA era do not demonstrate the capacity for metastatic spread at the time of radical prostatectomy. *Urology*. 2013;82:148-153.
49. Godtman RA, Holmberg E, Khatami A, Pihl CG, Stranne J, Hugosson J. Long-term Results of Active Surveillance in the Göteborg Randomized, Population-based Prostate Cancer Screening Trial. *Eur Urol*. 2016;70:760-766.
50. Loeb S, Vellekoop A, Ahmed HU, et al. Systematic review of complications of prostate biopsy. *Eur Urol*. 2013;64:876-892.
51. Loeb S, Walter D, Curnyn C, Gold HT, Lepor H, Makarov D V. How Active is Active Surveillance? Intensity of Followup during Active Surveillance for Prostate Cancer in the United States. *J Urol*. 2016;196:721-726.
52. Santeon. Zorg Voor Uitkomst. Uitkomstindicatoren Voor Kanker. Resultaten Santeon Ziekenhuizen. 2014.
53. Evans SM, Millar JL, Moore CM, et al. Cohort profile: The TrueNTH Global Registry - An international registry to monitor and improve localised prostate cancer health outcomes. *BMJ Open*. 2017;7(11):e017006.
54. Martini A, Gupta A, Lewis SC, et al. Development and internal validation of a side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer. *BJU Int*. 2018;122:1025-1033.

PART I: OPTIMISING ACTIVE SURVEILLANCE

CHAPTER 2

2

Active Surveillance for Prostate Cancer in a Real-Life Cohort: Comparing Outcomes PRIAS-eligible and PRIAS-ineligible Patients

T.F.W. Soeterik
H.H.E. van Melick
L.M. Dijksman
D.H. Biesma
J.A. Witjes
J.P.A. van Basten

EUROPEAN UROLOGY ONCOLOGY. 2018;1:231-237

ABSTRACT

Background: In daily practice, a wider range of patients with prostate cancer (PCa) are selected for active surveillance (AS) compared to those in AS trials, including higher-risk patients. However, less is known about the outcomes for off-protocol selected PCa patients who opt for AS. The aim of this study is to compare AS outcomes of higher-risk patients and very low-risk patients in a large cohort of patients diagnosed with PCa.

Methods: Patients diagnosed with PCa between 2008 and 2015, with clinical stage ≥T1c and treated with AS at six large teaching hospitals were included for analysis. AS constituted of regular prostate-specific antigen (PSA) testing (every 3-6 mo), combined with a confirmatory biopsy 1 yr after diagnosis and every 3 yr thereafter. Using the inclusion criteria of the Prostate Cancer Research International Active Surveillance (PRIAS) study, outcomes of PRIAS-eligible patients (clinical stage T1c-T2, Gleason sum score ≤6, ≤2 positive biopsy cores, PSA ≤10 ng/mL and PSA density <0.2 ng/ml/ml) were compared with outcomes of PRIAS-ineligible patients. Rates of unfavourable outcomes following deferred surgery, biochemical recurrence, and metastasis were established using univariate and multivariate Cox regression analysis.

Results: Of the 1000 patients included and treated with AS, almost half of the patients (49%) had higher-risk disease characteristics than the PRIAS inclusion criteria. PRIAS-ineligible patients discontinued AS, due to tumour progression, significantly earlier than PRIAS-eligible patients (HR 1.74, 95% confidence interval [CI] 1.44 - 2.11). PRIAS-ineligible patients also had a higher risk of positive surgical margins (odds ratio [OR] 2.15, 95% CI 1.11 - 4.17) and unfavourable pathological findings (OR 3.20, 95% CI 1.61 - 6.35) after deferred radical prostatectomy. PSA density ≥0.2 ng/ml/ml was the most important individual predictor and, in addition to a higher risk of tumour progression and unfavourable surgical outcomes, was associated with a significant higher risk of biochemical progression following deferred radical prostatectomy (OR 3.26, 95% CI 1.23 - 8.64). In the overall population, PSA density ≥0.2 ng/ml/ml was associated with a higher risk of metastasis (HR 2.71, 95% CI 1.23 – 5.96).

Conclusions: In this cohort, approximately half of the patients did not meet the inclusion criteria of the PRIAS study. These patients had a twofold higher risk of opting out of AS due to tumour progression and a threefold higher risk of unfavourable outcomes following deferred prostatectomy. PSA density is an important individual predictor of unfavourable outcomes and should be taken into account when selecting patients for AS.

BACKGROUND

To be selected for active surveillance (AS) in the largest prospective cohort studies, patients with prostate cancer (PCa) need to meet several inclusion criteria. However, men who harbour higher-risk disease and do not meet all these selection criteria, may also opt for AS. Selecting higher-risk patients for AS may be based on various reasons including, for example, a patient's strong motivation to preserve urinary continence and erectile function.[1] The AS course and outcomes in (very) low-risk patients included in the ongoing AS studies are well described and understood. By contrast, less is known about the outcomes for patients with PCa who would not qualify for inclusion in one of the AS studies because of a higher-risk profile. To adequately inform higher-risk patients opting for AS about their prognosis, more information on intermediate- and long-term outcomes of "off-protocol" selected patients is required.

In this study we evaluated the outcomes for a large cohort of men, including patients with higher-risk PCa, managed with AS. The inclusion criteria of the Prostate Cancer Research International Active Surveillance (PRIAS) study were used as a reference, since this protocol is acknowledged as standard practice in the Netherlands.[2]

METHODS

Study setting and data collection

This study was conducted within the Santeon consortium, which comprises seven large nonacademic teaching hospitals. Patients were selected from the Santeon prostate cancer database, which consists of data retrospectively collected as a part of quality improvement initiatives. Data from six out of seven hospitals were available during the study period, since the seventh hospital recently joined the collaboration and data from this centre were not yet available.

Included in the present study were patients diagnosed with PCa between January 1, 2008 and December 31, 2014 and treated with AS. Additional data on biochemical follow-up after radical treatment, histological and surgical outcomes of deferred radical prostatectomy, and metastasis rates were collected by abstractors with an academic background who were trained by the principal investigator (T.F.W.S.). Study data were collected and managed using the REDCap electronic data capture tool.[3] After data insertion by the data abstractors, all cases were verified by the principal investigator (T.F.W.S.).

Patient population and outcome measures

Patients were included if they were on AS, clearly distinguishing these patients from those who were managed conservatively without curative intent ("watchful waiting").[4] Excluded from the study were patients aged ≥80 yr at diagnosis, patients with incidental

tumours (T1a/b), patients with treatment strategies other than AS, and those initially treated at a hospital outside the Santeon group.

Active surveillance follow-up

AS was performed according to the PRIAS follow-up regimen. The protocol includes prostate-specific antigen (PSA) testing and digital rectal examination every 3-6 months, and a confirmatory biopsy one year after diagnosis and every three years thereafter.[5] AS could be discontinued due to disease progression (e.g. Gleason upgrading), or for non-oncological reasons (e.g. patient's preference, or onset of other diseases with a worse prognosis).

Outcome measures

The primary outcome measure was the proportion of patients who did not meet one or more of the PRIAS inclusion criteria: clinical stage T1c–T2, Gleason sum score >6, two or fewer positive biopsy cores, PSA <10 ng/ml, and PSA density (PSAD) <0.2 ng/ml/ml.

Secondary outcome measures evaluated in the overall population included the percentage of patients with metastasis (bone and/or local or distant lymph node) and rates of AS discontinuation because of tumour progression. In addition, outcomes for patients who underwent deferred radical prostatectomy were analysed separately. In this subgroup we evaluated the number of patients who underwent a non-nerve sparing procedure, had positive surgical margins and had unfavourable pathological findings (defined as a pathological T-stage ≥T3 and/or Gleason sum score ≥8). The incidence of biochemical recurrence was also assessed for these patients, defined as PSA ≥0.2 ng/ml two consecutive times after radical prostatectomy.[6]

Statistical analysis

Univariate and multivariate Cox regression analysis with backward elimination were used for comparisons between PRIAS-eligible and PRIAS-ineligible subgroups. Analysis of time to tumour progression (progression-free survival) was performed with the Kaplan-Meier method. To replace missing values, the multiple imputations package for SPSS V.24.0 (IBM Corp, Armonk, NY, USA) was used.

RESULTS

Patient population

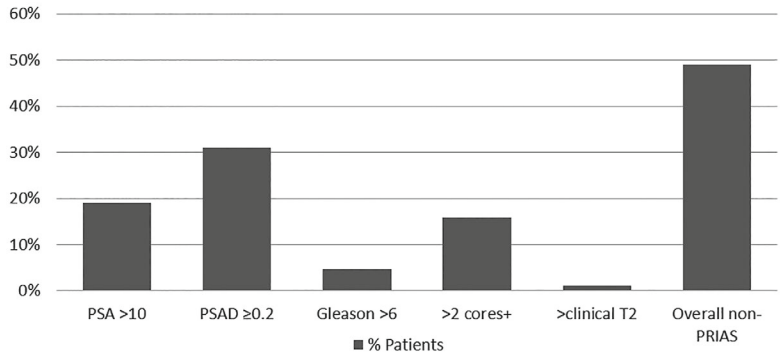
A total of 1181 patients were evaluated; this included 181 patients with incidental tumours who were subsequently excluded. The remaining 1000 patients with biopsy-detected PCa included 79 patients (8%) with one or more missing baseline values that were imputed. In this cohort, approximately half of the patients would not meet the PRIAS criteria: 490/1000 (49%). Among all the separate criteria, PSAD <0.2 ng/ml/ml was most often not met (31%; Figure 1). The baseline characteristics for PRIAS-eligible and PRIAS-ineligible patients are provided in Table 1.

TABLE 1. Baseline characteristics

	PRIAS-eligible	PRIAS-ineligible	*p*
No. of patients	510	490	
Age (years)	66.3 ± 0.3	68.2 ± 0.3	<0.001
PSA (ng/ml)	6.1 ± 1.9	10.2 ± 5.5	<0.001
PSA density[a] (ng/ml/ml)	0.12 ± 0.04	0.25 ± 0.14	<0.001
cT-stage[b] (%)			
T1c	407 (80)	381 (78)	
T2a	40 (8)	24 (5)	
T2b	4 (1)	10 (2)	
T2c	6 (1)	9 (2)	
T2	53 (10)	55 (11)	
T3	0 (0)	11 (2)	
Total no. biopsy cores	9.2 ± 1.9	8.9 ± 2.0	0.009
No. positive cores	1.3 ± 0.4	2.1 ± 1.3	<0.001
Gleason score			
2 + 2	2 (0)	3 (1)	
2 + 3 / 3 + 2	3 (1)	7 (1)	
3 + 3	505 (99)	434 (89)	
3 + 4	0	42 (9)	
4 + 3	0	4 (1)	

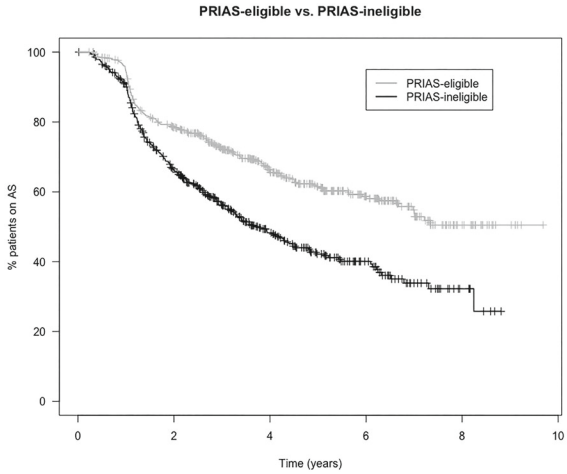
[a]PSA density: PSA divided by prostate volume (assessed with transrectal ultrasonography or MRI).
[b]Clinical stage based on digital rectal examination.
*All variables are in number (%) or mean ± SD.
*Percentages may not total 100% due to rounding.

FIGURE 1. Percentages of patients not meeting the overall and individual PRIAS selection criteria



PSA = prostate-specific antigen, PSAD = PSA density, PRIAS = Prostate Cancer Research International Active Surveillance.

FIGURE 2. Kaplan-Meier curves of time to tumour progression



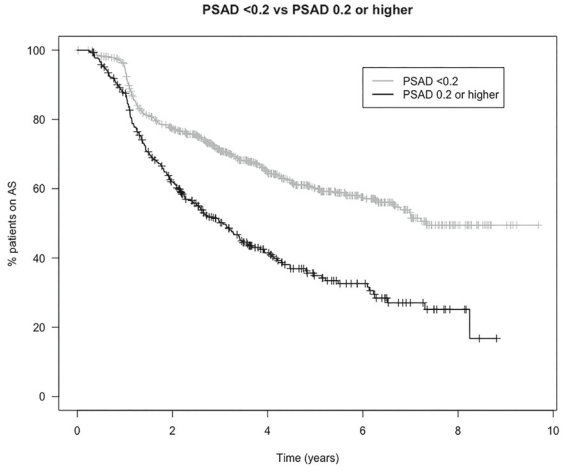


TABLE 2. Reasons for discontinuation of active surveillance in the overall population (*n* = 1000)

Reason	Category	Patients (%)
Tumour progression (*N* = 437)	PSA kinetics[a]	66 (11)
	DRE[b]	1 (0)
	Gleason upgrading	82 (14)
	Biopsy volume ↑[c]	65 (11)
	MRI findings[d]	19 (3)
	Other	10 (2)
	Combination of two or more	194 (33)
Non-oncological (*N* = 151)	Comorbidity / Age[e]	41 (7)
	Patient's preference	24 (4)
	Other[f]	30 (5)
	Lost to follow-up	56 (10)
	Total	588 (100)

[a]E.g.: PSADT <3 years and/or PSA >10.
[b]DRE = digital rectal examination.
[c]>50% tumour volume per core or ≥3 positive cores.
[d]E.g.: extracapsular extension, increased tumour volume.
[e]Comorbidity determining patient's prognosis or reaching an age whereby the clinical relevance of the tumour is not significant.
[f]Including urinary symptoms, or follow-up continued by general practitioner.

Time to tumour progression

In this cohort, 588 patients discontinued AS (59%), of whom 437 (74%) discontinued due to significant tumour progression. Reasons for discontinuation of AS are listed in Table 2. Figure 2 shows that PRIAS-ineligible patients experienced tumour progression significantly earlier than PRIAS-eligible patients. This increased risk was confirmed by univariate Cox regression analysis (hazard ratio [HR] 1.74, 95% confidence interval [CI] 1.44 - 2.11) (Table 3). In addition, PSAD ≥0.2 ng/ml/ml and positive biopsy cores >2 were both significant predictors of tumour progression. Of these two criteria, PSAD ≥0.2 ng/ml/ml was the strongest predictor in the multivariate Cox regression model (HR 2.01, 95 CI% 1.67 - 2.44) (Table 3).

TABLE 3. Correlation between clinical stage, PRIAS criteria and time to tumour progression

		Kaplan-Meier	Cox regression Univariate			Cox-regression Multivariate		
		Median AS time[a] (years)	Hazard ratio	95% CI	*p*	Hazard ratio	95% CI	*p*
Age	years		0.99	0.98–1.01	0.397	0.99	0.97–1.00	0.088
PSA	>10 ≤10	5.4 6.0	1.12	0.88–1.41	0.361	–	–	–
PSAD	≥0.2 <0.2	3.1 7.3	2.03	1.67–2.45	<0.001	2.01	1.67–2.44	<0.001
Gleason	≥7 ≤6	5.3 6.0	1.35	0.89–2.05	0.165	–	–	–
Positive cores	>2 ≤2	3.6 5.6	1.62	1.28–2.06	<0.001	1.57	1.24–2.00	<0.001
PRIAS[b]	No Yes	3.7 6.4[c]	1.74	1.44–2.11	<0.001	–	–	–
No. of criteria	≥2 ≤1	3.0 6.8	1.78	1.42–2.22	<0.001	–	–	–
T3 or higher	≥cT3 ≤cT2[d]	3.6 5.9	1.28	0.48–3.44	0.621	–	–	–

[a]Median follow-up time on active surveillance (time to event).
[b]PRIAS: complying with all PRIAS criteria (≤cT2, PSA ≤10, PSA density <0.2, Gleason score ≤6 and positive cores ≤2).
[c]Mean value; median value could not be calculated since no more than 50% of patients discontinued.

TABLE 4. Risk of unfavourable outcomes after delayed prostatectomy

	Non-nerve sparing[a] OR (95% CI)	*P*	Positive margin OR (95% CI)	*P*	Unfavourable histology[b] OR (95% CI)	*P*	Biochemical progression OR (95% CI)	*P*
Non-PRIAS vs. PRIAS	1.49 (0.73 – 3.03)	0.269	2.15 (1.11 – 4.17)	0.023	3.20 (1.61 – 6.35)	0.001	1.98 (0.75 – 5.23)	0.168
PSAD ≥0.2	1.23 (0.59 – 2.59)	0.586	1.97 (1.03 – 3.79)	0.042	3.61 (1.85 – 7.05)	<0.001	3.26 (1.23 – 8.64)	0.018
Positive cores >2	1.59 (0.61 – 4.17)	0.342	2.96 (1.26 – 6.94)	0.013	1.89 (0.83 – 4.34)	0.132	3.69 (1.32 – 10.35)	0.013

[a]Non-nerve sparing (including partial, unilateral and non-nerve sparing).
[b]Unfavourable pathology defined as presence of pT ≥T3 and/or pathological Gleason score ≥8.

Unfavourable outcomes after deferred radical prostatectomy

Of all 588 patients who discontinued AS, 420 (71%) underwent curative treatment. Of these, 182/420 (43%) underwent radical prostatectomy. PRIAS-ineligible patients had a significantly higher risk of unfavourable histological outcomes (pathological T-stage ≥T3 and/or Gleason sum >7) and positive surgical margins after deferred prostatectomy (Table 4). Patients with a PSAD ≥0.2 had a higher risk of unfavourable pathological findings (OR 3.61, 95% CI 1.85 - 7.05), positive surgical margins (OR 1.97, 95% CI 1.03 - 3.79) and biochemical progression (OR 3.26, 95% CI 1.23 - 8.64) (Table 4).

Association between off-protocol selection and risk of metastasis

Median follow-up (period between date of diagnosis and last hospital visit/contact) of the cohort was 5.0 (IQR 3.5 - 6.6) years. During this follow-up period, 25 of the 1000 patients (2.5%) developed metastatic PCa. In contrast to our earlier results, PRIAS-ineligible patients did not have a higher risk of metastasis compared with the PRIAS-eligible patients (HR 1.37, 95% CI 0.62 - 3.01).

However, the results for PSAD were in line with earlier outcomes, since patients with PSAD ≥0.2 also had a higher risk of metastasis compared with patients with a PSAD <0.2 ng/ml/ml (HR 2.71, 95% CI 1.23 - 5.96). Patients not complying with ≥2 PRIAS criteria also had a higher risk of metastasis (Table 5). In this cohort, at baseline, 57 patients had a PSAD ≥0.2 and >2 positive biopsy cores. In 7/57 (12%) of these patients, metastatic PCa was diagnosed.

TABLE 5. Metastasis rates per predictor

Predictor	HR	95% CI	*p*
PSA >10	1.33	0.53 – 3.33	0.546
PSAD ≥0.2	2.71	1.23 – 5.96	0.014
Gleason sum ≥7	2.25	0.53 – 9.56	0.273
Positive cores >2	2.17	0.90 – 5.19	0.083
Non-PRIAS	1.37	0.62 – 3.01	0.438
Two or more criteria[a]	3.04	1.37 – 6.78	0.006
PSAD ≥0.2 and positive cores >2	6.17	2.57 – 14.82	<0.001
PSAD ≥0.2 and PSA >10	2.11	0.84 – 5.28	0.112
PSA >10 and positive cores >2	3.48	1.04 – 11.64	0.043

[a]Not meeting two or more of the PRIAS selection criteria at initial diagnosis.
HR = hazard ratio.

DISCUSSION

In the present cohort of 1000 patients on AS, 49% did not meet one or more of the PRIAS criteria for (very) low-risk PCa. These PRIAS-ineligible patients experienced tumour progression significantly earlier compared with the PRIAS-eligible subgroup. In addition, PRIAS-ineligible patients had a higher risk of unfavourable histological and surgical outcomes following deferred radical prostatectomy. Baseline PSAD ≥0.2 ng/ml/ml was a strong individual predictor of (unfavourable) AS outcomes, and was associated with earlier tumour progression, a higher risk of unfavourable outcomes after deferred surgery, and a higher risk of metastasis.

A major issue related to AS is the extent to which AS eligibility criteria may be expanded.[7,8] Although there is limited evidence on the safety of AS for patients with higher-risk PCa, the present results indicate that the selection criteria are already being expanded in daily practice. These findings suggest that patients (and their treating clinicians) might be willing to trade off a survival benefit to avoid treatment-related comorbidities such as urinary incontinence and erectile dysfunction. It is also possible that urologists may not be fully aware of the potential consequences of selecting higher-risk patients. Either way, this study provides additional information that is highly relevant for clinicians counselling these patients.

Three previous studies evaluating AS outcomes for patients with lower- and higher-risk PCa reported a significantly higher rates of metastasis among intermediate-risk patients. For example, Bul et al reported a 10-year metastasis-free survival of 99.7% in low-risk patients and 96.4% in intermediate-risk patients (log-rank p = 0.03).[9] In the second study, Godtman et al reported that patients with low and intermediate-risk disease had a higher risk of AS failure (defined as death from PCa, PCa metastasis and/or biochemical progression and/or initiation of hormonal therapy after radical treatment) compared with very low-risk patients: HR 4.8 (95% CI 2.44 - 9.33).[10] A higher risk of metastasis in intermediate-risk patients was also observed in the Sunnybrook cohort. Musunuru et al reported an inferior 15-year metastasis-free survival in the intermediate-risk group compared with the low-risk group (HR 3.14, 95% CI 1.51 - 6.53).[11] Conflicting results were reported by Nyame et al, as they did not observe a significant difference in metastasis-free survival comparing men with intermediate/high-risk PCa and very low/low-risk (HR 1.50, 95% CI 0.16 - 14.5).[12]

The present study is unique in that we compared patients selected according to strict AS criteria based on the PRIAS protocol, with higher-risk off-protocol selected patients. In addition, we performed univariate analysis for all the separate criteria, which provided relevant information regarding the predictive value of each factor.

Comparison of PRIAS-eligible and PRIAS-ineligible patients showed no significant difference regarding rates of metastasis. However, when evaluating all the criteria separately, PSAD ≥0.2 ng/ml/ml was a strong individual predictor for the risk of metastasis and was also strongly associated with higher rates of other unfavourable outcomes. This suggests that in this patient cohort, using the PRIAS criteria classification overall was

less predictive compared to PSAD as an individual indicator. One out of three patients in the present cohort did not meet the PRIAS criterion PSAD <0.2 ng/ml/ml. This indicates that urologists may not regard this as an important AS eligibility criterion. Moreover, the vast majority of ongoing prospective AS studies (i.e. 7/9; 78%) did not include PSAD as a selection criterion.[13] However, given the strong association between PSAD and AS outcomes, the importance of PSAD as a predictor of AS outcomes might be currently underestimated. Furthermore, the higher number of patients not meeting the PSAD <0.2 ng/ml/ml criterion could be an alternative explanation for the strong association found for this clinical factor compared with the other PRIAS inclusion criteria. Because the numbers of patients were also lower in the other subgroups (e.g. patients with Gleason sum >6), this analysis might be susceptible to a higher probability of type II errors. In addition, owing to the limited statistical power, we cannot conclude that not meeting other criteria (including Gleason sum score ≤6 and clinical stage >T2), does not lead to a higher risk of unfavourable outcomes. We recommend further exploration of the potential correlations of the individual PRIAS criteria in AS cohorts including higher numbers of patients not fulfilling these criteria.

Nevertheless, we regard our findings regarding PSAD to be valid, since they are in line with results of previous studies. For instance, in a population-based study of predictors of adverse pathology after radical prostatectomy among AS candidates, a PSAD >0.15 ng/ml/ml was predictive for Gleason upgrading or pathological pT stage ≥T3 (OR 2.04, 95% CI 1.91 - 2.31).[14] In addition, another group found that PSAD >0.15 ng/ml/ml was strongly associated with adverse pathological findings on radical prostatectomy.[15] In four studies focusing on the association between baseline factors and disease progression in patients with PCa on AS, PSAD was also found to be a strong predictor for disease progression.[16–19] These previous results on the predictive value of PSAD for disease progression and prostatectomy outcomes in AS candidates, combined with our findings on the risk of metastasis, emphasise that PSAD should be taken into account when selecting patients with PCa for AS.

The question remains whether or not we can safely expand the eligibility criteria for AS based on earlier results and our findings and, if so, to what extent this might be safe. From the present data, we conclude that AS is less safe for patients with a PSAD ≥0.2 ng/ml/ml and >2 positive biopsy cores at baseline, since 12% of these patients eventually developed PCa metastases. In addition, since we found that 5% of patients with PSAD >0.2 ng/ml/ml developed metastasized PCa and thus exceeded the window of curability, we think that PSAD >0.2 ng/ml/ml should also be considered as a serious risk factor that should be taken into account during shared decision-making on treatment.

The present study has several strengths, most notably being a multicentre study with a large population and describing the real-life AS situation for patients with PCa. However, some limitations also need to be addressed. First, most of the data were collected retrospectively, which may have led to information bias. For instance, there is a possibility that because of our retrospective evaluation patients were labelled as AS patients while they actually were on watchful waiting. However, the percentages of

patients undergoing radical treatment following AS discontinuation were comparable between the PRIAS-eligible and PRIAS-ineligible subgroups (73% vs 73%). Some 11% of PRIAS-eligible patients who discontinued AS switched to watchful waiting, while this percentage was 14% in the PRIAS-ineligible group. Given the comparable percentages in both groups, we concluded that the presence of bias that would affect our study results was very unlikely. In addition, to guarantee data quality, all data inserted by abstractors were checked by the principal investigator. In case of uncertainty, the treating urologist involved in the clinical process was consulted. Furthermore, the number of missing values was within an acceptable range, since 8% of patients had one or more missing baseline values that were inserted using multiple imputation. Sensitivity analyses of the original dataset and imputed datasets showed no significant differences regarding outcomes and, therefore, did not alter our final conclusions. Second, the study has a risk of measurement bias, since not all patients underwent the same diagnostic testing during follow-up. Thus, metastasis could only be confirmed in patients undergoing, for example, pelvic lymph node dissection or imaging; therefore, the number of patients with metastasis may in fact be underestimated. Third, other factors that may influence AS outcomes, such as pursuing AS despite significant tumour progression and follow-up intensity, were beyond the scope of this study.

CONCLUSIONS

Weighing the advantages of potentially avoiding radical treatment against the risk of the tumour progressing to a stage with worse prognosis remains the most important dilemma when selecting (higher-risk) patients with PCa for AS. In the present cohort, half of all patients on AS did not meet the PRIAS criteria. These selected "off protocol" patients have earlier disease progression and a higher risk of unfavourable outcomes following deferred surgery. PSAD ≥0.2 ng/ml/ml was an important individual predictor of tumour progression and was associated with worse disease prognosis.

REFERENCES

1. Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: Progress and promise. *J Clin Oncol.* 2011;29:3669-3676.
2. Bul M, Zhu X, Valdagni R, et al. Active surveillance for low-risk prostate cancer worldwide: The PRIAS study. *Eur Urol.* 2013;63:597-603.
3. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) -A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42:377-381.
4. Matulewicz RS, Weiner AB, Schaeffer EM. Active surveillance for prostate cancer. *JAMA.* 2017;318:2152.
5. van den Bergh RCN, Roemeling S, Roobol MJ, Roobol W, Schröder FH, Bangma CH. Prospective Validation of Active Surveillance in Prostate Cancer: The PRIAS Study. *Eur Urol.* 2007;52:1560-1563.
6. Cookson MS, Aus G, Burnett AL, et al. Variation in the Definition of Biochemical Recurrence in Patients Treated for Localized Prostate Cancer: The American Urological Association Prostate Guidelines for Localized Prostate Cancer Update Panel Report and Recommendations for a Standard in the Reporting of surgical outcome. *J Urol.* 2007;177:540-545.
7. Dall'Era M, Klotz L. Active Surveillance for Intermediate Risk Prostate Cancer. *Curr Urol Rep.* 2017;18:1-6.
8. Perlis N, Klotz L. Contemporary Active Surveillance: Candidate Selection, Follow-up Tools, and Expected Outcomes. *Urol Clin North Am.* 2017;44:565-574.
9. Bul M, Van Den Bergh RCN, Zhu X, et al. Outcomes of initially expectantly managed patients with low or intermediate risk screen-detected localized prostate cancer. *BJU Int.* 2012;110:1672-1677.
10. Godtman RA, Holmberg E, Khatami A, Pihl CG, Stranne J, Hugosson J. Long-term Results of Active Surveillance in the Göteborg Randomized, Population-based Prostate Cancer Screening Trial. *Eur Urol.* 2016;70:760-766.
11. Musunuru HB, Yamamoto T, Klotz L, et al. Active Surveillance for Intermediate Risk Prostate Cancer: Survival Outcomes in the Sunnybrook Experience. *J Urol.* 2016;196:1651-1658.
12. Nyame YA, Almassi N, Haywood SC, et al. Intermediate-Term Outcomes for Men with Very Low/Low and Intermediate/High Risk Prostate Cancer Managed by Active Surveillance. *J Urol.* 2017;198:591-599.
13. Jeffrey J. Tosoian, H. Ballentine Carter, Abbey Lepor and SL. Active Surveillance for Prostate Cancer: Contemporary State of Practice. *Nat Rev Urol.* 2016;116:1477-1490.
14. Vellekoop A, Loeb S, Folkvaljon Y, Stattin P. Population based study of predictors of adverse pathology among candidates for active surveillance with Gleason 6 prostate cancer. *J Urol.* 2014;191:350-357.
15. Reese AC, Landis P, Han M, Epstein JI, Carter HB. Expanded criteria to identify men eligible for active surveillance of low risk prostate cancer at johns hopkins: A preliminary analysis. *J Urol.* 2013;190:2033-2038.
16. Tseng KS, Landis P, Epstein JI, Trock BJ, Carter HB. Risk Stratification of Men Choosing Surveillance for Low Risk Prostate Cancer. *J Urol.* 2010;183:1779-1785.
17. San Francisco IF, Werner L, Regan MM, Garnick MB, Bubley G, Dewolf WC. Risk stratification and validation of prostate specific antigen density as independent predictor of progression in men with low risk prostate cancer during active surveillance. *J Urol.* 2011;185:471-476.
18. Barayan GA, Brimo F, Bégin LR, et al. Factors influencing disease progression of prostate cancer under active surveillance: A McGill University Health Center cohort. *BJU Int.* 2014;114:E99-E104.
19. Mamawala MM, Rao K, Landis P, et al. Risk prediction tool for grade re-classification in men with favourable-risk prostate cancer on active surveillance. *BJU Int.* 2017;120:25-31.

CHAPTER 3

3

Follow-up in Active Surveillance for Prostate Cancer: Strict Protocol Adherence Remains Important for PRIAS-ineligible patients

T.F.W. Soeterik
H.H.E. van Melick
L.M. Dijksman
D.H. Biesma
J.A. Witjes
J.P.A. van Basten

EUROPEAN UROLOGY ONCOLOGY. 2019;2:483-489

ABSTRACT

Background: Active surveillance (AS) is a safe treatment strategy for men with very low-risk prostate cancer (PCa), when performed in a research setting using strict follow-up. However, less is known about the protocol adherence and outcomes of AS in real-world practice. In this study, we will evaluate the Prostate Cancer Research International Active Surveillance (PRIAS) protocol adherence in a real-world cohort and relate follow-up intensity to oncological safety.

Methods: Patients with biopsy-detected prostate cancer (PCa), diagnosed from 2008 until 2014, treated with active surveillance at six teaching hospitals in the Netherlands, were included. Proportions of patients complying with the PRIAS follow-up protocol (including PSA testing every 3-6 months combined with a confirmatory biopsy one year after diagnosis and every three years thereafter) were determined. We assessed if PRIAS-discordant follow-up was associated with a higher risk of metastasis compared with PRIAS-concordant follow-up using Cox regression analysis. Analysis was performed for separate risk groups (PRIAS-eligible and PRIAS-ineligible) on the basis of the PRIAS inclusion criteria.

Results: Of all patients on AS for >6 mo, 706/958 (74%) had PRIAS-concordant PSA monitoring. Overall concordant follow-up (PSA and repeat biopsy) was observed in 415/958 patients (43%). The percentage of patients with overall concordant follow-up varied between hospitals (range 34–60%; $p < 0.001$). Among PRIAS-ineligible patients, PRIAS-discordant PSA monitoring was associated with a higher risk of developing PCa metastases during AS compared with patients with concordant follow-up (hazard ratio 5.25, 95% confidence interval 1.02–27.1). In the PRIAS-eligible population, we found no significant differences regarding rates of metastases between patients with discordant and concordant follow-up.

Conclusions: We observed substantial variation in AS follow-up intensity between large urological practices in the Netherlands. Overall, 43% of patients on AS in daily clinical practice receive PRIAS-concordant follow-up. Noncompliance with the PRIAS follow-up protocol was associated with a higher rate of metastasis among PRIAS-ineligible patients, indicating that strict protocol adherence is important when these patients opt for AS.

BACKGROUND

The principle of active surveillance (AS) is to avoid overtreatment of clinically insignificant prostate cancer (PCa) and to defer treatment until objective evidence of disease progression.[1] Favourable outcomes of AS in a trial setting have led to its widespread acceptance.[2–4] However, it remains challenging to identify men with indolent disease among those with progressive PCa at risk of missing the window of curability.

The protocols of published AS studies adhere to the same principles: repeat prostate biopsies (intensity varying per protocol from yearly to every 3-4 yr) combined with regular prostate specific antigen (PSA) testing every 3–6 mo.[5] However, these monitoring protocols can be burdensome for patients, are time-consuming and costly. Moreover, repeat biopsies are unpleasant for the patient and bear a risk of bleeding and infection. It is therefore conceivable that AS protocols are not followed as strictly in daily practice as recommended by the prevailing guidelines. This hypothesis is supported by an inventory of real world practice patterns in the USA, which revealed that less than 13% of PCa patients undergo repeat biopsy beyond the first two years of AS.[6] Furthermore, a survey among European urologists indicated that 47% of those practicing AS do not use an official AS protocol nor is involved in a clinical AS trial.[7] Also, a nationwide survey in Japan showed that only 40.6% of the urologists performed a scheduled repeat biopsy at 1 yr after AS initiation.[8] The possible consequences of these AS protocol deviations with regard to oncological safety are for the most part unknown. This calls for research assessing the safety of lower-intensity AS monitoring.

In the present study, we evaluated AS follow-up strategies for PCa in six large Dutch teaching hospitals covering up to 15% of PCa patients in the Netherlands. We will determine the proportions of patients that underwent follow-up testing according to the Dutch guidelines, which are based on the follow-up protocol of the Prostate Cancer International Active Surveillance (PRIAS) study.[9] Furthermore, we assessed if patients with low-intensity monitoring had a higher risk of missing the window of curability because of the development of metastatic PCa during AS.

METHODS

Study setting and data collection

This study was conducted within the Santeon consortium, which consists of seven large nonacademic teaching hospitals in the Netherlands. During the study period, data for the AS cohorts from six of these seven hospitals were available. The study focuses on the same cohort of PCa patients on AS diagnosed between January 1, 2008 and December 31, 2014 on which we reported previously.[10] Data collection and analysis included initial age and tumour characteristics at diagnosis, dates of follow-up serum PSA tests, repeat biopsies, magnetic resonance imaging (MRI) of the prostate and metastasis rates.[11]

Variation in follow-up strategy and compliancy with PRIAS

The Dutch PCa guidelines recommend that AS follow-up should be in accordance with the study protocol of PRIAS.[12] This includes a PSA test every 3 months in the first 2yr and every 6 mo thereafter. Scheduled repeat biopsies should be performed at 1, 4, 7 and 10 yr following diagnosis. Definitions used for follow-up compliance with the PRIAS protocol were comparable to those published by the PRIAS study group.[13] PSA follow-up was regarded as concordant if a patient had undergone ≥75% of the recommended number of PSA tests for their follow-up duration. For example, a patient with an AS duration of 14 mo should have undergone three or more PSA tests to be regarded as compliant.

To assess compliance with repeat biopsy testing, we evaluated the percentage of patients who underwent the first (1 yr), second (4 yr), and third biopsy (7 yr) among men with follow-up of >1.5, >4.5, and >7.5 yr, respectively. We also determined the percentage of patients who received all scheduled biopsies according to the protocol, taking AS duration into account. Follow-up was scored as discordant if a patient should have undergone one or more biopsies according to the follow-up scheme, but missed one or more. A separate analysis was performed to determine in how many cases MRI of the prostate were performed instead of a prostate biopsy. Protocol adherent follow-up was assessed in patients with an AS duration of >6 mo. We assessed whether discordant follow-up was associated with an higher rate of metastasis during AS follow-up using risk classification based on the PRIAS inclusion criteria: PSA ≤10 ng/ml, PSA density (PSAD) <0.2 ng/ml/ml, Gleason ≤6, fewer than three positive biopsy cores, and clinical stage ≤T2.[9] Patients were classified as "PRIAS-eligible" if they met all these inclusion criteria at baseline and as "PRIAS-ineligible" if they did not.

Outcome measures

Our primary outcome measure was the total percentage of patients who received PSA monitoring, repeat biopsy testing, and overall follow-up (PSA and biopsies combined) concordant with the PRIAS follow-up protocol. Differences between hospitals in the proportion of patients with concordant follow-up were determined. A secondary outcome measure was the rate of metastasis (bone and/or lymph node) during AS monitoring. A patient was considered to have developed metastatic PCa during AS follow-up (time between the date of diagnosis and discontinuation of AS) if metastases were detected via diagnostic imaging (MRI, choline or prostate-specific membrane antigen positron emission tomography/computed tomography for lymph node metastasis and/or visceral metastasis and a bone scan for bone metastasis) or lymph node metastasis detected by lymph node dissection.

Statistical analysis

Possible significant differences in mean values between hospitals were assessed using using one-way analysis of variance. We evaluated differences in the proportions of patients using the Fisher's Exact Test. Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated using univariable cox-regression. Analysis was performed using SPSS for Windows v.24.0 (IBM Corp., Armonk, NY, USA). A *p*-value of <0.05 was considered significant.

RESULTS

Study population

A total number of 1181 patients were diagnosed with PCa and included for AS between January 1, 2008 and December 31, 2014 at six Santeon hospitals. This included an initial 181/1181 (15%) of patients with incidental tumours (cT1a/b), who were later excluded. Baseline characteristics for the 1000 patients with screen-detected PC on AS are presented in Table 1. The table shows the differences between hospitals with regard to baseline serum PSA, PSAD, number of positive biopsy cores and proportions of PRIAS-eligible and PRIAS-ineligible patients.

