

Improving Screening Strategies for Prostate Cancer

Tineke Wolters

Erasmus MC Rotterdam

Improving Screening Strategies for Prostate Cancer
©Tineke Wolters

Email: tinekewolters@gmail.com

ISBN/EAN: 978-90-9025121-9

Cover design: Tineke Wolters

Lay out: Legatron Electronic Publishing, Rotterdam

Printing: Ipskamp Drukkers, Enschede

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval center of any nature, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the permission of the author.

Improving Screening Strategies for Prostate Cancer

Verbetering van strategieën voor bevolkingsonderzoek
naar prostaatkanker

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam

op gezag van de rector magnificus
Prof.dr. H.G. Schmidt
en volgens het besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
Woensdag 28 april 2010 om 15.30 uur

door

Tineke Wolters

Geboren te Gouda



Promotiecommissie

Promotoren:

Prof.dr. F.H. Schröder

Prof.dr. C.H. Bangma

Overige leden:

Prof.dr. Th.H. van der Kwast

Prof.dr. P.F.A. Mulders

Prof.dr. L.A.L.M. Kiemeney

Co-promotoren:

Dr. M.J. Roobol

Dr. G.J.L.H. van Leenders

The studies reported in this thesis were performed at the departments of Urology and Pathology of the Erasmus Medical Center, Rotterdam, the Netherlands as an integrated part of the European Randomized Study of Screening for Prostate Cancer (ERSPC). The ERSPC was funded Europe-wide by grants of Beckman-Coulter-Hybritech Inc; European Union Grants; the 6th framework program of the EU; P-mark. The Dutch section of the ERSPC is supported by the Dutch Cancer Society; the Netherlands Organization for Health Research and Development (ZonMW).

The studies reported in this thesis were sponsored by an unconditional grant of Fonds NutsOhra.

The printing of this thesis was financially sponsored by:

Abbott, Astellas, AstraZeneca, Beckman Coulter, Ferring, GlaxoSmithKline, Eli-Lilly, Novartis, Olympus, Pfizer, Sanofi Aventis, Stichting Contactgroep Prostaatkanker, Stichting Urologisch Wetenschappelijk Onderzoek, Stichting Wetenschappelijk Onderzoek Prostaatkanker

Contents

Part I	General Introduction	7
Chapter 1	Prostate cancer	11
Chapter 2	Screening for prostate cancer: the European Randomized Study of Screening for Prostate Cancer (ERSPC) <i>The European Randomized Study of Screening for Prostate Cancer (ERSPC). Current clinical urology: Prostate cancer screening. Page 373-386. Editors D.P. Ankerst, C.M. Tangen and I.M. Thompson. Humana Press 2009, New York</i>	19
Chapter 3	Scope of the thesis	37
Part II	The disease specific mortality reduction due to screening for prostate cancer- the role of treatment	41
Chapter 4	The effect of study arm on prostate cancer treatment in a large screening trial (ERSPC) <i>Int J Cancer. 2009, epub ahead of print, doi:10.1002/ijc.24870</i>	43
Part III	How to screen for prostate cancer?	57
Chapter 5	Is prostate-specific antigen velocity selective for clinically significant prostate cancer in screening? ERSPC Rotterdam <i>Eur Urol 2009;55:385-92</i>	58
Chapter 6	Can non-malignant biopsy features identify men at increased risk of biopsy-detectable prostate cancer at re-screening after four years? (ERSPC Rotterdam) <i>BJU Int 2008;101:283-8</i>	73
Chapter 7	False-negative prostate needle-biopsies: frequency, histopathologic features and follow-up <i>Am J Surg Pathol 2010;34:35-43</i>	85
Part IV	Prognosis of prostate cancer: insignificant or aggressive?	101
Chapter 8	Should prostate tumour volume routinely be reported by the pathologist?: The prognostic value of tumour volume in prostate cancer <i>Eur Urol 2009, epub ahead of print, doi:10.1016/j.eururo.2009.07.027</i>	103
Chapter 9	Reassessment of the tumour volume threshold defining clinically insignificant prostate cancer; results from a randomized trial <i>Submitted</i>	119