PSA follow-up testing

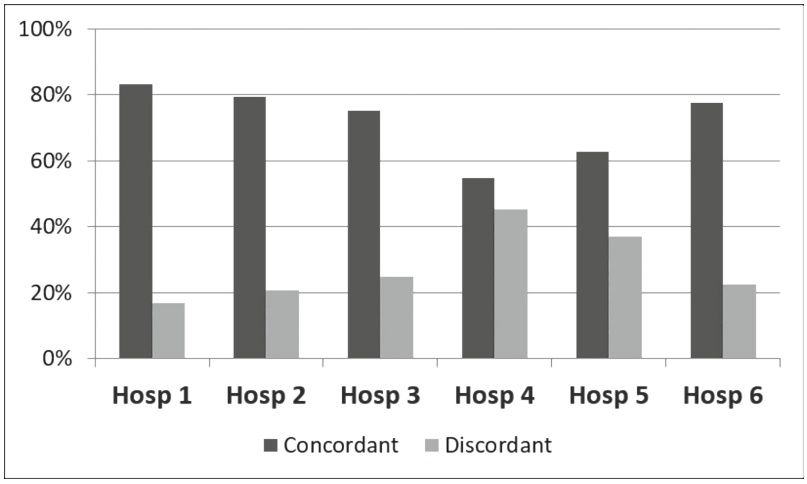
A total of 958 patients had treatment-free follow-up of more than 6 mo. The percentages of patients receiving PRIAS-concordant and -discordant PSA testing are presented in Figure 1. The variation between hospitals was considerable. Hospital 1 had the highest compliance, as 83% of patients had PRIAS-concordant PSA monitoring. The least strict monitoring occurred in hospital 4, where only 55% of the patients had PRIAS-concordant PSA monitoring. Overall, 706/958 patients (74%) had PRIAS-concordant PSA monitoring. The proportion of patients with a PSA doubling time (PSADT) of <3 yr did not differ significantly between the groups with discordant and concordant PSA follow-up (42% vs 48%; p = 0.156). The group of patients with discordant PSA testing was slightly younger (mean age 66.8 vs 68.7 yr; p < 0.001).

TABLE 1. Baseline characteristics of the AS patient population by hospital

	Hospital 1	Hospital 2	Hospital 3	Hospital 4	Hospital 5	Hospital 6	*p*
Patients (*N*)	248	144	166	178	78	186	
Age, SD (years)	67.9 ± 6.4	67.3 ± 6.5	68.1 ± 6.6	67.4 ± 5.6	67.6 ± 5.7	65.3 ± 6.8	<0.001
PSA (ng/ml)	8.2 ± 5.4	7.4 ± 3.4	7.0 ± 3.3	9.2 ± 4.8	8.2 ± 4.1	8.6 ± 4.9	<0.001
PSAD (ng/ml/ml)[a]	0.17 ± 0.14	0.18 ± 0.11	0.14 ± 0.07	0.21 ± 0.12	0.19 ± 0.11	0.18 ± 0.12	<0.001
cT stage, *N* (%)[b]							
T1c	200 (81)	110 (76)	104 (63)	144 (81)	74 (95)	156 (84)	
T2a	11 (4)	10 (7)	25 (15)	8 (5)	1 (1)	9 (5)	
T2b	2 (1)	1 (1)	4 (2)	7 (4)	0 (0)	0 (0)	
T2c	4 (2)	3 (2)	0 (0)	4 (2)	0 (0)	4 (2)	
T2	24 (10)	19 (12)	31 (19)	15 (8)	3 (3)	16 (9)	
T3/T4	7 (3)	1 (1)	2 (1)	0 (0)	0 (0)	1 (1)	
Total biopsy cores	9.4 ± 2.1	9.5 ± 1.7	8.9 ± 2.2	8.7 ± 2.4	9.9 ± 0.7	8.1 ± 1.2	<0.001
No. positive cores	1.8 ± 1.1	1.6 ± 1.0	1.6 ± 0.9	1.8 ± 1.2	1.6 ± 1.1	1.4 ± 0.8	<0.001
Gleason score *N* (%)							
4	0 (0)	3 (2)	0 (0)	0 (0)	0 (0)	2 (1)	
5	0 (0)	5 (4)	0 (0)	0 (0)	2 (3)	3 (2)	
3 + 3	230 (93)	126 (88)	165 (99)	171 (96)	74 (95)	169 (91)	
3 + 4	15 (6)	10 (6)	1 (1)	6 (3)	2 (3)	12 (7)	
4 + 3	3 (1)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)	
Risk Group *N* (%)[c]							
PRIAS-eligible	123 (50)	85 (59)	104 (63)	60 (34)	36 (46)	102 (55)	
PRIAS-ineligible	125 (50)	59 (41)	62 (37)	118 (66)	42 (54)	84 (45)	

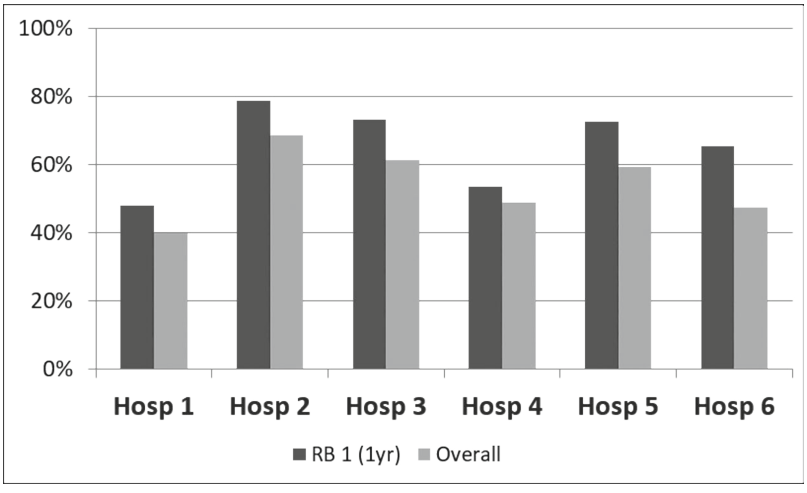
[a]PSAD; PSA density = PSA / prostate volume measured by TRUS or MRI.
[b]Clinical T stage based on digital rectal examination and transrectal ultrasonography.
[c]Risk classification based on PRIAS:
PRIAS-eligible: PSA≤10 ng/ml, PSAD <0.2, Gleason ≤6, <3 positive biopsy cores and clinical stage ≤T2.
PRIAS-ineligible: not complying to one or more of the abovementioned PRIAS inclusion criteria.

FIGURE 1. Percentages of patients with PRIAS compliant PSA testing



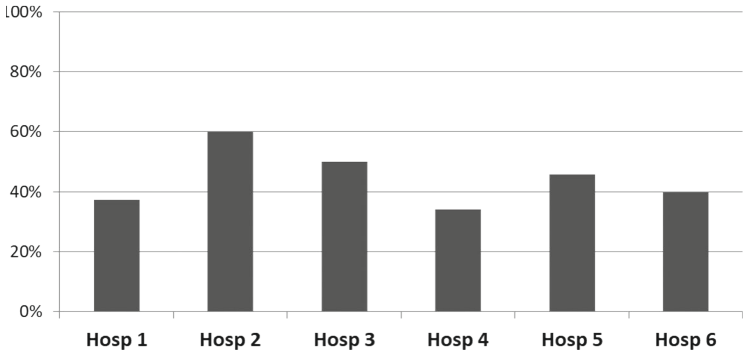
*PRIAS concordant PSA testing: patients underwent between 75% and 100% of the recommended. PSA tests. Discordant: patients underwent less than 75% of total PSA tests recommended by the PRIAS protocol.

FIGURE 2. Percentages of patients with PRIAS concordant follow-up biopsies



RB 1 = % of patients that underwent a first repeat biopsy at 1 year after diagnosis; Overall = % of patients that were compliant to all scheduled biopsies given their AS length.

FIGURE 3. Percentages of patients with overall PRIAS-concordant follow-up (PSA monitoring and biopsy testing)



*Patients underwent the minimum number of scheduled repeat biopsies depending on their follow-up time. PSA testing was regarded as PRIAS compliant if ≥75% of required PSA tests were performed depending on AS follow-up time.

Repeat biopsies and overall PRIAS concordant follow-up

Protocol adherence regarding scheduled prostate biopsies also differed between hospitals. As can be seen in Figure 2, the highest percentage of patient compliance with the first scheduled repeat biopsy was observed in hospital 2 (79%), while the lowest was in hospital 1 (48%). In the overall population, 473/912 patients (52%) were compliant with all scheduled repeat biopsies during their follow-up. At an institutional level, the highest percentage was found in hospital 2 (68%) and the lowest in hospital 1 (40%). There was also great variation between hospitals regarding the percentage of patients with overall PRIAS concordant follow-up; hospital 4 had the lowest proportion of patients with overall concordant follow-up (34%) and hospital 2 had the highest (60%; Figure 3).

MRI for follow-up

MRI of the prostate was performed during follow-up in 449/1000 patients (45%). The proportion of patients undergoing prostate MRI during AS varied significantly among institutions. The highest percentage was observed in hospital 6 (104/186, 56%) and the lowest in hospital 5 (22/78, 28%; $p < 0.001$). In some cases, MRI was performed instead of repeat biopsy. In a total of 41 patients, MRI of the prostate was performed instead of the first repeat biopsy. If we consider patients as compliant if MRI was performed instead of the scheduled biopsy, the total number of patients with PRIAS-concordant repeat biopsy (or MRI) follow-up would increase from 473/912 (52%) to 537/912 (59%). Prostate MRI increasingly replaced prostate biopsies over time. Of all patients diagnosed in 2008, 2/90 (2%) underwent MRI instead of the first repeat biopsy. Among patients diagnosed in 2014, this was performed in 15/126 cases (12%).

Follow-up intensity and oncological outcome

A total of 13 patients developed metastatic PCa (positive lymph nodes and/or bone metastasis) while on AS, and thus missed the window of curability during monitoring. The median duration of AS was 40.8 mo (interquartile range 16.8–59.1). The baseline characteristics for these patients are presented in Table 2. To evaluate the potential association between follow-up intensity and unfavourable oncological outcomes, data for men who developed metastatic PCa during AS follow-up were analysed. The rate of metastasis was significantly higher among men who had low-intensity PSA monitoring (<75% of recommended; Table 3). In the PRIAS-ineligible risk group, the HR for developing metastasis during AS was significantly higher among patients with discordant PSA monitoring than the concordant PSA monitoring group (HR 5.25, 95% CI 1.02 – 27.10). We found no significant correlation between discordant biopsy testing and the rate of metastasis for both the PRIAS-eligible and PRIAS-ineligible subgroups (p = 0.6). However, among all the patients with discordant biopsy testing, 35% had rapidly increasing PSA levels (defined as PSADT <3 yr), which is significantly lower than for the concordant biopsy group (54%; p < 0.001). In the PRIAS-ineligible group, overall discordant follow-up (discordant PSA monitoring and/or discordant biopsy testing) was also associated with a higher rate of metastatic PCa (2% vs 0%; p = 0.047).

TABLE 2. Patients who developed metastasized prostate cancer during AS follow-up

Hosp	cT	PSA	Positive cores	Grade Group	Risk group	PRIAS Eligible	PSADT (years)	Meta-stasis site[a]	AS FU (months)	PSA[b] PRIAS	RB[c] PRIAS
1	T2	23.0	1	1	High risk	No	3.4	LN	26	No	No
	T2	11.4	7	2	Intermed	No	1.4	LN	19	Yes	No
2	T2	7.4	2	1	Low-risk	Yes	0.9	Bone	13	Yes	Yes
	T1c	6.3	1	1	Low-risk	Yes	9.3	LN	84	Yes	Yes
3	T1c	9.0	1	1	Low-risk	No	4.6	Bone	49	No	Yes
4	T1c	16.0	1	1	Intermed	No	0.3	LN	25	No	Yes
	T2b	5.9	2	1	Low-risk	Yes	1.9	Bone, LN	39	No	No
	T1c	17.0	1	1	Intermed	No	0.9	Bone, LN	11	No	Yes
5	T1c	10.0	3	1	Low-risk	No	1.3	LN	13	No	Yes
	T1c	5.8	1	1	Low-risk	Yes	1.2	LN	7	Yes	Yes
	T1c	4.9	1	1	Low-risk	Yes	1.1	LN	4	Yes	Yes
	T1c	7.0	1	1	Low-risk	No	7.5	Bone	73	Yes	No
6	T2	7.0	1	1	Low-risk	Yes	2.1	LN	24	No	No

[a]Location where metastastized prostate cancer was found (no distinction between distant or pelvic lymphnodes); [b]PSA testing performed concordant with the PRIAS protocol; [c]Repeat biopsies performed concordant with the PRIAS protocol. PSADT = PSA doubling time, FU = follow-up, LN = lymph node.

TABLE 3. Risk of metastasis due to low intensity monitoring per risk group

		Follow-up	Patients who developed metastasized PCa during AS (%)	*p*	OR (95% CI)
PSA	PRIAS-eligible (*N* = 495)	Discordant PSA monitoring[a]	2/104 (2)	0.611	2.39 (0.42 – 13.50)
		Concordant PSA monitoring	4/391 (1)		
	PRIAS-ineligible (*N* = 463)	Discordant PSA monitoring	5/148 (3)	0.037	5.25 (1.02 – 27.10)
		Concordant PSA monitoring	2/315 (1)		
Biopsy	PRIAS-eligible (*N* = 481)	Discordant repeat biopsy testing	2/210 (1)	0.701	0.36 (0.06 – 2.08)
		Concordant repeat[b] biopsy testing	4/271 (2)		
	PRIAS-ineligible (*N* = 431)	Discordant repeat biopsy testing	3/229 (1)	0.711	0.32 (0.07 – 1.51)
		Concordant repeat biopsy testing	4/202 (2)		
Overall	PRIAS-eligible (*N* = 495)	Overall discordant[c] follow-up	2/257 (1)	0.434	0.30 (0.05 – 1.69)
		Overall concordant follow-up	4/238 (2)		
	PRIAS-ineligible (*N* = 463)	Overall discordant follow-up	7/286 (2)	0.047	32.9 (0.02 – 51163)
		Overall concordant follow-up	0/177 (0)		

[a]Patients underwent ≥75% of recommended PSA tests according to the PRIAS protocol given their AS duration.

[b]Patients underwent all scheduled repeat biopsies given their AS duration.

[c]Patients underwent PRIAS concordant PSA monitoring as well as concordant repeat biopsy testing.

DISCUSSION

In this study, we observed significant variation in AS protocol adherence between six large teaching hospitals in the Netherlands. In the overall population, 43% of the patients on AS received follow-up concordant with PRIAS. These are important findings, as they confirm our hypothesis that AS follow-up is less strict in daily clinical practice.

Results from a large number of prospective AS cohorts have been published in the literature.[12,14–21] However, results for compliance rates with the respective follow-up protocols are limited. We identified two other study groups that performed comparable research. Our findings are in line with their results, as Luckenbaugh et al[22] also observed limited protocol adherence in a real-world cohort treated with AS at collaborating urological practices in Michigan (MUSIC). The authors reported that 26.5% of the patients who remained on AS for a minimum of 2 yr had follow-up compliant with the National Comprehensive Cancer Network guidelines.[22] The PRIAS study group reported that 91% of patients complied with all PSA visits and that 81% of men complied with the 1-yr prostate biopsy.[13] Overall percentages in our cohort were lower, as we observed that 74% of patients (706/958) had concordant PSA testing and 63% (570/912) complied with the first repeat biopsy. The differences between the PRIAS cohort and the real-world populations described in the present study and the MUSIC study indicate that there is a substantial gap regarding AS protocol adherence between the research setting and daily practice. The size of this gap may also differ at an institutional level, as significant differences between institutions were observed in both studies.

Noncompliance with AS protocols is understandable, as biopsies are often considered painful by patients.[13] Moreover, biopsies are associated with several complications such as pain, hematuria, urinary tract infections and even urosepsis.[23] Given the significant burden, costs, and time associated with frequent monitoring, we have to deliberate on the intensity of AS follow-up schedules. However, it remains challenging to determine for whom and to what extent the intensity of AS follow-up schedules can be reduced. Given the higher risk of missing the window of curability, low-intensity monitoring may not be a safe option for PRIAS-ineligible patients. However, we found no significant difference in the rate of metastasis between discordant and concordant monitoring in the PRIAS-eligible group. This suggests that patients who can be classified with the lowest risk at diagnosis might be candidates for a less intensive follow-up schedule without being at risk of worsening their prognosis.

We did not observe a higher rate of metastasis among patients noncompliant with the PRIAS repeat biopsy schedule in comparison to compliant patients. The lack of association between biopsy noncompliance and prognosis can be partly explained by the fact that most of these men did not have rapidly increasing PSA (65% had a PSADT of >3 yr). Of the patients who were compliant, 54% had a PSADT of <3 yr, indicating that repeat biopsies were performed more frequently in cases with faster rising PSA. This indicates that the decision to repeat prostate biopsy was partly based on serial serum PSA results instead of what the protocol advised. This can be further explained by taking hospital 1 as an

example. Hospital 1 had the lowest percentage of patients complying with 1-yr repeat biopsy (48%) but the highest percentage of patients with concordant PSA monitoring (83%). The rate of metastasis was relatively low in this hospital (2/248, 1%) in comparison to the other clinics. These findings suggest that deviating from the repeat biopsy protocol may be safe as long as PSA kinetics are monitored closely.

The strengths of our study include the large sample size, a study population representing the real-world clinical situation, and evaluation of AS management strategies including a wide range of follow-up tests. Besides the strengths of the study, some limitations should be acknowledged. First, the retrospective nature of the study is a limitation and carries a risk of bias due to confounding by indication, especially concerning allocation of patients to a low- or high-intensity monitoring strategy. However, given the fact that it is possible that patients with more aggressive tumours received closer monitoring than patients with less aggressive tumours, we still found significant differences regarding rates of metastasis. Thus, this form of bias has not altered our conclusions. Second, we only evaluated whether AS follow-up was concordant with the PRIAS follow-up guidelines. We did not evaluate deviations from the PRIAS protocol regarding recommendations for discontinuation of AS (ie, Gleason sum score ≥7 on repeat biopsy or >2 cores positive). Therefore, we cannot assess the potential impact of this on our observed outcomes, as it was unclear which patients remained on AS despite Gleason upgrading or a substantial increase in tumour volume. However, we did collect data for an individual's PSA course, which also provides important information on tumour aggressiveness.

CONCLUSIONS

Compliance with the PRIAS protocol in a real-world cohort is low: 43% of patients on AS in daily clinical practice receive PRIAS-concordant follow-up. Noncompliance with the PRIAS follow-up protocol was associated with a higher rate of metastasis among PRIAS-ineligible patients. Higher rates of metastasis were found in patients with discordant PSA monitoring, but not among patients not complying with the biopsy schedule. This suggests that low biopsy compliance seems to be "compensated" by high PSA compliance. The fact that discordant follow-up was not associated with a higher rate of metastasis among PRIAS-eligible patients suggests that less strict monitoring may be safe in this subgroup.

REFERENCES

1. Bruinsma SM, Bokhorst LP, Roobol MJ, Bangma CH. How Often is Biopsy Necessary in Patients with Prostate Cancer on Active Surveillance? *J Urol*. 2016;195:11-12.
2. Loeb S, Berglund A, Stattin P. Population based study of use and determinants of active surveillance and watchful waiting for low and intermediate risk prostate cancer. *J Urol*. 2013;190:1742-1749.
3. Weerakoon M, Papa N, Lawrentschuk N, et al. The current use of active surveillance in an Australian cohort of men: A pattern of care analysis from the Victorian Prostate Cancer Registry. *BJU Int*. 2015;115:50-56.
4. Loeb S, Folkvaljon Y, Curnyn C, Robinson D, Bratt O, Stattin P. Uptake of active surveillance for very-Low-risk prostate cancer in Sweden. *JAMA Oncol*. 2017;3:1393-1398.
5. Jeffrey J. Tosoian, H. Ballentine Carter, Abbey Lepor and SL. Active Surveillance for Prostate Cancer: Contemporary State of Practice. *Nat Rev Urol*. 2016;116:1477-1490.
6. Loeb S, Walter D, Curnyn C, Gold HT, Lepor H, Makarov D V. How Active is Active Surveillance? Intensity of Followup during Active Surveillance for Prostate Cancer in the United States. *J Urol*. 2016;196:721-726.
7. Azmi A, Dillon RA, Borghesi S, et al. Active surveillance for low-risk prostate cancer: diversity of practice across Europe. *Ir J Med Sci*. 2015;184:305-311.
8. Mitsuzuka K, Koga H, Sugimoto M, et al. Current use of active surveillance for localized prostate cancer: A nationwide survey in Japan. *Int J Urol*. 2015;22:754-759.
9. Bul M, Zhu X, Valdagni R, et al. Active surveillance for low-risk prostate cancer worldwide: The PRIAS study. *Eur Urol*. 2013;63:597-603.
10. Soeterik TFW, van Melick HHE, Dijksman LM, Biesma DH, Witjes JA, van Basten JPA. Active Surveillance for Prostate Cancer in a Real-life Cohort: Comparing Outcomes for PRIAS-eligible and PRIAS-ineligible Patients. *Eur Urol Oncol*. 2018;1:231-237.
11. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) -A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377-381.
12. Bokhorst LP, Valdagni R, Rannikko A, et al. A Decade of Active Surveillance in the PRIAS Study: An Update and Evaluation of the Criteria Used to Recommend a Switch to Active Treatment. *Eur Urol*. 2016;70:954-960.
13. Bokhorst LP, Alberts AR, Rannikko A, et al. Compliance rates with the Prostate Cancer Research International Active Surveillance (PRIAS) protocol and disease reclassification in noncompliers. *Eur Urol*. 2015;68:814-821.
14. Barnett CL, Auffenberg GB, Cheng Z, et al. Optimizing active surveillance strategies to balance the competing goals of early detection of grade progression and minimizing harm from biopsies. *Cancer*. 2018;124:698-705.
15. Kearns JT, Faino A V., Newcomb LF, et al. Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study[Formula presented]. *Eur Urol*. 2018;73:706-712.
16. Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. *J Clin Oncol*. 2015;33:272-277.
17. Godtman RA, Holmberg E, Khatami A, Pihl CG, Stranne J, Hugosson J. Long-term Results of Active Surveillance in the Göteborg Randomized, Population-based Prostate Cancer Screening Trial. *Eur Urol*. 2016;70:760-766.
18. Selvadurai ED, Singhera M, Thomas K, et al. Medium-term outcomes of active surveillance for localised prostate cancer. *Eur Urol*. 2013;64:981-987.
19. Thompson JE, Hayen A, Landau A, et al. Medium-term oncological outcomes for extended vs saturation biopsy and transrectal vs transperineal biopsy in active surveillance for prostate cancer. *BJU Int*. 2015;115 (6):884-891.

20. Tosoian JJ, Mamawala M, Epstein JI, et al. Intermediate and longer-term outcomes from a prospective active-surveillance program for favorable-risk prostate cancer. *J Clin Oncol.* 2015 ;33:3379-3385.
21. Eggener SE, Mueller A, Berglund RK, et al. A Multi-Institutional Evaluation of Active Surveillance for Low Risk Prostate Cancer. *J Urol.* 2013;189:S19-25.
22. Luckenbaugh AN, Auffenberg GB, Hawken SR, et al. Variation in Guideline Concordant Active Surveillance Followup in Diverse Urology Practices. *J Urol.* 2017;197:621-626.
23. Loeb S, Vellekoop A, Ahmed HU, et al. Systematic review of complications of prostate biopsy. *Eur Urol.* 2013;64:876-892.

PART II: IMPACT OF MRI ON PROSTATE CANCER RISK CLASSIFICATION

CHAPTER 4

4

Multiparametric Magnetic Resonance Imaging Should Be Preferred Over Digital Rectal Examination for Prostate Cancer Local Staging and Disease Risk Classification

T.F.W. Soeterik
H.H.E. van Melick
L.M. Dijksman
D.H. Biesma
J.A. Witjes
J.P.A. van Basten

UROLOGY. 2020;38:372-378

ABSTRACT

Background: The currently recommended prostate cancer risk classification system, including digital rectal examination (DRE) for local staging, is the cornerstone to guide clinical decision-making. Although multiparametric magnetic resonance imaging (mpMRI) has potential benefits over DRE with regard to local tumour staging, its impact on risk group classification and treatment selection is unclear. Therefore, we will assess the impact of mpMRI local tumour staging on prostate cancer risk stratification and choice of treatment.

Methods: Prostate cancer patients, newly diagnosed from 2017 to 2018 at 7 Dutch teaching hospitals were included. Risk group classification was done twice, using either DRE or mpMRI information. Risk group migration and rates of treatment intensification associated with mpMRI upstaging were established. Diagnostic accuracy measures for the detection of non-organ-confined disease (stage ≥T3a), for both DRE and mpMRI, were assessed in patients undergoing robot-assisted radical prostatectomy.

Results: A total of 1683 patients were included. Upstaging due to mpMRI staging occurred in 493 of 1683 (29%) patients and downstaging in 43 of 1683 (3%) patients. Upstaging was associated with significant higher odds for treatment intensification (odds ratio [OR]: 3.5 95% confidence interval [CI] 1.9 - 6.5). Stage ≥T3a on mpMRI was the most common reason for risk group upstaging (77%). Sensitivity for the detection of stage ≥T3a was higher for mpMRI compared to DRE (51% vs 12%, $p < 0.001$), whereas specificity was lower (82% vs 97%, $p < 0.001$). mpMRI resulted in a significantly higher cumulative rate of true positive and true negative stage ≥T3a predictions compared with DRE (67% vs 58%, $p < 0.001$).

Conclusions: Use of mpMRI tumour stage for prostate cancer risk classification leads to migration to a higher risk group in 1 of 3 patients. mpMRI enables superior detection of non-organ-confined disease compared with DRE and should be the preferred tool for determining clinical tumour stage.

BACKGROUND

In 1998, D'Amico and colleagues proposed the D'Amico classification, stratifying patients with prostate cancer into low-, intermediate-, or high-risk disease to predict disease progression following radical treatment based on pretreatment clinical T-stage, Gleason score and serum prostate-specific antigen (PSA) level.[1] The classification forms the basis for both the National Comprehensive Cancer Network and the European Association of Urology staging systems and is therefore commonly used by urologists to assist patient counselling.[2,3]

Multiparametric magnetic resonance imaging (mpMRI) of the prostate is one of the most influential emerging diagnostic modalities within the field of prostate cancer of the last 2 decades, establishing an important role in the detection of clinically significant disease[4,5] and local tumour staging.[6,7] With regard to local staging of prostate cancer, mpMRI potentially enables more accurate estimation of the local tumour size and extent compared with digital rectal examination (DRE),[8] and thus may improve risk stratification and treatment selection. However, it is recommended that imaging alone should not replace the DRE for defining clinical tumour stage due to lack of interobserver reproducibility, issues with patient selection and contradictory results.[9,10]

Previous studies have shown that use of mpMRI information for risk group classification leads to substantial upstaging of patients to a higher risk group.[11,12] In addition, these studies have demonstrated that mpMRI use may lead to alterations in treatment advice. However, these studies regarded both relatively small, single-centre populations, including solely patients that underwent radical prostatectomy. Therefore, results may not be generalisable to the overall population of newly diagnosed prostate cancer patients. In addition, the possible stage migration due to mpMRI could impact the prognostic characteristics of the original D'Amico classification system, due to migration of patients with more favorable characteristics to higher risk groups. For instance, it remains unclear whether mpMRI stage ≥T3 is prognostically equivalent to clinical stage ≥T3 assessed by DRE.[12]

In this study, we aim to establish the impact of mpMRI use for local staging on prostate cancer risk stratification and treatment assignment in a large real-world multi-centre cohort of newly diagnosed prostate cancer patients. In addition, we will evaluate the accuracy of both DRE and mpMRI for the detection of non-organ-confined disease in a subset of patients undergoing robot-assisted radical prostatectomy (RARP).

METHODS

Study population and data collection

This study was conducted within the Santeon consortium, which consists of 7 teaching hospitals in the Netherlands. After receiving institutional review board approval, data was extracted from the prospectively maintained Santeon database containing baseline tumour characteristics (baseline PSA, biopsy International Society of Urological Pathology [ISUP] grade,[13] clinical T-stage assessed by DRE, mpMRI T-stage, prostate volume, choice of treatment, surgical outcomes and PSA follow-up). Patients newly diagnosed with clinically localised or locally advanced prostate cancer (any T, any N, M0), between January 1st, 2017 and December 31st, 2018 who underwent mpMRI were included for analysis.

Variable and outcome definitions

The DRE was performed by a senior resident or urologist during the primary staging work-up. MpMRI T-stage was retrieved from the radiology reports. MpMRI was performed either pre-biopsy or a minimum of 6 weeks after biopsy. The MRI protocols of the seven Santeon hospitals are presented in Table S1. Radiological reporting was done according to the PI-RADS Version 2.[14] Presence of non-organ-confined disease included reported suspicion of extraprostatic extension, seminal vesicle invasion and/or invasion to organs adjacent to the prostate (stage ≥T3a). Surgical pathology reporting was performed according to the ISUP guidelines.[15]

Impact of mpMRI on risk group classification and treatment intensification

Patients were classified into risk groups according to the European Association of Urology (EAU) classification system, using clinical T-stage assessed by DRE: low-risk (T1 or T2a, PSA ≤ 10 ng/mL and ISUP 1), intermediate-risk (T2b, PSA 10 – 20 ng/mL, ISUP 2 or 3), localised high-risk (T2c or higher, PSA > 20 ng/mL or ISUP > 3) and locally advanced high-risk (T3a or higher and PSA > 20 ng/mL or ISUP > 3). Patients were also classified using the mpMRI T-stage. Number of patients that migrated to either a lower or higher risk group based on mpMRI findings was established. In patients who underwent external beam radiation therapy (EBRT), the consequence of upstaging by mpMRI results for treatment advice was defined as addition of androgen deprivation therapy to EBRT. Rates of treatment intensification were compared for patients with and without upstaging due to mpMRI information.

Accuracy of DRE and MRI for detection of non-organ-confined disease

Sensitivity, specificity, negative predictive value and positive predictive value were assessed for DRE as well as mpMRI for prediction of non-organ-confined disease (pathological tumour stage ≥T3a). Analysis was done in the subset of patients that underwent RARP as primary treatment (within 6 months of performed DRE as well as MRI), using histopathological evaluation of the prostate specimen as the gold standard.

Prognostic impact of non-organ-confined disease observed on mpMRI

Potential prognostic impact of non-organ-confined disease (stage ≥T3a) on mpMRI was assessed in patients undergoing RARP, between January 1, 2017 and July 1, 2018. Patients treated from July and onwards were excluded due to limited follow-up time. Biochemical recurrence (BCR) rates were assessed per D'Amico risk group, stratifying for presence of stage ≥T3a on mpMRI per risk group. BCR after RARP was defined as a postoperative PSA ≥ 0.2 ng/ml.

Statistical analysis

Diagnostic accuracy measures (sensitivity, specificity, positive predictive value [PPV] and negative predictive value [NPV]) were established for the detection of stage ≥T3a for both DRE and mpMRI, using McNemar's Test to assess statistical significance of differences. Diagnostic accuracy of both modalities, in terms of area under the receiver operating curve (AUC), was also evaluated. Biochemical progression-free survival was assessed by Kaplan-Meier using the log-rank test for assessing statistical significant differences. A p-value of < 0.05 was considered significant. Statistical analysis was done using IBM SPSS for windows, Version 24.0 (IBM Corp., Armonk, N.Y., USA).

RESULTS

Study population

A total of 2383 patients were identified fulfilling the inclusion criteria. In total, 1740 (73%) underwent mpMRI during diagnostic work-up, of which 1683 of 1740 (97%) had complete baseline data, who were included for analysis (Figure S1). Patient baseline characteristics are presented in Table 1.

TABLE 1. Baseline characteristics of the overall population

	***N* (%)**
Patients	1683
Age (median, IQR)	68 (64 - 73)
PSA (mean)	8.5 (6.0 – 13.7)
Clinical T-stage	
T1	928 (55)
T2a	488 (29)
T2b	46 (3)
T2c	62 (4)
T3/T4	159 (9)
Highest biopsy ISUP grade	
1	664 (40)
2	448 (27)
3	255 (15)
4	274 (16)
5	42 (3)
Risk group (DRE)	
Low	447 (27)
Intermediate	658 (39)
High (localized)	350 (21)
High (locally advanced)	228 (14)
MRI T-stage	
T1c	239 (14)
T2a	682 (41)
T2b	51 (3)
T2c	164 (10)
T3a	363 (17)
T3b	159 (3)
T4	25 (2)
Risk Group (mpMRI)	
Low	368 (22)
Intermediate	435 (26)
High (localized)	301 (18)
High (locally advanced)	579 (34)
Clinical N stage	
N0	1327 (79)
N1	103 (6)
NX	253 (15)

TABLE 2 Risk group migration patterns in patients undergoing mpMRI in the overall population

		Risk group based on MRI				
		Low *N* (%)	**Intermediate *N* (%)**	**High (localized) *N* (%)**	**High (locally advanced) *N* (%)**	**Total *N* (%)**
Based on DRE	**Low**	357 (80)	9 (2)	39 (9)	43 (10)	448 (27)
	Intermediate	5 (1)	409 (62)	66 (10)	178 (27)	658 (39)
	High (localized)	3 (1)	7 (2)	181 (52)	158 (45)	349 (21)
	High (locally advanced)	3 (1)	10 (4)	15 (7)	200 (88)	228 (14)
		368 (22)	435 (26)	301 (18)	579 (34)	1683 (100)

EAU risk group migration

Percentages of patients classified according to the EAU risk classification based on DRE as well as mpMRI are presented in Table 2. Risk group migration based on mpMRI findings occurred in 536 of 1683 (32%) of patients. Respectively 493 (29%) patients migrated to a higher risk group and 43 (3%) patients to a lower risk group. Upstaging occurred respectively among 21% of patients initially classified as low-risk, 37% of the intermediate-risk and 45% of the localised high-risk. Presence of stage ≥T3a disease on mpMRI (and subsequent migration to locally advanced high-risk) was the reason for upstaging in 378 of 493 (77%). Migration to locally advanced high-risk occurred in respectively 10% of the low-risk, 27% intermediate-risk and 45% of the localised high-risk patients (Table 2).

Accuracy of DRE and MRI for detection of non-organ-confined disease

RARP was the primary choice of treatment in 552 of 1683 (33%) patients. Of these, 509 of 552 (92%) underwent mpMRI within 6 months before surgery, including 3 patients that underwent surgery elsewhere with unavailable pathology reports. A total of 506 patients with complete study data treated with RARP were included for the analysis (Table S2). Stage migration patterns in the RARP subset where comparable to those observed in the overall population (Table S3). Sensitivity of mpMRI and DRE for the detection of stage ≥T3a in this subset were respectively 51% and 12%, $p < 0.001$ (Table 3). Specificity rates were respectively 97% and 82% for DRE and mpMRI ($p < 0.001$). Cumulative rate of true positive and true negative stage ≥T3a predictions was higher for mpMRI compared to DRE (341 of 506 [67%] vs 294 of 506 [58%], $p < 0.001$). Diagnostic accuracy measures of both DRE and mpMRI for detection of nonorgan-confined disease (stage ≥T3a) per EAU risk group are presented in Table 3.

TABLE 3. Diagnostic accuracy measures of DRE and mpMRI for the detection of non-organ-confined disease

DRE	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (95% CI)
Low	0/15 (0)	54/54 (100)	N/A	54/69 (78)	0.50 (0.33 – 0.67)
Intermediate	8/116 (7)	159/164 (97)	8/13 (62)	159/267 (60)	0.52 (0.34 – 0.59)
High	21/103 (20)	52/54 (96)	21/23 (91)	52/134 (39)	0.58 (0.49 – 0.67)
Overall	29/234 (12)	265/272 (97)	29/36 (81)	205/470 (44)	0.55 (0.50 – 0.60)
MRI	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Accuracy (95% CI)**
Low	6/15 (40)	44/54 (82)	6/16 (38)	44/53 (83)	0.61 (0.44 – 0.78)
Intermediate	59/116 (51)	136/164 (83)	59/87 (68)	136/193 (71)	0.67 (0.50 – 0.74)
High	54/103 (52)	42/54 (78)	54/66 (82)	42/91 (46)	0.65 (0.56 – 0.74)
Overall	119/234 (51)	222/272 (82)	119/169 (70)	222/337 (66)	0.66 (0.61 – 0.71)

DRE = digital rectal examination, MRI = magnetic resonance imaging, Low: ISUP 1 and PSA ≤10 ng/mL, intermediate: ISUP 2 or 3 and/or PSA between 10 – 20 ng/mL, high: ISUP >3 and/or PSA > 20 ng/mL.

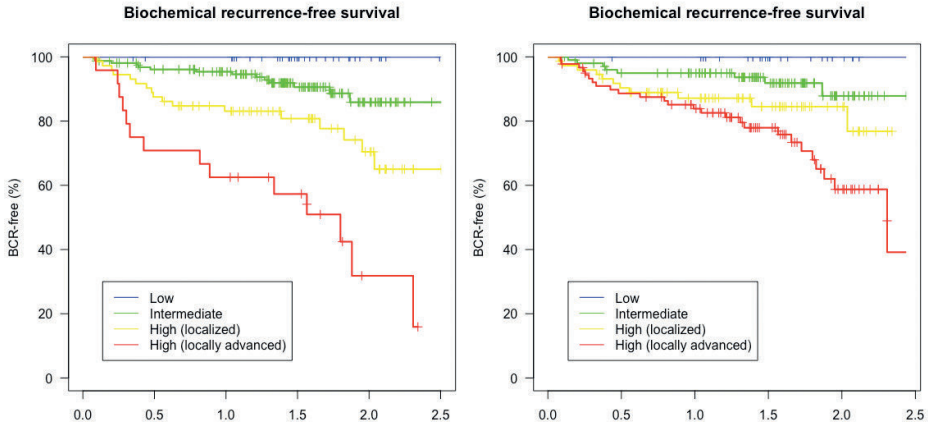
Impact of mpMRI on Treatment Intensification

The cohort included 193 patients low- and intermediate risk patients treated with EBRT. Among these, 101 of 193 (53%) migrated to a higher risk group after mpMRI. Upstaging was also associated with a higher percentage of patients with treatment intensification (addition of androgen deprivation therapy): 52 of 101 (52%) of patients with upstaging versus 23 of 92 (25%) patients without upstaging, p <0.001. Upstaging was also associated with a higher OR for treatment intensification (OR: 3.5 95% CI 1.9-6.5).

Association of DRE- and mpMRI T-stage with biochemical recurrence after RARP

A total of 300 patients underwent RARP between 2017 and mid-2018. Of these, PSA follow-up data was available in 293 (98%) patients. BCR occurred in 46 (16%) of these patients. Median overall follow-up was 567 days (IQR 432-706 days). BCR-free survival patterns for both DRE- and mpMRI-based risk classification are presented in Figure 1A and B. As shown, migration of patients to different risk groups led to a change in BCR-free survival patterns. Although log-rank tests showed statistically significant difference for both risk group stratification systems (p < 0.001 for both DRE and mpMRI), relative differences regarding progression-free survival between risk groups decreased if mpMRI T-stage was used. (Figure 1B). Stratifying for organ-confined (stage ≤T2) and stage ≥T3a disease on mpMRI within the EAU risk groups did not show any significant differences regarding BCR-free survival (Figure S2).

FIGURE 1. Impact of DRE and mpMRI based risk classification on biochemical-free survival patterns in 293 patients undergoing RARP. (A) DRE-based classification, Log-rank test: p < 0.001. (B) mpMRI-based classification, Log-rank test: p < 0.001.



A. B.

DISCUSSION

Our study demonstrates that use of mpMRI for local staging has a major impact on prostate cancer risk group classification, as stage migration to a higher disease stage occurred in one-third of the evaluated patients. In addition, mpMRI information impacts treatment assignment, as upstaging due to mpMRI findings was associated with a 3 times higher odds for treatment intensification in patients treated with EBRT. A comparative analysis of DRE and mpMRI for the detection of tumour extension outside the prostate showed that with mpMRI more cases of non-organ-confined disease were detected (sensitivity: 51% vs 12%, $p < 0.001$), countered by higher rates of false positive test results (specificity: 82% vs 97%, $p < 0.001$). Since mpMRI outperformed DRE in terms of cumulative rates of true negative and positive test results for non-organ-confined disease (67% vs 58%, $p < 0.001$), it should be preferred over DRE for local staging. The present findings results are consistent with those reported by Draulans et al, who evaluated the impact of mpMRI on the EAU risk classification in 180 patients undergoing RARP.[11] The authors reported comparable upstaging (31%) and downstaging (4%) of risk grouping, as well as the need for treatment intensification due to mpMRI information (27%).[11] Marcus et al also previously evaluated the impact of mpMRI on risk stratification, adhering to the NCCN risk classification.[12] In their study including 71 patients, they observed upstaging in 17% of the patients, as well as significant differences between pre-MRI and post-MRI risk group classifications. Impact on change of management as a direct result of mpMRI was observed in 13 of 71 (18%) patients. Main limitations of both studies were their single-centre nature, including a relatively small sample of patients undergoing RARP and a relatively low number of patients with low-risk disease (4% and 16%, respectively).

What our study adds is that we have validated their previously reported findings in a large multi-institutional cohort of newly diagnosed prostate cancer patients, including a far larger number of low-risk patients (*N* = 447, 27%). Our data confirms that additional use mpMRI for risk stratification results in stage migration to higher risk groups in one-third of the patients, and also assignment of patients to intensified treatment regimens. Furthermore, our analysis showed that the subsequent stage migration of patients to higher risk groups resulted in a more favorable prognosis of the high-risk subgroup as a whole, in terms of prolonged recurrence-free survival (Figure 1). This observed evolvement of cancer prognosis, due to change of tumour stage as a consequence of use of newly developed diagnostic modalities, is described as the "Will Rogers phenomenon."[16]

Our additional findings regarding the diagnostic accuracy of mpMRI and DRE for the detection of non-organ-confined disease may further guide clinicians in the way they should incorporate mpMRI information for risk group classification. As shown, the positive predictive value of mpMRI for stage ≥T3a increases in case of a higher PSA and ISUP biopsy grade, reaching 82% in patients with either PSA above 20 ng/mL or ISUP > grade 3. Opposite findings were observed regarding the NPV, as the NPV declined inversely proportional to the risk classification from low to locally advanced high-risk, respectively from 83% to 46%.

Our study shows that solely relying on mpMRI potentially leads to overtreatment given the relatively high false positive rates for non-organ-confined disease compared with DRE, especially in patients with a PSA below 10 ng/mL and biopsy ISUP grade 1. Presence of non-organ-confined disease on mpMRI is extremely rare in ISUP grade 1 prostate cancer, with a reported prevalence of 7 of 7817 (0.3%) in a series of patients undergoing radical prostatectomy.[17] Thus, if stage ≥T3a is present on mpMRI in patients with ISUP 1 on biopsy, this may either indicate a false positive test result or biopsy sampling error. In addition, our study revealed that lower specificity of mpMRI for the detection of stage ≥T3a is countered by a far higher sensitivity, compared with DRE. Among patients undergoing RARP with a serum PSA <20 ng/mL and ISUP grade ≤ 3, sensitivity and specificity for detection of stage ≥T3a were respectively 8 of 131 (6%) and 213 of 218 (98%) for DRE versus 65 of 131 (50%) and 180 of 218 (83%) for mpMRI. Use of mpMRI thus resulted in correctly classifying an additional 57 of 131 (44%) patients as high-risk, countered by additionally over staging solely 38 of 218 (17%) patients with organ-confined disease as high-risk.

These results confirm that mpMRI outperforms DRE with regard to local staging, and reduces the risk of understaging and potential undertreating those who harbor non-palpable stage ≥T3a prostate cancer. However, the lower specificity of mpMRI for the detection of stage ≥T3a tumours, and thus its increased risk overstaging, shows the technique is still not perfect. Combining mpMRI information with other baseline factors including detailed (target) biopsy information and baseline PSA can further improve prediction of unfavorable histopathology.[18] At present, several nomograms have been developed which can be used for this purpose.[19,20]

As extraprostatic extension is a strong independent predictor for biochemical recurrence following radical prostatectomy,[21,22] presence of stage ≥T3a on mpMRI can potentially be a valuable prognostic factor indicating the potential need for treatment intensification (eg, non-nerve sparing prostatectomy and additional pelvic lymph node dissection). In this cohort, mpMRI stage ≥T3a was not associated with decreased progression-free survival when compared with mpMRI stage ≤T2. Graphically, most noticeable difference was seen in the intermediate-risk group (Figure S2B). However, this difference was not statistically significant (p = 0.099). Absence of an association may be explained by either a lack of prognostic relevance of non-organ-confined disease on mpMRI or the fact that patients with stage ≥T3a on mpMRI were accurately assigned to a more intensive treatment regimen (eg, non-nerve sparing during RARP). The second hypothesis is the most attainable explanation, as this is an observational study wherein patients are assigned to certain treatments based on their prognostic risk factors. In addition, present study was limited by a relatively small sample size and short follow-up duration, and significant findings could still occur following study prolongation. Moving forward, further research is needed regarding the harms and benefits of mpMRI as a staging tool.

Despite the relevance of our findings, some important limitations need to be acknowledged. First, a limitation of our study is that we did not account for the potential impact of mpMRI target biopsy on stage migration. Only the highest ISUP grade found on either systematic or either mpMRI-targeted biopsy was registered. Thereby, there were differences regarding use of prebiopsy MRI between the hospitals. Fact that prebiopsy MRI was not performed standard in all patients may limit the generalisability of our results to cohorts in which prebiopsy MRI is standardly performed, as upstaging occurs less in hospitals who use prebiopsy mpMRI potentially resulting in lower rates of biopsy sampling error compared to only systematic biopsies.[23] Other limitations of our study are the heterogeneity of mpMRI protocols used, the wide variety of urologists and residents performing the DRE and radiologists performing mpMRI reading. However, the heterogeneity reflects a real-world situation and as we used multi-centre data, our results might provide an accurate overview of the clinical impact of mpMRI local staging in newly diagnosed patients.

CONCLUSIONS

Use of mpMRI information for prostate cancer risk stratification leads to upstaging in one-third of the patients, mainly due to superior detection of non-organ-confined disease compared with DRE. Since mpMRI results in a higher overall diagnostic accuracy for the detection of non-organ-confined disease, it should be preferred over DRE for determining clinical tumour stage. However, due to its limited specificity for the detection of non-organ-confined disease clinicians should be aware that mpMRI can lead to overstaging and potential overtreatment of patients who harbor genuine low-risk disease.

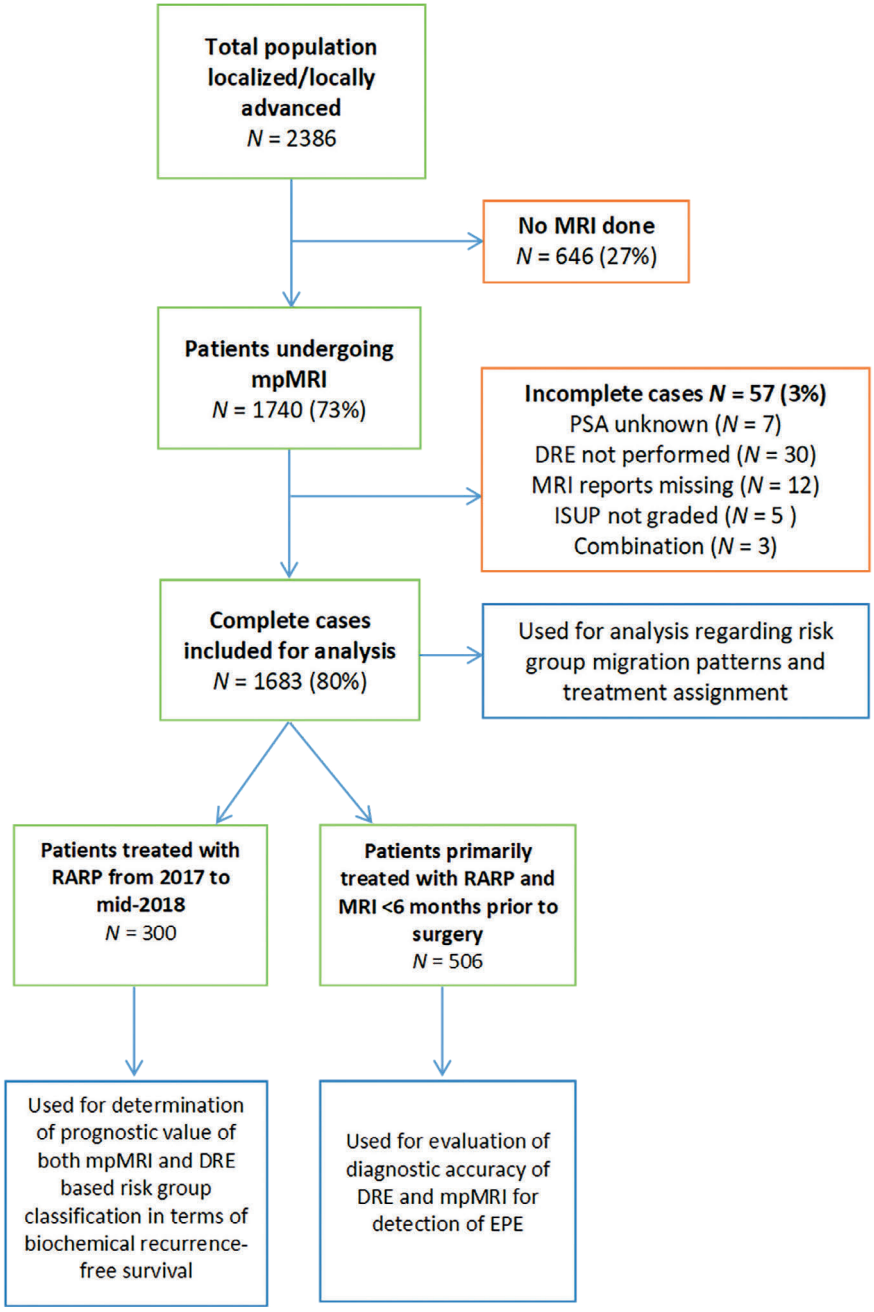
REFERENCES

1. D'Amico AV, Whittinton R, Malkowicz B, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA*. 1998;280:969–974.
2. Mohler JL, Armstrong AJ, Bahnson RR, et al. Prostate cancer, version 1.2016. *J Natl Compr Canc Netw*. 2016;14:19–30.
3. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol*. 2017;71:618–629.
4. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet*. 2017;389:815–822.
5. van der Leest M, Cornel E, Israel B, et al. Head-to-head comparison of transrectal ultrasound-guided prostate biopsy versus multiparametric prostate resonance imaging with subsequent magnetic resonance-guided biopsy in biopsy-naive men with elevated prostate specific antigen: a large prospective multicenter clinical study. *Eur Urol*. 2018;75:570–578.
6. Marenco J, Orczyk C, Collins T, Moore C, Emberton M. Role of MRI in planning radical prostatectomy: what is the added value? *World J Urol*. 2019;37:1289–1292.
7. Boesen L, Chabanova E, Logager V, Balslev I, Mikines K, Thomsen HS. Prostate cancer staging with extracapsular extension risk scoring using multiparametric MRI: a correlation with histopathology. *Eur Radiol*. 2015;25:1776–1785.
8. Caglic I, Kovac V, Barrett T. Multiparametric MRI - local staging of prostate cancer and beyond. *Radiol Oncol*. 2019;53:159–170.
9. de Rooij M, Hamoen EH, Witjes JA, Barentsz JO, Rovers MM. Accuracy of magnetic resonance imaging for local staging of prostate cancer: a diagnostic meta-analysis. *Eur Urol*. 2016;70:233–245.
10. Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW, Amin MB. Updates in the eighth edition of the tumor-node-metastasis staging classification for urologic cancers. *Eur Urol*. 2018;73:560–569.
11. Draulans C, Everaerts W, Isebaert S, et al. Impact of magnetic resonance imaging on prostate cancer staging and European Association of urology risk classification. *Urology*. 2019;130:113–119.
12. Marcus DM, Rossi PJ, Nour SG, Jani AB. The impact of multiparametric pelvic magnetic resonance imaging on risk stratification in patients with localized prostate cancer. *Urology*. 2014;84:132– 137.
13. Epstein JI, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol*. 2016;40:244–252.
14. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS prostate imaging - reporting and data system: 2015, version 2. *Eur Urol*. 2016;69:16–40.
15. Fine SW, Amin MB, Berney DM, et al. A contemporary update on pathology reporting for prostate cancer: biopsy and radical prostatectomy specimens. *Eur Urol*. 2012;62:20–39.
16. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med*. 1985; 312:1604–1608.
17. Anderson BB, Oberlin DT, Razmaria AA, et al. Extraprostatic extension is extremely rare for contemporary gleason score 6 prostate cancer. *Eur Urol*. 2017;72:455–460.

18. Gandaglia G, Ploussard G, Valerio M, Mattei A, Fiori C, Roumiguie M, et al. The key combined value of multiparametric magnetic resonance imaging, and magnetic resonance imaging-targeted and concomitant systematic biopsies for the prediction of adverse pathological features in prostate cancer patients undergoing radical prostatectomy. *Eur Urol*. 2019. Online ahead of print.
19. Martini A, Gupta A, Lewis SC, et al. Development and internal validation of a side-specific, multiparametric mag-netic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer. *BJU Int*. 2018;122:1025–1033.
20. Nyarangi-Dix J, Wiesenfarth M, Bonekamp D, et al. Combined clinical parameters and multiparametric magnetic resonance imaging for the prediction of extraprostatic disease-A risk model for patient-tailored risk stratification when planning radical prostatectomy. *Eur Urol Focus*. 2018. Online ahead of print.
21. Ball MW, Partin AW, Epstein JI. Extent of extraprostatic extension independently influences biochemical recurrence-free survival: evidence for further pT3 subclassification. *Urology*. 2015;85:161–164.
22. Jeong BC, Chalfin HJ, Lee SB, et al. The relationship between the extent of extraprostatic extension and survival following radical prostatectomy. *Eur Urol*. 2015;67:342–346.
23. Ploussard G, Beauval JB, Lesourd M, et al. Added value of concomitant systematic biopsies for grade group prediction based on radical prostatectomy final pathology in magnetic resonance imaging positive cases undergoing fusion targeted biopsies. *J Urol*. 2019; 202:1182–1187.

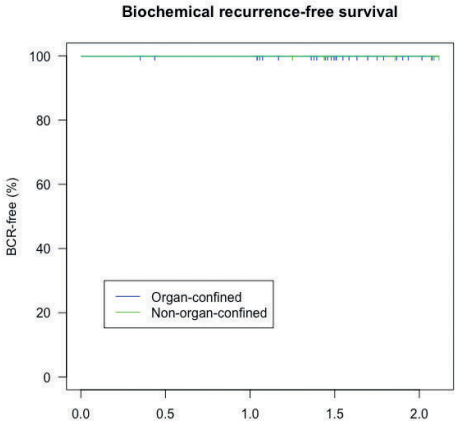
SUPPLEMENTAL SECTION

SUPPLEMENTARY FIGURE S1. Patient inclusion flowchart



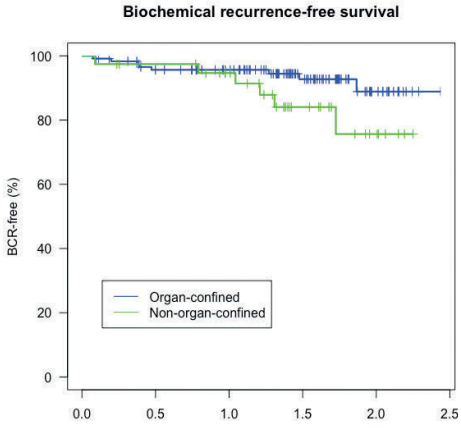
SUPPLEMENTARY FIGURE S2. Impact of mpMRI stage ≥T3a (EPE) on biochemical recurrence-free survival per risk group

A. Low-risk (N = 36)



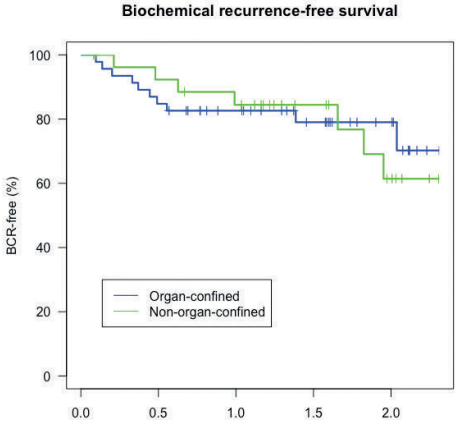
Log-rank test: N/A (no events)

B. Intermediate-risk (N = 158)



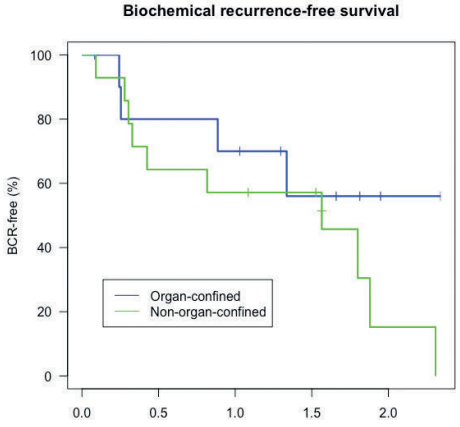
Log-rank test: p = 0.099

C. High-risk (localized) (N = 74)



Log-rank test: p = 0.83

D. High-risk (locally advanced) (N = 25)



Log-rank test: p = 0.23

SUPPLEMENTARY TABLE S1. MRI prostate protocols used in the seven hospitals during the study period

Hospital	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Field of strength of magnet	3T	3T	3T	3T	3T	3T	3T
Coils use	Body coil	Body coil	Body coil	Body coil	Body coil	Body-array coil	Body coil
Slice thickness	3mm	3mm	3mm	3mm	3mm	3mm	3mm
Acquired planes for T2-weighted imaging	Axial, sagittal, coronal	Axial, sagittal, coronal	Axial, sagittal, coronal	Axial, sagittal, coronal	Axial, sagittal, coronal	Axial, sagittal, coronal	Axial, sagittal, coronal
For DWI, the b-values, image sets analyses	B 50-400-800-1400 and ADC map	B 0-50-400-800-2000 and ADC map	B 0-50-100-200-250-800 and ADC map	B 0-50-400-800-1500 and ADC map	B 0-50-1000-1500 and ADC map	B 0-50-400-800-1400 and ADC map	B 0-50-1000-1500 and ADC map
Scoring system used	PI-RADS v2	PI-RADS v2	PI-RADS v2	PI-RADS v2	PI-RADS v2	PI-RADS v2	PI-RADS v2
Experience of radiologists reporting prostate MRI in the period	2->5 years	>5 years	6 years	2->5 years	2-10 years	2->5 years	>5 years
Intravenous administration	Gadolinium 1.0 mg/kg (Dotarem)	Gadolinium 1.0 mg/kg (Dotarem)	Gadolinium 1.0 mg/kg (Dotarem)	No	Gadolinium 1.0 mg/kg (Dotarem)	No	No

SUPPLEMENTARY TABLE S2. Patient characteristics of RARP subpopulation included for analysis

	***N* (%)**
Patients	506
Age (mean, SD)	65.8 (6.1)
PSA (mean, SD)	11.9 (15.9)
Clinical T-stage	
T1c/T2	470 (93)
≥T3	36 (7)
Risk group (DRE)	
Low	63 (12)
Intermediate	252 (50)
High (localized)	132 (26)
High (locally advanced)	59 (12)
MRI T-stage	
Organ-confined	337 (66)
≥T3	169 (34)
Pathological T-stage	
T2	272 (54)
≥T3a	234 (46)

SUPPLEMENTARY TABLE S3. Risk group migration patterns in patients undergoing mpMRI in the RARP subset

		Risk group based on MRI				
		Low ***N* (%)**	**Intermediate** ***N* (%)**	**High (localized)** ***N* (%)**	**High (locally advanced)** ***N* (%)**	**Total** ***N* (%)**
Based on DRE	**Low**	46 (73)	0 (0)	6 (10)	11 (17)	63 (13)
	Intermediate	0 (0)	144 (57)	33 (13)	75 (30)	252 (50)
	High (localized)	0 (0)	3 (2)	83 (63)	46 (35)	132 (26)
	High (locally advanced)	0 (0)	3 (5)	5 (9)	51 (86)	59 (12)
		46 (9)	150 (30)	127 (25)	183 (36)	506 (100)

SUPPLEMENTARY TABLE S4. Baseline characteristics of patients that did not undergo mpMRI

	***N* (%)**
Patients	646
Age (median, IQR)	73 (67 - 78)
PSA (median, IQR)	10.5 (6.6 – 25.0)
Clinical T-stage	
T1	256 (40)
T2a	177 (27)
T2b	38 (6)
T2c	41 (6)
T3/T4	128 (20)
Unknown	6 (1)
Highest biopsy ISUP grade	
1	206 (32)
2	159 (25)
3	87 (13)
4	161 (25)
5	25 (4)
Unknown	8 (1)
Risk group (DRE)	
Low	137 (21)
Intermediate	191 (30)
High (localized)	148 (23)
High (locally advanced)	151 (23)
Unknown	19 (3)
Clinical N stage	
N0	222 (34)
N1	57 (9)
NX	365 (57)
Unknown	2 (0)

SUPPLEMENTARY TABLE S5. Diagnostic accuracy for the detection of non-organ-confined disease of pre- and post-biopsy mpMRI

mpMRI timepoint	Sensitivity (%)	Specificity (%)	PPV (%)	*NPV* (%)	Accuracy (95% CI)
Pre-biopsy	63/133 (47)	140/210 (85)	63/87 (72)	140/210 (67)	0.66
Post-biopsy	56/82 (55)	82/127 (76)	56/82 (68)	82/127 (65)	0.66

SUPPLEMENTARY TABLE S6. Concordance of risk group classification with final surgical pathology

Risk group based on surgical pathology (*N* = 426)		Based on DRE *N* (%)	Based on mpMRI *N* (%)
Low-risk (*N* = 2)	Low-risk	1 (50)	1 (50)
	Intermediate	1 (50)	1 (50)
	Localized high-risk	0	0
	Locally advanced high-risk	0	0
Intermediate-risk (*N* = 31)	Low-risk	4 (13)	3 (10)
	Intermediate	24 (77)	18 (58)
	Localized high-risk	3 (10)	6 (19)
	Locally advanced high-risk	0 (0)	4 (13)
Localized high-risk (*N* = 151)	Low-risk	30 (20)	21 (14)
	Intermediate	80 (53)	55 (36)
	Localized high-risk	33 (22)	49 (33)
	Locally advanced high-risk	8 (5)	26 (17)
Locally advanced high-risk (*N* = 242)	Low-risk	13 (5)	10 (4)
	Intermediate	102 (42)	45 (19)
	Localized high-risk	80 (33)	54 (22)
	Locally advanced high-risk	47 (19)	133 (55)

*Total population of patients undergoing RARP analyzed in this study included 506 patients, however, in respectively 80 patients no pT2 sub classification (T2a, T2b or T2c) was reported.

PART III: ASSOCIATION BETWEEN NERVE SPARING AND POSITIVE SURGICAL MARGINS

CHAPTER 5

5

Nerve Sparing during Robot-Assisted Radical Prostatectomy Increases the Risk of Ipsilateral Positive Surgical Margins

T.F.W. Soeterik
H.H.E. van Melick
L.M. Dijksman
S. Stomps
J.A. Witjes
J.P.A. van Basten

THE JOURNAL OF UROLOGY. 2020;204:91-95

ABSTRACT

Background: Available published studies evaluating the association between nerve sparing robot-assisted radical prostatectomy (RARP) and risk of ipsilateral positive surgical margins were subject to selection bias. In this study we aim to overcome these limitations by using multivariable regression analysis.

Methods: Patients undergoing RARP for prostate cancer at 4 institutions from 2013 to 2018 were included in the study. A multilevel logistic random intercept model, including covariates on patient level and side-specific factors on prostate lobe level, was used to evaluate the association between nerve sparing and risk of ipsilateral positive margins.

Results: A total of 5148 prostate lobes of 2574 patients who underwent RARP were analysed. Multivariable analysis showed nerve sparing was an independent predictor for ipsilateral positive margins (OR 1.42, 95% CI 1.14 - 1.82). Other significant predictors for positive margins were prostate specific antigen density (OR 3.64, 95% CI 2.36 - 5.90) and side-specific covariates including highest preoperative biopsy International Society of Urological Pathology (ISUP) grade (OR 1.58, 95% CI 1.13 - 2.53; OR 1.62, 95% CI 1.13 - 2.69; OR 2.11, 95% CI 1.39 - 3.59 and OR 4.43, 95% CI 3.17 - 10.12 for ISUP grade 2, 3, 4 and 5, respectively), presence of extraprostatic extension on magnetic resonance imaging (OR 1.42, 95% CI 1.03 - 1.91) and percentage of positive cores on systematic biopsy (OR 3.82, 95% CI 2.50 - 5.86).

Conclusions: Nerve sparing was associated with an increased risk of ipsilateral positive surgical margins. The increased risk of positive margins should be taken into account when counselling patients who opt for nerve sparing RARP.

BACKGROUND

Radical prostatectomy (RP) is a treatment modality for localised prostate cancer shown in a prospective randomized trial to significantly increase life expectancy compared with conservative management.[1] Together with radiation therapy, it is one of the most established treatment options for patients with localised prostate cancer and a life expectancy of more than 10 years.[2]

Erectile dysfunction and urinary incontinence are unfortunately common consequences of RP that have a severe impact on quality of life, affecting approximately 80% and 20% of patients, respectively.[3,4] Preservation of the neurovascular bundles can potentially decrease the risk of erectile dysfunction and to a lesser extent urinary incontinence.[5, 6] Because the neurovascular bundles are adjacent to the prostate, it is highly possible that nerve sparing increases the risk of positive surgical margins.

Although the European Association of Urology guidelines state that nerve bundle preservation is contraindicated in case of tumours with a high risk of extracapsular disease, it is assumed that it can be performed safely in most men with localised disease.[2] A systematic review and meta-analysis confirmed the safety of nerve sparing in patients with localised prostate cancer, as it was not associated with an increased risk of positive margins among patients with pT2 tumours (RR 0.92, 95% CI 0.72 - 1.13). Remarkably, in patients with pT3 disease nerve sparing was even associated with a decreased risk of positive margins (RR 0.83, 95% CI 0.71 - 0.96).[7] However, these results should be interpreted with caution as the observational studies previously performed on this subject were susceptible to selection bias.

The available published studies that support current guidelines may have insufficiently accounted for case-mix differences due to patient selection. Therefore, these confounders may have consequently masked the actual association between nerve sparing and the risk of positive margins. Since a positive margin is associated with a higher risk of biochemical recurrence and even cancer specific mortality,[8,9] proper surgical planning for nerve sparing with a minimum risk of positive margins should be undertaken.

Given the importance of the issue and the limitations of the previous research, there is a remaining need for studies of higher methodological quality on this subject. Obviously, a randomised controlled trial would be the most methodologically sound approach. However, randomising patients for nerve sparing and non-nerve sparing surgery would not likely be done based on ethical grounds as assignment to non-nerve sparing could be regarded as unnecessarily harmful for patients randomised into the non-nerve sparing arm. As there are several studies reporting the benefits of nerve sparing during RP, these patients will not have the opportunity to retain erectile function.[10] Therefore, the aim of this study is to evaluate the association between nerve sparing RP and the risk of a positive surgical margin by retrospectively analysing a large multicentre patient population, adjusting for a large number of patient-related and prostate side-specific covariates using multivariable regression analysis.

METHODS

Patients and data collection

Patients diagnosed with prostate cancer undergoing robot-assisted radical prostatectomy (RARP) as primary treatment at 4 Dutch teaching hospitals (Martini Hospital Groningen, Hospital Group Twente, St. Antonius Hospital Nieuwegein/Utrecht and the Canisius Wilhelmina Hospital Nijmegen) from 2013 to 2018 were included in the study (IRB No. Z18.023). Data were captured in a prospective manner. Patients were excluded from analysis if they underwent salvage RARP or were treated with up-front androgen deprivation therapy. Baseline characteristics (age, clinical T-stage based on DRE, radiological T-stage based on multiparametric magnetic resonance imaging (mpMRI), preoperative serum PSA, total biopsy cores taken and number of positive cores at diagnosis, biopsy International Society of Urological Pathology (ISUP) grade, prostate volume measured using transrectal ultrasonography (TRUS) or mpMRI, treatment information (date of surgery, surgeon, nerve sparing as mentioned in the surgical report) and definitive pathology data (pathological T-stage, Gleason score, margin status) were documented. In addition, prostate side-specific radiological, surgical and pathological data were retrospectively collected.

Predictor and outcome definitions

The most recent preoperative PSA and prostate volume measured by TRUS or MRI were used to calculate PSAD (serum PSA [ng/ml] divided by prostate volume [ml]). DRE was subdivided into the 3 stages of T1 (benign), T2 (nodule) or T3 (EPE). All radical prostatectomies performed in the study period were robot-assisted. Interfascial nerve sparing was performed using an antegrade approach. After the upward traction of the vas deferens and seminal vesicles, the prostatic pedicle was observed and controlled athermally at the base of the prostate. Then the prostate was pulled to the opposite side and the lateral pelvic fascia was exposed. The triangular space between the lateral pelvic fascia, Denonvilliers' fascia and the prostate was observed and the neurovascular bundle was defined. Subsequently, the lateral pelvic fascia was exposed and the interfascial dissection was performed. The non-nerve sparing technique included dissection posterior to Denonvilliers' fascia and incision on to the perirectal fat lateral to the neurovascular bundles. RARP was performed by 14 surgeons. Surgical experience per surgeon varied from 0 procedures (least experienced) to 500 procedures (most experienced) at the beginning of the study period. For the analysis the most experienced surgeon was used as the reference category. Prostatic carcinoma was graded using the 2014 ISUP grading system.[11] A positive surgical margin, assessed by dedicated uropathologists, was defined as tumour cells present at the inked margin.[12]

Statistical analysis

Each prostate lobe was considered as a separate case. A multilevel regression model was used to evaluate the association between nerve sparing and positive surgical margins. Side-specific factors included nerve sparing, DRE, mpMRI local stage (organ confined vs EPE), highest ISUP grade found on biopsy and percentage of positive cores. Covariates available on patient level included PSAD, surgeon, hospital and age. To adjust for the consequential data clustering on patient level a random intercept was included in the model. Missing data were assumed to be missing at random, based on the missing data patterns, and were imputed using multiple imputations.[13] Analysis was performed using R Studio.[14]

RESULTS

Study population

A total number of 2574 patients underwent RARP from 2013 to 2018 at the 4 hospitals. The baseline characteristics and surgical outcomes on prostate lobe level are presented in Table 1. Patient-level baseline characteristics and surgical outcomes are presented in Supplementary Table S1 and S2. Positive surgical margin rates were observed in 844 (33%) of 2574 cases. The positive margin rate was 23% in pT2 (353 of 1533) and 47% in pT3 or greater (491 of 1041) tumours. A total of 1774 (69%) patients underwent interfascial nerve sparing surgery (unilateral or bilateral).

Nerve sparing vs Non-nerve sparing

Nerve sparing status was not available in 97 patients and, thus, these patients could not be categorized. Baseline characteristics of 4954 prostate lobes of the remaining 2477 patients with known nerve sparing status are presented in Table 1. Overall, the nerve sparing group had relatively more favourable tumour characteristics compared with the non-nerve sparing group (Table 1).

Missing data

Of the 2574 patients who underwent RARP during the study period, data relevant for analysis were missing in 889. This was mainly attributable to the fact that 364 patients (14%) did not undergo preoperative mpMRI and 263 (10%) underwent targeted biopsies without systematic biopsies. Thus, radiological T-stage and prostate side-specific percentage of positive cores were not available in these cases. In addition, preoperative prostate volume was not determined in 86 (3%) cases. Extensive information regarding missing data is given in a patient flow chart (Supplementary Figure).

TABLE 1. Baseline characteristics of 4954 prostate lobes of 2477 patients

	Nerve sparing (no. of lobes [%], mean [SD])	**Non-nerve sparing (no. of lobes [%], mean [SD])**
No. of prostate lobes	2711	2243
Age	63.9 (6.0)	66.4 (5.4)
PSA (ng/ml)	9.4 (7.3)	11.9 (12.2)
PSAD (ng/ml/ml)	0.23 (0.19)	0.29 (0.27)
DRE[a]		
T1	2356 (87)	1465 (65)
T2	273 (10)	572 (26)
T3	10 (0)	126 (6)
Unknown	72 (3)	80 (4)
EPE on MRI[a]		
EPE absent	2182 (80)	1431 (64)
EPE present	108 (4)	451 (20)
Unknown	421 (16)	361 (16)
ISUP biopsy grade[a]		
Benign	829 (31)	264 (12)
1	1009 (37)	570 (25)
2	444 (16)	618 (28)
3	127 (5)	291 (13)
4	52 (2)	248 (11)
5	12 (0)	161 (7)
Unknown	238 (9)	91 (4)
Percentage positive cores[a]	0.29 (0.31)	0.54 (0.35)
Surgical margin status[a]		
Negative	2201 (81)	1716 (77)
Positive	484 (18)	505 (23)
Unknown	26 (1)	22 (1)

[a]Side-specific covariates on prostate lobe level.

Abbreviations: PSA = prostate-specific antigen, SD = standard deviation, MHG = Martini Hospital Groningen, PSAD = prostate-specific antigen density, DRE = digital rectal examination, MRI = magnetic resonance imaging, TRUS = transrectal ultrasonography, ISUP = International Society of Urological Pathology.

Evaluation of Predictors for Positive Surgical Margins

The results of the multivariable analysis predicting positive margins are presented in Table 2. Model 1 included solely complete cases. Additional analysis was done after accounting for missing data using multiple imputation (model 2). Overall, model 2 resulted in more precise estimation of coefficients, with narrower 95% confidence intervals compared with the complete case analysis (model 1). In both models, nerve sparing was associated with significantly higher odds of ipsilateral positive margins. Other covariates found to be significant predictors for positive margins in models 1 and 2 were PSAD, highest ipsilateral biopsy ISUP grade 2 and higher, percentage of positive cores on systematic biopsy and presence of EPE on preoperative mpMRI (Table 2).

TABLE 2. Multivariable logistic regression analysis predicting positive surgical margins

	Model 1[a] (*N* = 3325) OR (95% CI)	***p***	**Model 2[b] (*N* = 5148) OR (95% CI)**	***p***
Age	0.98 (0.96 – 1.0)	0.038	0.98 (0.97 – 1.00)	0.105
PSAD	2.72 (1.57 – 4.72)	<0.001	3.64 (2.36 – 5.90)	<0.001
ISUP grade				
Benign	Referent		Referent	
1	1.15 (0.79 – 1.68)	0.4675	1.24 (0.93 – 1.81)	0.204
2	1.48 (0.97 – 2.27)	0.0687	1.58 (1.13 – 2.53)	0.015
3	1.65 (0.99 – 2.73)	0.053	1.62 (1.13 – 2.69)	0.037
4	2.09 (1.20 – 3.66)	0.0097	2.11 (1.39 – 3.59)	0.002
5	5.56 (2.90 – 10.63)	<0.001	4.43 (3.17 – 10.12)	<0.001
DRE				
T1	Referent		Referent	
T2	1.33 (0.99 – 1.79)	0.0618	1.21 (0.93 – 1.62)	0.173
T3	1.45 (0.81 – 2.60)	0.2143	1.66 (0.93 – 3.01)	0.075
MRI stage				
Organ-confined	Referent		Referent	
EPE	1.48 (1.05 – 2.07)	0.024	1.42 (1.03 – 1.91)	0.031
% Positive cores	3.50 (2.23 – 5.49)	<0.001	3.82 (2.50 – 5.86)	<0.001
Nerve sparing				
Non-nerve sparing	Referent		Referent	
Nerve sparing	1.53 (1.15 – 2.03)	<0.001	1.42 (1.14 – 1.82)	0.005

[a]Model 1: complete case analysis.

[b]Model 2: imputed case analysis using multiple imputations.

*The analysis also included the covariates: hospital (N=4) and surgeon (N=14), ORs are not shown. Abbreviations: PSAD = prostate-specific antigen density, DRE = digital rectal examination, EPE = extraprostatic extension.

DISCUSSION

In this study, we explored the association between side-specific nerve sparing RP and the risk of ipsilateral positive margins using a large, multi-institutional, real-world patient cohort. On multivariable logistic regression analysis nerve sparing was associated with significantly higher odds of positive margins compared with non-nerve sparing (OR 1.42, 95% CI 1.14 – 1.82). Our study results call into question the classic assumption that nerve sparing is not associated with an increased risk of positive surgical margins.

Our main findings are relevant for clinical practice as patients and their urologists need to be aware of the fact that nerve sparing does increase the risk of positive margins. This effect was masked in previous studies, apparently due to methodological limitations and insufficient unadjusted residual confounding by indication. Also, as nerve sparing does not guarantee preservation of erectile function,[10] patients unlikely to benefit from nerve sparing should not be unnecessarily exposed to its risks.

Several studies on this topic have been performed previously, reporting conflicting results. Coelho et al reported comparable positive margin rates of 876 patients regardless of nerve sparing type.[15] For bilateral, unilateral and non-nerve sparing, respectively, the positive margin rates in pT2 tumours were 8.2%, 6.1% and 8.5% (p = 0.93) and 27.7%, 26.7% and 30.8% (p = 0.93) in pT3. Comparable findings were reported in a study by Moore et al, including 945 patients.[16] The authors reported no significant differences in positive margin rates between nerve sparing groups on multivariable analysis adjusting for age, PSA, Gleason score, percentage of positive biopsy cores and clinical stage. The reported relative risks were 0.58 (95% CI 0.30 - 1.40, p = 0.11) for unilateral nerve sparing and 0.64 (95% CI 0.35 - 1.17) for bilateral nerve sparing. Choi et al evaluated functional outcomes and positive margin rates in their series of 602 consecutive RARPs.[17] Nerve sparing improved 24-month urinary control without an increase in positive margin rates compared to non-nerve sparing RARP. Lastly, a study on the SEARCH (Shared Equal Access Regional Cancer Hospital) database including 1018 cases echoed the previously stated findings, and reported that neither bilateral nor unilateral nerve sparing techniques were associated with a higher risk of a positive margin.[18]

Our findings are inconsistent with those reported in previous studies, for which we have two possible explanations. The potential confounders adjusted for during analysis in previous studies were prostate-specific and not prostate side-specific. To determine causality between a nerve sparing approach and ipsilateral positive margins, each prostate lobe should be considered as a separate case. For example, it is likely that among patients in whom unilateral nerve sparing was performed the ipsilateral side had favourable tumour characteristics compared to the contralateral side. Disregarding the side-specific factors in the analysis limits the ability to evaluate the causality between nerve sparing and an ipsilateral positive surgical margin and, therefore, the effects of side-specific covariates remain masked. The second reason regards the type and number of covariates for which was adjusted during multivariable analysis in previous studies. In this study, the large sample size and side-specific nature of the majority of covariates

enabled inclusion of a large number of potential confounders in the multivariable analysis, including the influence of the individual surgeon (and, thus, experience) on the occurrence of positive margins. To our knowledge, none of the previous studies performed an analysis including all of the most important potential predictors, including MRI stage, for positive margins.