Chapter 10	The value of EZH2, p27 ^{kip1} , BMI-1 and MIB-1 on biopsy specimens with low risk PC in selecting men with significant prostate cancer at prostatectomy <i>BJU Int 2009, epub ahead of print, doi:10.1111/j.1464-410x.2009.08998</i>	131
Part V	General discussion	145
Chapter 11	General discussion	147
Chapter 12	Summary	168
	Samenvatting (Nederlands)	171
	Curriculum vitae	173
	List of publications	175
	Dankwoord	179
	Appendix: color images	181
	PhD portfolio	185

PART I

GENERAL

INTRODUCTION

Chapter 1

Prostate cancer

Chapter 2

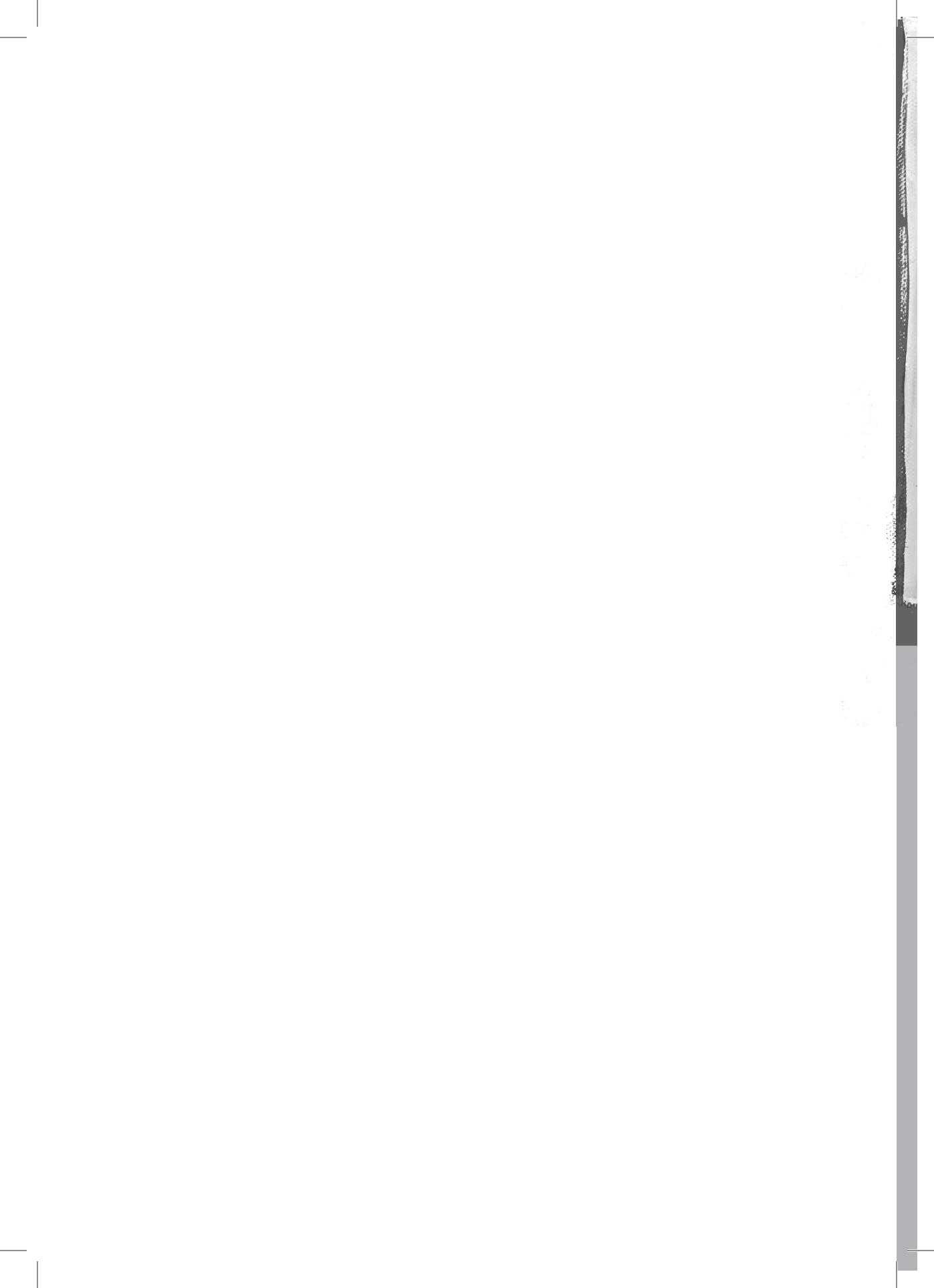
**Screening for prostate cancer: the European Randomized Study of
Screening for Prostate Cancer (ERSPC)**