In two previously performed studies on this subject comparable conclusions were reported.[19,20] Zorn et al reported significantly higher posterolateral positive margin rates among patients with pT3 tumours who underwent interfascial nerve preservation compared to patients with pT3 tumours undergoing non-nerve sparing RARP (73% vs 33%, $p = 0.05$).[19] Fact this study had comparable results may be explained by the methodological approach, as their analysis was also done on the prostate lobe-level. In addition, the nerve sparing technique performed was comparable to ours, as interfascial nerve sparing was performed.[19] Liss et al also reported nerve sparing to be associated with an increased risk of positive margins on multivariable analysis (OR 5.58, 95% CI 1.176 - 26.46).[20] However, the calculated ORs (and large corresponding 95% CIs) on multivariable analysis should be interpreted with caution as the number of events was relatively low (21) for the total number of covariates included (6).[20]

The positive margin rates, especially those observed for pT2 tumours (23%), were relatively high compared to those reported in other series. In a recent meta-analysis by Nguyen et al an absolute risk of positive margins of 8.1% for any nerve sparing and 7.7% for non-nerve sparing was reported.[7] The higher rates of positive margins observed in this study may be explained by the selection of higher-risk patients for surgery, with relatively higher biopsy Gleason scores (65% Gleason 7 or higher) and relatively high pT3 rates (40%) compared to those reported in other series (42% Gleason 7 or higher and 19% pT3).[10] Surgeon experience was previously reported to be associated with positive margins after RARP, and could also explain the higher positive margin rates in our cohort.[21] Of all surgeons performing RARP in this study a large proportion were novice, with 8 of 14 (57%) having performed fewer than 50 RARPs.

Our study has a number of strengths, as it is a multicentre study with a large sample size, enabling inclusion of a relatively large number of covariates into the multivariable logistic regression model. However, some potential limitations must be acknowledged. Our study lacks central review regarding histopathological findings on prostate biopsy and final pathology after RARP. However, we assume this has no large impact on our main findings as positive surgical margin interpretation by uropathologists generally shows a high degree of interobserver agreement.[22] In addition, data regarding the degree of interfascial nerve sparing were lacking in the surgery reports, which could have led to measurement bias. Finally, inclusion of the location of positive surgical margins was outside the scope of the present study. Evaluation of the specific locations of the positive margins should be the subject of future research as the association between location and nerve sparing remains poorly understood.

CONCLUSIONS

Preservation of the neurovascular bundles during robot-assisted radical prostatectomy is associated with an increased risk of ipsilateral positive surgical margins when adjusting for patient and side-specific covariates on multivariate analysis. The increased risk of ipsilateral positive margins should be taken into account when counselling patients who opt for nerve sparing RARP.

REFERENCES

1. Bill-Axelson A, Holmberg L, Garmo H et al. Radical prostatectomy or watchful waiting in prostate cancer - 29-year follow-up. *N Engl J Med.* 2018;379:2319.
2. Mottet N, Bellmunt J, Bolla M et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol.* 2017;71:618.
3. Johansson E, Steineck G, Holmberg L et al. Long-term quality-of-life outcomes after radical prostatectomy or watchful waiting: the Scandinavian prostate cancer group-4 randomised trial. *Lancet Oncol.* 2011;12:891.
4. Donovan JL, Hamdy FC, Lane JA et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. *N Engl J Med.* 2016375:1425.
5. Walsh PC, Donker PJ: Impotence following radical prostatectomy: insight into etiology and prevention. *J Urol.* 2017;197:S165.
6. Suardi N, Moschini M, Gallina A et al. Nerve sparing approach during radical prostatectomy is strongly associated with the rate of postoperative urinary continence recovery. *BJU Int.* 2013;111:717.
7. Nguyen LN, Head L, Witiuk K et al. The risks and benefits of cavernous neurovascular bundle sparing during radical prostatectomy: a systematic review and meta-analysis. *J Urol.* 2017;198:760.
8. Chalfin HJ, Dinizo M, Trock BJ et al. Impact of surgical margin status on prostate-cancer-specific mortality. *BJU Int.* 2012;110:1684.
9. Preisser F, Mazzone E, Knipper S et al. Rates of positive surgical margins and their effect on cancer-specific mortality at radical prostatectomy for patients with clinically localized prostate cancer. *Clin Genitourin Cancer.* 2019;17:e130.
10. Ficarra V, Novara G, Ahlering TE et al. Systematic review and meta-analysis of studies reporting potency rates after robot-assisted radical prostatectomy. *Eur Urol.* 2012;62:418.
11. Epstein JI, Egevad L, Amin MB et al. The 2014 International Society of Urological Pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol.* 2016;40:244.
12. Fine SW, Amin MB, Berney DM et al. A contemporary update on pathology reporting for prostate cancer: biopsy and radical prostatectomy specimens. *Eur Urol.* 2012;62:20.
13. Van Buuren S, Groothuis-Oudshoorn K. Multivariate imputation by chained equations in R. *J Stat Soft.* 2011; 45:1-67.
14. RStudio Team (2015). RStudio: Integrated development for R. RStudio, Inc., Boston, MA URL: http://www.rstudio.com/
15. Coelho RF, Chauhan S, Orvieto MA et al. Predictive factors for positive surgical margins and their locations after robot-assisted laparoscopic radical prostatectomy. *Eur Urol.* 2010;57:1022.
16. Moore BM, Savdie R, PeBenito RA et al. The impact of nerve sparing on incidence and location of positive surgical margins in radical prostatectomy. *BJU Int.* 2012;109:533.
17. Choi WW, Freire MP, Soukup JR et al. Nerve sparing technique and urinary control after robot-assisted laparoscopic prostatectomy. *World J Urol.* 2011;29:21.
18. Nelles JL, Freedland SJ, Presti JC,Jr et al. Impact of nerve sparing on surgical margins and biochemical recurrence: results from the SEARCH database. *Prostate Cancer Prostatic Dis.* 2009;12:172.
19. Zorn KC, Gofrit ON, Orvieto MA et al. Robotic-assisted laparoscopic prostatectomy: functional and pathologic outcomes with interfascial nerve preservation. *Eur Urol.* 2007;51:755.
20. Liss M, Osann K, Ornstein D. Positive surgical margins during robotic radical prostatectomy: a contemporary analysis of risk factors. *BJU Int.* 2008;102:603.

21. Bravi CA, Tin A, Vertosick E et al. The impact of experience on the risk of surgical margins and biochemical recurrence after robot-assisted radical prostatectomy: a learning curve study. *J Urol.* 2019;202:108.

22. Evans AJ, Henry PC, Van der Kwast TH et al. Interobserver variability between expert urologic pathologists for extraprostatic extension and surgical margin status in radical prostatectomy specimens. *Am J Surg Pathol.* 2008;32:1503.

SUPPLEMENTAL SECTION

SUPPLEMENTARY TABLE S1. Baseline characteristics on patient level

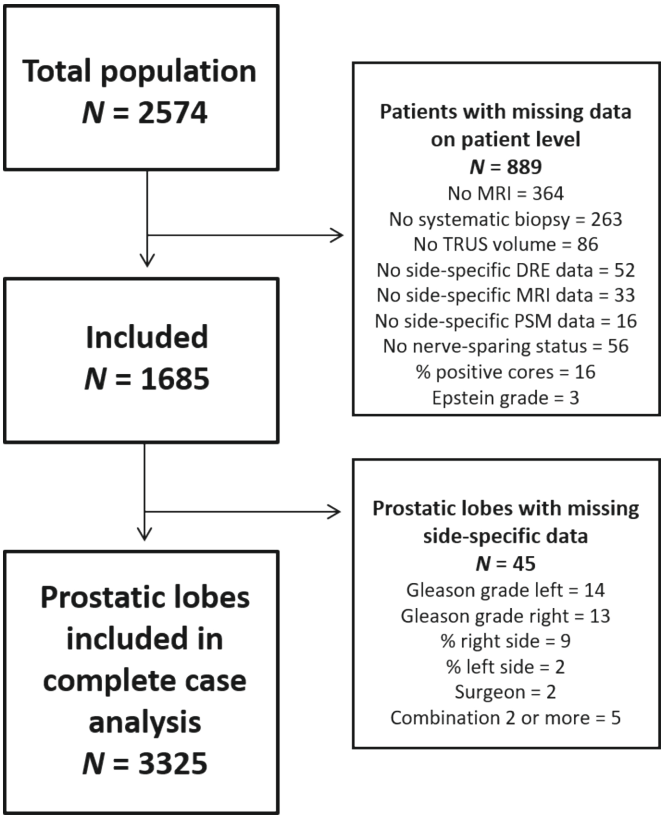
	***N* (%), mean (SD)**
No. of patients	2574
Age	65.0 (5.9)
Preoperative PSA (ng/ml)	10.6 (10)
PSA density (ng/ml/ml)	0.26 (0.19)
Hospital	
MHG	261 (10)
HGT	515 (20)
SAH	746 (29)
CWH	1052 (41)
Clinical T stage	
T1a/b/c	1533 (60)
T2/T2a	686 (27)
T2b	133 (5)
T2c	66 (3)
T3	134 (5)
Unknown	22 (1)
Preoperative MRI	
Yes	2210 (86)
No	364 (14)
Radiological T stage (N=2210)	
T0	212 (10)
T2a	1017 (46)
T2b	102 (5)
T2c	278 (13)
T3a	466 (21)
T3b	96 (4)
T4	5 (0)
Unknown	34 (2)
Biopsy type	
TRUS-guided systematic	1799 (70)
MRI-guided	265 (10)
TRUS + MRI-guided	496 (19)
Incidental	9 (0)
Unknown	5 (0)
Highest biopsy ISUP Grade	
1	887 (35)
2	875 (34)
3	370 (14)
4	276 (11)
5	148 (6)
Unknown	18 (0)

SD = standard deviation, PSA = prostate-specific antigen, MRI = magnetic resonance imaging, TRUS = transrectal ultrasonography, MHG = Martini Hospital Groningen, HGT = Hospital Group Twente, SAH = St. Antonius Hospital Nieuwegein, CWH = Canisius Wilhelmina Hospital Nijmegen.

SUPPLEMENTARY TABLE S2. Surgical outcomes on patient level

	***N* (%)**
No. of cases	2574
Year of surgery	
2013	307 (12)
2014	271 (11)
2015	357 (14)
2016	492 (19)
2017	539 (21)
2018	608 (24)
Nerve sparing	
Non-nerve sparing	703 (27)
Unilateral	837 (33)
Bilateral	937 (36)
Unknown	97 (4)
Pathological stage	
pT2	96 (4)
pT2a	234 (9)
pT2b	55 (2)
pT2c	1148 (45)
pT3a	751 (29)
pT3b	283 (11)
pT4	7 (0)
Gleason sum score	
≤6	486 (19)
7	1732 (67)
8	171 (7)
9	179 (7)
10	6 (0)
Surgical margin status	
pT2	
Negative	1177 (77)
Positive	353 (23)
Unknown	3 (0)
pT3	
Negative	550 (53)
Positive	491 (47)
Unknown	0 (0)

SUPPLEMENTARY FIGURE. Missing data flowchart



PART IV: INCORPORATION OF MRI INTO CLINICAL PREDICTION MODELS

CHAPTER 6

6

External Validation of the Martini Nomogram for Prediction of Side-Specific Extraprostatic Extension of Prostate Cancer in Patients Undergoing Robot-Assisted Radical Prostatectomy

T.F.W. Soeterik
H.H.E. van Melick
L.M. Dijksman
H.V.N. Küsters-Vandevelde
D.H. Biesma
J.A. Witjes
J.P.A. van Basten

UROLOGIC ONCOLOGY: SEMINARS AND ORIGINAL INVESTIGATIONS. 2020;38:372-378

ABSTRACT

Background: To establish oncological safe nerve sparing robot-assisted radical prostatectomy (RARP), accurate assessment of extraprostatic extension (EPE) is critical. A recently developed nomogram including magnetic resonance imaging parameters accurately predicted side-specific EPE in the development cohort. The aim of this study is to assess this model's performance in an external patient population.

Methods: Model fit was assessed in a cohort of 550 patients who underwent robot-assisted radical prostatectomy in 2014 to 2017 for prostate cancer. Model calibration was evaluated using calibration slopes. Discriminative ability was quantified using the area under the receiver operating characteristic curve. Model updating was done by adjusting the linear predictor to minimize differences in predicted and observed risk for EPE.

Results: A total of 792 prostate lobes were included for model validation. Discriminative ability expressed in terms of area under the receiver operating characteristic curve was 0.78, 95% confidence interval 0.75 - 0.82. Graphical evaluation of the calibration showed poor fit with high disagreement between predicted and observed probabilities of EPE in the population. Model updating resulted in excellent agreement between mean predicted and observed probabilities. However, calibration plots showed substantial miscalibration; including both under- and overestimation.

Conclusions: External validation of the novel nomogram for the prediction of side-specific EPE developed by Martini and co-workers showed good discriminative ability but poor calibration. After updating, substantial miscalibration was still present. Use of this nomogram for individualized risk predictions is therefore not recommended.

BACKGROUND

In patients with prostate cancer who undergo robot-assisted radical prostatectomy (RARP), the trade-off between optimal oncological results versus an attempt to retain erectile function using nerve sparing techniques is an important feature in medical decision-making.[1] The decision to use a nerve sparing approach is particularly based on the probability of present extraprostatic extension (EPE), as this increases the risk of a positive surgical margin during nerve sparing on this ipsilateral side. To exclude the presence of EPE and to ensure the oncological safety of a nerve sparing approach, multiparametric magnetic resonance imaging (mpMRI) of the prostate for local staging is recommended by the European Association of Urology Guidelines, especially in high-risk prostate cancer.[1] However, the sensitivity of mpMRI for detection of EPE is low, which potentially makes a nerve sparing procedure, planned using solely mpMRI information, oncologically unsafe.[2]

To improve preoperative estimation of the presence or absence of side-specific EPE and thus the safety of an ipsilateral nerve sparing procedure, several nomograms have been developed.[3–5] The predictors included per model vary, but generally serum prostate-specific antigen (PSA), clinical tumour stage assessed by digital rectal examination, highest International Society of Urological Pathology (ISUP) grade and other biopsy characteristics (e.g., highest percent tumour involvement, number of positive cores) are included. In addition, these models include specific biopsy features of transrectal ultrasonography-guided systematic biopsies. However, due to current developments and recommendations regarding the use of prebiopsy MRI with subsequent target biopsies, these nomograms are becoming less applicable in modern urological practice.[6–9]

In a recent study of Martini et al, a nomogram with a potentially higher applicability in daily practice was developed.[10] Besides mpMRI information, the model includes oncological parameters such as serum PSA level, highest ISUP, and highest percent tumour involvement of the corresponding biopsy core. Due to these features, the model can be applied in urological practices that use either transrectal ultrasonography- guided systematic, solely MRI-guided target biopsies or the combination of both techniques. The authors observed that inclusion of these parameters into a prediction model, combined with mpMRI results, improves EPE prediction.[10] Reported discriminative ability after internal validation was good, with an area under the receiver operating characteristic curve (AUC) of respectively 82.11% (95% confidence interval [CI] 78.49 - 85.73). Also, calibration was reported to be good given the excellent concordance of predicted probabilities and observed prevalence of EPE.

To determine the performance of the Martini model in another population cohort, and thus its generalisability, we performed an external validation study using data of a large cohort of Dutch prostate cancer patients that underwent RARP.

METHODS

Study population and setting

The validation cohort was derived from the population of 625 prostate cancer patients that underwent RARP in 2014 to 2017 at the Canisius Wilhelmina Hospital in Nijmegen, The Netherlands. Data from the prospectively collected local RARP database including preoperative characteristics such as age, initial serum PSA level, highest ISUP grade on biopsy, and postoperative characteristics such as pathologic tumour stage, ISUP grade, and margin status were used. Additional variables specifically needed for the validation (e.g. side of EPE on MRI, side of EPE during pathological evaluation) were collected retrospectively.

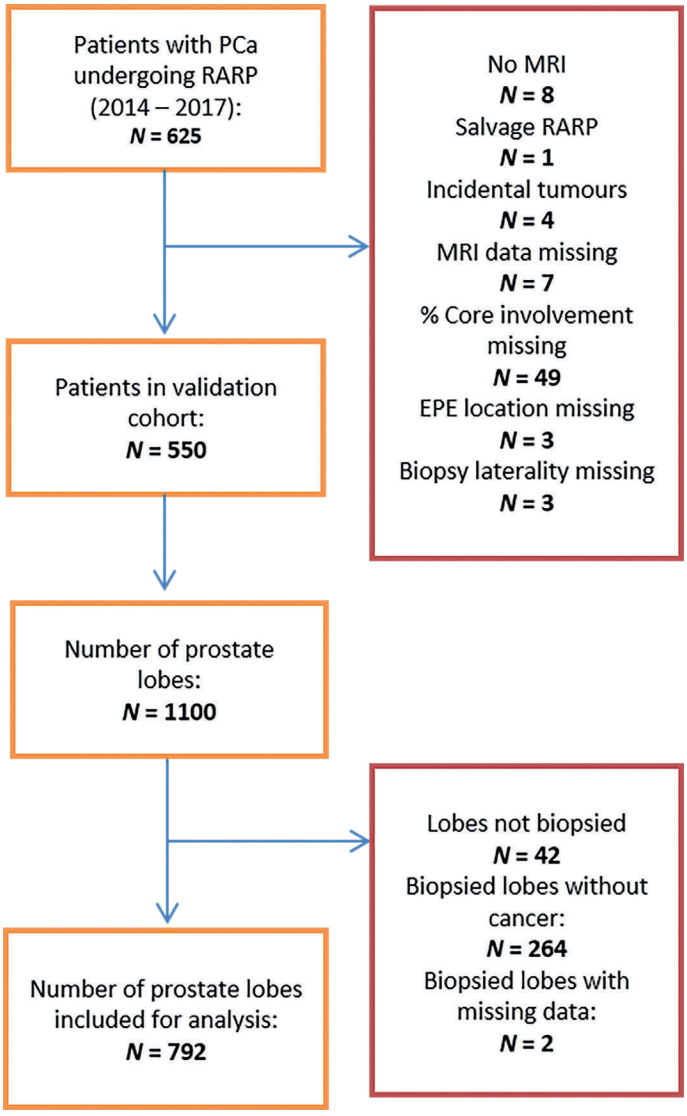
Radiological reporting was done according to the European Society of Uro-Radiology (ESUR) guidelines and from 2015 PI-RADS v2.[11,12] Presence of EPE on MRI was defined as loss of prostate capsule and irregularity of the capsule. Presence of reported "uncertain EPE"; e.g. presence of solely capsule abutment or broad capsular contact, were regarded as EPE-negative. These definitions were consistent with those used for model development.[10]

The cohort consisted of patients diagnosed locally and referred patients (tertiary centre and other regional [teaching] hospitals), reflecting a real-world population. Patients with complete cases were included for analysis. Derivation of the study sample from the overall population was further clarified using a patient inclusion flow chart (Figure 1).

Assessment of present EPE

All radical prostatectomy specimens were formaldehyde fixed and completely imbedded in paraffin. All specimens were standardized inked in 3 ink colors for optimal orientation and cut in a standardized way. From the apical region and bladder neck region a 5 mm thick section was taken and sectioned perpendicular to the inked surface and totally embedded. EPE was defined as tumour that bulges beyond the prostatic contour, tumour that is admixed with periprostatic fat tissue or in the posterolateral area, as tumour within connective tissue or between nerve bundles of the neurovascular bundle. In the bladder neck region, the presence of tumour between thick smooth muscle bundles in the absence of benign prostatic glands was considered as bladder neck invasion.[13] In case information on EPE was missing, slides were revised by a single uropathologist (H.K.V.).

FIGURE 1. Patient inclusion flowchart



Model performance, recalibration and clinical usefulness

Model calibration, which refers to the agreement between observed endpoints and predictions, was assessed using calibration-in-the-large and the Hosmer-Lemeshow goodness-of-fit test. Calibration was further evaluated in a graphical matter using calibration plots, wherein the agreements between predicted probabilities and observed outcomes in the dataset were visualised. Discrimination, which refers to the ability of the model to distinguish a case with the endpoint (EPE) from a case without EPE, was quantified using the AUC.[14] In case of poor model fit, the potential need for adjustment of the intercept and/or slope (and if needed, the degree of adjustment necessary) was determined by inserting the linear predictor as the only predictor in the logistic regression formula.[15] To determine clinical usefulness, sensitivity and specificity were determined for different risk thresholds (0.07, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40 and 0.50). We also calculated the net benefit for a range of threshold probabilities, using decision curve analysis. The net benefit was calculated as the proportion of "net" true positives (true positives corrected for the false positives weighted by the odds of the risk cut-off, divided by the sample size).[16,17] Statistical analysis was performed using RStudio Version 1.1.456.

RESULTS

Patient population

As pointed out in Figure 1, 625 patients underwent RARP, and a total of 792 prostate lobes were derived for analysis. EPE was reported on pathological evaluation in 250/792 (32%) lobes, resulting in an adequate number of events for validation (events per variable [EPV] = 41).[18] Baseline characteristics on a patient level are presented in Table 1. Descriptive statistics of the predictors and outcome variable of all prostate lobes included for analysis are reported in Table 2.

TABLE 1. Baseline characteristics of the validation cohort (N = 550)

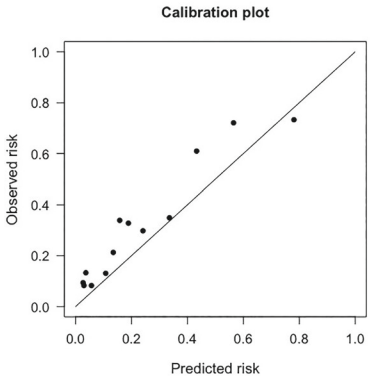
	Validation cohort Mean (SD, min/max) or *N* (%)
No. of patients	550
Age (years)	64.6 (5.6, 42 - 76)
Preoperative PSA (ng/ml)	10.4 (8.6, 0.5 – 77)
Clinical T-stage[a]	
T1c	317 (58)
T2a	107 (18)
T2b	16 (3)
T2c	17 (3)
T2	44 (8)
T3	43 (8)
Missing	6 (1)
MRI T-stage	
T0	49 (9)
T2a	147 (27)
T2b	27 (5)
T2c	74 (14)
T2	39 (7)
T2/T3 (EPE uncertain)	55 (10)
T3a	127 (23)
T3b	26 (5)
T4	4 (1)
Missing	2 (0)
Biopsy type	
TRUS-guided systematic	335 (61)
MRI-guided target	53 (10)
TRUS + MRI-guided	162 (29)
Pathological T-stage	
T2	1 (0)
T2a	62 (11)
T2b	7 (1)
T2c	208 (38)
T3a	202 (37)
T3b	69 (13)
T4	1 (0)
Surgical Margin	
Positive	187 (34)
Negative	363 (66)

[a]Based on digital rectal examination, [b]In case of uncertain EPE. Percentages may not add up to 100% due to rounding.

TABLE 2. Baseline characteristics of included cases (N=792)

	Mean (min/max) or *N* (%)
Preoperative PSA (ng/ml)	9.3 (0.5 – 20)
ISUP biopsy grade	
ISUP 1	271 (34)
ISUP 2	274 (35)
ISUP 3	104 (13)
ISUP 4+5	143 (18)
Percent highest biopsy tumour involvement	
>50%	352 (44)
≤50%	440 (56)
EPE on MRI	
Present	160 (20)
Absent	632 (80)
EPE on pathological evaluation	
Present	250 (32)
Absent	542 (68)

FIGURE 2 AND TABLE 3 - Calibration plot and table of the original Martini Nomogram

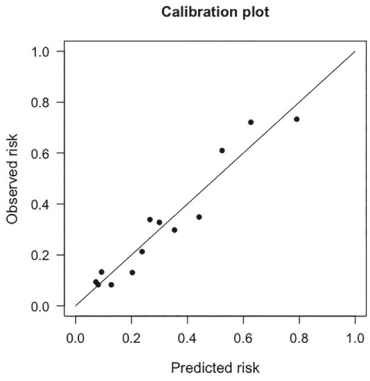


Group	**Predicted risk (%)**	**Patients (N)**	**Mean predicted (%)**	**Mean observed (%)**
1	2 – 3	64	3	9
2	3 – 3	60	3	8
3	3 – 4	60	4	13
4	4 – 9	60	6	8
5	10 – 12	61	11	13
6	12 – 15	61	14	21
7	15 – 17	62	16	34
8	17 – 21	64	19	33
9	21 – 29	57	24	30
10	29 – 39	63	34	35
11	39 – 49	59	43	61
12	49 – 65	61	57	72
13	65 – 88	60	78	73

Model performance

The area under the receiver operating characteristic curve (AUC) was 0.78 (95% CI 0.75 – 0.82), indicating fair discriminative ability of the model in the study population. Calibration-in-the large showed substantial underestimation of mean predicted risk: 24% vs. 32% mean observed risk. The underestimation of present EPE of the model was confirmed in a graphical matter using the calibration curve (Figure 2). As shown in both calibration plot and table, the underestimation of the predicted risk was systematic and was present in 12 of 13 groups (Table 3). The outcome of the Hosmer-Lemeshow test was in line with these findings regarding calibration of the model, as it was statistically significant ($p < 0.001$); indicating poor model fit.

FIGURE 3 AND TABLE 4 - Calibration plot and table of the updated Martini nomogram



Group	Predicted risk (%)	Patients (N)	Mean predicted (%)	Mean observed (%)
1	7 – 8	64	7	9
2	8 – 9	60	8	8
3	9 – 10	60	9	13
4	10 – 19	60	13	8
5	19 – 22	61	20	13
6	22 – 25	61	24	21
7	25 – 28	62	27	34
8	28 – 32	64	30	33
9	32 – 40	57	35	30
10	40 – 49	63	44	35
11	49 – 57	59	52	61
12	57 – 69	61	63	72
13	69 – 87	60	79	73

Recalibration

Insertion of the original linear predictor as the only variable in the original model resulted in the following coefficients; 1.42 for the intercept and 0.8258 for the beta coefficients. Adding 1.42 to the original intercept and multiplying all separate beta coefficients with 0.8258 corrected the original linear predictor, leading to a new model:

Ln (odds EPE) = -2.77 + 0.049*PSA + 1.18*ISUP2 + 1.44*ISUP3 + 2.12*ISUP4 + 0.31*Tumour involvement >50% + 1.21*EPE present on MRI

Recalculation of the AUC using the update formula showed no change in discriminative ability, since the AUC remained 0.78 (95% CI 0.75 - 0.82). As shown in the new calibration plot, agreement between predicted probability and prevalence of the outcome improved (Figure 3). Also, calibration-in-the-large showed similar mean values of predicted and observed probabilities (respectively, 32% and 32%). Although calibration improved graphically after the model adjustment, the Hosmer-Lemeshow test was still statistically significant (p = 0.0007). As shown in Figure 3 and Table 4, agreement between predicted and observed probability was excellent in group 1 and 2 and fair in groups 6 and 8. In the other groups we observed substantial miscalibration; both under- and overestimation of the predicted risk (Table 4).

Clinical usefulness

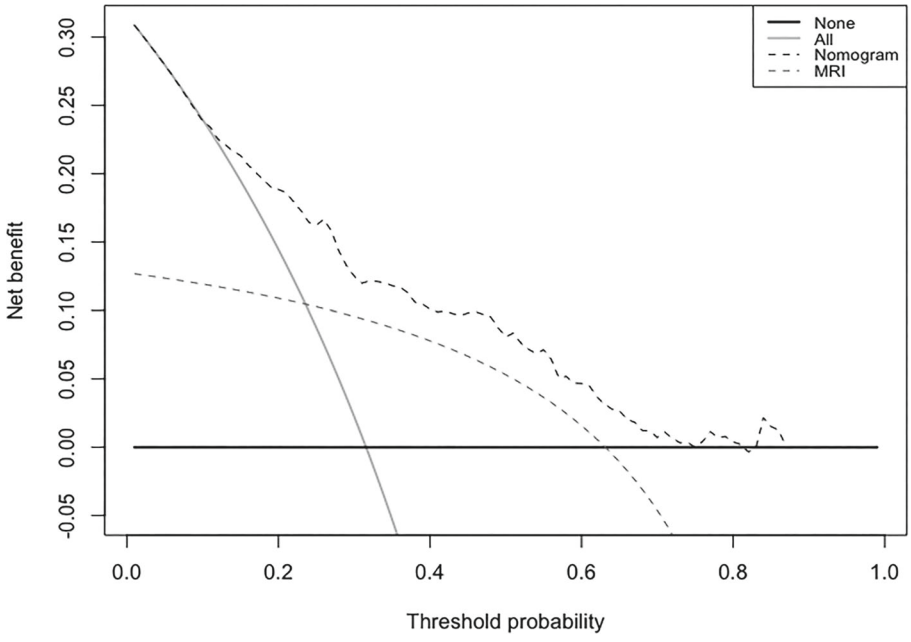
Sensitivity and specificity per risk threshold were presented in Table 5. As shown in this table, sensitivity was excellent if a threshold of 7% was used (prostate lobes with predicted risk ≥7% were regarded as EPE positive). However, this would lead to high false positive rates as specificity was low (2%). Highest cumulative optimum (respectively 140%) would be reached if risk thresholds of either 25% or 40% were used.

Decision curve analysis showed that use of the updated version of the nomogram would be superior to relying on MRI, with regard to prediction of side-specific EPE, for all clinically relevant thresholds (Figure 4). Highest net benefit would be achieved if risk thresholds between 20% and 30% would be used (Figure 4).

TABLE 5. Sensitivity and specificity of the updated nomogram per risk threshold

Risk threshold (%)	Sensitivity (%)	Specificity (%)
7	100	2
10	92	30
15	91	39
20	90	44
25	82	58
30	67	70
40	58	82
50	46	90

FIGURE 4. Decision curve analysis of the updated nomogram



DISCUSSION

In this study, we assessed the performance of a nomogram predicting side-specific EPE, using data from an external patient population undergoing RARP. The discriminative ability of this nomogram observed in our study was fair, given the AUC of 0.78. However, calibration of the nomogram was poor, as the nomogram resulted in systematic underestimation of side-specific EPE risk. Although model updating improved the agreement between overall mean predicted and observed risk, substantial and unsystematic disagreement between predicted and observed probabilities on the calibration plot persisted. Use of this model for individualized patient EPE risk prediction is therefore not recommended. Based on the superior net benefit of this nomogram compared to use of MRI alone, the nomogram could potentially be of value if risk thresholds were used to determine side-specific present EPE. Based on the high net benefit and acceptable sensitivity of 82%, the most optimal results would be reached if a risk threshold of 25% is used. However, clinicians who aim to use this model in daily practice should realize that specificity for this threshold is low (58%). Using a 25% threshold would thus lead to overtreatment, as the nomogram would advise against nerve sparing in a large number of cases with ipsilateral organ-confined disease.

Our findings concerning calibration (underestimation) of the original nomogram were consistent with those reported by Sighinolfi et al.[19] These authors also observed poor model fit of the Martini nomogram in their external validation study.[19] Findings regarding discrimination were inconsistent, as their reported AUC was substantially lower (0.68 versus 0.78). It should be noted that this previous external validation study consisted of a retrospective series of only 106 patients, accounting for a total of 137 biopsy-positive lobes of which 40 lobes contained EPE. Given this relatively small population with low number of events and subsequent low EPV (<7), sample size could be too small for a reliable validation.

Generalisability of a prediction model strongly depends on a number of factors, of which case mix is crucial. When comparing this validation cohort with the derivation cohort, a number of important differences can be addressed that may explain the poor fit. First, prevalence of EPE on final pathology was much higher in the validation cohort compared to the development cohort (250/792 [32%] vs. 142/829 [17%]). This may partially explain the systematic underestimation of the predicted probabilities on initial calibration.

Also, evaluation of the distribution of the predictors revealed that presence of EPE on mpMRI was more common in the external validation cohort (20% vs. 14%), as well as the number of cases with maximum % core involvement >50 (44% vs. 34%), compared with the development cohort. The prevalence of predictors in the study sample is important for the generalisability of the model since they establish the total variance that is being explained by the model. Compared with the validation cohort, the prevalence of predictors were lower in the development cohort. The relatively low prevalence of the predictors also explains the relatively wide 95% CI around the odds ratios of the original model,

(e.g. Gleason grade group 3) OR: 26.7 95% CI: 13.5 - 53.1 and EPE on MRI present OR: 7.27 95% CI 4.7 – 11.2), indicating a lack of precision of the estimated coefficients. This lack of precision and the fact that the validation cohort comprises of more patients with higher-risk disease could explain the observed miscalibration.

The lack of precision of this model may be also explained from another methodological viewpoint. What characterizes the study cohort is that it comprises clustered data. Each prostate lobe is considered as an independent case. However, if prostate cancer is found bilaterally, both lobes derived from a single patient are included in the sample, leading to clustering within the study population. Prediction models based on clustered data require a slightly different development approach and may require use of a random intercept. If this methodological approach was used for the development of the model, it may have been more precise.[20]

The poor model fit can also be explained by the definition of the selected predictors and the used outcome. First, our patient cohort comprises of patients diagnosed and staged at our own centre and referred patients staged elsewhere. There was no central review of MRI and pathological evaluation (the endpoint EPE), leading to a wide range of different radiologists and pathologists evaluating MRI/prostate specimen, which could induce interobserver variability. Lastly, the outcome of EPE on MRI was binary. In our population, radiologists used a subclass in their reports defined as "indefinite" or "uncertain" EPE in a substantial percentage of patients (10%). By dichotomizing this predictor, a lot of explained variance is lost. To overcome this problem, the use of a Likert scale for the probability of EPE presence may improve the model's accuracy.[21,22]

The strengths of this study include a large sample size with a high EPV rate, resulting in a study population suitable for external validation. The limitations of this study are its retrospective nature and the lack of central review regarding MRI and histological features. However, we assume that the present case mix variation reflects a real-world clinical situation.

CONCLUSIONS

External validation of the novel nomogram developed by Martini and co-workers in a large real-world cohort showed fair discriminative ability of this model, but poor calibration. After updating, substantial miscalibration was still present. Use of this nomogram for individualized risk predictions is therefore not recommended.

REFERENCES

1. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol*. 2017;71:618-629.
2. de Rooij M, Hamoen EHJ, Witjes JA, Barentsz JO, Rovers MM. Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis. *Eur Urol*. 2016;70:233-245.
3. Ohori M, Kattan MW, Koh H, et al. Predicting the presence and side of extracapsular extension: A nomogram for staging prostate cancer. *J Urol*. 2004;171:1844-1849.
4. Sayyid R, Perlis N, Ahmad A, et al. Development and external validation of a biopsy-derived nomogram to predict risk of ipsilateral extraprostatic extension. *BJU Int*. 2017;120:76-82.
5. Patel VR, Sandri M, Grasso AAC, et al. A novel tool for predicting extracapsular extension during graded partial nerve sparing in radical prostatectomy. *BJU Int*. 2018;121:373-382.
6. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet*. 2017;389:815-822
7. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *N Engl J Med*. 2018:1767-1777.
8. van der Leest M, Cornel E, Israël B, et al. Head-to-head Comparison of Transrectal Ultrasound-guided Prostate Biopsy Versus Multiparametric Prostate Resonance Imaging with Subsequent Magnetic Resonance-guided Biopsy in Biopsy-naïve Men with Elevated Prostate-specific Antigen: A Large Prospective Mu. *Eur Urol*. 2019;75:570-578.
9. Wegelin O, Exterkate L, van der Leest M, et al. The FUTURE Trial: A Multicenter Randomised Controlled Trial on Target Biopsy Techniques Based on Magnetic Resonance Imaging in the Diagnosis of Prostate Cancer in Patients with Prior Negative Biopsies. *Eur Urol*. 2019;75:582-590.
10. Martini A, Gupta A, Lewis SC, et al. Development and internal validation of a side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer. *BJU Int*. 2018;122:1025-1033.
11. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol*. 2012;22:746-757.
12. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol*. 2016;69:16-40.
13. Fine SW, Amin MB, Berney DM, et al. A contemporary update on pathology reporting for prostate cancer: Biopsy and radical prostatectomy specimens. *Eur Urol*. 2012;62:20-39.
14. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015;162:W1 -W73.
15. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2009.
16. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*. 2016;352.
17. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol*. 2018;74(6):796-804.
18. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: Seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925-1931.

19. Sighinolfi MC, Sandri M, Torricelli P, et al. External validation of a novel side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer: preliminary outcomes on a series diagnosed with multiparametric magnetic resonance imaging-targeted plus systematic saturation biopsy. *BJU Int*. 2019;124:192-194.
20. Bouwmeester W, Twisk JW, Kappen TH, Van Klei WA, Moons KG, Vergouwe Y. Prediction models for clustered data: Comparison of a random intercept and standard regression model. *BMC Med Res Methodol*. 2013;13(1).
21. Costa DN, Passoni NM, Leyendecker JR, et al. Diagnostic Utility of a Likert Scale Versus Qualitative Descriptors and Length of Capsular Contact for Determining Extraprostatic Tumor Extension at Multiparametric Prostate MRI. *AJR Am J Roentgenol*.2018;210:1066-1072.
22. Freifeld Y, A DDL, Xi Y, et al. Diagnostic Performance of Prospectively Assigned Likert Scale Scores to Determine Extraprostatic Extension and Seminal Vesicle Invasion With Multiparametric MRI of the Prostate. *AJR Am J Roentgenol*. 2019;212:576-581.

CHAPTER 7

Development and External Validation of a Novel Nomogram to Predict Side-Specific Extraprostatic Extension in Patients with Prostate Cancer Undergoing Radical Prostatectomy

T.F.W. Soeterik
H.H.E. van Melick
L.M. Dijksman
H.V.N. Küsters-Vandevelde
S. Stomps
I.G. Schoots
D.H.Biesma
J.A. Witjes
J.P.A. van Basten

EUROPEAN UROLOGY ONCOLOGY. 2020;S2588-9311(20)30133-4. ONLINE AHEAD OF PRINT

ABSTRACT

Background: Prediction of side-specific extraprostatic extension (EPE) is crucial in selecting patients for nerve sparing radical prostatectomy (RP). The aim of this study is to develop and externally validate nomograms including multiparametric magnetic resonance imaging (mpMRI) information to predict side-specific EPE.

Methods: A retrospective analysis of 1870 consecutive prostate cancer patients that underwent robot-assisted RP from 2014 to 2018 at three institutions was performed. Four multivariable logistic regression models were established, including combinations of patient-based and side-specific variables including: prostate-specific antigen (PSA)-density, highest ipsilateral biopsy International Society of Urological Pathology (ISUP) grade, ipsilateral percentage of positive cores on systematic biopsy and side-specific clinical stage assessed by both digital rectal examination and mpMRI. Discrimination (area under the curve [AUC]), calibration and net benefit of these models were assessed in the development cohort and two external validation cohorts.

Results: On external validation, AUCs of the four models ranged from 0.80 (95% CI 0.68 – 0.88) to 0.83 (95% CI 0.72 – 0.90) in validation cohort 1 and from 0.77 (95% CI 0.62 – 0.87) to 0.78 (95% CI 0.64 – 0.88) in validation cohort 2. The three models including mpMRI staging information resulted in relatively higher AUCs compared with the model without mpMRI information. No major differences between the four models regarding net benefit were established. The model based on PSA-density, biopsy ISUP grade and mpMRI T-stage was superior in terms of calibration. Using this model with a cut-off of 20%, 1980/2908 (68%) of all prostatic lobes without EPE would be found eligible for nerve sparing, whereas non-nerve sparing would be advised in 642/832 (77%) lobes with EPE.

Conclusions: Our analysis resulted in a simple and robust nomogram for the prediction of side-specific EPE which should be used to select patients for nerve sparing RP.