Chapter 3

Scope of the thesis

General introduction

This thesis describes research on screening for prostate cancer. To improve understanding of the thesis, some background information will be provided in this introduction. First, a short description of the prostate and of prostate cancer will be given in **Chapter 1**, followed by more detailed background information on screening for prostate cancer in **Chapter 2**. The final part of this introduction, **Chapter 3**, will outline the scope of this thesis.





CHAPTER 1

Prostate cancer

1.1 | The prostate

The prostate is a gland which is located beneath the urine bladder, surrounding the proximal urethra in men. It produces a fluid which constitutes part of the semen and prolongs the lifespan of sperm.

1.2 | Prostate cancer

1.2.1 | Diagnosis

There are several ways to examine the prostate for the presence of prostate cancer: a serum prostate specific antigen (PSA) test, a digital rectal examination (DRE), transrectal ultrasonography (TRUS) and prostate biopsy.

PSA is a protein produced by prostate cells, which may leak into the blood stream. An increased serum PSA level indicates an increased prostate cancer risk. However, an increase in the serum PSA level may also be caused by other causes: 1) an increase in normal prostate glands like in benign prostate hyperplasia (BPH) or 2) an increased leakage of PSA into the bloodstream due to infectious processes or obstruction. An increased PSA level thus is not specific for prostate cancer. Moreover, a low PSA level does not exclude prostate cancer: even in very low PSA ranges, a considerable prostate cancer detection rate has been described.¹ Despite the important limitations of lack of specificity and lack of what might be considered a “normal” level of PSA, the discovery of PSA has drastically changed prostate cancer care. In general, PSA levels increase prior to the occurrence symptoms of metastasis, while prostate cancer mostly manifests itself clinically only at the time of (bone) metastasis. This shifts the diagnosis to an earlier and possibly curable stage of prostate cancer. Furthermore, PSA plays an important role in the follow-up of prostate cancer patients.

A DRE entails the palpation of the prostate through the rectal wall (Figure 1). The size of the prostate can be evaluated as well as the presence of nodules or indurations suspicious for prostate cancer. TRUS produces an image of the prostate, which may show signs of prostate cancer and can provide an accurate measure of prostate volume. TRUS is often used for guiding prostate biopsies.

PSA, DRE and TRUS all give an indication about the risk of prostate cancer, but the definite diagnosis can only be made with histological prostate tissue examination. During a prostate biopsy, generally under ultrasound guidance, tissue cores are taken from the prostate with a needle. These tissue samples are examined by a pathologist and may confirm the diagnosis of prostate cancer.

More diagnostic tests are available, for example PSA derivatives², and prostate cancer antigen 3 (PCA3)⁴ have been described amongst others to show promising results. However, those tests have not yet been incorporated in standard clinical care.

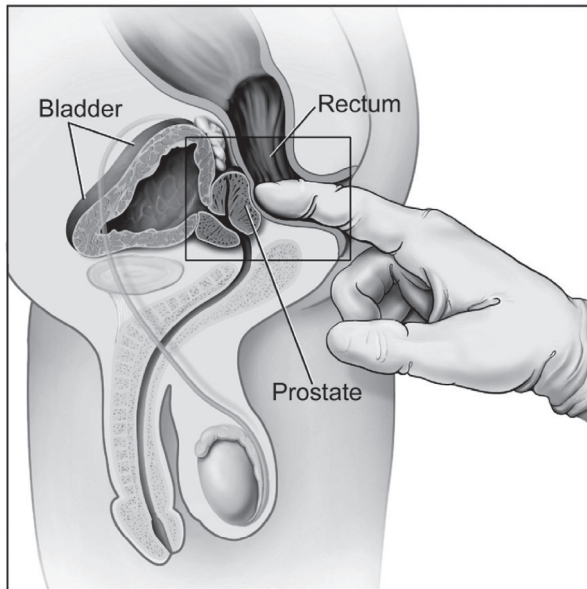


Figure 1 | Digital rectal examination, adapted from³

1.2.2 | Staging

The extent of the disease is classified according to the Tumour/Node/Metastasis (TNM) classification (Table 1). The stage of the disease is predictive for the prognosis and is useful in selecting treatment. The local stage of the tumour is determined by DRE and TRUS, known as clinical tumour stage. The definite stage, or pathological tumour stage, can only be obtained after radical prostatectomy.

1.2.3 | Grading

The grade of the tumour expresses the degree of abnormality of the tissue and thus the aggressiveness of the tumour. Prostate cancer is graded using the Gleason grading system.⁴ The growth pattern is scored 1 (well differentiated) to 5 (poorly differentiated tumour) (see figure 2). The most common and second most common growth pattern observed by the pathologist are summed to give a Gleason score, ranging from 2 to 10. A Gleason score is assigned to biopsy tissue and prostatectomy specimens and both are highly prognostic for patient outcome.

Table 1 | TNM classification, 2002

T-primary tumour	Tx	Primary tumour cannot be assessed		
	T0	No evidence of primary tumour		
	T1	Clinically unapparent tumour neither palpable nor visible by imaging	T1a	Tumour incidental histological finding in 5% or less of tissue resected
			T1b	Tumour incidental histological finding in more than 5% of tissue resected
			T1c	Tumour identified by needle biopsy (e.g. because of elevated PSA)
	T2	Tumour confined within the prostate	T2a	Tumour involves one-half of one lobe or less
			T2b	Tumour involves more than one-half of one lobe, but not both lobes
T2c			Tumour involves both lobes	
T3	Tumour extends through the prostatic capsule	T3a	Extracapsular extension in periprostatic tissue	
		T3b	Invasion of seminal vesicle(s)	
T4	Tumour is fixed or invaded adjacent structures other than the seminal vesicles: bladder neck, external sphincter, rectum, levator muscles, or pelvic wall			
N- regional lymph nodes	Nx	Regional lymph nodes cannot be assessed		
	N0	No regional lymph node metastasis		
	N1	Metastasis in regional lymph nodes		
M- distant metastasis	Mx	Distant metastasis cannot be assessed		
	M0	No distant metastasis		
	M1	Distant metastasis	M1a	Non-regional lymph nodes
			M1b	Bone(s)
M1c			Other sites	

1.2.4 | Treatment

Treatment of prostate cancer depends on patient and tumour characteristics. Overall, treatment can be divided into 4 main modalities: surgical treatment, radiation therapy, hormonal treatment and monitoring, i.e. active surveillance or watchful waiting.



Figure 5.1: Appearance of cancer cells, ranked according to the Gleason grade.

Figure 2 | Gleason grading system, adapted from⁶

1.3 | Prostate cancer as a major health problem

In Western countries, prostate cancer is the most common non-skin cancer type diagnosed amongst men.⁷ Based on rates of 2004-2006, a man in the US has a life-time risk of being diagnosed with prostate cancer of almost 16%.⁸ This indicates that 1 in 6 men will develop prostate cancer. The incidence of prostate cancer has been rising since the early nineties (Figure 4). Several factors may have induced this increase, amongst which aging of the population, increased awareness of prostate cancer, and the possibility of early detection through PSA.