BACKGROUND

A challenging aspect of performing the radical prostatectomy (RP) for prostate cancer includes balancing the risk of positive margins versus optimisation of quality of life by maximising the probability of retaining the patient's erectile function and urinary continence. In 1982, the first purposeful nerve sparing radical prostatectomy was performed, resulting in a normal postoperative sexual function and retained quality of life of the patient.[1] Following the introduction of this technique, its therapeutic effect has been evaluated in several other studies. A recent meta-analysis showed nerve sparing to be associated with a lower risk of postoperative incontinence (relative risk [RR] 0.75, 95% confidence interval [CI] 0.65 - 0.85) and erectile dysfunction (RR 0.77, 95% CI 0.70 – 0.85).[2]

Preoperative assessment of extraprostatic extension (EPE) is a long-established strategy to guide patient selection for nerve sparing. If there is a high risk of EPE, nerve sparing should be discouraged due to the increased risk of positive surgical margins.[3] EPE risk prediction is often done using nomograms such as the Partin tables and the MSKCC nomogram.[4,5] However, these models do not provide information on the laterality of EPE. Since EPE is mostly one-sided (85%), localisation is essential as unilateral nerve sparing surgery remains possible in the majority of patients.[6] Nomograms predicting side-specific EPE have also been developed. However, these models lack the inclusion of multiparametric magnetic resonance imaging (mpMRI) information.[6–8]

Adoption of mpMRI to guide clinical decision-making in prostate cancer has drastically increased recent years.[9] MpMRI alone has limited ability to guide patient selection for nerve sparing, due to a low per-patient sensitivity for the detection of EPE of 57%.[10] However, its predictive potential when combined with other clinical parameters remains poorly understood. Previous studies have shown that combining mpMRI information with traditional preoperative clinical parameters including biopsy information and serum prostate-specific antigen, can improve the prediction of adverse surgical pathology including EPE.[11,12] The number of available nomograms including a combination of both mpMRI and clinical parameters for the prediction of side-specific EPE, however, is scarce. The need for further exploration of the additional value of using mpMRI information for the prediction of side-specific EPE is emphasised by the results of a recent external validation study; showing that mpMRI-naive nomograms are inaccurate when applied in external populations.[13]

Therefore, we aim to develop a nomogram that enables accurate prediction of side-specific EPE, applicable in current state of clinical practice, including readily available clinical and MRI input parameters. Generalisability of the tool will be assessed by performing external validation using two separate hospital populations.

METHODS

Patient population and study data

After receiving institutional review board approval, data from 1871 consecutive patients diagnosed with prostate cancer that underwent robot-assisted radical prostatectomy (RARP) at three teaching hospitals were extracted from prospectively maintained databases. Of these, one patient was excluded due to prior treatment with androgen deprivation therapy. The cohort of patients that underwent RARP from 2014-2018 at the Canisius Wilhelmina Hospital (CWH) Nijmegen, was used for nomogram development. This cohort was selected for model derivation due to the population size and its multi-centre nature. Since 2013, regional prostate cancer surgery has been centralised and patients from two other hospitals (Catharina Hospital Eindhoven and Radboud University Medical Centre Nijmgen) all undergo RARP at CWH. The cohorts of patients undergoing RARP at the Hospital Group Twente in Almelo-Hengelo (validation cohort 1) and St. Antonius Hospital Nieuwegein-Utrecht (validation cohort 2) from 2015 – 2018 were used for external validation.

Predictor selection

We used a clinically driven, evidence-based approach for prediction selection. First, a very recent literature review was used to identify significant predictors for side-specific EPE.[13] Second, three consensus meetings were organized with clinical experts including urologists (HvM, JW, SS, and JPvB), an expert uro-radiologist (IS) and uro-pathologist (HKV). Predictors were selected based on relevance, availability and usefulness. Patient-based (prostate-specific antigen [PSA] density [PSAD]) and side-specific covariates (digital rectal examination [DRE] local staging, mpMRI-based local staging, highest International Society of Urological Pathology (ISUP) biopsy grade and percentage of positive systematic cores) were included.

MRI protocol

MRI was performed using 3 Tesla scanners and a body coil. Gadolinium (1 mg/kg) was intravenously administrated. Radiological reporting was done by dedicated radiologists with at least two years of experience with prostate MRI reading. MRI reporting in 2013 and 2014 was done according to the European Society of Urogenital Radiology (ESUR) guidelines.[14] From 2015 and onwards, the principles of PI-RADS v2 were followed.[15] Imaging-based T-stage was defined according to the American Joint Committee on Cancer TNM classification.[16]

TABLE 1. Baseline characteristics of the separate patient cohorts

	Development	**Validation 1**	**Validation 2**
No. of patients *N* (%)	887	513	470
Age (median, IQR)	66 (61 – 69)	66 (61 – 70)	66 (62 – 70)
PSA (ng/ml)			
Mean (SD)	7.9 (5.9 – 11.0)	8.0 (5.9 – 11)	8.3 (5.9 – 12.5)
PSA density (ng/ml/ml)			
Mean (SD)	0.18 (0.12 – 0.27)	0.17 (0.12 – 0.28)	0.20 (0.13 – 0.32)
Clinical T-stage *N* (%)			
T1c	509 (57)	338 (66)	288 (61)
T2a	240 (27)	93 (18)	148 (32)
T2b	34 (4)	45 (9)	9 (2)
T2c	29 (3)	16 (3)	8 (2)
T3	64 (8)	19 (4)	14 (3)
Unknown	11 (1)	2 (0)	3 (0)
Preoperative MRI *N* (%)			
Yes	879 (99)	496 (97)	387 (82)
No	8 (1)	17 (3)	83 (18)
Radiological T-stage *N* (%)			
T0	66 (7)	57 (11)	38 (8)
T2/T2a	285 (32)	130 (25)	216 (46)
T2b	46 (5)	27 (5)	10 (2)
T2c	117 (13)	60 (12)	48 (10)
T2/T3 (uncertain EPE)	94 (11)	62 (12)	12 (3)
T3a	200 (23)	133 (26)	53 (11)
T3b	48 (5)	25 (5)	3 (1)
T4	5 (1)	0 (0)	0 (0)
Unknown	26 (3)	19 (4)	90 (19)
Biopsy type *N* (%)			
TRUS-guided systematic	497 (56)	313 (61)	380 (81)
MRI-guided	140 (16)	66 (13)	17 (4)
TRUS + MRI-guided	250 (28)	134 (26)	73 (15)
Pathological stage *N* (%)			
T2	36 (4)	1 (0)	35 (8)
T2a	86 (10)	49 (10)	37 (8)
T2b	8 (1)	11 (2)	16 (3)
T2c	303 (34)	256 (50)	237 (50)
T3a	338 (38)	142 (28)	99 (21)
T3b	112 (13)	53 (10)	46 (10)
T4	4 (0)	1 (0)	0 (0)

IQR = interquartile range, PSA = prostate specific antigen, SD = standard deviation, MRI = magnetic resonance imaging, EPE = extraprostatic extension, TRUS = transrectal ultrasonography.

Predictors and outcome definitions

Patient-based PSA and prostate volume, necessary for calculation of PSAD, were based on the most recent available measurements preoperatively. Prostate volume was measured by transrectal ultrasonography or mpMRI. Side-specific DRE staging information was collected before biopsy by the treating urologist during routine clinical care. Side-specific DRE and mpMRI staging information were both subdivided in three subclasses. These included non-palpable disease (T1), organ-confined nodal disease (T2) and EPE (T3) for DRE. As for mpMRI, these included non-visible lesions (T1), organ-confined lesions (T2) and lesions with EPE (T3).

Imaging features used to assess EPE included thickening or suspicion for invasion of the neurovascular bundle, bulging of the prostatic contour, capsule irregularity, obliteration of the recto-prostatic angle, presence of a hypo-intensive signal in a periprostatic area and length of tumour contact with the capsule.[17-19] Explicit statements about presence or absence of EPE in the radiological report were scored accordingly. In less explicit cases a strong suspicion of EPE was classified as positive. Cases in which EPE could not be ruled out were classified as negative.[19]

Side-specific biopsy information, including highest percentage of positive cores on systematic biopsy and highest ISUP grade, were documented during routine clinical care for both right and left lobe separately. Final surgical histopathological information including pathological tumour stage and highest ISUP grade found in the resected prostate specimen was documented on a whole-gland level. If EPE was observed, the laterality (left, right or both lobes) was reported. EPE was defined as a tumour that bulges beyond the prostate contour, tumour that is admixed with periprostatic fat tissue or, in the posterolateral area, as tumour within connective tissue or between nerves of the neurovascular bundle. EPE was distinct from microscopic bladder neck invasion (presence of tumour between thick smooth muscle bundles in the absence of benign prostate glands) and seminal vesicle invasion, which were not considered as EPE in our study.[20] The RP specimens were processed with conventional sections in 1810 (97%) cases and using whole-mount sections in 60 (3%) cases.

Model building

Four models were built according to the "full model" principle,[21] including combinations of five predictors corresponding with different staging work-up strategies. Model 1 consisted of PSAD, DRE, biopsy ISUP grade and percentage positive cores on systematic biopsy. Model 2 included PSAD, mpMRI and biopsy ISUP grade. Model 3 included PSAD, mpMRI, DRE and biopsy ISUP grade. Model 4 included all five predictors. For analysis purposes, the right and left prostatic lobe of each patient were regarded as separate cases. That is, for calculating the probability of EPE in the right lobe: patient-based PSAD, right-sided biopsy information, right-sided DRE staging information and right-sided mpMRI staging information were used.

Model performance, external validation and clinical usefulness

Performance of all models was assessed in the development cohort and two validation cohorts. Discrimination, which refers to the ability of the model to distinguish a prostate lobe with the endpoint (EPE) from a lobe without EPE, was quantified using the area under the receiver operating characteristic curve (AUC).[14] Model calibration, which refers to the agreement between observed endpoints and predictions, was assessed using calibration in the large and calibration slopes.[14] The net benefit per risk threshold was determined using decision-curve analysis. The net benefit is calculated as the proportion of "net" true positives (true positives corrected for the false positives weighted by the odds of the risk cut-off, divided by the sample size).[22]

Missing data

Missing data patterns were explored using response matrix and correlation plots. Missing data was assumed to be missing at random, as their missingness was related to the diagnostic work-up (e.g. selection of patients for mpMRI and biopsy protocols) of the hospitals. Missing data were handled by using multivariate imputation by chained equations including pooling using Rubin's rules.[23]

RESULTS

Patient populations

Overall, respectively 887 patients (development cohort), 513 patients (validation cohort 1) and 470 patients (validation cohort 2) were included. The values of EPE prevalence on prostatic lobe level of these cohorts were respectively 458/1774 (26%), 225/1026 (21%) and 148/940 (16%). Baseline characteristics on patient and prostatic lobe level are presented in respectively Table 1 and 2.

Performance of the four multivariable logistic regression models in the development cohort

At multivariable analyses, PSAD, DRE staging, mpMRI staging, ISUP grades 3-5, and percentage positive cores were all found to be significant predictors of EPE (Table 3). Model 4, which includes all available predictors, resulted in the highest AUC (0.82). The AUCs of the other three models ranged from 0.80 to 0.81 (Table 3).

TABLE 2. Baseline characteristics on prostate lobe level

	Development (*N* = 1774)		**Validation 1 (*N* = 1026)**		**Validation 2 (*N* = 938[a])**	
	No EPE at histopathology	**EPE at histopathology**	**No EPE at histopathology**	**EPE at histopathology**	**No EPE at histopathology**	**EPE at histopathology**
No. of lobes	1316	458	801	225	790	148
PSA density						
Median	0.17	0.21	0.16	0.23	0.19	0.21
(IQR)	(0.12 - 0.17)	(0.14 – 0.33)	(0.11 – 0.25)	(0.14 - 0.43)	(0.13 – 0.30)	(0.14 – 0.33)
Unknown (%)	19 (1)	9 (2)	14 (2)	8 (4)	13 (2)	5 (3)
ISUP grade *N* (%)						
Benign	341 (26)	31 (7)	213 (26)	10 (4)	204 (26)	9 (6)
1	365 (28)	65 (14)	315 (39)	54 (24)	306 (39)	38 (26)
2	296 (22)	147 (32)	119 (15)	79 (35)	151 (19)	41 (28)
3	93 (7)	75 (16)	54 (7)	33 (15)	55 (7)	20 (14)
4	58 (4)	58 (13)	22 (3)	27 (12)	27 (3)	21 (14)
5	36 (3)	71 (16)	12 (2)	13 (6)	17 (2)	17 (11)
Unknown	127 (10)	11 (2)	66 (8)	9 (4)	30 (4)	2 (1)
Percentage positive cores						
Median	0.20	0.67	0.20	0.80	0.33	0.75
(IQR)	(0 – 0.50)	(0.33 – 1.0)	(0 – 0.60)	(0.40 -1.0)	(0 – 0.60)	(0.40 – 1.0)
Unknown *N* (%)	242 (18)	57 (12)	112 (14)	29 (13)	47 (6)	5 (6)
Clinical stage assessed by DRE *N* (%)						
T1	1107 (84)	236 (51)	697 (87)	134 (60)	644 (81)	85 (57)
T2	134 (10)	133 (29)	71 (9)	61 (27)	101 (13)	39 (26)
T3	20 (2)	54 (12)	7 (1)	14 (6)	6 (1)	7 (5)
Unknown	55 (4)	35 (8)	26 (3)	16 (7)	39 (5)	17 (12)
Clinical stage assessed by MRI *N* (%)						
No lesion visible	622 (47)	66 (14)	395 (50)	28 (13)	315 (40)	23 (16)
Lesion but no EPE	551 (42)	209 (46)	306 (38)	92 (41)	265 (33)	68 (46)
EPE	101 (8)	169 (37)	66 (8)	93 (41)	31 (4)	28 (19)
Unknown	42 (3)	14 (3)	34 (4)	12 (5)	179 (23)	29 (19)

[a]In this cohort presence of EPE was unknown in one patient / two prostatic lobes. PSA = prostate specific antigen, DRE = digital rectal examination, EPE = extraprostatic extension, IQR = interquartile range, ISUP = International Society of Urological Pathology.

TABLE 3. Multivariable logistic regression outcomes of four different models

	Model 1 OR (95% CI)	***p***	**Model 2 OR (95% CI)**	***p***	**Model 3 OR (95% CI)**	***p***	**Model 4 OR (95% CI)**	***p***
PSA density	1.64 (0.93 – 2.90)	0.086	2.27 (1.31 – 3.94)	0.004	1.76 (0.99 – 3.10)	0.052	1.70 (0.96 – 3.02)	0.071
DRE T-stage								
T1	Referent						Referent	
T2	2.66 (1.96 - 3.60)	<0.001					1.97 (1.43 – 2.72)	<0.001
T3	5.08 (2.71 - 9.53)	<0.001	-	-	-	-	3.32 (1.70 – 6.48)	<0.001
MRI								
No lesion			Referent		Referent		Referent	
Organ-confined			2.36 (1.71 - 3.25)	<0.001	2.22 (1.60 – 3.08)	<0.001	1.96 (1.40 – 2.73)	<0.001
EPE	-	-	8.45 (5.79 - 12.32)	<0.001	6.90 (4.67 - 10.21)	<0.001	5.22 (3.47 – 7.87)	<0.001
ISUP Grade								
Benign	Referent		Referent		Referent		Referent	
1	0.89 (0.52 – 1.54)	0.7	1.48 (0.92 - 2.38)	0.11	0.74 (0.43 – 1.27)	0.3	0.80 (0.46 - 1.38)	0.4
2	1.99 (1.17 – 3.39)	0.012	3.03 (1.94 - 4.74)	<0.001	1.30 (0.76 – 2.22)	0.4	1.44 (0.83 – 2.49)	0.19
3	2.60 (1.44 - 4.70)	0.001	4.46 (2.66 - 7.46)	<0.001	1.94 (1.08 – 3.48)	0.027	1.90 (1.04 – 3.48)	0.036
4	3.36 (1.86 - 6.05)	<0.001	5.99 (3.47 – 10.3)	<0.001	2.79 (1.52 – 5.14)	0.001	2.62 (1.41 – 4.86)	0.002
5	4.94 (2.53 - 9.62)	<0.001	9.92 (5.62 - 17.5)	<0.001	3.66 (1.88 - 7.13)	<0.001	3.63 (1.92 – 7.21)	<0.001
% Positive cores	4.75 (2.83 - 7.94)	<0.001	-	-	4.77 (2.83 – 8.00)	<0.001	3.84 (2.25 – 6.53)	<0.001
AUC	0.80 (0.69 – 0.87)		0.80 (0.70 – 0.87)		0.81 (0.71 – 0.88)		0.82 (0.72 – 0.89)	
Hosmer and Lemeshow *p*	0.65		0.59		0.78		0.99	

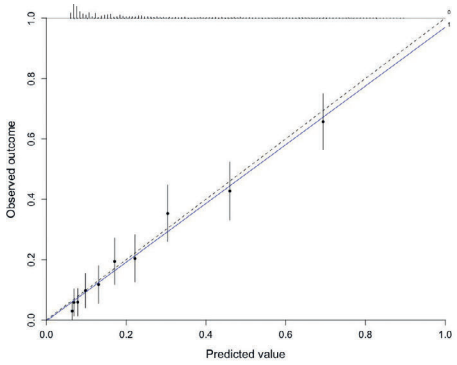
Performance of the four models when applied in two external patient cohorts

Overall, higher AUCs were observed when applied in validation cohort 1 compared with those observed in cohort 2 (Table 4). As shown in Table 4. and Figure 1., discrimination and calibration of all models were both excellent when applied in cohort 1. In cohort 2, model 2 had the best performance with both fair AUC and relatively highest agreement between predicted and observed probabilities (Figure 1 and 2). In cohort 2, substantial miscalibration was observed for the other three models (Figure 2A, 2C and 2D).

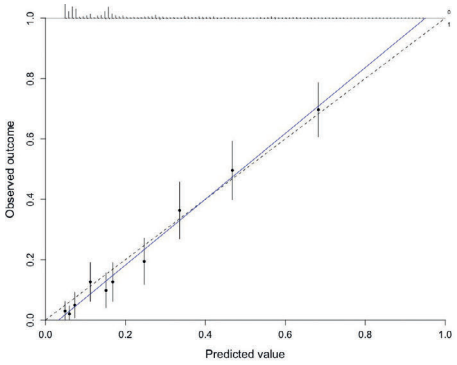
TABLE 4. Performance of all models when applied in two external cohorts

	Validation cohort 1				**Validation cohort 2**			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
AUC (95%CI)	0.80 (0.68–0.88)	0.83 (0.71–0.90)	0.83 0.71–0.90)	0.83 (0.72–0.90)	0.77 (0.62–0.87)	0.77 (0.64–0.87)	0.78 (0.64–0.88)	0.78 (0.64–0.88)

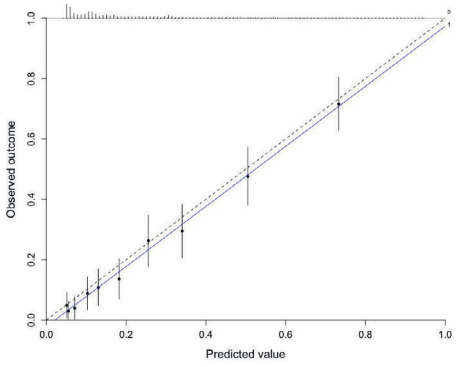
FIGURE 1. Calibration plots of all 4 models when applied in validation cohort 1



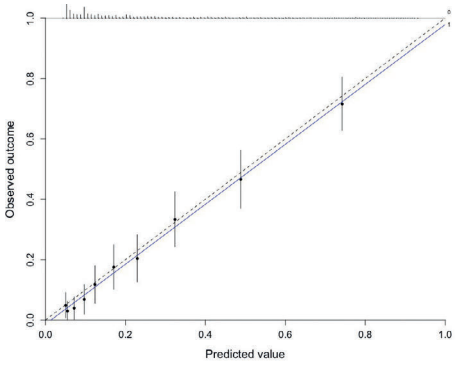
A. Model 1



B. Model 2

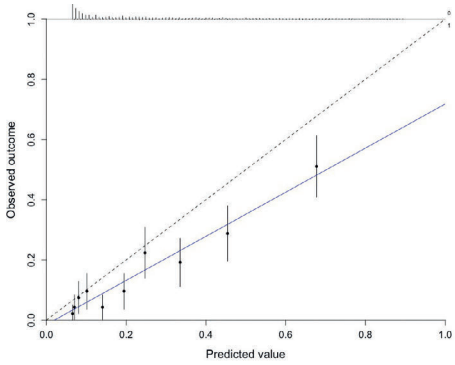


C. Model 3

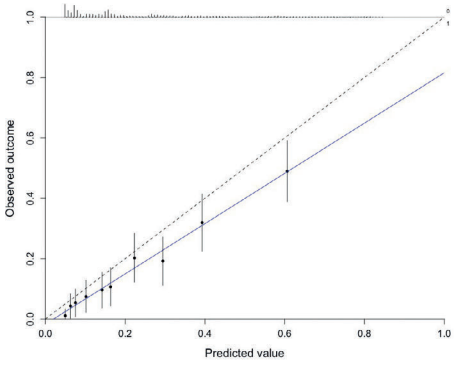


D. Model 4

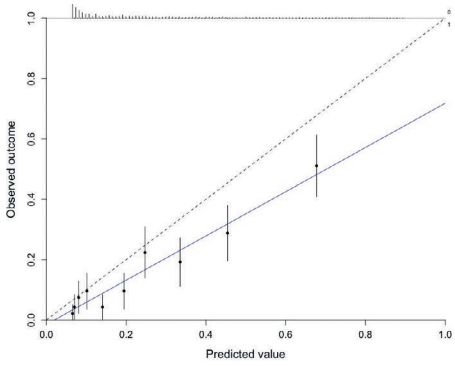
FIGURE 2. Calibration plots of all 4 models when applied in validation cohort 2



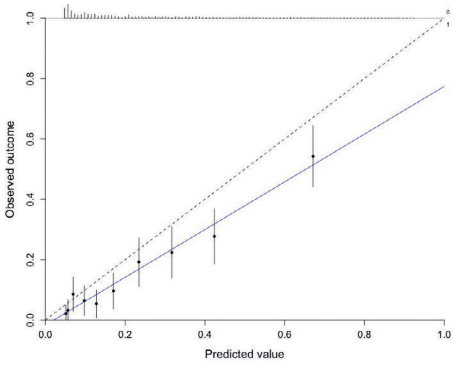
A. Model 1



B. Model 2



C. Model 3

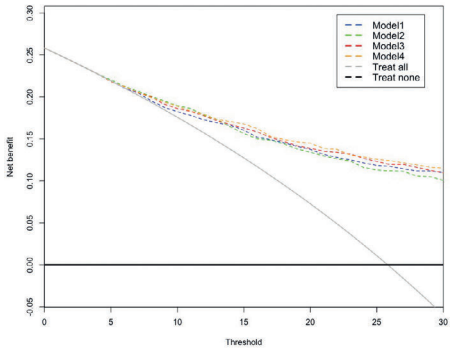


D. Model 4

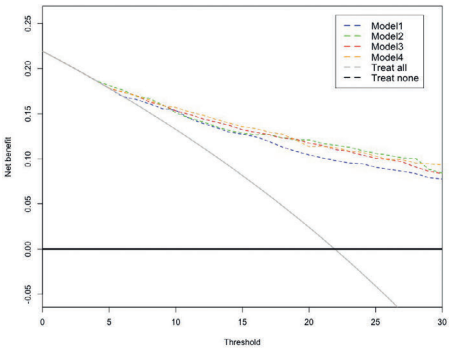
TABLE 5. Systematic analysis of outcomes for model 2 per threshold using all three cohorts

Threshold (%)	Below the cut-off (Nerve sparing recommended)			Above the cut-off (Nerve sparing not recommended)		
	Total (%)	Without EPE (%)	With EPE (%)	Total (%)	Without EPE (%)	With EPE (%)
5	256 (7)	251 (98)	5 (2)	3484 (93)	2657 (76)	827 (24)
7.5	922 (17)	883 (96)	39 (4)	2818 (83)	2025 (72)	793 (28)
10	1141 (31)	1082 (95)	59 (5)	2599 (69)	1826 (70)	773 (30)
12.5	1346 (36)	1271 (94)	75 (6)	2394 (64)	1637 (68)	757 (32)
15	1612 (43)	1497 (93)	115 (7)	2128 (57)	1411 (66)	717 (34)
17.5	2044 (55)	1876 (92)	168 (8)	1696 (45)	1032 (61)	664 (39)
20	2170 (58)	1980 (91)	190 (9)	1570 (42)	928 (59)	642 (41)
22.5	2235 (60)	2035 (91)	200 (9)	1505 (40)	873 (58)	632 (42)
25	2306 (62)	2085 (90)	221 (10)	1434 (38)	823 (57)	611 (43)
27.5	2508 (67)	2250 (90)	258 (10)	1232 (33)	658 (53)	574 (47)
30	2692 (72)	2379 (88)	313 (12)	1048 (28)	529 (50)	519 (50)

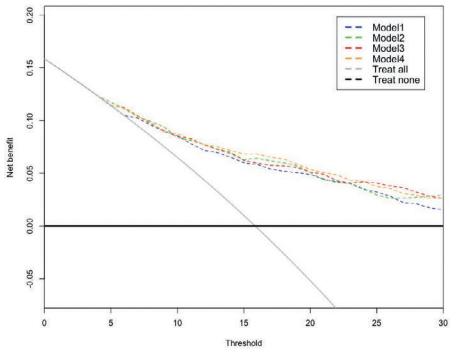
FIGURE 3. Net benefit of model 2 determined in all three cohorts using decision curve analysis



A. Development



B. Validation



C. Validation 2

Clinical usefulness

A systematic analysis of the event status of patients who would fall above and below the risk threshold between 5% and 30% is provided in Table 5. At a risk threshold of 20%, a non-nerve sparing approach would be advised in 642/832 (77%) of the prostatic lobes with EPE. Nerve sparing would be recommended in 1980/2908 (68%) of all prostatic lobes without EPE.

Risk thresholds ranging from 0% to 30% were regarded as clinically most relevant, for which net benefits of all four models are presented in Figure 3. All models can be regarded clinically useful for risk thresholds between 6% and 30%, as net benefits were found to be higher compared to the "treat all" and "treat none" approach. On external validation, DCA revealed relatively lower net benefits for model 1 compared with models 2,3 and 4, respectively. (Figure 3B and 3C).

DISCUSSION

Our analysis showed that the three nomograms (model 2, 3 and 4) based on clinical information combined with mpMRI staging information outperformed the nomogram without mpMRI staging information (model 1), in terms of AUC, calibration and net benefit. Among these three nomograms, discrimination and net benefit were comparable. However, model 2 outperformed both model 3 and 4 in terms of agreement between predicted and observed probabilities. Therefore, this nomogram should be the preferred tool for side-specific EPE risk prediction. Besides the fact model 2 outperformed all other models in terms of calibration, it was also the most minimalistic, since the model consists solely of three predictor variables (PSAD, highest biopsy ISUP grade and mpMRI T-stage). In addition, this model is applicable in a wide range of clinical situations, independent of prostate biopsy protocol used. The model can be accessed online at https://www.evidencio.com/models/show/2142.

A common explanation for the miscalibration observed when a nomogram is applied in an external population is the case-mix severity. In this study, this is also the most likely cause of the systematic overestimation of the predicted EPE risk in validation cohort 2. As an example, suspicion of EPE on mpMRI was reported in 15% of the lobes in the development cohort, compared with 6% in validation cohort 2. In addition, the prevalence of EPE on final pathology was substantially higher among cases in the development cohort compared to validation cohort 2: 26% versus 16%. As stated previously, highest agreement between predicted and observed probabilities was achieved when using model 2. However, overestimation of the predicted risk was still observed when applied in validation cohort 2. Overestimation was predominantly observed for predicted risks >30%. This was possibly due to the fact that a substantial lower number of patients with relatively high risk for EPE were selected for RARP in validation cohort 2, compared to the development cohort. For example, only the relative proportion of patients with low

suspicion of (extensive) EPE, on mpMRI or DRE were selected for RARP. Whereas patients with a high risk of (extensive) EPE, were more likely to be treated with radiation therapy. This assumption is supported by positive predictive value (PPV) rates for EPE established by DRE and mpMRI. The PPV for EPE assessed by DRE was respectively 54% in validation cohort 2, whereas this was 73% in the development cohort. PPV of mpMRI for EPE detection was 48% in validation cohort 2, compared to 63% in the development cohort.

Interestingly, model discrimination was found to be higher for model 2, 3 and 4 when applied in validation cohort 1, compared with the development cohort (0.83 vs. 0.80, 0.81 and 0.81). These differences might be explained by the heterogeneity of the patient cohort used for model development. As mentioned previously, a large proportion of the patients undergoing RARP at CWH underwent diagnostic staging work-up elsewhere. Owing to the referral pattern, there was a larger variation in used prostate biopsy protocols, mpMRI readings and histopathological biopsy evaluation as patients came from different hospitals. However, we assume that the multi-centre nature of this cohort enabled accurate model estimation leading to robust tools which can be applied in different patient settings. Another explanation for the observed improved discrimination could stem from the fact a large prospective prostate biopsy trial (4M study) was ongoing in validation cohort 1 during the study period.[24] As of part of the protocol, higher rates of patients underwent MRI target biopsy as well as concomitant systematic biopsies in a protocolled trial setting, potentially leading to more accurate tumour sampling.

To our knowledge, two other nomograms for prediction of side-specific EPE including mpMRI features have been developed previously. One of these was derived using data of 264 consecutive men undergoing RP between 2012 and 2015. The authors reported excellent model discrimination (AUC: 0.86) and excellent calibration.[25] The model, however, includes a number of complex features, which may not always be readily available in a real-world clinical setting, such as ESUR classification for EPE and capsule contact length on MRI. In addition, this model has not yet been externally validated and thus the performance when applied in other populations remains unclear. The other nomogram, developed by Martini et al, was based on data from 589 patients who underwent RARP between February 2014 and October 2015. The authors reported excellent discrimination in terms of AUC (0.82) and high agreement between predicted and observed probabilities.[26] Sighinolfi et al also externally validated The Martini model.[27] In this external validation study, moderate-low discriminative ability (AUC 0.68) and low sensitivity (20%) and specificity (54%) at the 20% cutoff were reported.[27] In another recently published external validation study, good discrimination in terms of AUC (0.78) but poor calibration, even after model updating, were reported.[28] What our study adds to this previous work is that we have shown that our developed nomogram not only provides accurate EPE risk prediction in the development cohort, but also when applied in external patient populations.

Implementation of tools that facilitate shared decision-making may improve quality of prostate cancer care, as active involvement of patients is associated with less decision conflict and less decision regret.[29] Our proposed nomograms can facilitate this,

as demonstrated by the net benefit over a range of risk thresholds within a suitable range for clinical decision-making. Besides the potential of improving quality of care in terms of patient experience, our nomogram may also improve quality of care in terms of clinical outcomes. Using the nomogram with a risk threshold set at 20%, accurate patient selection for nerve sparing RP is possibly leading to the relatively highest clinical benefit. With a 20% risk threshold, 2170/3740 (58%) of the prostatic lobes of the development cohort would fall below the cut-off and nerve sparing would be advised for these. Of these however, 190 cases (9%) would have EPE. Although nerve sparing can be safely performed in the majority of patients with risk of EPE below this threshold, it remains critical to relate these risks to the patient's preferences and willingness to trade-off between the potential quality of life benefit of nerve sparing and the increased risk of positive surgical margins. In addition to optimising preoperative staging, surgeons could consider other tools, such as intraoperative frozen section technology (NeuroSafe), to further optimise surgical outcomes.[30] Moreover, since NeuroSafe is a time-consuming and costly procedure, our nomogram as a triage test for NeuroSafe could contribute to the cost-effective deployment of NeuroSafe.

Although this study has a number of strengths, such as a large number of cases and external validation in two separate patient cohorts, some limitations have to be acknowledged. Firstly, the majority of the study data was derived from daily clinical practice and there was no central histopathologic or radiologic review. However, fact real-world clinical data were used for model development and validation could also be a potential strength since these features reflect the real-world clinical situation, for which this nomogram is designed. Secondly, although accounted for using multiple imputations, the percentage of prostatic lobes with one or more missing covariates (27%) in this study is a limitation. However, results from our additional analysis performed using complete-case data (data not shown) would not alter the study's main conclusions. Lastly, model development and external validation were both performed by the same study group and solely concerned Dutch patients. We therefore encourage other (international) study groups to also externally validate our nomogram as well.

CONCLUSIONS

We developed a simple and robust nomogram, including mpMRI information and readily available clinical parameters, for the prediction of side-specific EPE. This nomogram can be safely used to optimise patient selection for nerve sparing radical prostatectomy.

REFERENCES

1. Walsh PC. The discovery of the Cavernous Nerves and Development of Nerve Sparing Radical Retropubic Prostatectomy *J Urol*. 2007;177:1632-1635.
2. Nguyen LN, Head L, Witiuk K et al. The Risks and Benefits of Cavernous Neurovascular Bundle Sparing during Radical Prostatectomy: A Systematic Review and Meta-Analysis. *J Urol*. 2017;198:760-769.
3. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol*. 2017;71:618-629.
4. Memorial Sloan Kettering Cancer Center. Pre-radical prostatectomy tool to predict probability of extraprostatic extension in prostate cancer patients. www.mskcc.org/nomograms/prostate/pre_op.
5. Tosoian JJ, Chappidi M, Feng Z, Humphreys EB, Han M, Pavlovich CP, et al. Prediction of pathological stage based on clinical stage, serum prostate-specific antigen, and biopsy Gleason score: Partin Tables in the contemporary era. *BJU Int*. 2017;119:676-683.
6. Ohori M, Kattan MW, Koh H, et al. Predicting the presence and side of extracapsular extension: a nomogram for staging prostate cancer. *J Urol*. 2004;171:1844-1849.
7. Sayyid R, Perlis N, Ahmad A, et al. Development and external validation of a biopsy-derived nomogram to predict risk of ipsilateral extraprostatic extension. *BJU Int*. 2017;120:76-82.
8. Patel VR, Sandri M, Grasso AAC, et al. A novel tool for predicting extracapsular extension during graded partial nerve sparing in radical prostatectomy. *BJU Int*. 2018;121:373-382.
9. Caglic I, Kovac V, Barrett T. Multiparametric MRI - local staging of prostate cancer and beyond. *Radiol Oncol*. 2019;53:159-170.
10. de Rooij M, Hamoen EH, Witjes JA, et al. Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis. *Eur Urol*. 2016;70:233-245.
11. Rayn KN, Bloom JB, Gold SA, et al. Added Value of Multiparametric Magnetic Resonance Imaging to Clinical Nomograms for Predicting Adverse Pathology in Prostate Cancer. *J Urol*. 2018;200:1041-1047.
12. Gandaglia G, Ploussard G, Valerio M, et al. The Key Combined Value of Multiparametric Magnetic Resonance Imaging, and Magnetic Resonance Imaging-targeted and Concomitant Systematic Biopsies for the Prediction of Adverse Pathological Features in Prostate Cancer Patients Undergoing Radical Prostatectomy. *Eur Urol*. 2020;77:733-741.
13. Rocco B, Sighinolfi MC, Sandri M, et al. Is Extraprostatic Extension of Cancer Predictable? A Review of Predictive Tools and an External Validation based on a Large and a Single Center Cohort of Prostate Cancer Patients. *Urology*. 2019;129:8-20.
14. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol*. 2012;22:746-57.
15. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol*. 2016;69:16-40.
16. Buyyounouski MK, Choyke PL, McKenney JK, et al. Prostate cancer - major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017;67:245-253.
17. Baco E, Rud E, Vlatkovic L, et al. Predictive value of magnetic resonance imaging determined tumor contact length for extracapsular extension of prostate cancer. *J Urol*. 2015;193:466-72.
18. Cornud F, Rouanne M, Beuvon F, et al. Endorectal 3D T2-weighted 1mm-slice thickness MRI for prostate cancer staging at 1.5Tesla: should we reconsider the indirects signs of extracapsular extension according to the D'Amico tumor risk criteria?. *Eur J Radiol*. 2012;81:591-597.

19. Somford DM, Hamoen EH, Futterer JJ, et al. The predictive value of endorectal 3 Tesla multiparametric magnetic resonance imaging for extraprostatic extension in patients with low, intermediate and high risk prostate cancer. *J Urol.* 2013;190:1728-1734.
20. Fine SW, Amin MB, Berney DM, et al. A contemporary update on pathology reporting for prostate cancer: biopsy and radical prostatectomy specimens. *Eur Urol.* 2012;62:20-39.
21. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162:1-73.
22. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol.* 2018;74:796-804.
23. Van Buuren S, Groothuis-Oudshoorn K,. Multivariate Imputation by Chained Equations in R. *J Stat Soft.* 2011;45:1-67.
24. van der Leest M, Cornel E, Israel B, et al. Head-to-head Comparison of Transrectal Ultrasound-guided Prostate Biopsy Versus Multiparametric Prostate Resonance Imaging with Subsequent Magnetic Resonance-guided Biopsy in Biopsy-naive Men with Elevated Prostate-specific Antigen: A Large Prospective Multicenter Clinical Study. *Eur Urol.* 2019;75:570-578.
25. Nyarangi-Dix J, Wiesenfarth M, Bonekamp D, et al. Combined Clinical Parameters and Multiparametric Magnetic Resonance Imaging for the Prediction of Extraprostatic Disease-A Risk Model for Patient-tailored Risk Stratification When Planning Radical Prostatectomy. *Eur Urol Focus.* 2020;6:1205-1212.
26. Martini A, Gupta A, Lewis SC, et al. Development and internal validation of a side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer. *BJU Int.* 2018;122 1025-33.
27. Sighinolfi MC, Sandri M, Torricelli P, et al. External validation of a novel side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer: preliminary outcomes on a series diagnosed with multiparametric magnetic resonance imaging-targeted plus systematic saturation biopsy. *BJU Int.* 2019;124(2) 192-194
28. Soeterik TFW, van Melick HHE, Dijksman LM, et al. External validation of the Martini nomogram for prediction of side-specific extraprostatic extension of prostate cancer in patients undergoing robot-assisted radical prostatectomy. *Urol Oncol* 2020 [epub ahead of print].
29. van Stam MA, Pieterse AH, van der Poel HG, et al. Shared Decision Making in Prostate Cancer Care-Encouraging Every Patient to be Actively Involved in Decision Making or Ensuring the Patient Preferred Level of Involvement?. *J Urol.* 2018;200(3) 582-589.
30. Schlomm T, Tennstedt P, Huxhold C, et al. Neurovascular structure-adjacent frozen-section examination (NeuroSAFE) increases nerve sparing frequency and reduces positive surgical margins in open and robot-assisted laparoscopic radical prostatectomy: experience after 11,069 consecutive patients. *Eur Urol.* 2012;62 333-40.

SUPPLEMENTAL SECTION

SUPPLEMENTARY TABLE S1. Coefficients of all four models

	Model 1	Model 2	Model 3	Model 4
Intercept	-2.717	-3.065	-2.976	-2.988
PSA density	0.497	0.819	0.563	0.530
Clinical stage at DRE				
T1				
T2	0.978			0.679
T3	1.625			1.200
MRI				
No lesion				
Organ-confined		0.857	0.796	0.671
EPE		2.134	1.932	1.653
ISUP Grade				
Benign				
1	-0.111	0.393	-0.301	-0.227
2	0.687	1.109	0.261	0.364
3	0.957	1.494	0.661	0.644
4	1.211	1.790	1.026	0.963
5	1.597	2.295	1.297	1.289
% Positive cores	1.557		1.561	1.345

SUPPLEMENTARY TABLE S2. Association between patient-based EPE on MRI and EPE at final histopathology

	No EPE at histopathology (%)	EPE at histopathology (%)	Total (%)
No EPE on MRI	981 (71)	410 (29)	1391 (100)
EPE on MRI	181 (38)	298 (62)	479 (100)

SUPPLEMENTARY TABLE S3. Association between side-specific EPE on MRI and side-specific EPE at final histopathology (prostate lobe level)

	No EPE at histopathology (%)	EPE at histopathology (%)	Total (%)
No EPE on MRI	2693 (84)	525 (16)	3218 (100)
EPE on MRI	215 (41)	307 (59)	522 (100)

CHAPTER 8

8

External validation of the MSKCC and Briganti Nomograms For Prediction of Lymph Node Involvement of Prostate Cancer Using Clinical Stage Assessed by Magnetic Resonance Imaging

T.F.W. Soeterik
T.A. Hueting
B. Israël
H.H.E. van Melick
L.M. Dijksman
S. Stomps
D.H. Biesma
H. Koffijberg
M. Sedelaar
J.A. Witjes
J.P.A. van Basten

BRITISH JOURNAL OF UROLOGY INTERNATIONAL. 2021;128:236-243.