Prostate cancer is the second leading cause of death from malignancy.⁷ When one compares the incidence and mortality rates a striking difference between those rates is observed (Figure 4). This implies that more men die with prostate cancer than from prostate cancer. Especially in the last decades, incidence and mortality rates have diverged (Figure 4). In addition to the rise in prostate cancer detection, a more recent decline in mortality has occurred. Explanations for this decrease may be found in improved prostate cancer treatment, wide-spread use of statins and possibly for a part in screening for prostate cancer.^{9,10}

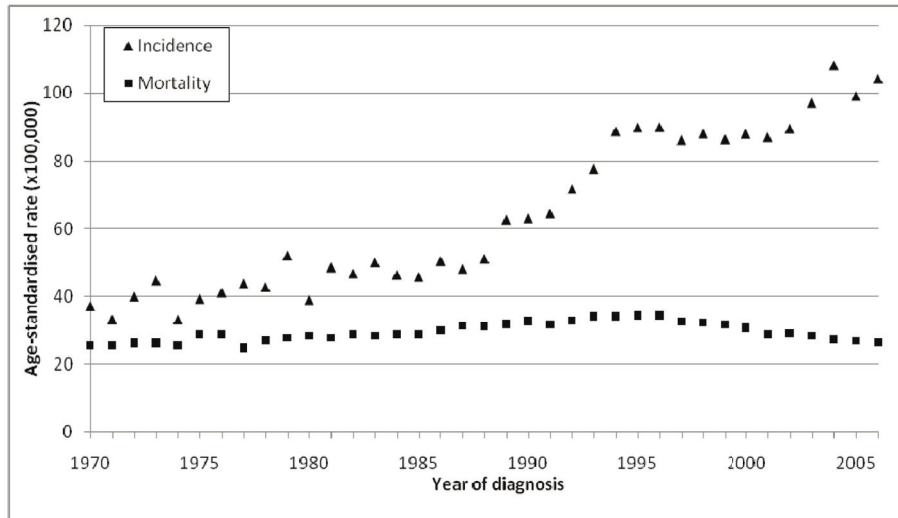
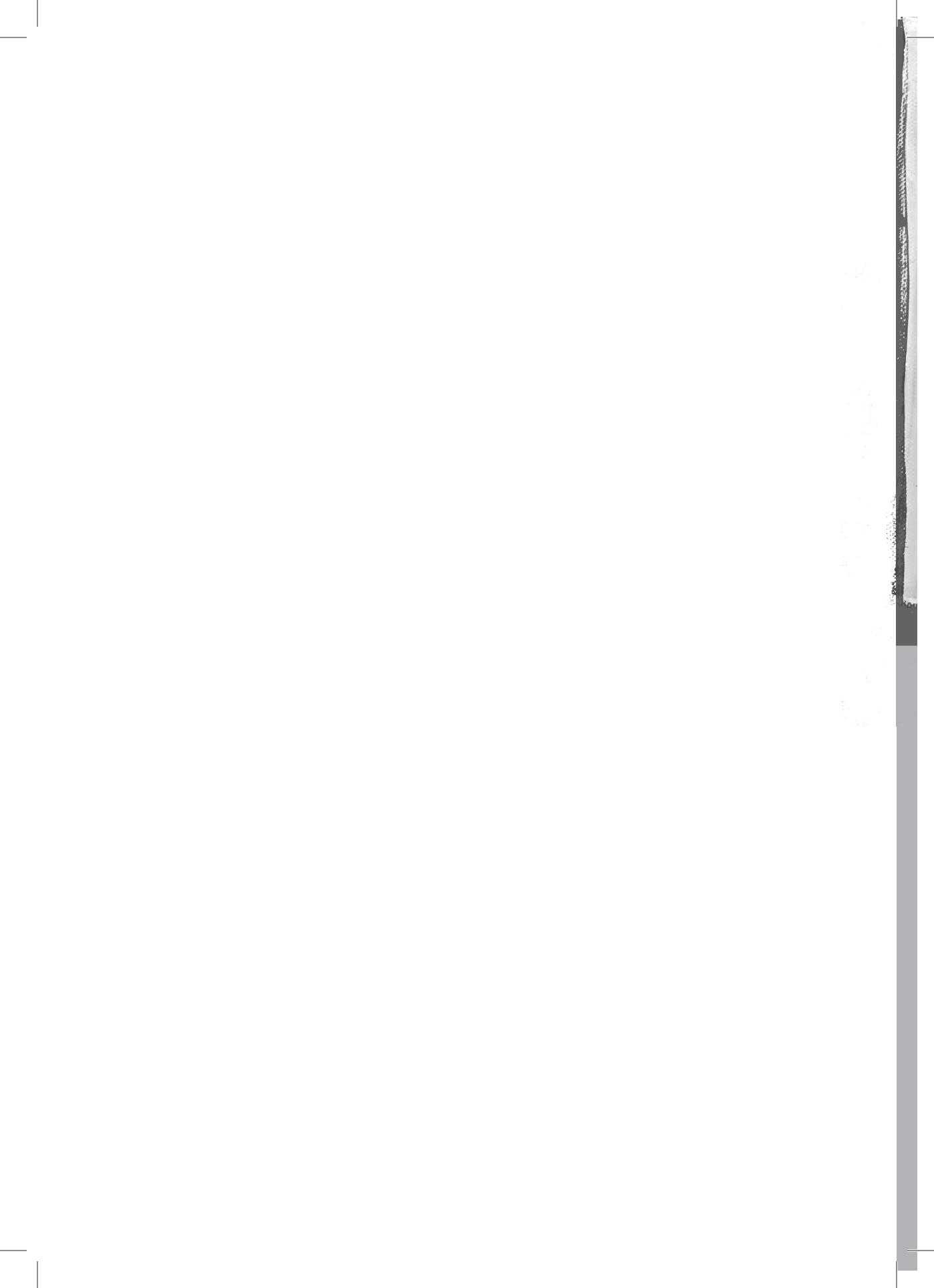