ABSTRACT

Background: Two established prediction tools that can be used to assess the probability of pelvic lymph node involvement (LNI) in patients with primary prostate cancer are the Memorial Sloan Kettering Cancer Centre (MSKCC; 2018) and Briganti (2012) nomograms. However, these nomograms include DRE-based clinical tumour stage (T-stage) as impute parameter. It is unclear if using T-stage assessed by multiparametric magnetic resonance imaging (mpMRI) improves predictive performance. Therefore, the aim of this study is to assess the impact of mpMRI T-stage on the model performance of two well-established nomograms for the prediction of LNI.

Methods: Patients undergoing robot-assisted extended pelvic lymph node dissection (ePLND) from 2015 to 2019 at three teaching hospitals were retrospectively evaluated. Risk of pelvic LNI was calculated four times for each patient, using T-stage assessed by digital rectal examination (DRE) and by mpMRI, in the MSKCC (2018) and Briganti (2012) nomograms. Discrimination (area under the curve [AUC]), calibration, and net benefit of these four strategies were assessed and compared.

Results: A total of 1062 patients were included, of whom 301 (28%) had histologically proven LNI. Using DRE T-stage resulted in AUCs of 0.71 (95% CI 0.70 - 0.72) for the MSKCC and 0.73 (95% CI 0.72 – 0.74) for the Briganti nomogram. Using mpMRI T-stage, the AUCs were 0.72 (95% CI 0.71 – 0.73) for the MSKCC and 0.75 (95% CI 0.74 – 0.76) for the Briganti nomogram. The mpMRI T-stage resulted in equivalent calibration compared with DRE T-stage. Combined use of mpMRI T-stage and the Briganti 2012 nomogram was shown to be superior in terms of AUC, calibration, and net benefit. Use of mpMRI T-stage led to increased sensitivity for the detection of LNI for all risk thresholds in both models, countered by a decreased specificity, compared with DRE T-stage.

Conclusions: Clinical T-stage assessed by mpMRI is an appropriate alternative for T-stage assessed by DRE to determine nomogram-based risk of LNI in patients with prostate cancer, and was associated with improved model performance of both the MSKCC 2018 and Briganti 2012 nomograms.

BACKGROUND

Assessment of pelvic lymph node involvement (LNI) by extended pelvic lymph node dissection (ePLND) is an essential component of the general staging work-up in patients with newly diagnosed prostate cancer selected for radical prostatectomy, and is indicated in patients with a risk of LNI above 5%.[1] Even minimal tumour involvement of the lymphatic system is thought to have pivotal impact on disease prognosis, and should be established to identify patients with an increased risk of disease recurrence.[2]

Currently, the field of clinical imaging, particularly prostate-specific membrane antigen (PSMA) positron-emission tomography/computer tomography (PET-CT), is rapidly evolving. However, since the sensitivity of PSMA PET-CT for the detection of LNI in primary prostate cancer is only moderate, it cannot yet replace ePLND to exclude LNI.[3,4] Thus, ePLND remains the preferred option for nodal staging in primary prostate cancer.[1]

However, performing ePLND in patients undergoing radical prostatectomy is associated with unfavourable intraoperative and perioperative outcomes, including symptomatic lymphocele development (in up to 18%), bleeding (2.7%), infections (3.6%), and ureteral damage (0.8%), whereas there is no high-level evidence for a direct therapeutic effect.[5,6] Therefore, ePLND should be reserved for carefully selected patients.

Both the European Association of Urology (EAU) and the National Comprehensive Cancer Network (NCCN) guidelines recommend the use of nomograms to guide patient selection for ePLND.[1,7] Several of these prediction tools have been developed over the years.[8] The Memorial Sloan Kettering Cancer Centre (MSKCC) pre-radical prostatectomy (update 2018) and Briganti 2012 nomograms are the two most established models.[9,10] In a recent validation study using a contemporary cohort of patients with Prostate cancer, the 2012 Briganti and the 2018 MSKCC nomograms were identified as the most accurate prediction tools available, with a reported area under the curve (AUC) of 0.76 and 0.75, respectively.[8] Both the MSKCC 2018 and Briganti 2012 nomograms include clinical tumour (T-stage) assessed by digital rectal examination (DRE) as one of the input parameters.[9,10] However, recent guideline updates include the recommendation for performing multiparametric MRI (mpMRI) prior to prostate biopsy.[1,11] As a result, mpMRI staging information will become standardly available in newly diagnosed patients. In addition, mpMRI potentially enables a more accurate estimation of local tumour extent compared with DRE.[12] However, it is not clear if the use of T-stage assessed by mpMRI (mpMRI T-stage) results in more accurate nomogram-based LNI risk prediction.

In the present study, therefore, we evaluate whether replacing DRE T-stage by mpMRI T-stage results in a more accurate LNI risk prediction by the MSKCC 2018 and Briganti 2012 nomograms.

METHODS

Study population

After receiving institutional review board approval, patients diagnosed with prostate cancer undergoing ePLND between January 2015 and September 2019 at three Dutch teaching hospitals (St. Antonius Hospital Nieuwegein/Utrecht, Hospital Group Twente Almelo/Hengelo, and Canisius Wilhelmina Hospital Nijmegen), were included. Patients underwent ePLND combined with radical prostatectomy or prior to radiation therapy. In general, patients with a risk of LNI >5% (based on DRE T-stage), calculated using the MSKCC web calculator, were considered as candidates for ePLND.[9] However, deviations were allowed at the discretion of the treating urologist. DRE T-stage, mpMRI T-stage, preoperative prostate-specific antigen (PSA), highest International Society of Urological Pathology (ISUP) grade observed on most recent preoperative biopsy, total number of biopsy cores taken on systematic biopsy and the relative number of cores containing prostate cancer on systematic biopsy were collected. Patients were included if they underwent systematic biopsies with or without MRI-guided target biopsy and mpMRI for local staging prior to ePLND. Patients undergoing salvage ePLND or those who received androgen deprivation therapy prior to ePLND were excluded.

Covariates and endpoints

Prostate-specific antigen, DRE T-stage, mpMRI T-stage, total number and relative number of positive biopsy cores as well as pathological lymph node status were collected during standard clinical practice. Biopsy grading was performed according to the new Gleason Grade group classification.[13]

Digital rectal examination was performed during the primary diagnostic evaluation by urologists with >5 years of experience in diagnosing and staging prostate cancer. DRE consisted of systematic palpation of all prostate regions including both lateral sides, the posterior region and the sulcus. DRE was performed in either the dorsal lithotomy or lateral position. Findings were reported according to the clinical classification of the American Joint Committee on Cancer.[14]

During the study period, 3-Tesla MRI scanners were used at the three institutions. Radiological reporting was performed by dedicated uro-radiologists. Reporting was done according to the Prostate Imaging – Reporting and Data System (PI-RADS) v2 guidelines.[15] MpMRI T-stages were defined as T1c (non-visible lesion), T2a (unilateral suspicious lesion, involving <50% of the prostatic lobe), T2b (unilateral suspicious lesion, involving >50% of the prostatic lobe) T2c (bilateral suspicious lesion), T3a (definite or high degree of suspicion for extraprostatic extension, T3b (definite or high-degree of suspicion of seminal vesicle invasion) and T4 (invades adjacent structures). The MRI protocols used at the three institutions are presented in the Supplemental Section (Table S1).

The ePLND template included removal of nodes overlying the external iliac vessels, internal iliac artery, and the nodes located within the obturator fossa.[16]

All resected nodal tissue was submitted for pathologic evaluation, performed by experienced uro-pathologists. The total number of lymph nodes found in the tissue, as well as the number of nodes containing prostate cancer metastasis were assessed. Histopathological evaluation was performed in accordance with the ISUP consensus statement on handling and staging of radical prostatectomy specimens.[17]

Statistical analysis

The risk of LNI was estimated a total of four times per patient: using both the MSKCC 2018 and Briganti 2012 nomograms, with both DRE- and mpMRI T-stage. Other covariates used for LNI risk calculation included most recent preoperative serum PSA level, highest ISUP grade found on either systematic or target biopsy, as well as the number of positive cores and the total number of cores taken on systematic biopsy. Model discrimination was quantified using the AUC, and refers to the probability of a random patient with histologically proven LNI (pN1) having a higher predicted risk than a random patient without histologically proven LNI (pN0).[18] Classification plots showing the true and false positive rates per risk threshold were used to visualize discriminatory ability.[19] Model calibration, which refers to the agreement between observed and predicted LNI, was assessed by plotting calibration curves and by determining calibration-the-large and calibration slopes.[18] Calibration-in-the-large indicates whether predicted probabilities are systematically too low or too high. Perfect calibration is characterized by an intercept of 0, and a calibration slope of 1.[18] The scaled Brier score, which is the average squared difference between the actual outcomes (i.e. LNI) and predicted probabilities, was also determined. A scaled Brier score close to 1 shows overall poor predictive ability, whereas a scaled Brier score of 0 corresponds with perfect risk prediction of the model.[18] Decision-curve analysis was performed to determine net benefit of the models over multiple clinically relevant thresholds. The calculated net benefit of the models was compared to the scenarios of treating either all or no patients.[20] A systematic analysis was performed to determine the number of patients (with or without LNI) in whom ePLND would be advised, for LNI risk thresholds between 1% and 15%. Missing data were handled by using multiple imputations by chained equations.[21] A total of 10 imputed datasets were created. Model performance measures were estimated by bootstrapping each imputed dataset 500 times. To select the best performing approach, the different approaches were compared head-to-head by estimating in how many bootstrap samples a specific approach resulted in the highest pooled AUC measure. Statistical analysis was performed using R v3.6.3. (R Project for Statistical Computing, www.r-project.org).

RESULTS

Study population

A total of 1062 patients fulfilled the inclusion criteria. Overall, 301 (28%) patients had histologically confirmed LNI. The median number of lymph nodes removed was 20 (interquartile range 13-25). A total of 21 patients (2%) had one or more covariates missing including DRE T-stage (*N* = 11), mpMRI T-stage (*N* = 2), and biopsy data (*N* = 8). Baseline characteristics of the study cohort are presented in Table 1.

Model performance using DRE or mpMRI T-stage

Initial validation included use of DRE T-stage. Discrimination in terms of AUC was 0.71 (95% CI 0.70 – 0.72) for the MSKCC and 0.73 (95% CI 0.72 – 0.74) for the Briganti nomogram. Mean predicted probability for LNI was respectively 24% for the MSKCC and 21% for the Briganti nomogram, where the observed LNI rate was 28% (Table S2). On visual exploration of calibration plots, we also observed systematic underestimation of the predicted risk of both nomograms, particularly for risk thresholds between 0% and 30% (Figures 1a and 1b).

FIGURE 1. Calibration plots of both nomograms based on DRE T-stage (A and B) and mpMRI T-stage (C and D). MSKCC, Memorial Sloan Kettering Cancer Centre. A histogram displayed at the top of each calibration plot shows the distribution of predicted risks for pN0 and pN1 cases. pN0 is indicated by the 0 (top side of the histogram), and pN1 is indicated by the 1 (bottom side of the histogram)

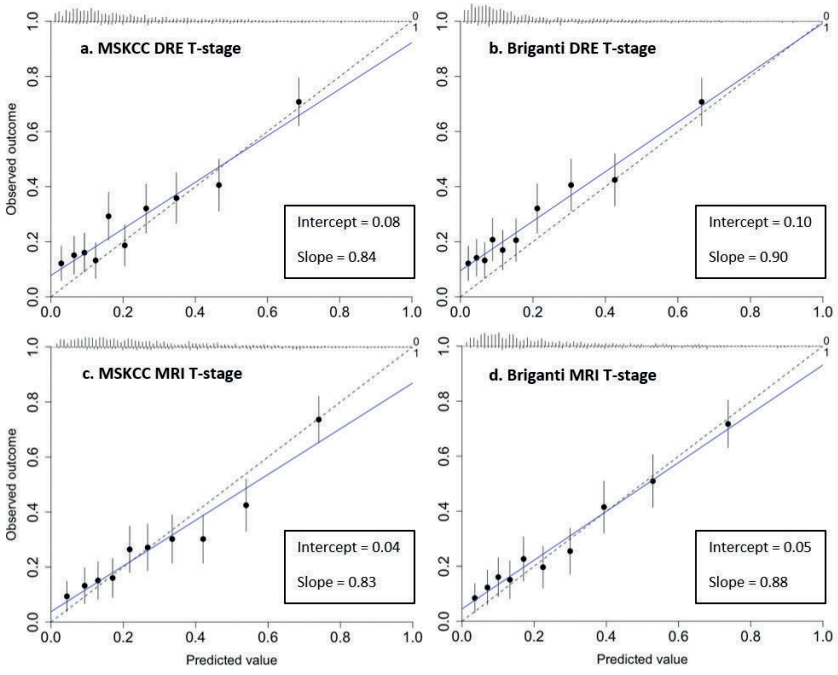


TABLE 1. Baseline characteristics of the validation cohort

	Overall	**pN0**	**pN1**
No. of patients	1062 (100)	761 (72)	301 (28)
Hospital *N* (%)			
SAH	258 (24)	186 (24)	72 (24)
HGT	246 (23)	159 (21)	87 (29)
CWH	558 (53)	416 (55)	142 (47)
Median (IQR) age, years	67 (63 - 71)	68 (63 - 71)	67 (63 - 71)
Median (IQR) PSA, ng/mL	10 (6.6 – 18)	9.3 (6.2 – 16)	13 (7.8 – 22)
Median (IQR) total cores	10 (10 - 12)	10 (9 – 12)	10 (10 – 12)
Data on total cores missing *N* (%)	5 (0)	4 (1)	1 (0)
Total positive cores	5 (3 – 8)	5 (3 – 7)	7.1 (3.3)
Data on positive cores missing *N* (%)	6 (1)	5 (1)	1 (0)
Percentage of positive cores (%)	0.50 (0.33 – 0.75)	0.50 (0.25 – 0.67)	0.75 (50 - 100)
Data on percentage of positive cores missing *N* (%)	7 (1)	6 (1)	1 (0)
Biopsy ISUP grade *N* (%)			
1	78 (7)	65 (9)	13 (4)
2	245 (23)	191 (25)	54 (18)
3	280 (26)	202 (27)	78 (26)
4	253 (24)	189 (25)	64 (21)
5	201 (19)	110 (14)	91 (30)
Missing	5 (0)	4 (0)	1 (0)
DRE T-stage *N* (%)			
T1c	384 (36)	301 (40)	83 (28)
T2a	328 (31)	248 (33)	80 (27)
T2b	84 (8)	57 (7)	27 (9)
T2c	77 (7)	52 (7)	35 (12)
T3a	169 (16)	103 (14)	66 (22)
T3b	7 (1)	3 (0)	4 (1)
T4	3 (0)	1 (0)	2 (0)
Missing	10 (1)	6 (1)	4 (1)
mpMRI T-stage *N* (%)			
T1c	40 (4)	36 (5)	4 (1)
T2a	301 (28)	250 (33)	51 (17)
T2b	41 (4)	29 (4)	12 (4)
T2c	120 (11)	103 (14)	17 (6)
T3a	376 (35)	261 (34)	115 (38)
T3b	160 (15)	69 (9)	91 (30)
T4	22 (2)	12 (2)	10 (3)
Missing	2 (0)	1 (0)	1 (0)

TABLE 1. Continued

	Overall	pN0	pN1
Biopsy type *N* (%)			
TRUS-guided SB	694 (65)	479 (63)	215 (71)
TRUS-SB+ MRI-TB	368 (35)	282 (37)	86 (29)
Median (IQR) total nodes resected	20 (13 - 25)	17 (12 - 24)	20 (14 - 28)

SAH: St. Antonius Hospital, HGT: Hospital Group Twente, CWH: Canisius Wilhelmina Hospital, SD: standard deviation, IQR: interquartile range, TRUS: transrectal ultrasonography, SB: systematic biopsy, MRI-TB: magnetic resonance image-guided target biopsy. Percentages may not total 100 due to rounding.

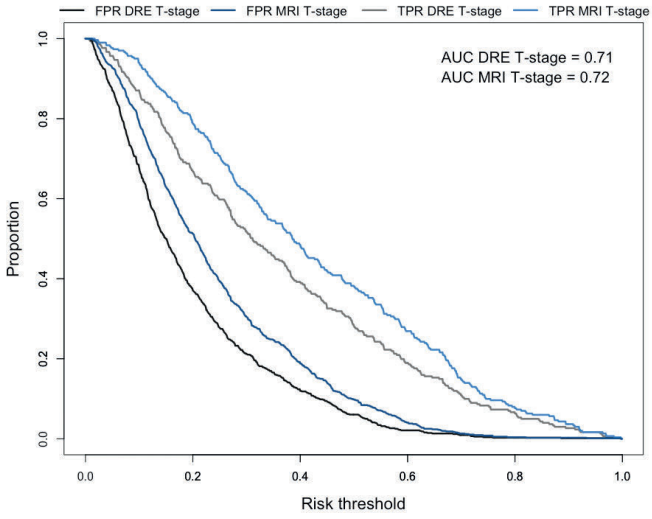
When using mpMRI T-stage, discrimination in terms of AUC, increased to 0.72 (95% CI 0.71 – 0.73) for the MSKCC and 0.75 (95% CI 0.74 – 0.76) for the Briganti nomogram (Table S2). Mean predicted probability for LNI changed to 30% for the MSKCC and 27% for the Briganti nomogram (Table S2). As shown in the calibration plots, the agreement between predicted and observed probabilities was comparable (both moderate calibration) for both DRE T-stage and mpMRI T-stage. Calibration intercepts were closer to 0 when using mpMRI instead of DRE for both the MSKCC (0.04 [95% CI: 0 – 0.08] versus 0.08 [95% CI: 0.04 – 0.12]), and the Briganti nomogram (0.05 [95% CI: 0 – 0.08] versus 0.10 [95% CI: 0.06 – 0.13]); (Table S2). In a head-to-head comparison, calculating the LNI risk using mpMRI T-stage with the Briganti nomogram led to higher AUCs in all bootstrap samples, compared with Briganti DRE T-stage as well as both DRE T-stage and mpMRI T-stage with the MSKCC nomogram.

Clinical usefulness

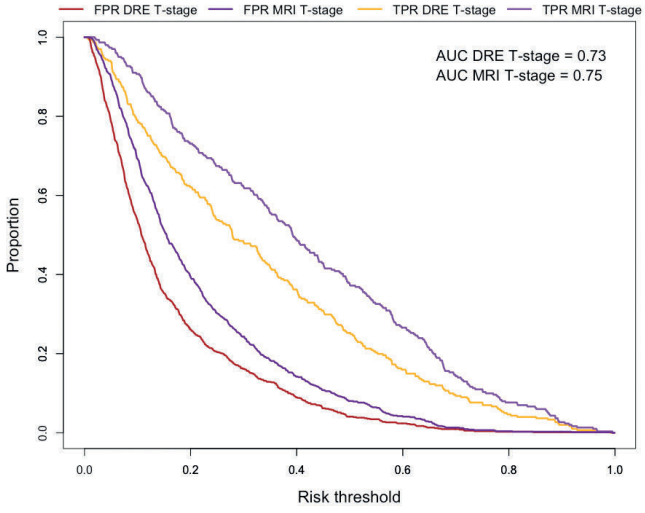
Using mpMRI T-stage resulted a in higher true-positive rate and a higher false-positive rate for the detection of positive lymph nodes for all risk thresholds, compared to using DRE T-stage (Figure 2). Use of mpMRI T-stage led to increased sensitivity for the detection of LNI for all risk thresholds in both models, countered by a lower specificity, compared with DRE T-stage. In Tables S3 and S4, total numbers of missed LNI cases per risk threshold are presented, combined with rates of performed ePLND and number of positive LNI cases. For all thresholds, the number of missed LNI cases was lower when mpMRI T-stage was used, countered by higher rates of unnecessary ePLND (pN0).

Decision curve analysis revealed that use of mpMRI T-stage in both nomograms resulted in higher net benefits, compared with DRE T-stage, for the risk thresholds between 5% and 20%. Net benefits for both the MSKCC and Briganti nomograms, using mpMRI T-stage, were comparable for this range of risk thresholds. For risk thresholds ranging from 20% and 30%, the combined use of mpMRI T-stage with the Briganti nomogram would lead to the highest net benefit (Figure 3).

FIGURE 2. Classification plots of both nomograms displaying false positive rate (FPR) and true positive rate (TPR) established using both DRE T-stage and multiparametic MRI (mpMRI) T-stage with (A) the Memorial Sloan Kettering Cancer Centre 2018 and (B) the Briganti 2012 nomogram

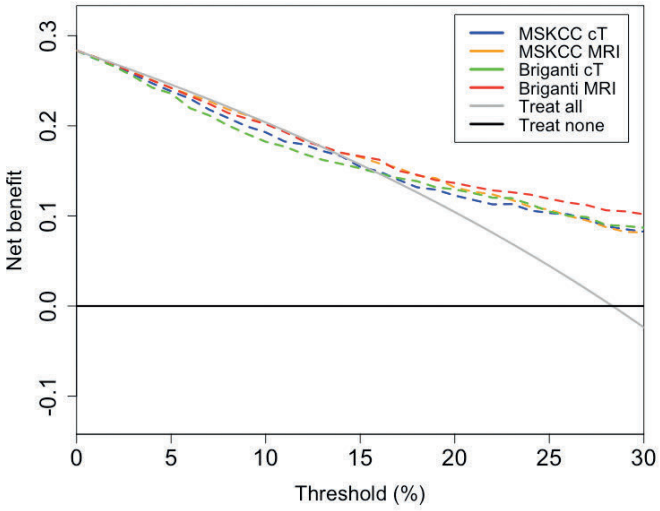


A.



B.

FIGURE 3. Decision curves for the four performed validation scenarios compared to the default strategies



DISCUSSION

Use of mpMRI T-stage for nomogram-based LNI risk assessment resulted in higher AUC, comparable agreement between predicted and observed probabilities, and higher net benefit compared with DRE T-stage, in both the MSKCC 2018 and Briganti 2012 nomograms. In our study population, use of DRE T-stage would lead to overall LNI risk underestimation in the clinically relevant range of risk thresholds (0-30%). In the head-to-head comparison, combined use of the mpMRI T-stage with the Briganti 2012 nomogram resulted in the most accurate LNI risk prediction.

Our study acknowledges the robustness of both the MSKCC 2018 and Briganti 2012 nomograms, since model performance was still fair to good, even when the model was applied in a patient population with a considerably higher prevalence of the predicted outcome compared with the development populations. In our cohort, LNI prevalence (28%) was substantially higher compared to both MSKCC (7% [internal communication MSKCC research team]) and Briganti (8%) populations.[10] Therefore, our results show both models are applicable in a contemporary patient cohort. In addition, our analysis confirmed that mpMRI T-stage can be safely used as impute parameter for these nomograms, even leading to improved accuracy of the predicted LNI risk compared with DRE T-stage.

The present study's main findings add up to the available body of literature supporting the additional value of mpMRI information for predicting presence of LNI in prostate cancer. For example, Porpiglia et al. showed MRI has an important role in LNI risk prediction in patients with a nomogram-predicted risk <5%.[22] Huang et al. demonstrated

that addition of PI-RADS score improved model discrimination in terms of AUC for both nomograms, increasing from 75% to 86% for Briganti and from 79% to 88% for MSKCC, respectively.[23]

Recently, two new nomograms have been introduced, including mpMRI and target biopsy features such as maximum diameter of the index lesion and maximum percentage of tumour involvement in one core.[24,25] Of these, the 2019 Briganti nomogram was recently externally validated showing excellent characteristics, including an AUC of 79% and high agreement between predicted and observed probabilities for risk thresholds below 35%.[26] In their head-to-head comparison, the 2019 Briganti nomogram outperformed the Briganti 2017 and MSKCC 2018 in terms of discrimination, calibration and net benefit.[26] Although these new nomograms potentially enable improved LNI risk prediction due to the addition of mpMRI guided target biopsy, they both include complex features which may not be always available in clinical practice, such as maximum diameter of the index lesion on mpMRI and highest tumour length in millimetres of all biopsy cores taken.[24,25] In addition, more external validation studies are warranted to confirm the accuracy of these new nomograms in external patient populations, as model transportability needs to be adequate to prevent systematic wrong decision-making.

Our results do not support the statements in a recent position paper on prostate cancer staging by Paner et al., who suggested that DRE should not be replaced by mpMRI for establishing clinical T-stage.[27] In the present study, mpMRI outperformed DRE in terms of AUC for nomogram-based LNI risk prediction, as the use mpMRI T-stage resulted in higher AUCs for all bootstrap samples. This is most likely the consequence of the main advantage that mpMRI has over DRE for determining local tumour extent, which is the visualisation of the prostate gland as a whole and improved detection of non-organ-confined disease. Our study group has confirmed this in a recent study, as the reported sensitivity for the detection of non-organ-confined disease was significantly lower for DRE compared with mpMRI (12% vs. 51%, $p < 0.001$).[28]

Although our main study results favour the use of mpMRI T-stage for nomogram-based LNI risk prediction, there are arguments against replacing DRE with mpMRI T-stage that should be mentioned. First, disadvantages of MRI include reader interobserver variability and quality differences regarding mpMRI reading.[27] However, a previous study by Angulo et al showed interobserver inconsistency also to be an issue for DRE, resulting in a low ability to reproduce clinical staging by DRE among multiple examiners.[29]

Second, use of mpMRI compared with DRE would lead to upstaging of clinical T-stage in one-third of the patients.[28] Although mpMRI can provide valuable prognostic information for specific patients, including those with non-organ-confined disease which was not detected during DRE, the high upstaging rates bear the risk of overstaging and hence overtreatment in patients with genuine low-risk disease.[28]

To select patients for ePLND, it remains important to take into account patient's preferences, age and prognostic tumour parameters other than those included in the nomograms to distinguish the patients who would benefit from additional ePLND, from those in whom this intervention would potentially do more harm than good.

In addition, the trade-off between subjecting node-negative patients to the concomitant risks of ePLND vs the potential advantages of ePLND in the specific node-positive subgroup, remains to be explored. Future studies should focus on finding the optimum risk threshold at which the benefits of ePLND, at best, outweigh the harms.

Although this study has several strengths, such as the inclusion of a multicentre cohort representing a real-world prostate cancer population and a large study sample with a sufficient number of events for adequate external validation, it also has some limitations. Firstly, the data used in the study were derived from routine clinical practice, and no central review of DRE, mpMRI and histopathological evaluation was performed. Secondly, the majority of the data were collected retrospectively, which could have led to measurement bias. Lastly, the indication to perform an ePLND in this patient cohort was done using nomogram-based LNI risk estimation (risk of LNI >5%). Even though this is according to current EAU guideline recommendations, and reflects contemporary clinical practice, this could have introduced bias due to the selection of patients for ePLND with higher risk of LNI (reflected by the relatively high LNI prevalence). For instance, selecting patients with higher risk of LNI (and prevalence) could explain the counterintuitive finding on decision-curve analysis, showing that a "treat all" approach would lead to higher net benefit compared with nomogram-based selection for risk thresholds between 0% and 15%.

CONCLUSIONS

The MSKCC 2018 and Briganti 2012 nomograms were found to be adequate models for the prediction of LNI in patients with prostate cancer when using either mpMRI T-stage or DRE T-stage. The use of mpMRI T-stage led to improved model discrimination, equal calibration, and lower rates of missed LNI cases. Using the mpMRI T-stage with the Briganti 2012 nomogram was shown to be the most accurate strategy for LNI risk prediction.

REFERENCES

1. Mottet N, van den Bergh RCN, Briers E, et al. EAU - ESTRO - ESUR - SIOG Guidelines on Prostate Cancer 20202; presented at the EAU Annual Congress Amsterdam 2020. Arnhem, The Netherlands: European Association of Urology Guidelines Office.
2. Wilczak W, Wittmer C, Clauditz T, et al. Marked Prognostic Impact of Minimal Lymphatic Tumor Spread in Prostate Cancer. *Eur Urol*. 2018;74:376-386.
3. Luiting HB, van Leeuwen PJ, Busstra MB, et al. Use of gallium-68 prostate-specific membrane antigen positron-emission tomography for detecting lymph node metastases in primary and recurrent prostate cancer and location of recurrence after radical prostatectomy: an overview of the current literature. *BJU Int*. 2020;125:206-214.
4. Perera M, Papa N, Roberts M, et al. Gallium-68 Prostate-specific Membrane Antigen Positron Emission Tomography in Advanced Prostate Cancer—Updated Diagnostic Utility, Sensitivity, Specificity, and Distribution of Prostate-specific Membrane Antigen-avid Lesions: A Systematic Review and Meta-. *Eur Urol*. 2020;77:403-417.
5. Fossati N, Suardi N, Gandaglia G, et al. Identifying the Optimal Candidate for Salvage Lymph Node Dissection for Nodal Recurrence of Prostate Cancer: Results from a Large, Multi-institutional Analysis. *Eur Urol*. 2019;75:176-183.
6. Oderda M, Diamand R, Albisinni S, et al. Indications for and complications of pelvic lymph node dissection in prostate cancer: accuracy of available nomograms for the prediction of lymph node invasion. *BJU Int*. 2021;127:318-325.
7. Carroll PR, Parsons JK, Andriole G, et al. NCCN Guidelines Insights: Prostate Cancer Early Detection, Version 2.2016. *J Natl Compr Cancer Netw*. 2016;14:509-519.
8. Hueting TA, Cornel EB, Somford DM, et al. External Validation of Models Predicting the Probability of Lymph Node Involvement in Prostate Cancer Patients. *Eur Urol Oncol*. 2018;1:411-417.
9. Memorial Sloan Kettering Cancer Center. Pre-radical prostatectomy tool to predict probability of lymph node involvement in prostate cancer patients. www.mskcc.org/nomograms/prostate/pre_op. Published 2018, Accessed June 15 2020.
10. Briganti A, Larcher A, Abdollah F, et al. Updated nomogram predicting lymph node invasion in patients with prostate cancer undergoing extended pelvic lymph node dissection: The essential importance of percentage of positive cores. *Eur Urol*. 2012;61:480-487.
11. Bjurlin MA, Carrol PR, Eggener S, et al. Update of the Standard Operating Procedure on the Use of Multiparametric Magnetic Resonance Imaging for the Diagnosis, Staging and Management of Prostate Cancer. *J Urol*. 2020;203:706-712.
12. de Rooij M, Hamoen EHJ, Witjes JA, Barentsz JO, Rovers MM. Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis. *Eur Urol*. 2016;70:233-245.
13. Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. *Eur Urol*. 2016;69:428-435.
14. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol*. 2016;40:244-252.
15. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur Urol*. 2016;69:16-40.

16. Mattei A, Fuechsel FG, Bhatta Dhar N, et al. The Template of the Primary Lymphatic Landing Sites of the Prostate Should Be Revisited: Results of a Multimodality Mapping Study. *Eur Urol*. 2008;53:118-125.
17. Berney DM, Wheeler TM, Grignon DJ, et al. International society of urological pathology (ISUP) consensus conference on handling and staging of radical prostatectomy specimens. working group 4: Seminal vesicles and lymph nodes. *Mod Pathol*. 2011;24:39-47.
18. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015;162:W1-W73.
19. Verbakel JY, Steyerberg EW, Uno H, et al. ROC curves for clinical prediction models part 1. ROC plots showed no added value above the AUC when evaluating the performance of clinical prediction models. *J Clin Epidemiol*. 2020;126:207-216.
20. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Mak*. 2006;26:565-574.
21. Van Buuren S, Groothuis-Oudshoorn K. Multivariate Imputation by Chained Equations in R. *J Stat Softw*. 2011;45:1-67.
22. Porpiglia F, Manfredi M, Mele F, et al. Indication to pelvic lymph nodes dissection for prostate cancer: The role of multiparametric magnetic resonance imaging when the risk of lymph nodes invasion according to Briganti updated nomogram is <5%. *Prostate Cancer Prostatic Dis*. 2018;21:85-91.
23. Huang C, Song G, Wang H, et al. Preoperative PI-RADS Version 2 scores helps improve accuracy of clinical nomograms for predicting pelvic lymph node metastasis at radical prostatectomy. *Prostate Cancer Prostatic Dis*. 2020;23:116-126.
24. Gandaglia G, Ploussard G, Valerio M, et al. A Novel Nomogram to Identify Candidates for Extended Pelvic Lymph Node Dissection Among Patients with Clinically Localized Prostate Cancer Diagnosed with Magnetic Resonance Imaging-targeted and Systematic Biopsies. *Eur Urol*. 2019;75:506-514.
25. Draulans C, Everaerts W, Isebaert S, et al. Development and External Validation of a Multiparametric Magnetic Resonance Imaging and International Society of Urological Pathology Based Add-On Prediction Tool to Identify Prostate Cancer Candidates for Pelvic Lymph Node Dissection. *J Urol*. 2020;203:713-718.
26. Gandaglia G, Martini A, Ploussard G, et al. External Validation of the 2019 Briganti Nomogram for the Identification of Prostate Cancer Patients Who Should Be Considered for an Extended Pelvic Lymph Node Dissection. *Eur Urol*. 2020;78:138-142.
27. Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW, Amin MB. Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers. *Eur Urol*. 2018;73:560-569.
28. Soeterik TFW, van Melick HH., Dijksman LM, Biesma DH, Witjes JA, van Basten J-PA. Multiparametric Magnetic Resonance Imaging Should Be Preferred Over Digital Rectal Examination for Prostate Cancer Local Staging and Disease Risk Classification. *Urology*. 2020:1-8.
29. Angulo JC, Montie JE, Bukowsky T, et al. Interobserver consistency of digital rectal examination in clinical staging of localized prostatic carcinoma. *Urol Oncol Semin Orig Investig*. 1995;1:199-205.

SUPPLEMENTAL SECTION

SUPPLEMENTARY TABLE S1. MRI protocols used at each hospital

	SAH	CWH	HGT
Field of strength of magnet	3T	3T	3T
Coils used?	Body coil	Body coil	Pelvic coil
Slice thickness	3mm	3mm	3mm
For T2-weighted imaging, which planes were acquired?	Axial, sagittal, coronal	Axial, sagittal, coronal	Axial, sagittal, coronal
For DWI, the b-values, image sets analyses (e.g. high b-value image, ADC map, or both)	B 50-400-800 and calculated B1400 and ADC map	B 0-50-100-200-250-800 and ADC map	B 50-400-800 and calculated B1400 and ADC map
Scoring system used (if different were used, indicate per year)	PI-RADS v2	PI-RADS v2	PI-RADS v2
Experience of radiologists reporting prostate MRI in the period	2-5 years	6 years	>5 years
Intravenous administration? (dosage, type e.g. Gadolinium + manufacturer)	Gadolinium 0.1 mmol/kg (Dotarem)	Gadolinium 1.0 mg/kg intravenous (Dotarem)	Gadolinium 0.1 mmol/kg (Dotarem)

SUPPLEMENTARY TABLE S2. Model performance parameters after pooling for DRE T- and mpMRI T- stage

	MSKCC 2018		Briganti 2012	
	DRE T-stage	mpMRI T-stage	DRE T-stage	mpMRI T-stage
AUC (95% CI)	0.71 (0.70 – 0.72)	0.72 (0.71 – 0.73)	0.73 (0.72 – 0.74)	0.75 (0.74 – 0.76)
Mean predicted probability (%)	24	30	21	27
Calibration slope	0.84 (0.71 – 0.99)	0.83 (0.69 – 0.95)	0.90 (0.77 – 1.02)	0.88 (0.76 – 1.00)
Calibration intercept	0.08 (0.04 – 0.12)	0.04 (0 – 0.08)	0.10 (0.06 – 0.13)	0.05 (0 – 0.08)
Scaled Brier score	0.13 (0.08 – 0.18)	0.14 (0.08 – 0.19)	0.13 (0.07 – 0.18)	0.18 (0.12 – 0.24)

* Overall prevalence of LNI on final histopathological evaluation was 28%.