Figure 3 | Age-standardised rates (European Standardised Rate) for incidence and mortality of prostate cancer in the Netherlands 1970-2006 (incidence rates 1970-1988: data Comprehensive Cancer Centre South; incidence rates 1989-2006: data Netherlands Cancer Registry - no differences between CCCS and NCR data in period 1989-2006 -; mortality rates 1970-2006: Netherlands Cancer Registry)

References

1. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, Minasian LM, Ford LG, Lippman SM, Crawford ED, Crowley JJ, Coltman CA Jr. Prevalence of prostate cancer among men with a prostate-specific antigen level $< \text{or} = 4.0$ ng per milliliter. *N Engl J Med.* 2004;350:2239-46
2. Jansen FH, Roobol MJ, Schröder FH, Bangma CH. Screening for Prostate Cancer in 2008 II: The Importance of Molecular Subforms of Prostate-Specific Antigen and Tissue Kallikreins. *Eur Urol* 2008;55:562-574
3. www.yourhealthinformation.com, provided by the National Cancer Institute.
4. Hessels D, Schalken JA. The use of PCA3 in the diagnosis of prostate cancer. *Nat Rev Urol.* 2009;6:255-61
5. Gleason DF. Histologic grading of prostate cancer: a perspective. *Hum Pathol.* 1992;23:273-9
6. www.prostatecancerfoundation.org
7. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin.* 2009;59:225-49
8. Horner MJ, Ries LAG, Krapcho M, Neyman N, Aminou R, Howlader N, Altekruse SF, Feuer EJ, Huang L, Mariotto A, Miller BA, Lewis DR, Eisner MP, Stinchcomb DG, Edwards BK (eds). *SEER Cancer Statistics Review, 1975-2006*, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2006/
9. Etzioni R, Tsodikov A, Mariotto A, Szabo A, Falcon S, Wegelin J, DiTommaso D, Karnofski K, Gulati R, Penson DE, Feuer E. Quantifying the role of PSA screening in the US prostate cancer mortality decline. *Cancer Causes Control.* 2008;19:175-81
10. Hamilton RJ, Freedland SJ. Review of recent evidence in support of a role for statins in the prevention of prostate cancer. *Curr Opin Urol.* 2008;18:333-9





CHAPTER 2

Screening for prostate cancer: the European Randomized Study of Screening for Prostate Cancer (ERSPC)

*The European Randomized Study of Screening for Prostate Cancer (ERSPC).
Current clinical urology: Prostate cancer screening. Page 373-386. Editors D.P. Ankerst,
C.M. Tangen and I.M. Thompson. Humana Press 2009, New York*

Wolters T. and Schröder F.H.

2.1 | Introduction

The European Randomized Study of Screening for Prostate Cancer (ERSPC) is a large randomized study which aims to show or exclude a reduction in prostate cancer specific mortality due to screening. In addition, the feasibility of such a program for population-wide use, the cost-effectiveness and quality of life are studied. Recently, after the third interim analysis, the first selected data end-points of the ERSPC have been published: PSA-based screening reduced prostate cancer specific mortality by 20%.¹ In this chapter, an attempt is made to describe the ongoing ERSPC.

2.2 | Background

2.2.1 | The start of the ERSPC

With emerging insights about the use of PSA as a screening test², an idea of conducting a randomized controlled study of screening for prostate cancer originated in Belgium and the Netherlands during 1990-1991.^{3,4} No evidence of the effectiveness of screening for prostate cancer existed at that time and the only way of obtaining such evidence seemed to be by conducting a prospective randomized controlled trial (RTC). Several uncertainties existed related to randomization, acceptance and value of screening tests, follow-up and others. Therefore, between 1991 and 1994 a series of pilot studies were carried out in Belgium and in the Netherlands. Summaries of the results of these pilots were published in 1995.^{5,6} The main conclusion was that a European RCT of screening for prostate cancer seemed feasible. Yet, the expense of such a trial made international co-operation a prerequisite, as no single country could afford such a study. The ERSPC formally started on July 1 1994 in Belgium and the Netherlands. After successfully conducting pilot studies⁷, Finland became the third partner in the ERSPC during 1995. Furthermore, Sweden, Spain, Italy, Switzerland and France joined the ERSPC.

A publication of Adami et al⁸ in 1994 gave rise to a public discussion about the ethical justification of such a RTC. Some felt that prerequisites for performing screening studies, for example knowing the natural history and effectiveness of treatment, were not met. This controversy has been subject of an extensive discussion.^{9,10,11} However, in all participating countries, ethical approval was obtained and the ERSPC started in 1994.

2.2.2 | Purpose and structure of the study

The main goal of the ERSPC is to show or exclude a prostate cancer mortality reduction through screening and early treatment. In this large randomized controlled trial screening is offered to the intervention group and the control group is managed according to regional health care policies. The trial aims at showing or excluding a 20% difference in prostate cancer mortality

with a power of 90%. This was decided at a consensus workshop held in 1994 on the basic elements of screening in a RCT.¹²

Other important decisions were made at the consensus workshop, like the determination of 4.0 ng/ml as a PSA threshold for recommending a biopsy and initially including Digital Rectal Examination (DRE) and Transrectal Ultrasound (TRUS) as screening tests. An age range of 55-70 years was determined as being the core age range on which power is calculated. Inclusion of higher or lower age ranges could be chosen by the individual centers. Initial sample size calculations were made without consideration of possible contamination in the control arm by opportunistic PSA-based screening and it was calculated that 65,000 men per arm and a follow-up of 10 years would be needed. Re-calculations considering this contamination showed that a sample size of 85,000 men per arm would be needed.¹³

Furthermore, basic requirements for participation in the ERSPC and the content of the future database were discussed during this workshop. This resulted in the establishment of the following committees which run and control the ERSPC: an Epidemiology Committee, a Pathology Committee, a PSA Committee, a Quality Control Committee, a Causes of Death Committee, an independent Data Monitoring Committee (DMC) and the supervisory body of the study as a whole, the Scientific Committee (SC). All important decisions are made by the voting members of the SC, which consists of two representatives of each center. Every participating center has accepted the authority of the controlling committees (i.e. the Quality Control Committee and Data Monitoring Committee) and signed the set of basic requirements as a cooperative contract. Because of the relative autonomy of the participating centers, the ERSPC is conducted in a decentralized fashion. Centralized data collection is in the hands of an independent center located in the UK (the Central Database).