SUPPLEMENTARY TABLE S3. Systematic analysis of MSKCC 2018 for nomogram-derived cut-offs

MSKCC 2018										
Threshold (%)	**DRE T-stage**					**mpMRI T-stage**				
	Patients in whom ePLND is not / is recommended (%)	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without / with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**	**Patients in whom ePLND is not / is recommended (%)**	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without/with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**
1	5 (0.05) / 1057 (99.5)	3 (60) / 2 (40)	758 (72) / 299 (28)	5 (0.05)	2 (0.2)	0 (0) / 1062 (100)	0 (0) / 0 (0)	761 (72) / 301 (28)	0 (0)	0 (0)
2	33 (3.1) / 1029 (96.9)	31 (94) / 2 (6.1)	730 (71) / 299 (29)	33 (3.1)	2 (0.2)	9 (0.8) / 1053 (99.2)	8 (89) / 1 (11)	753 (71) / 300 (29)	9 (0.9)	1 (0.1)
3	56 (5.4) / 985 (95)	50 (89) / 6 (11)	711 (71) / 295 (29)	56 (5.3)	6 (0.6)	31 (3.0) / 1031 (97)	28 (90) / 3 (10)	733(71) / 298 (29)	31 (2.9)	3 (0.3)
4	87 (8.3) / 955 (92)	77 (89) / 10 (11)	684 (70) / 291 (30)	87 (8.2)	10 (1.0)	51 (4.9) / 1011 (95)	46 (90) / 5 (10)	715 (71) / 296 (29)	51 (4.8)	5 (0.5)
5	114 (11) / 948 (89)	101 (89) / 13 (11)	660 (70) / 288 (30)	114 (11)	13 (1.2)	63 (6.1) / 999 (93.9)	56 (89) / 7 (11)	705 (71) / 294 (29)	63 (5.9)	7 (0.7)
6	143 (13) / 919 (87)	126 (88) / 17 (12)	635 (69) / 284 (31)	143 (13)	17 (1.6)	80 (7.5) / 982 (92.5)	72 (90) / 8 (10)	689 (70) / 293 (30)	80 (7.5)	8 (0.8)
7	184 (17) / 878 (83)	159 (86) / 25 (14)	602 (69) / 276 (31)	184 (17)	25 (2.3)	102 (9.7) / 960 (90.3)	93 (91) / 9 (9)	668 (70) / 292 (30)	102 (9.6)	9 (0.8)
8	225 (21) / 822 (79)	195 (87) / 30 (13)	566 (68) / 271 (32)	225 (21)	30 (2.8)	129 (10) / 933 (90)	116 (90) / 13 (10)	645 (69) / 288 (31)	129 (12.1)	13 (1.2)
9	259 (24) / 788 (75)	223 (86) / 36 (14)	538 (67) / 265 (33)	259 (24)	36 (3.4)	146 (14) / 916 (86)	132 (90) / 14 (10)	629 (69) / 287 (31)	146 (14)	14 (1.3)

SUPPLEMENTARY TABLE S3. Continued

MSKCC 2018

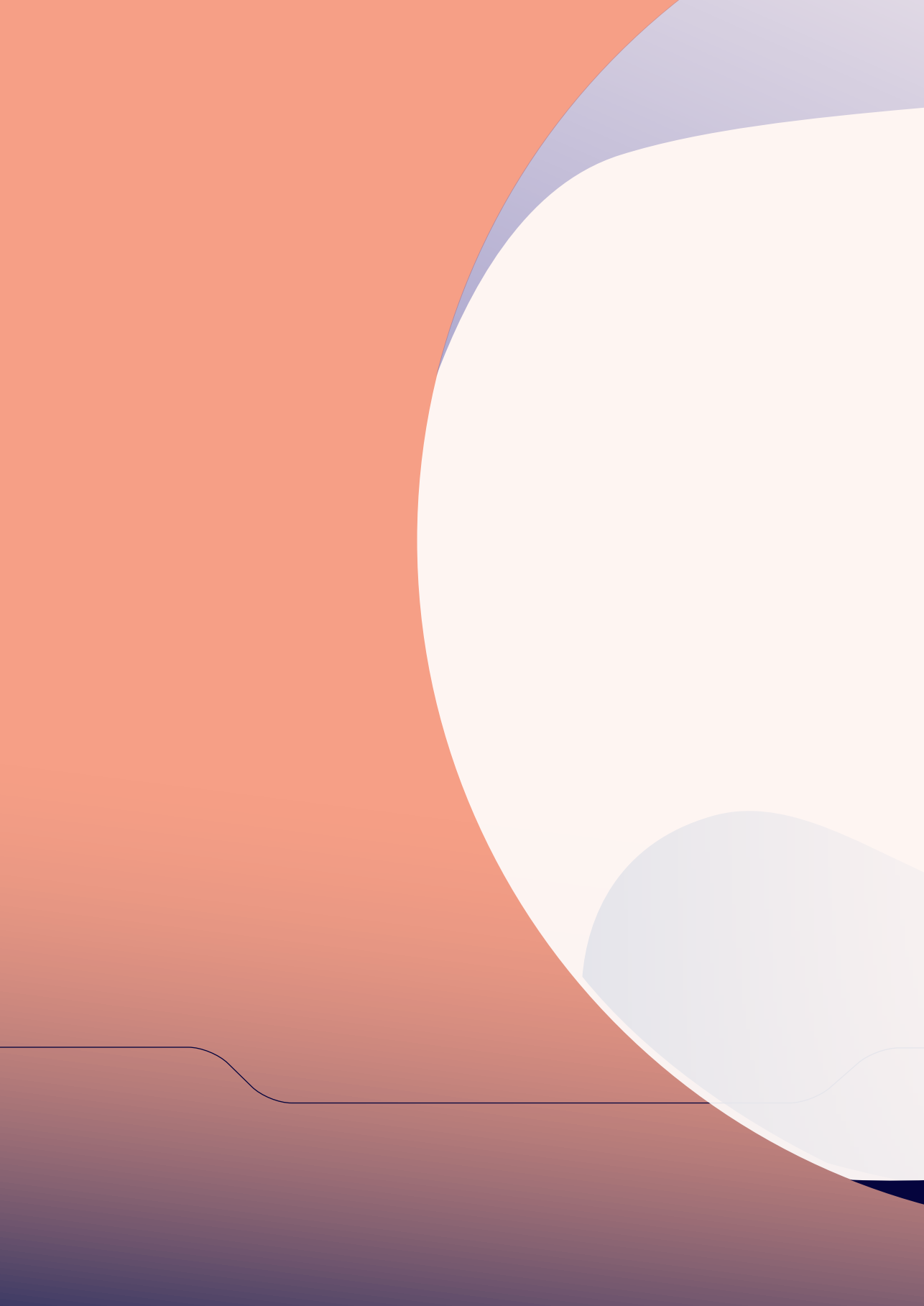
Threshold (%)	**DRE T-stage**					**mpMRI T-stage**				
	Patients in whom ePLND is not / is recommended (%)	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without / with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**	**Patients in whom ePLND is not / is recommended (%)**	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without/with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**
10	286 (27) / 776 (73)	247 (86) / 39 (14)	514 (66) / 262 (34)	286 (27)	39 (3.7)	179 (17) / 883 (83)	160 (89) / 19 (11)	601 (68) / 282 (32)	179 (17)	19 (1.8)
11	323 (30) / 739 (70)	276 (85) / 47 (15)	485 (66) / 254 (34)	323 (30)	47 (4.4)	209 (20) / 853 (80)	186 (89) / 23 (11)	575 (67) / 278 (33)	186 (18)	23 (2.2)
12	360 (34) / 702 (66)	311 (86) / 49 (14)	450 (64) / 252 (36)	360 (34)	49 (4.6)	241 (23) / 821 (77)	212 (88) / 29 (12)	549 (67) / 272 (33)	212 (20)	29 (2.7)
13	393 (37) / 669 (63)	338 (86) / 55 (14)	423 (63) / 246 (37)	393 (37)	55 (5.2)	264 (25) / 798 (75)	230 (87) / 34 (13)	531 (67) / 267 (33)	230 (22)	34 (3.2)
14	426 (40) / 636 (60)	366 (86) / 60 (14)	395 (62) / 241 (38)	426 (40)	60 (5.6)	294 (28) / 768 (72)	256 (87) / 38 (13)	505 (66) / 263 (34)	256 (24)	38 (3.6)
15	451 (42) / 611 (58)	381 (85) / 70 (16)	380 (62) / 231 (38)	451 (42)	70 (6.6)	323 (30) / 739 (70)	282 (87) / 41 (13)	479 (65) / 260 (35)	282 (27)	41 (3.9)

SUPPLEMENTARY TABLE S4. Systematic analysis of Briganti 2012 for nomogram-derived cut-offs

Briganti 2012										
Threshold (%)	**DRE T-stage**					**MRI T-stage**				
	Patients in whom ePLND is not / is recommended (%)	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without/with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**	**Patients in whom ePLND is not / is recommended (%)**	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without/with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**
1	7 (0.7) / 1055 (99.3)	5 (71) / 2 (29)	756 (72) / 299 (28)	7 (0.7)	2 (0.2)	1 (0.1) / 1061 (99.9)	1 (100) / 0 (0)	761 (72) / 301 (28)	1 (0.1)	0 (0)
2	44 (4.1) / 1018 (95.9)	40 (90.9) / 4 (9.1)	721 (71) / 297 (29)	44 (4.2)	4 (0.4)	15 (1.4) / 1047 (98.6)	14 (93.3) / 1 (6.7)	747 (71) / 300 (29)	15 (1.4)	1 (0.1)
3	84 (7.9) / 978 (92.1)	75 (89) / 9 (11)	686 (70) / 292 (30)	84 (8.0)	9 (0.8)	39 (3.7) / 1023 (96.3)	35 (89.7) / 4 (10.3)	726 (71) / 297 (29)	39 (3.7)	4 (0.4)
4	142 (13) / 920 (87)	125 (88) / 17 (12)	636 (69) / 284 (31)	142 (13)	17 (1.6)	66 (6.2) / 996 (93.8)	60 (91) / 6 (9)	701 (70) / 295 (30)	66 (6.3)	6 (0.6)
5	181 (17) / 881 (83)	162 (90) / 19 (11)	599 (68) / 282 (32)	181 (17)	19 (1.8)	87 (8.2) / 975 (91.8)	79 (91) / 8 (9)	682 (70) / 293 (30)	87 (8.2)	8 (0.8)
6	233 (22) / 829 (78)	201 (86) / 32 (14)	560 (68) / 269 (32)	233 (22)	32 (3.0)	121 (11) / 941 (89)	109 (90) / 12 (10)	652 (69) / 289 (31)	121 (12)	12 (1.1)
7	278 (26) / 784 (74)	241 (87) / 37 (13)	520 (66) / 264 (34)	278 (26)	37 (3.5)	158 (15) / 904 (85)	142 (90) / 16 (10)	619 (68) / 285 (32)	158 (15)	16 (1.5)
8	338 (32) / 724 (68)	291 (86) / 47 (14)	470 (65) / 254 (35)	338 (32)	47 (4.4)	193 (19) / 869 (81)	171 (89) / 22 (11)	590 (68) / 279 (32)	193 (18)	22 (2.1)

SUPPLEMENTARY TABLE S4. Continued

Briganti 2012										
Threshold (%)	**DRE T-stage**					**MRI T-stage**				
	Patients in whom ePLND is not / is recommended (%)	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without/with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**	**Patients in whom ePLND is not / is recommended (%)**	**Patients below cut-off without / with LNI (%)**	**Patients above cut-off without/with LNI (%)**	**ePLND spared (%)**	**LNI missed (%)**
9	381 (36) / 681 (64)	325 (85) / 56 (15)	436 (64) / 245 (36)	381 (36)	56 (5.3)	229 (22) / 833 (78)	204 (89) / 25 (11)	557 (67) / 276 (33)	229 (22)	25 (2.4)
10	416 (39) / 646 (61)	353 (85) / 63 (15)	408 (63) / 238 (37)	416 (39)	63 (5.9)	262 (25) / 800 (75)	234 (89) / 28 (11)	527 (66) / 273 (34)	262 (25)	28 (2.6)
11	456 (43) / 606 (57)	388 (84) / 68 (15)	373 (62) / 233 (38)	456 (43)	68 (6.4)	305 (29) / 757 (71)	270 (89) / 35 (11)	491 (65) / 266 (35)	305 (29)	35 (3.3)
12	493 (46) / 569 (54)	418 (85) / 75 (15)	343 (60) / 226 (40)	493 (46)	75 (7.1)	330 (31) / 732 (69)	288 (87) / 42 (13)	473 (65) / 259 (35)	330 (31)	42 (4.0)
13	520 (49) / 542 (51)	440 (85) / 80 (15)	321 (59) / 221 (41)	520 (49)	80 (7.5)	362 (34) / 700 (66)	316 (87) / 46 (13)	445 (64) / 255 (36)	362 (34)	46 (4.3)
14	555 (52) / 507 (48)	469 (85) / 86 (15)	292 (58) / 215 (42)	555 (52)	86 (8.1)	394 (37) / 668 (63)	342 (87) / 52 (13)	419 (63) / 249 (37)	394 (37)	52 (4.9)
15	582 (55) / 480 (45)	491 (84) / 91 (16)	270 (56) / 210 (44)	582 (55)	91 (8.6)	429 (40) / 633 (60)	373 (87) / 56 (13)	388 (61) / 245 (39)	429 (40)	56 (5.3)



PART V: SUMMARY AND GENERAL DISCUSSION

CHAPTER 9

9

General Summary

Prostate cancer is a common disease, which can present in a wide variety of stages. Accurate staging of prostate cancer is crucial for determining the most suitable treatment strategy, resulting in the best possible outcomes. Optimising primary tumour staging and consequent treatment indication for either invasive or non-invasive therapeutic strategies are therefore crucial to ensure patients are offered the best treatment.

In this thesis, we explored a number of scientific topics within the context of prostate cancer, including: the evaluation of real-world active surveillance (AS) outcomes, focusing on patient selection criteria and follow-up intensity, the impact and value of magnetic resonance imaging (MRI) on staging as well as consequences of staging such as extent of surgery and implementation of MRI information into existing and novel prediction tools. In the following chapter, the findings of the performed studies presented in this thesis are summarized.

PART I. OPTIMISING ACTIVE SURVEILLANCE

In **Chapter 2**, we conclude that patient selection for AS in the real-world clinical setting substantially differs from the selection of patients in a trial setting. Our analysis revealed that 49% of the patients in a real-world Dutch AS cohort did not meet one or more of the inclusion criteria used in the largest ongoing AS study (Prostate Cancer Research International Active Surveillance Study [PRIAS]). We observed large differences in AS outcomes comparing PRIAS-eligible and PRIAS-ineligible patients. PRIAS-ineligible patients experienced tumour progression during AS significantly earlier and had a three-fold increased risk of unfavourable surgical pathology after deferred radical prostatectomy, compared with PRIAS-eligible patients. Of all PRIAS selection criteria, PSA density (PSAD) was found to be a significant predictor for developing metastasized prostate cancer. Patients with PSAD ≥0.2 ng/ml/ml at diagnosis had a three-fold higher risk of metastasis, compared with patients with PSAD <0.2 ng/ml/ml. The novel insights regarding AS outcomes of patients who are selected "off-protocol", can help guide patients and their treating physicians to determine if AS is a suitable treatment option, taking into account the patient's baseline characteristics and preferences.

In **Chapter 3**, we report that AS follow-up in daily practice is also less stringent than advised by the prevailing AS follow-up protocols. Overall, less than half of all patients (43%) evaluated underwent PSA testing and repeat biopsies concordant with the PRIAS protocol. In patients meeting all PRIAS criteria at baseline (PRIAS-eligible), less frequent PSA monitoring as well as less frequent repeat biopsy were not associated with an increased risk of metastasis. Whereas among patients who were PRIAS-ineligible, concordant monitoring (especially PSA testing) is strongly advised, as discordant follow-up was associated with an increased risk of developing metastasized prostate cancer. Our findings suggest that it may be possible to loosen follow-up intensity in selected patients

fulfilling all PRIAS inclusion criteria, whereas stringent follow-up remains advised in PRIAS-ineligible patients.

PART II. IMPACT OF MRI ON PROSTATE CANCER RISK CLASSIFICATION

In the study described in **Chapter 4,** we observed that multiparametric magnetic resonance imaging (mpMRI) has superior detection of non-organ-confined disease (stage ≥T3a), compared with DRE. Use of mpMRI T-stage for prostate cancer risk group classification leads to upstaging in one-third of patients. Overall, use of mpMRI resulted in a 9% higher cumulative rate of true positive and true negative stage ≥T3a test results compared with DRE. Since mpMRI results in a higher overall diagnostic accuracy for the detection of non-organ-confined disease, it should be preferred over DRE for determining clinical tumour stage. However, due to its limited specificity for the detection of non-organ-confined disease, clinicians should be aware that mpMRI can lead to overstaging and potential overtreatment of patients who harbor genuine low-risk disease.

PART III. ASSOCIATION BETWEEN NERVE SPARING AND POSITIVE SURGICAL MARGINS

Chapter 5 includes the description of an extensive multivariable analysis, using side-specific covariates, evaluating the association between nerve sparing robot-assisted radical prostatectomy (RARP) and the risk of positive margins. In this multi-centre cohort, including analysis of 5148 prostate lobes derived from 2574 patients, nerve sparing was an independent predictor for ipsilateral positive margins (OR 1.42, 95% CI 1.14 – 1.82) on multivariable analysis. Our study results call into question the classic assumption that nerve sparing is not associated with an increased risk of positive surgical margins. The study's main finding is relevant for clinical practice, as patients and their urologists need to be aware of the fact that nerve sparing does increase the risk of positive margins. The potential benefits of nerve sparing should therefore be carefully weighed against the concomitant risks when consulting patients opting for nerve sparing radical prostatectomy.

PART IV. INCORPORATION OF MRI INTO CLINICAL PREDICTION MODELS

Since MRI has the potential to improve the predictive ability of nomograms, it should be evaluated if incorporation of MRI information into nomograms combined with traditional characteristics such as PSA and biopsy data can improve side-specific

extraprostatic extension (EPE) risk prediction. Martini and colleagues have developed such a nomogram, and concluded that it can lead to reliable prediction of side-specific EPE in patients undergoing RARP. To determine if this nomogram can be safely applied in other populations, we performed an external validation of the Martini nomogram for the prediction of side-specific EPE (**Chapter 6**). Although external validation showed good discrimination in terms of AUC (0.78), the model showed substantial disagreement between predicted and observed probabilities on the calibration plots. Use of a miscalibrated nomogram (either underestimation or overestimation) can be potentially harmful for patients. Therefore, we concluded that based on this external validation study, the Martini nomogram may not be a suitable prediction tool to predict side-specific EPE in patients undergoing RARP.

In **Chapter 7,** we describe the outcomes of a study in which we developed and externally validated four nomograms for the prediction of side-specific EPE, using combinations of PSAD, highest ipsilateral biopsy International Society of Urological Pathology (ISUP) grade, ipsilateral percentage of positive cores on systematic biopsy and side-specific clinical stage assessed by both digital rectal examination and mpMRI. The three models including mpMRI staging information resulted in relatively higher AUCs compared with the model without mpMRI information. The model based on PSAD, ISUP grade and mpMRI T-stage was superior in terms of model calibration. Using this model with a cut-off of 20%, 1980/2908 (68%) of all prostatic lobes without EPE would be found eligible for nerve sparing, whereas non-nerve sparing would be advised in 642/832 (77%) lobes with EPE. Our study resulted in a simple and robust nomogram, including mpMRI information and readily available clinical parameters, for the prediction of side-specific EPE. This nomogram can be safely used in clinical practice and can potentially improve patient selection for nerve sparing radical prostatectomy.

In **Chapter 8**, we evaluated the impact of using mpMRI T-stage on the diagnostic accuracy of two commonly used nomograms (MSKCC 2018 and Briganti 2012) for the prediction of pelvic lymph node involvement (LNI) in patients with prostate cancer undergoing extended pelvic lymph node dissection. We concluded that use of mpMRI T-stage improved model discrimination of both nomograms, compared with DRE T-stage. Model calibration was comparable for both modalities. Use of mpMRI T-stage and the Briganti 2012 nomogram resulted in the overall best model performance in terms of discrimination, calibration and net benefit. These results are important, since scientific data regarding the use of mpMRI T-stage for nomogram-based LNI risk prediction by existing nomograms are scarce. Based on the study's main results, we can conclude that mpMRI T-stage can be safely used as an input parameter for nomogram-based LNI risk prediction.

CHAPTER 10

10

General Discussion and Future Perspectives

In this thesis, we have evaluated patient selection and follow-up for active surveillance in a real-world patient cohort, and related selection and follow-up intensity to oncological outcomes. Additionally, we evaluated the impact of incorporating prostate MRI information into risk classification systems and nomograms to assist primary staging of prostate cancer, and treatment consequences of MRI including extent of surgery. This chapter is a discussion of the main findings and their implications, and future perspectives.

EVALUATION OF ACTIVE SURVEILLANCE IN REAL-WORLD CLINICAL PRACTICE

Our analysis revealed that in real-world practice, 49% of patients on AS for prostate cancer are selected "off-protocol"; not fulfilling the traditional eligibility criteria used in the landmark AS studies such as the Prostate Cancer International Active Surveillance Study (PRIAS) (Chapter 2).[1] Although number of studies performed describing the outcomes of "off-protocol" (PRIAS-ineligible) selected patients are scarce, our findings implicate that AS is frequently advised in these patients, even though outcomes of AS in these patients are relatively unknown.

In order to fill this scientific gap, we established rates of metastasis, risk of biochemical recurrence and outcomes of deferred surgery for both PRIAS-eligible and PRIAS-ineligible patients. These results show that PRIAS-ineligible patients experienced tumour progression significantly earlier than PRIAS-eligible patients, and had a three-fold increased risk of unfavourable surgical pathology after deferred radical prostatectomy. Our study also resulted in the identification of PSA density (PSAD) as an important prognostic parameter for unfavourable outcomes. For instance, PSAD ≥0.2 ng/ml/ml at diagnosis was associated with a three-fold higher risk of metastasis, compared with patients with a PSAD <0.2 ng/ml/ml. Identification of such clinically relevant predictors of AS outcomes is important, since these parameters and their related outcomes enable more individualized treatment decision-making.

A limitation of the study is that the vast majority of patients (97%) were diagnosed by transrectal ultrasonography guided systematic biopsy, without prebiopsy MRI and target biopsy.[1] As systematic biopsy is associated with missing one-third of clinically significant tumours on initial biopsy, this may limit the generalisability of the results to the contemporary population of patients on AS in whom pre-biopsy MRI and subsequent target biopsy are increasingly used.[2] Another important limitation of this study includes the lack of comparative analysis of PRIAS-ineligible patients on AS with a matched patient cohort undergoing immediate radical treatment. Consequently, we were not able to conclude if better overall oncological outcomes would be achieved if this subset would undergo immediate active treatment. Future studies should focus on revealing the most optimal treatment strategy for PRIAS-ineligible, low to favourable intermediate-risk

disease patients, preferably by a randomised controlled trial assigning patients to either AS or immediate active treatment.

In addition to adherence to less stringent selection criteria for AS, our analysis revealed that AS follow-up in routine clinical care is also less strict compared to that protocolled in most AS cohort studies (Chapter 3). Overall, less than half (43%) of all patients received follow-up testing concordant with the PRIAS protocol.[3] Our study's main finding was that in PRIAS-eligible patients, discordant monitoring was not associated which an increased risk of metastasis. However, our analysis also revealed that discordant monitoring in those not meeting the PRIAS criteria at baseline, did increase the risk of metastasis. These results contribute to our joint effort to reduce burden of patients with prostate cancer opting for AS, and specific recommendations can be extracted from the study. For instance, our study's main findings suggest follow-up schedules in PRIAS-eligible patients may be loosened; whereas strict follow-up is recommended in patients not meeting one or more PRIAS inclusion criteria.

A limitation of the study includes the lack of standard use of MRI during AS follow-up. At present, it is advised that MRI should be performed at least once at some point during AS follow-up.[4] Our study was performed before these recommendations were published, resulting in the relatively high percentage of patients on AS who were monitored without application of MRI (55%).[3] This could impact the generalisability of our outcomes to the contemporary state of practice. In addition, evaluation of the use of MRI and/or subsequent MRI target biopsy during follow-up were not included in the analysis. As we assume that MRI in this retrospective cohort was done in case of an indication (e.g. signs of tumour progression), application of MRI during follow-up in selected patients could impact the association between discordant follow-up and AS outcome, and could have masked the true impact of discordant follow-up.

IMPACT OF MAGNETIC RESONANCE IMAGING ON PROSTATE CANCER RISK CLASSIFICATION

Use of mpMRI T-stage, instead of DRE T-stage, would lead to migration of one-third of newly diagnosed patients to a higher prostate cancer risk group (Chapter 4). The stage migration is consequential to the higher detection rates of non-organ-confined disease of mpMRI, compared with DRE. The higher detection rates are countered by higher rates of false positive test results, especially in patients with low-risk disease. However, given the overall 9% higher cumulative rate of true positive and true negative detected non-organ-confined disease cases at diagnosis, mpMRI should be regarded as a superior to DRE as a staging and risk classification tool.

Question remains if these results provide sufficient evidence for fully abandoning DRE for local tumour staging. The upstaging of one-third of patients could translate into unnecessary treatment intensification in patients low- and intermediate-risk cancer, especially in those with false-positive stage ≥T3a test results. Although DRE is

an unreliable tool to exclude stage ≥T3a cancer, it still remains a valuable tool with a high specificity and thus high true positive rate for non-organ-confined disease, and should remain to be used in concert with mpMRI.
Moving forward, novel strategies to further optimise mpMRI interpretation should be explored to decrease false positive stage ≥T3a test results and thus improve its prognostic value. Areas of opportunity include the reporting of the likelihood of extraprostatic extension (EPE) by using a 5-point Likert scale and standardization of the degree of EPE when definite EPE on mpMRI is established.[5–8] In addition, future studies should focus on evaluating whether local staging and risk stratification assessed by mpMRI improves oncological as well as patient-reported outcomes.

Association between nerve sparing and positive surgical margins

In Chapter 5 we report that side-specific nerve sparing during robot-assisted radical prostatectomy (RARP) was associated with a 40% increased odds of ipsilateral positive margins versus non-nerve sparing. Reason this association was not found in the majority of previously performed studies on this subject could be explained by insufficient adjustment for (side-specific) confounding factors.[9,10]

Although these results can contribute to patient counselling and indication for nerve sparing, the study has some limitations. The retrospective nature of our study, including the performance of nerve sparing based on patient selection, is not the most ideal study design to evaluate this issue. To further evaluate the association, prospective randomised trials are needed wherein patients are randomised into nerve sparing versus non-nerve sparing. However, this design may not be justified from an ethical perspective as retaining urinary continence and erectile function is withheld in those randomised into the non-nerve sparing group.

In addition to revealing the association between nerve sparing and risk of positive margins, our study led to the identification of other significant predictors of positive margins including presence of EPE on mpMRI and surgeon. To further improve the safety of nerve sparing surgery and thus improving outcomes for patients, quality improvement initiatives should focus on surgical training and surgical outcome feedback programmes, creating an environment wherein surgeons evaluate their procedures and learn from each other. Moreover, clinicians should focus on optimising their diagnostic and staging modalities. Local multi-disciplinary quality improvement programmes involving urologists, pathologists and radiologists can improve mpMRI reading and interpretation, further enhancing the detection and exclusion of EPE on pre-operative MRI enabling optimal preoperative surgical planning.

Incorporating MRI information into clinical prediction models

The main finding of this section is that incorporation of mpMRI information into preoperative prostate cancer nomograms positively impacts their predictive accuracy

(Chapter 7 & Chapter 8). However, inherent to use of prediction tools in general, assessment of performance of these models in external patient populations remains important to ensure their safety. This was emphasized by the results of our study described in Chapter 6, showing disagreement between predicted and observed rates of EPE when using the Martini EPE nomogram for the prediction of side-specific EPE in an external hospital population. These findings were concordant with those reported by Sighinolfi et al.[11] Based on the findings of these studies we conclude that the Martini nomogram may not be a suitable prediction tool to predict side-specific EPE in patients undergoing RARP.

The major strength of our study in which we propose an alternative side-specific EPE nomogram, is that we used two separate hospital patient populations for external validation (Chapter 7). This nomogram is the first side-specific EPE prediction tool including mpMRI information, showing good performance when applied in other hospital populations. Therefore, this tool should be recommended for side-specific EPE risk prediction in clinical practice. The nomogram is currently publicly accessible as an online web calculator.[12]

Although our side-specific EPE nomogram has shown to be an accurate tool, two important limitations of the study should be mentioned. First, there was a wide variety of biopsy protocols used in the development cohort. Overall 56% were diagnosed using systematic biopsy, 16% using MRI target biopsy without concomitant systematic biopsy and 28% underwent target biopsy and concomitant systematic biopsy. Due to the increased adoption of MRI-guided target biopsy, it can be expected that sampling error may further decrease, leading to more precise biopsy ISUP grading. Therefore, future updating of the nomogram is crucial to ensure the applicability in the contemporary era of clinical practice. It is advised that updating occurs by using data of a more contemporary cohort including men all receiving target and concomitant systematic biopsies. The second limitation is the lack of central review of MRI, which could have improved the diagnostic accuracy of MRI for the detection of EPE.[13] However, variations in MRI reading (quality and interobserver variability) are inherent to real-world clinical care, and central review is mostly not incorporated into daily clinical practice. Also, by using a multi-centre hospital population for model development, these variations were already partially accounted for.

In Chapter 8, we reported that use of mpMRI T-stage improved model discrimination of two established nomograms used for the prediction of pelvic lymph node invasion (LNI) in patients with primary prostate cancer, compared with DRE T-stage. These results are important, since studies evaluating the added value of using mpMRI T-stage for LNI risk prediction are scarce. Based on the study's main results, we can conclude that mpMRI T-stage can safely replace DRE T-stage as input parameter for nomogram-based LNI risk prediction. Use of mpMRI T-stage led to relatively higher probabilities of LNI, due to migration of patients initially staged as cT1c or cT2 disease by DRE, to cT2 and higher. When using mpMRI T-stage, lower rates of pelvic LNI would be missed (higher sensitivity), compared with DRE T-stage. However, this was at the cost of performing

more lymph node dissections in patients without nodal metastasis (higher false positive rate). Although our analysis showed that mpMRI T-stage is a valuable alternative to DRE T-stage for nomogram-based risk prediction, clinicians should be aware of the stage migration use of mpMRI T-stage initiates, potentially resulting in increased number of unnecessary lymph node dissections. The selection of patients for extended pelvic lymph node dissection (ePLND), using set thresholds (with variation allowed between hospitals and urologists, between 5% to 15%) for risk of LNI, forms an important limitation of the present study; leading to a study population with a relatively high prevalence of LNI. This aspect could limit the generalisability of our study results to patients with a predicted LNI risk for the lower thresholds (0 – 10%). In addition, the selection of patients can also explain the counterintuitive findings regarding the net benefit; as decision curve analysis showed that a "treat all" approach (subjecting all patients to ePLND), would be superior compared to nomogram-based selection. Nevertheless, although the study regards a selected subgroup, we assume the population to be suitable for testing the primary hypothesis; confirming mpMRI T-stage is at least equal to DRE T-stage when used for nomogram-based LNI risk prediction.

FUTURE PERSPECTIVES

Optimisation of active surveillance in the present MRI era

Data of studies with MRI-based AS protocols will reach greater maturity in the near future, yielding valuable information regarding the added value of MRI application in AS. The preliminary outcomes indicate that baseline MRI can further discriminate indolent tumours from those likely to become clinically significant.[14] In addition, performing MRI before entering AS was associated with lower rates of discontinuation at 15 months (14% versus 26-28% in MRI-naïve men).[15] Use of MRI-based risk stratification tools and scoring systems such as PI-RADS and PRECISE can help to reduce the number of necessary repeat biopsies, decreasing AS burden for the patient.[14,16,17] The mid-term results of an AS cohort including full MRI-led monitoring are promising, showing comparable discontinuation, mortality and metastasis rates compared with those reported in standard AS, while drastically reducing the number of repeat biopsies.[18] Although the long-term results are to be awaited, these results show the potential of full MRI-based AS monitoring.

Impact of MRI on prostate cancer pre-treatment risk prediction tools

Application of MRI and target biopsy information are expected to further shift the paradigm of predictive modelling in prostate cancer. It has been established previously that application of MRI-guided target biopsies, combined with concomitant systematic biopsies, leads to decreased ISUP upgrading rates at final pathology.[19,20] The more accurate the ISUP grade can be assessed preoperatively, the more accurate risk prediction by

nomograms incorporating ISUP grade as a predictor can be established. The latter was confirmed by a literature review of Dell'Oglio and colleagues.[21] The authors provided an overview of several studies confirming that MRI and MRI-guided target biopsy information improve prostate cancer risk calculators developed for the prediction of pathological outcomes after surgery (EPE, SVI and LNI) and the risk of biochemical recurrence after radical prostatectomy.[21]

These findings were confirmed in a recent study by Gandaglia and colleagues, showing that novel MRI-based pre-operative tools improved the prediction of unfavourable pathological outcomes (EPE and seminal vesicle invasion), in terms of model discrimination.[22] In addition, the same study group proposed a novel nomogram for the prediction of LNI, based on MRI target biopsy and MRI information. In their development study, they observed substantial higher model accuracy compared with conventional MRI-naive LNI prediction tools.[23] The latter was confirmed in an additional external validation study, showing that use of the mpMRI and mpMRI target biopsy based nomogram resulted in the highest AUC compared to the 2012 and 2017 Briganti nomograms and the MSKCC 2018 pre-radical prostatectomy nomogram (79% vs. 75% vs. 65% vs. 74%).[24] Another novel nomogram including MRI information for prediction of LNI was developed bij Draulans and colleagues.[25] This nomogram also showed favourable characteristics on internal validation (AUC 0.80), and acceptable characteristics on external validation (AUC: 0.73).[25] Main advantage of this nomogram is it can be used independently of biopsy technique, since the nomogram does not include MRI target or systematic biopsy specific variables. It can be expected that the number of mpMRI-based nomograms for the prediction of prostate cancer treatment outcomes and pathological features will continue to grow. The most important question that future studies should attempt to answer in this context is whether use of these improved pre-treatment tools also leads to better clinical and patient-reported outcomes.

PSMA PET-CT Imaging

Rapid evolvements currently ongoing in the field of prostate membrane specific positron emission computer tomography (PSMA PET-CT) are promising, and are expected to greatly impact prostate cancer staging. In a study by Hofman and colleagues, PSMA PET-CT outperformed conventional imaging (CT and bone scan) with regard to detection of metastasis at initial prostate cancer staging.[26] The use of PSMA PET-CT in high-risk patients led to 27% greater accuracy for the detection of pelvic nodal or distant metastatic disease compared with CT and bone scan.[26] PSMA PET-CT also resulted in significant higher rates in change of management (23% vs. 7%).[26]

Unarguably, adoption of PSMA PET-CT and standard implementation within the diagnostic pathway will result in stage migration of the general patient population; especially those classified with high-risk disease. Given the widespread adoption of PSMA PET-CT, it is to be awaited if nomograms predicting the probability of LNI remain relevant for clinical practice. For instance, in case PSMA PET-CT shows pelvic

lymph node metastasis at primary staging, use of a nomogram to predict the risk of LNI will not impact change of management. However, since sensitivity of PSMA PET-CT at present remains moderate (40%), more than half of patients with pelvic LNI remain node-negative on PSMA PET-CT.[27–29] For these patients, nomograms can still be of value to decide whether ePLND should be performed. Furthermore, recent data of Meijer and colleagues showed that combining mpMRI information with PSMA PET-CT outcomes can improve the negative predictive value.[30] In case the number of patients undergoing PSMA PET-CT increases, sufficient data will become available to develop novel nomograms including PSMA PET-CT information. Combined with mpMRI data, such as the PI-RADS score and the mpMRI T-stage, this can potentially lead to promising new prediction tools further improving the staging of primary prostate cancer.

Artificial intelligence and predictive modelling

The traditional risk prediction tools are mostly based using on multivariable logistic regression analysis. Advanced statistical and mathematical techniques, such as machine learning and deep learning algorithms, can further enhance risk prediction. The main limitation of a logistic regression based prediction tool is that the outcome is rather static.[31] For instance, one can calculate the risk of pelvic LNI in newly diagnosed patients based on clinical stage, PSA, relative number of positive cores and the highest ISUP grade on biopsy. However, if information derived from other staging modalities, not included in the model becomes available, the nomogram-calculated risk should be interpreted differently. For example, if the nomogram-calculated risk of LNI in a patient is 10%, but subsequent PSMA PET-CT shows high suspicion for LNI, this risk of LNI has drastically increased.

The advantage of machine learning tools is that they can be updated faster and more efficient by delivering novel data; compared to traditional logistic regression models. For developing a novel model; specific steps should be undertaken every time which are time-consuming.[32] A machine learning tool enabling immediate model updating and validation based on recent data collected during routine clinical care, can substantially speed up the development of novel highly accurate prediction tools.

Value-Based Healthcare

The Santeon Value-Based Healthcare Initiative (VBHC) for prostate cancer formed an excellent basis for the research presented in this thesis. This quality improvement initiative, started in 2012, had resulted in interesting insights as well as accumulated data proven to be very useful for scientific research. Although it remains difficult to define the added value of such quality improvement initiatives, the collaboration itself has proven to be a successful starting point for several research initiatives. In general, the adoption of VBHC as a concept is growing on a national and international level. Within Santeon, number of staff members involved in VBHC teams are growing. During the study period

of this thesis, substantial progression has been made regarding the data-infrastructure and data quality. It can be expected that the interaction between the quality improvement initiatives and research projects will remain successful in the future. The studies described in this thesis are exemplary for the way VBHC and research can complement each other. By using the novel insights presented in this thesis for quality improvement initiatives, quality of prostate cancer care will hopefully further improve and become more efficient. Prolongation of VBHC and its included multidisciplinary meetings ensures reflection on the quality of clinical prostate cancer care, and enables the identification of areas requiring improvement. The insights gained by these evaluations will hopefully keep inspiring other researchers and lead to novel valuable prostate cancer studies.

REFERENCES

1. Soeterik TFW, van Melick HHE, Dijksman LM, Biesma DH, Witjes JA, van Basten JPA. Active Surveillance for Prostate Cancer in a Real-life Cohort: Comparing Outcomes for PRIAS-eligible and PRIAS-ineligible Patients. *Eur Urol Oncol*. 2018;1:231-237.
2. Mottet N, van den Bergh RCN, Briers E, et al. EAU - ESTRO - ESUR - SIOG Guidelines on Prostate Cancer 20202; presented at the EAU Annual Congress Amsterdam 2020. Arnhem, The Netherlands: European Association of Urology Guidelines Office.
3. Soeterik TFW, van Melick HHE, Dijksman LM, Biesma DH, Witjes JA, van Basten JPA. Follow-up in Active Surveillance for Prostate Cancer: Strict Protocol Adherence Remains Important for PRIAS-ineligible Patients. *Eur Urol Oncol*. 2019;2:483-489.
4. Lam TBL, MacLennan S, Willemse PPM, et al. EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel Consensus Statements for Deferred Treatment with Curative Intent for Localised Prostate Cancer from an International Collaborative Study (DETECTIVE Study). *Eur Urol*. 2019;76:790-813.
5. Freifeld Y, Diaz de Leon A, Xi Y, et al. Diagnostic Performance of Prospectively Assigned Likert Scale Scores to Determine Extraprostatic Extension and Seminal Vesicle Invasion With Multiparametric MRI of the Prostate. *AJR Am J Roentgenol*. 2019;212:576-581.
6. Ball MW, Partin AW, Epstein JI. Extent of extraprostatic extension independently influences biochemical recurrence-free survival: Evidence for further PT3 subclassification. *Urology*. 2015;85(1):161-164.
7. Jeong BC, Chalfin HJ, Lee SB, et al. The relationship between the extent of extraprostatic extension and survival following radical prostatectomy. *Eur Urol*. 2015;67:342-346.
8. Onay A, Ertas G, Vural M, Colak E, Esen T, Bakir B. The role of T2-weighted images in assessing the grade of extraprostatic extension of the prostate carcinoma. *Abdom Radiol*. 2020;45:3293-3300.
9. Coelho RF, Chauhan S, Orvieto MA, Palmer KJ, Rocco B, Patel VR. Predictive Factors for Positive Surgical Margins and Their Locations After Robot-Assisted Laparoscopic Radical Prostatectomy. *Eur Urol*. 2010;57:1022-1029.
10. Lavery HJ, Nabizada-Pace F, Carlucci JR, Brajtbord JS, Samadi DB. Nerve sparing robotic prostatectomy in preoperatively high-risk patients is safe and efficacious. *Urol Oncol Semin Orig Investig*. 2012;30:26-32.
11. Sighinolfi MC, Sandri M, Torricelli P, et al. External validation of a novel side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer: preliminary outcomes on a series diagnosed with multiparametric magnetic resonance im. *BJU Int*. 2019;124:192-194.
12. Santeon mpmri-based nomogram to predict side-specific extraprostatic extension. https://www.evidencio.com/models/show/2142. Published 2020.
13. Wibmer A, Vargas HA, Donahue TF, et al. Diagnosis of extracapsular extension of prostate cancer on prostate MRI: Impact of second-opinion readings by subspecialized genitourinary oncologic radiologists. *Am J Roentgenol*. 2015;205:73-78.
14. Wang AZ, Oonnor LP, Yerram NK, et al. PIRADS Category as a Predictor of Progression to Unfavorable Risk Prostate Cancer in Men on Active Surveillance. *J Urol*. 2020;204:1229-1235.
15. Luiting HB, van Leeuwen PJ, Busstra MB, et al. Use of gallium-68 prostate-specific membrane antigen positron-emission tomography for detecting lymph node metastases in primary and recurrent prostate cancer and location of recurrence after radical prostatectomy: an overview of the current literature. *BJU Int*. 2020;125:206-214.
16. Alberts AR, Roobol MJ, Drost FJH, et al. Risk-stratification based on magnetic resonance imaging and prostate-specific antigen density may reduce unnecessary follow-up biopsy procedures in men on active surveillance for low-risk prostate cancer. *BJU Int*. 2017;120:511-519.

17. Moore CM, Giganti F, Albertsen P, et al. Reporting Magnetic Resonance Imaging in Men on Active Surveillance for Prostate Cancer: The PRECISE Recommendations—A Report of a European School of Oncology Task Force. *Eur Urol*. 2017;71:648-655.
18. Stavrinides V, Giganti F, Trock B, et al. Five-year Outcomes of Magnetic Resonance Imaging–based Active Surveillance for Prostate Cancer: A Large Cohort Study. *Eur Urol*. 2020;78:443-451.
19. Ploussard G, Beauval JB, Lesourd M, et al. Added Value of Concomitant Systematic and Fusion Targeted Biopsies for Grade Group Prediction Based on Radical Prostatectomy Final Pathology on Positive Magnetic Resonance Imaging. *J Urol*. 2019;202:1182-1187.
20. Ahdoot M, Wilbur AR, Reese SE, et al. MRI-Targeted, Systematic, and Combined Biopsy for Prostate Cancer Diagnosis. *N Engl J Med*. 2020;382:917-928.
21. Dell'Oglio P, Stabile A, Dias BH, et al. Impact of multiparametric MRI and MRI-targeted biopsy on pre-therapeutic risk assessment in prostate cancer patients candidate for radical prostatectomy. *World J Urol*. 2019;37:221-234.
22. Gandaglia G, Ploussard G, Valerio M, et al. The Key Combined Value of Multiparametric Magnetic Resonance Imaging, and Magnetic Resonance Imaging–targeted and Concomitant Systematic Biopsies for the Prediction of Adverse Pathological Features in Prostate Cancer Patients Undergoing Radical Prostatect. *Eur Urol*. 2020;77:733-741.
23. Gandaglia G, Ploussard G, Valerio M, et al. A Novel Nomogram to Identify Candidates for Extended Pelvic Lymph Node Dissection Among Patients with Clinically Localized Prostate Cancer Diagnosed with Magnetic Resonance Imaging-targeted and Systematic Biopsies. *Eur Urol*. 2019;75:506-514.
24. Gandaglia G, Martini A, Ploussard G, et al. External Validation of the 2019 Briganti Nomogram for the Identification of Prostate Cancer Patients Who Should Be Considered for an Extended Pelvic Lymph Node Dissection. *Eur Urol*. 2020;78:138-142.
25. Draulans C, Everaerts W, Isebaert S, et al. Development and External Validation of a Multiparametric Magnetic Resonance Imaging and International Society of Urological Pathology Based Add-On Prediction Tool to Identify Prostate Cancer Candidates for Pelvic Lymph Node Dissection. *J Urol*. 2020;203:713-718.
26. Hofman MS, Lawrentschuk N, Francis RJ, et al. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multicentre study. *Lancet*. 2020;395:1208-1216.
27. Yaxley JW, Raveenthiran S, Nouhaud F-X, et al. Yaxley - Outcomes of primary lymph node staging of intermediate and high risk prostate cancer with gallium PSMA.pdf. *J Urol*. 2019;201:815-820.
28. Van Kalmthout LWM, Van Melick HHE, Lavalaye J, et al. Prospective validation of Gallium-68 PSMA PET/CT in primary staging of prostate cancer patients. *J Urol*. 2020;203:537-545.
29. Jansen BHE, Bodar YJL, Zwezerijnen GJC, et al. Pelvic lymph-node staging with 18F-DCFPyL PET/CT prior to extended pelvic lymph-node dissection in primary prostate cancer - the SALT trial -. *Eur J Nucl Med Mol Imaging*. 2021;48:509-520.
30. Meijer D, De Barros H, Bodar YJL, et al. The Predictive Value of Preoperative Negative Prostate Specific Membrane Antigen Positron Emission Tomography Imaging for Lymph Node Metastatic Prostate Cancer. *J Urol*. 2021;205:1655-1662.
31. Vickers AJ, Fearn P, Scardino PT, Kattan MW. Why Can't Nomograms Be More Like Netflix? *Urology*. 2010;75:511-513.
32. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015;162:1-73.

APPENDICES

NEDERLANDSE SAMENVATTING

Prostaatkanker is een veel voorkomende ziekte. Een belangrijke maat voor het beloop van deze aandoening is het stadium waarin de patiënt zich bevindt. Accurate stadiëring van prostaatkanker is daarom van cruciaal belang voor het bepalen van de best passende behandelstrategie. In dit proefschrift hebben we de uitkomsten geëvalueerd van een grote patiëntenpopulatie uit de dagelijkse klinische praktijk. Hierbij hebben wij gekeken naar de uitkomsten van actief afwachtend beleid, gericht op tumorkenmerken ten tijde van diagnose en de intensiteit van de vervolgcontrole, de impact en waarde van *magnetic resonance imaging* (MRI) op stadiëring en de gevolgen van stadiëring zoals de uitgebreidheid van de chirurgische behandeling. Tevens hebben we onderzocht of het toevoegen van MRI informatie aan bestaande en nieuwe predictiemodellen leidt tot nauwkeurigere voorspellingen van relevante tumoruitkomsten. Dit hoofdstuk bevat de samenvatting van de bevindingen van de onderzoeken waaruit dit proefschrift bestaat.

DEEL I. OPTIMALISATIE VAN ACTIEF AFWACHTEND BELEID

Het is reeds bekend dat, wanneer patiënten aan bepaalde tumorkarakteristieken voldoen, er veilig een actief afwachtend beleid kan worden ingezet na de diagnose van prostaatkanker. Echter, de veiligheid van actief afwachtend beleid is tot op heden alleen aangetoond bij patiënten die aan specifieke karakteristieken voldeden, duidend op een laag-risico vorm van prostaatkanker. Uit ons onderzoek beschreven in Hoofdstuk 2 bleek dat in de dagelijkse praktijk echter breder geselecteerd wordt voor actief afwachtend beleid, dan tot op heden in de lopende onderzoeken werd gedaan. Uit onze analyse bleek dat ongeveer de helft van de patiënten (49%) niet voldeed aan een of meer van de inclusiecriteria die werden gebruikt in de grootste lopende studie over actief afwachtend beleid, de zogenaamde "PRIAS" studie. Patiënten die niet aan de inclusiecriteria voor deze studie voldeden (PRIAS-ongeschikte), maar waarbij wel een actief afwachtend beleid was geïnitieerd, ervaarden gemiddeld twee keer sneller tumorprogressie tijdens de vervolgcontrole, én hadden een drievoudig verhoogd risico op ongunstige pathologische uitkomsten na chirurgische prostaatverwijdering. Van alle afzonderlijke PRIAS-selectiecriteria bleek prostaat-specifieke antigeen densiteit (PSAD) de sterkste voorspeller te zijn voor het ontwikkelen van uitzaaiingen gedurende actief afwachtend beleid. De uitkomsten van deze studie dragen bij aan een beter begrip van het beloop van actief afwachtend beleid bij patiënten met hoger-risico kenmerken. De studieresultaten kunnen zorgverleners helpen om samen met de patiënt te bepalen of actief afwachtend beleid, op basis van individuele tumorkenmerken en voorkeuren, een veilige en geschikte behandelkeuze betreft.

In Hoofdstuk 3 rapporteren we dat de vervolgcontrole in de dagelijkse klinische praktijk óók minder streng is dan wordt geadviseerd in de protocollen van de lopende studies naar actief afwachtend beleid. In deze studie werd het studieprotocol van de PRIAS-

studie wederom als referentie gebruikt. Er werd onderzocht welk percentage van de patiënten actief werden vervolgd conform de richtlijnen uit de PRIAS-studie; uitgedrukt in aantal uitgevoerde PSA serum tests en prostaatbiopten. In totaal onderging minder dan de helft van de patiënten (43%) het aantal PSA tests en prostaatbiopten die zij conform het PRIAS-protocol zouden moeten ondergaan. Bij PRIAS-geschikte patiënten was discordante vervolgcontrole niét geassocieerd met een verhoogd risico op uitzaaiingen, wat suggereert dat de intensiteit van de controle schema's kan worden verlaagd in deze subgroep. Anderzijds blijkt het geobserveerde verhoogde risico op uitzaaiingen bij discordante vervolgcontrole onder PRIAS-ongeschikte patiënten een argument om in deze subgroep wél strikt het protocol te blijven volgen.

DEEL II. IMPACT VAN MRI OP DE PROSTAATKANKER RISICO-CLASSIFICATIE

Bij de diagnose van prostaatkanker wordt een patiënt op basis van bepaalde tumorfactoren ingedeeld in één van de drie risicogroepen: laag-, intermediair- of hoog-risico. Een belangrijk kenmerk waarop deze risicoclassificatie wordt gebaseerd is het tumorstadium. Van oudsher wordt het tumorstadium bepaald via rectaal toucher van de prostaat. Het gebruik van MRI ter bepaling van het tumorstadium heeft mogelijk voordelen ten opzichte van het rectaal toucher, aangezien de prostaat in volledigheid kan worden afgebeeld en dus geëvalueerd. Middels rectaal toucher kan echter alleen de achterzijde van de prostaat worden beoordeeld. In hoofdstuk 4 hebben we onderzocht wat het effect is van gebruik van MRI ter bepaling van het tumorstadium, en daarmee ook het effect op de risicoclassificatie. We vonden dat middels MRI uitbreiding van de tumor buiten de prostaat vaker kon worden gedetecteerd vergeleken met het rectaal toucher. Doordat tumoruitbreiding váker werd gedetecteerd, migreerde door gebruik van het MRI tumorstadium voor de risicoclassificatie, één op de drie patiënten naar een hogere risicogroep, vergeleken met wanneer het tumorstadium via rectaal toucher werd bepaald. Een nadeel van het gebruik van MRI voor het bepalen van het tumorstadium betrof het risico op een overschatting van het stadium. Zo werd er in enkele gevallen ten onrechte de aanwezigheid tumoruitbreiding buiten de prostaat vastgesteld, terwijl dit bij pathologische evaluatie niet aanwezig bleek te zijn. Echter, aangezien MRI over het algemeen resulteerde in een hoger totaal aantal correcte observaties van zowel aan- als afwezigheid van tumoruitbreiding vergeleken met het rectaal toucher, concludeerden we dat MRI de voorkeur geniet om het tumorstadium te bepalen. Zorgverleners dienen zich er wel van bewust te zijn dat in enkele gevallen MRI kan leiden tot overschatting van het tumorstadium met dientengevolge het inzetten van een te intensief behandelschema.

DEEL III. RELATIE TUSSEN ZENUWSPARING EN POSITIEVE SNIJVLAKKEN

Wanneer ter behandeling van prostaatkanker de prostaat operatief wordt verwijderd, kan er voor gekozen worden om de zenuwbundels die aan weerszijden van de prostaat lopen, intact te laten. Het "sparen" van deze zenuwen is mogelijk niet zonder risico. Doordat de zenuwen aan weerszijden vlak langs de prostaat lopen, wordt er dicht op de prostaat geopereerd en bestaat er de kans dat de tumor niet volledig wordt verwijderd; resulterend in een "positief" snijvlak. Het niet volledig verwijderen van de tumor kan ertoe leiden dat er vervolgbehandeling nodig is. In de overgrote meerderheid van de uitgevoerde wetenschappelijke onderzoeken, waarin de associatie tussen zenuwsparing en positieve snijvlakken werd onderzocht, werd geconcludeerd dat zenuwsparing niét geassocieerd was met een verhoogd risico op positieve snijvlakken. Deze contra-intuïtieve conclusie kan echter veroorzaakt worden door beperkingen in de methodologie van deze studies. In hoofdstuk 5 beschrijven we de resultaten van een uitgebreide analyse van de associatie tussen zenuwsparing en positieve snijvlakken, waarin een groot aantal factoren werden meegenomen. In de studie waarin data van 2574 patiënten uit vier ziekenhuizen werden geanalyseerd, observeerden wij dat zenuwsparing wél geassocieerd was met een anderhalf keer hóger risico op een positief snijvlak, vergeleken met een niet-zenuwsparende procedure. Onze studieresultaten trekken de klassieke suggestie, namelijk dat het sparen van zenuwen niet geassocieerd is met een verhoogd risico op positieve snijvlakken, in twijfel. Deze bevinding is relevant voor de klinische praktijk, aangezien patiënten en behandelend urologen zich bewust moeten zijn van het feit dat een zenuwsparende operatieve prostaatverwijdering wel degelijk gepaard kan gaan met verhoogd risico op positieve snijvlakken. Een accurate preoperatieve evaluatie van het tumorstadium is derhalve cruciaal om te verzekeren dat zenuwsparing veilig kan worden verricht.

DEEL IV. IMPLEMENTATIE VAN MRI IN KLINISCHE PREDICTIEMODELLEN

Aangezien MRI het potentieel heeft om het voorspellende vermogen van predictiemodellen te verbeteren, moet worden geëvalueerd of opname van MRI-informatie in bestaande predictiemodellen in combinatie met traditionele kenmerken zoals PSA en biopt gegevens, de voorspelling van lokale uitbreiding van de tumor buiten de prostaat kan verbeteren. Martini en collega's hebben een nomogram met gebruik van MRI-informatie ontwikkeld en geconcludeerd dat dit kan leiden tot een betrouwbare voorspelling van lokale tumoruitbreiding buiten de prostaat. Om te bepalen of dit nomogram veilig kan worden toegepast in andere ziekenhuispopulaties, hebben we de nauwkeurigheid van dit predictiemodel getoetst met behulp van data van patiënten met prostaatkanker die zijn geopereerd in het Canisius Wilhelmina Ziekenhuis Nijmegen (Hoofdstuk 6). Hoewel de evaluatie een goed discriminerend vermogen liet zien (het vermogen om een patiënt zónder tumoruitbreiding te onderscheiden van een patiënt mét tumoruitbreiding),

observeerden wij aanzienlijke discrepanties tussen door het model voorspelde kans op tumoruitbreiding en het daadwerkelijke procentuele voorkomen ervan in de patiëntpopulatie; de zogenaamde kalibratie. Het gebruik van een verkeerd gekalibreerd nomogram (onderschatting of overschatting van de voorspelde kans) kan schadelijk zijn voor patiënten wanneer deze gebruikt wordt voor de ondersteuning van medische besluitvorming. Op basis van deze studie lijkt het Martini-nomogram geen geschikt predictiemodel om accuraat tumoruitbreiding te voorspellen.

Gezien het belang van het betrouwbaar inschatten van de aanwezigheid van tumoruitbreiding, en daarmee de resterende behoefte aan betrouwbare predictiemodellen, hebben we een studie uitgevoerd waarin we een alternatief predictiemodel hebben ontwikkeld en getoetst in twee externe populaties. In Hoofdstuk 7 presenteren we vier verschillende predictiemodellen, bestaande uit een combinatie van tumorkenmerken die standaard bij iedere patiënt worden bepaald, waaronder de PSA densiteit, prostaatbiopsie informatie en het tumorstadium. Alhoewel het discriminerende vermogen vrijwel gelijkwaardig was onder de modellen, toonde het model gebaseerd op MRI tumorstadium PSA densiteit en biopsie tumor gradering de beste kalibratie. Het ontwikkelde nomogram kan derhalve veilig worden gebruikt om in te schatten of de tumor buiten de contour van de prostaat zich uitbreidt, en is daarmee zeer geschikt om het plan voor de operatie mee te ondersteunen.

Naast het voorspellen van tumoruitbreiding buiten de prostaat, is het voorspellen van de aanwezigheid van lymfklieruitzaaiingen ten tijde van diagnose een belangrijke component voor het bepalen van de behandelstrategie. De reeds bestaande predictie-modellen waarmee de kans op lymfklieruitzaaiingen kan worden voorspeld, bevatten informatie over het tumorstadium bepaald middels het rectaal toucher. Zoals eerder in Hoofdstuk 4 is beschreven, kan MRI helpen het tumorstadium nauwkeuriger te bepalen. We hebben daarom onderzocht of de voorspelling van de kans op lymfklieruitzaaiingen nauwkeuriger kan worden gedaan als het MRI tumorstadium wordt toegepast, in plaats van het rectaal toucher tumorstadium. We concludeerden dat wanneer het MRI tumorstadium wordt gebruikt voor de risicovoorspelling, er mínder lymfklieruitzaaiingen worden gemist, vergeleken met het gebruik van het rectaal toucher tumorstadium. Deze resultaten zijn belangrijk, aangezien er nog maar beperkt informatie over de impact van het toepassen van MRI informatie in bestaande predictiemodellen beschikbaar is. Op basis van deze onderzoeksresultaten kunnen we concluderen dat het MRI tumorstadium veilig kan worden gebruikt als invoerparameter voor de voorspelling van de kans op lymfkliermetastasen met behulp van een predictiemodel.

AUTHORS AND AFFILIATIONS

St. Antonius Hospital Nieuwegein/Utrecht, The Netherlands
Department of Value-Based Healthcare
Biesma, Douwe H.
Dijksman, Lea M.

Department of Urology
Van Melick, Harm H.E.

Canisius Wilhelmina Hospital, Nijmegen, The Netherlands
Department of Urology
Van Basten, Jean-Paul A.

Department of Pathology
Küsters-Vandevelde, Heidi V.N.

Radboud University Medical Centre, Nijmegen, The Netherlands
Department of Urology
Witjes, J. Alfred
Sedelaar, Michiel

Department of Radiology & Nuclear Medicine
Israël, Bas

Erasmus University Medical Centre, Rotterdam, The Netherlands
Department of Radiology and Nuclear Medicine
Schoots, Ivo G.

Technical Medical Centre, University of Twente, Enschede, The Netherlands
Department of Health Technology & Services Research
Hueting, Tom A.
Koffijberg, Erik

Hospital Group Twente, Almelo/Hengelo, The Netherlands
Department of Urology
Stomps, Saskia P.

REVIEW COMMITTEE

Prof. dr. L.M.O. de Kort
Professor of Urology
University Medical Centre Utrecht, The Netherlands

Prof. dr. M.A.A.J. van den Bosch
Professor of Radiology
University Medical Centre Utrecht, The Netherlands

Prof. dr. H.P. Beerlage
Professor of Urology
Amsterdam University Medical Centre location AMC, Amsterdam, The Netherlands

Prof. dr. M. Roobol
Professor of decision making in Urology
Erasmus University Medical Centre Rotterdam, The Netherlands

Prof. dr. H.M. Verkooijen
Professor of Evaluation of Image-Guided Treatment
University Medical Centre Utrecht, The Netherlands

Prof. dr. ir. J.J.W. Lagendijk
Professor of Radiation Oncology Physics
University Medical Centre Utrecht, The Netherlands

Dr. R.C.N. van den Bergh
Department of Urology
St. Antonius Hospital Nieuwegein, The Netherlands

LIST OF PUBLICATIONS

1. Martini A, **Soeterik TFW**, Haverdings H, Rahota RG, Checcucci E, De Gillis S, Hermanns T, Fankhauser CD, Afferi L, Moschini M, Mattei A, Kesch C, Heidegger I, Preisser F, Zattoni F, Marquis A, Marra G, Gontero P, Briganti A, Montorsi F, Porpiglia F, Van Basten JPA, Van den Bergh RCN, Van Melick HHE, Ploussard G, Gandaglia G, Valerio M. An Algorithm to Personalize Nerve Sparing in Men with Unilateral High-Risk Prostate Cancer. *J Urol.* 2021 Aug 23, Online ahead of print.

2. Mazzone E, Gandaglia G, Ploussard G, Marra G, Valerio M, Campi R, Mari A, Minervini A, Serni S, Moschini M, Marquis A, Beauval JB, van den Bergh R, Rahota RG, **Soeterik TFW**, Roumiguiè M, Afferi L, Zhuang J, Tuo H, Mattei A, Gontero P, Cucchiara V, Stabile A, Fossati N, Montorsi F, Briganti A. Risk Stratification of Patients Candidate to Radical Prostatectomy Based on Clinical and Multiparametric Magnetic Resonance Imaging Parameters: Development and External Validation of Novel Risk Groups. *European Urology.* 2021 Aug 13, Online ahead of print.

3. Sampurno F, Cally J, Opie JL, Kannan A, Millar JL, Finelli A, Vickers AJ, Moore CM, Kowalski C, Foster C, Barocas DA, Galvin D, Van Basten JP, Gore JL, Ferencz J, Lawson KA, Ghani KR, Kwan L, Saarela O, Connor SE, Dieng S, Linsell S, **Soeterik TFW**, Villanti P, Litwin MS, Evans SM. Establishing a Global Quality of Care Benchmark Report. *Health Informatics J.* 2021 June 3, Online ahead of print.

4. Marra G, **Soeterik TFW**, Oreggia D, Tourinho-Barbosa R, Moschini M, Filippini C, van Melick HHE, van den Bergh RCN, Gontero P, Cathala N, Macek P, Sanchez-Salas R, Cathelineau X. Long-term Outcomes of Focal Cryotherapy for Low- to Intermediate-risk Prostate Cancer: Results and Matched Pair Analysis with Active Surveillance. *Eur Urol Focus.* 2021 Apr 26, Online ahead of print.

5. **Soeterik TFW**, Hueting TA, Israel B, van Melick HHE, Dijksman LM, Stomps S, Biesma DH, Koffijberg H, Sedelaar M, Witjes JA, van Basten JA. External validation of the Memorial Sloan Kettering Cancer Centre and Briganti nomograms for the prediction of lymph node involvement of prostate cancer using clinical stage assessed by magnetic resonance imaging. *BJU Int.* 2021;128:236-243.

6. **Soeterik TFW**, van Melick HHE, Dijksman LM, Biesma DH, Witjes JA, van Basten JA. Multiparametric Magnetic Resonance Imaging Should Be Preferred Over Digital Rectal Examination for Prostate Cancer Local Staging and Disease Risk Classification. *Urology.* 2021 Jan;147:205-212.

7. **Soeterik TFW**, van Melick HHE, Dijksman LM, Küsters-Vandevelde H, Stomps S, Schoots IG, Biesma DH, Witjes JA, van Basten JA. Development and External Validation of a Novel Nomogram to Predict Side-specific Extraprostatic Extension in Patients with Prostate Cancer Undergoing Radical Prostatectomy. *Eur Urol Oncol*. 2020 Sep 21 ;S25889311 (20) 30133-4. Online ahead of print.

8. **Soeterik TFW**, Van Melick HHE, Dijksman LM, Küsters-Vandevelde HV, Biesma DH, Witjes JA, Van Basten JPA. External Validation of the Martini Nomogram for Prediction of Side-Specific Extraprostatic Extension of Prostate Cancer in Patients Undergoing Robot-Assisted Radical Prostatectomy. *Urol Oncol*. 2020;38:372-378.

9. **Soeterik TFW**, Van Melick HHE, Dijksman LM, Stomps S, Witjes JA, Van Basten JPA. Nerve Sparing During Robot-Assisted Radical Prostatectomy Increases the Risk of Ipsilateral Positive Surgical Margins. *J Urol*. 2020;204:91-95.

10. **Soeterik TFW**, Van Melick HHE, Dijksman LM, Biesma DH, Witjes JA, Van Basten JPA. Follow-up in Active Surveillance for Prostate Cancer: Strict Protocol Adherence Remains Important for PRIAS-ineligible Patients. *Eur Urol Oncol*. 2019;2:483-489.

11. **Soeterik TFW**, Van Melick HHE, Dijksman LM, Biesma DH, Witjes JA, Van Basten JPA. Active Surveillance for Prostate Cancer in a Real-life Cohort: Comparing Outcomes for PRIAS-eligible and PRIAS-ineligible Patients. *Eur Urol Oncol*. 2018;1:231-237.

12. **Soeterik TFW**, Veenboer PW, Oude-Ophuis RJA, Lock MTWT. Electroejaculation in Patients With Spinal Cord Injuries: A 21-year, Single-Center Experience. *Int J Urol*. 2017;24:157-161.

13. Ottenhof SR, De Graaf P, **Soeterik TFW**, Neeter LMFH, Zilverschoon M, Spinder M, Bosch JLHR, Bleys RLAW, De Kort LMO. Architecture of the Corpus Spongiosum: An Anatomical Study. *J Urol*. 2016;196:919-925.

14. **Soeterik TFW**, Veenboer PW, Lock MTWT. Electroejaculation in Psychogenic Anejaculation. *Fertil Steril*. 2014;101:1604-1608.

CURRICULUM VITAE AUCTORIS

Timo Soeterik was born on the 28th of September 1991 in Weert, the Netherlands. He was raised in Haelen, together with his older brother and sister. He graduated from Gymnasium at the Scholengemeenschap St. Ursula in Horn.

In 2009, he started his study Medicine at the University of Utrecht. During his study time, he was an active member of several student associations. In his third year of his Bachelor of Medicine, he got engaged in scientific research within the field of andrology, under supervision of drs. M.T.W.T. Lock. He received several research grants while being a student researcher, including the "Van Walree Travel grant" of the "KNAW Fonds Medische Wetenschappen", and travel awards for two of the yearly conferences of the American Society for Andrology. During his Master study period, he did two of his clinical clerkships abroad, including Gynaecology in s' Lands Hospitaal, Paramaribo, Surinam and his social medicine intership at Ndlovu Care Group, Groblersdal, South Africa. In his final year, he participated in a study on the anatomy of the *corpus spongiosum,* at the Department of Regenerative medicine and Reconstructive Urology, under supervision of dr. P. de Graaf and prof. dr. L.M.O. de Kort.

Once graduated as a Master of Science in Medicine, he started as a PhD Candidate at the Santeon hospital group, working at the St. Antonius Hospital Nieuwegein. During his PhD period he completed the Master "Evidence Based Practice" at the University of Amsterdam, to fulfill his registration as an epidemiologist (Epidemioloog B) under the supervision of prof. dr. L.A.L.M. Kiemeney. Timo was rewarded research grants and awards from a number of institutions and associations, including the "Vlietstra Prize for Best Abstract" of the Dutch Urological Association, "Stichting Kwaliteitsgelden Medisch Specialisten", the "st. Antonius Ziekenhuis Onderzoeksfonds" and "ZonMw".

Timo currently works as a resident not in training at the Department of Urology of the St. Antonius Hospital in Nieuwegein (September 2021), under the supervision of dr. H.H.E. van Melick. He will start his urology residency training program in January 2022. He lives happily together with Evelien Klijn in Utrecht.

DANKWOORD

THE JOB IS DONE!

Het was een waar voorrecht om de afgelopen jaren aan dit proefschrift te werken. De mate van voldoening, persoonlijke groei en ontwikkeling die ik gedurende mijn promotietraject heb mogen doormaken overtroffen werkelijk al mijn verwachtingen. Echter, het meest bijzonder vond ik alle belangeloze hulp, inzet en steun die ik heb mogen ontvangen van iedereen die een bijdrage heeft geleverd aan de totstandkoming van dit proefschrift! Mijn dank is groot!

Heel graag zou ik nog mijn dank willen uitspreken aan een aantal personen in het bijzonder.

Geachte dr. J.P.A. van Basten, beste Jean-Paul, dank voor het grenzeloze vertrouwen dat je gedurende dit gehele traject in mij hebt gehad. De manier waarop je me vrij liet in het kiezen van de onderwerpen, werkzaamheden en planning was grandioos. Ik kon mijn creativiteit de vrije loop laten, en kon daarnaast altijd op je steun en raad rekenen wanneer nodig; ongeacht tijd en plaats. Jouw leiderschapskwaliteiten en je aanstekelijke enthousiasme voor de wetenschap zijn ongekend, maar ook als mentor heb je veel voor mij betekend. Dank voor alle moeite die je hebt gedaan om de juiste omstandigheden te creëren die ervoor zorgden dat ik dit promotietraject met zoveel plezier heb kunnen doorlopen.

Geachte dr. L.M. Dijksman, beste Lea, het gehele promotietraject was jij dé onmisbare spin in het web. Met al jouw onderzoekservaring wist je me naadloos door de uitdagingen waarmee ik als beginnend onderzoeker te maken kreeg heen te loodsen. Ik ben je nog steeds heel dankbaar voor de wijze waarop jij me wist te enthousiasmeren en stimuleren om mijn kennis van de epidemiologie en statistiek te vergroten. Jouw inzicht en ervaring in het opzetten van onderzoeken, interpretatie van resultaten en counteren van ingewikkeld statistisch review commentaar zijn werelds en hebben me geweldig geholpen. Dank!

Geachte dr. H.H.E. van Melick, beste Harm, ons wetenschappelijke avontuur begon met de wekelijkse brainstormsessies op de woensdagmiddag. Inmiddels is de samenwerking uitgegroeid tot een wetenschappelijke hoge snelheidstrein met tal van succesvolle tussenstops. Ik ben trots op alle projecten die we samen hebben opgestart, begeleid en afgerond. Naast een geweldige wetenschappelijke begeleider heb je me ook klinisch op een onwijs fijne manier de weg gewezen. Dank voor al je wijze raad, adviezen en de kansen die je me hebt gegeven. Ik hoop in mijn opleiding tot uroloog nog veel van je te mogen leren.

Geachte prof. D.H. Biesma, beste Douwe, vanaf moment één besefte ik dat de kennis en ervaring die jij mee zou brengen met jouw achtergrond als hoogleraar Interne Geneeskunde en bestuurder van grote waarde zouden zijn tijdens dit traject. Ik heb enorm

veel respect voor de wijze waarop jij je bestuurlijke verantwoordelijkheden combineert met topklinisch wetenschappelijk onderzoek. De manier waarop jij je visie en kennis op het gebied van kwaliteitsverbetering in de zorg weet over te brengen heb ik als enorm inspirerend en motiverend ervaren. Dank voor alles.

Geachte prof. J.A. Witjes, beste Fred, jouw enorme berg aan kennis en onderzoekservaring op het gebied van prostaatkanker waren van onschatbare waarde. Het was een eer om met iemand met zo'n bewonderingswaardige staat van dienst als jij, op zowel klinisch als wetenschappelijk vlak, samen te werken. Het feit dat je werkzaam was in een ander ziekenhuis en regio, deed niets af aan je betrokkenheid. Ik ben je zeer erkentelijk voor het feit dat je al mijn mails binnen een dag beantwoordde en me telkens weer voorzag van een antwoord of advies waarmee ik verder kon.

Geachte leden van de promotie- en beoordelingscommissie, dank voor het beoordelen van dit proefschrift.

Alle co-auteurs van de in deze thesis opgenomen artikelen, in het bijzonder: Dr. H.V.N. Küsters-Vandevelde, drs. S. Stomps en drs. T.A. Hueting. Heidi, jij hebt een weergaloze inspanning geleverd door alle prostatectomie preparaten te reviseren. Hierdoor hebben we het aantal missende waarden enorm kunnen inperken, hetgeen de data kwaliteit een grote boost heeft gegeven. Dank voor al je inspanningen. Saskia, het was geweldig om het ZGT bij de studies te kunnen betrekken. Dank voor je interesse en enthousiasme voor de participatie aan de studies. Tom, dank voor de prettige samenwerking op het gebied van predictiemodellering. Dank voor al je hulp met de statistische analyse en R code.

De urologen van de Santeon Ziekenhuizen en de Ziekenhuisgroep Twente, en alle andere medewerkers verbonden aan de afdelingen urologie van deze ziekenhuizen die hebben bijgedragen aan het faciliteren van de ideale omstandigheden om de data te verzamelen. Dank voor de altijd prettige ontvangst en alle hulp met het regelen van de logistiek rondom de datacollectie.

Ellen Tromp, Hans Kelder, Pieter Zanen en Martijn Stuiver dank voor het in behandeling nemen van al mijn statistische vraagstukken en zeer welkome hulp bij de analyses.

De pathologen van het St. Antonius Ziekenhuis Nieuwegein en het Martini Ziekenhuis Groningen: dank voor alle tijd en energie die jullie hebben gestoken in het beantwoorden van mijn vragen over het verrichte pathologisch onderzoek en alle revisies die jullie hebben uitgevoerd om missende data op te halen.

Alle dossiermedewerkers die hebben bijgedragen aan het ophalen van de data. Door jullie inspanningen en zorgvuldigheid waren alle analyses mogelijk!

Collega's van het Canisius Wilhelmina Ziekenhuis. Urologen, dank voor de altijd welkome ontvangst als ik mijn onderzoeksdagen in het Nijmeegse doorbracht. Joost de Baaij en Diederik Baas, jullie in het bijzonder bedankt voor het aanleggen en bijhouden van die gigantische prostaatkanker database. Mede daardoor was er veel mogelijk. Leden van het VBHC team: Judith, Bart en Charlotte, dank voor de gezellige lunches en gastvrijheid. Voelde me altijd enorm welkom in het Jonkerbosch. Christine, wat een eer om na jou de tweede "Santeon promovendus" te zijn. Dank voor al je bruikbare adviezen ten aanzien van PROMS, SPSS en Excel. Bovenal bedankt voor de gezellige koffiemomenten.

The Movember, Monash University and UCLA family: thank you for including me in the TrueNTH Prostate Cancer Registry team. It was an amazing experience to meet such dedicated and inspiring fellow researchers from all over the world. Thank you for all the assistance regarding the Santeon data infrastructure, the inspiring working group sessions and of course the unforgettable LA data manager meetings.

Dr. P.B. van der Nat. Beste Paul, toen ik begon als arts-onderzoeker voelde de afdeling Value Based Healthcare, mede door jou, als een warm bad. Veel dank voor jouw bijdrage aan de prettige werkomstandigheden. Alle andere leden van het Value-Based Healthcare team in het St. Antonius: bedankt voor de prettige samenwerking en het goede gezelschap tijdens alle uitjes.

De VBHC projectleiders en data-analisten van de Santeon ziekenhuizen. Dank voor het beantwoorden van al mijn data- en EPD-gerelateerde vragen. Het was erg leerzaam om met jullie op het raakvlak VBHC en wetenschap samen te werken.

De collega's van het Santeon programmabureau. Dank voor de prettige werkdagen die ik bij jullie in Utrecht heb mogen doorbrengen, ik voelde me altijd enorm welkom. In het bijzonder bedankt Leonique Niessen, voor het mede mogelijk maken van dit promotietraject. Pieter de Bey, dank voor de interesse en betrokkenheid.

Drs. M.T.W.T. Lock, Tycho, toen ik je in mijn derde studiejaar benaderde om wetenschappelijk onderzoek te komen doen had ik nooit gedacht dat je ervoor zou zorgen dat ik er zó verknocht aan zou raken. Dank voor het aanwakkeren van het wetenschappelijke vuur.

Promovenski maatjes Nynke, Nina, Milad, Mirjam, Anna-Roos, Kirsten en Daan. Door jullie goede gezelschap fietste ik dagelijks fluitend naar mijn werk. Dank voor de yoga sessies (die soms bijna m'n rug kostte). Legendarische skireizen (vooral de TLC die ik nodig had wanneer ik weer eens gestrekt op de piste lag of zelfs op krukken de baan verliet en de bijdrage aan mijn zelfoverschatting "Timo, deze piste kun jij makkelijk"). De nachtelijke redding als ik weer iets doms had gedaan: wie heeft er nou collega's die

midden in de nacht uit bed komen om je te helpen je huissleutels uit een put te vissen? Hey jongens het was onvergetelijk!

Urologen van het St. Antonius Ziekenhuis. Ik had geen betere plek kunnen wensen om mijn eerste klinische stappen te zetten. Dank voor het creëren van de fijne werksfeer, alle uitleg en nuttige supervisie momenten. Super dat jullie me de ruimte boden om naast de klinische taken mijn promotie af te ronden. Collega arts-assistenten met wie ik heb mogen samenwerken: Robin, Lieke, Franklin, Pepijn, Olivier, Robert, Berend, en Sanne dank voor al jullie geduld en uitleg tijdens mijn eerste klinische jaar. Adriaan, Hajar, Laura Manou en Fieke, dank voor de fijne samenwerking en goede werksfeer in ons huidige assistenten team. Lieke en Joris, zo gaaf dat jullie nu ook met je promotietraject gestart zijn. Ben super trots en dankbaar dat jullie bereid zijn het werk voort te zetten.

Studievrienden van het eerste uur: Claudia, Vincent, Joris, Ingmar en Carlie. Ik krijg een brede lach op mijn gezicht als ik terugdenk aan alle hilarische momenten uit onze tijd in de collegebanken (en vooral de momenten daarbuiten). Geweldig dat we elkaar na al die jaren nog regelmatig binnen en buiten het ziekenhuis om zien.

De jongens van de Donkerstraat 27bis, Eelco en Daan. Het was super om ten tijde van de eindfase van mijn studie en de start van mijn promotie jullie als huisgenoten te hebben. Dank voor de mooie tijd.

Heren van het Dispuut Typhoon. Het kan niet anders dan dat jullie gedeelde fascinatie voor het urogenitale stelsel me indirect heeft geïnspireerd om binnen dit thema te gaan promoveren, waarvoor dank! Tijdens mijn promotietijd vormden alle weekenden, borrels en gala's in jullie gezelschap een heerlijke afwisseling op het onderzoekswerk. Ook al moest ik er soms een week van bijkomen, ik had ze voor geen goud willen missen. VVOV!

Mannen van Huize Ruisenstein. Geweldig dat al die mooie jaren samen onder een dak hebben geleid tot zulke hechte vriendschap. Ik ben blij dat ik ook deze mijlpaal weer met jullie kan vieren. Ooooooo...

De "Echte Vrienden", Nando, Jasper, Joris-Jan, Ruud, Frank en Maarten. Hoe moeders er toch soms zó naast kunnen zitten ("Jullie krijgen toch andere vrienden!"). We kennen reeds een prachtig lange historie met tal van mooie herinneringen; en de lijst wordt alleen maar langer. Nu nog een mijlpaal die ik met jullie kan delen. Dank voor de vele jaren hechte vriendschap.

Jaarclub Crisis. Al ruim tien jaar vormen jullie een rode draad door m'n leven. Een van onze hoogtepunten was natuurlijk de lustrumreis naar Nicaragua, waarbij heelhuids met jullie uit dit land terugkeren misschien wel de grootste uitdaging was die ik in mijn promotietijd heb gehad. Wat was het goed!

Mijn paranimfen; Jos, van lief en leed delen op de Donkerstraat tot nu samen in het Academiegebouw, geweldig dat jij mijn paranimf wilt zijn. Sjon, wie anders dan jij als ware scriptiekoning weet hoe het is om ergens je tanden in te zetten, ben trots je aan mijn zijde te hebben.

Mijn lieve schoonfamilie. Jan en Annemarie, dank voor jullie steun en interesse. De heerlijke weekendjes weg en vakanties waar ik van heb mogen genieten waren telkens weer een welkome afwisseling op de doordeweekse uurtjes in de boeken. Lieve Anneloes en Gijs; jullie zijn toppers.

Lieve Josje, zo leuk om Feddo en jij samen zo gelukkig te zien. Ben enorm blij met jou als schoonzusje. Dear Jamal, so special to be a witness of the J & J wedding. Welcome to the family!

Lieve broer, Fed, jouw enthousiaste reacties en verwondering wanneer ik iets vertelde over mijn onderzoek werkten super inspirerend. Nog geweldiger was onze Amerika reis met pa, een absoluut hoogtepunt in mijn promotietijd! Pien, lieve zus. Creatief supertalent. Van het ontwerpen van mijn eerste wetenschappelijke poster, tot nu aan de vormgeving van deze thesis: jij stond voor me klaar. Je hebt me zo ontzettend goed geholpen, dankjewel!

Lieve pa en ma, zoals altijd kon ik ook tijdens dit avontuur op jullie onvoorwaardelijke support, raad en daad rekenen. Dank voor jullie warmte, wijze lessen en liefde. Het was heel fijn dat ik me terug kon trekken in "Hotel Soeterik" gedurende de laatste loodjes van deze thesis; en onderwijl kon genieten van jullie gezelschap en gastvrijheid. Lieve pa, fantastisch dat Fed en jij me kwamen opzoeken tijdens een congres in Chicago. Onze gezamenlijke trip naar Detroit, de plek waar je opgroeide, was onvergetelijk! Lieve ma, ik denk nog vaak terug aan het moment dat je aan me vroeg of de studie geneeskunde niet iets voor mij was. Nooit gedacht dat ik uiteindelijk als arts zou promoveren! *Dit proefschrift is voor jullie.*

De laatste woorden van deze thesis zijn voor jou, allerliefste Eef. Mijn stilte in de storm. Zonder jouw onuitputtelijke liefde, steun en positiviteit was het me nooit gelukt. Jij maakt me zó gelukkig! Ik koester alles wat we tot nu toe hebben mogen delen en kijk uit naar al het moois dat we nog mogen meemaken.