2.3 | Randomization

Due to differences in legal requirements for running a RTC in the participating countries, two different randomization schemes are used. In Belgium, Spain, Switzerland and the Netherlands, informed consent is required before randomization (Figure 1). In the other countries, informed consent is only required for those men who are randomized to the screen arm.

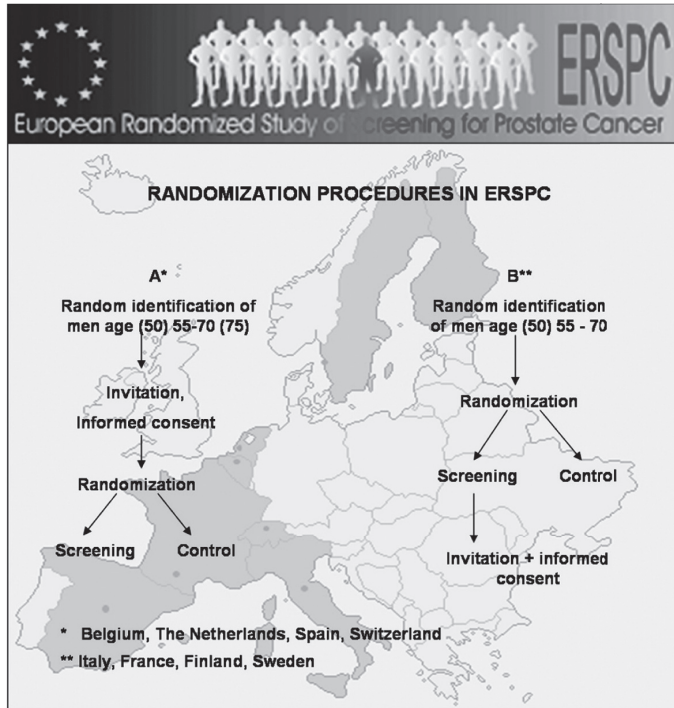


Figure 1

2.4 | Screening tests

2.4.1 | PSA, DRE and TRUS

At the time the ERSPC started, PSA, DRE and TRUS were potential screening tests. Other screening methods, such as the use of PSA derivatives (e.g. PSA velocity, PSA density, PSA doubling time) and biomarkers are possibly proper screening tests as well, which were to be evaluated in a later phase.

Between ERSPC centers next to the differences in randomization schemes, slightly different screening protocols are used. Except one center all use a four-year screening interval (Sweden uses two years). A PSA cut-off value of 3.0 ng/ml is applied after the prevalence screen and was suggested to all centers as of 1997. The core age group is 55-69 and all centers use sextant biopsies, at least initially.

Up to May 1997, a PSA \geq 4.0 ng/ml and/or abnormal findings on DRE and/or TRUS was an indication for prostate biopsy in the Rotterdam center of the ERSPC.

For a proper screening algorithm, it is very important to find a delicate balance between sensitivity and specificity of the screening tests, as lead-time and over-diagnosis are inevitable in prostate cancer screening. Therefore, evaluation of screening procedures was and is an

essential part of the ERSPC, and prior to the initiation of the study, it was agreed that screening procedures would be adjusted one time if necessary.

The first contribution to this evaluation came from the Italian group¹⁴, and was followed by major investments made by other centers as well to clarify the role of the different screening tests. Improvements of test characteristics, potential reduction of proportions of men to be biopsied, the loss of otherwise diagnosed cancers and the numbers of biopsies needed per prostate cancer were investigated and simulated by Bangma et al.^{15,16} These studies indicated the future direction of test evaluation: main goal was to improve specificity (avoidance of unnecessary testing and biopsies) and still maintain the detection rate of the first round (4-5%). The search for methods for improvement of specificity was continued and led to many suggestions.^{17,18} However, great reluctance existed to change the protocol with the acceptance of the loss of a proportion of prostate cancers, whose final outcomes could not be judged. Nevertheless, a major change in the screening protocol was implemented as a result of the evaluation outcomes. From February 1996, men with a PSA value of 0-0.9 ng/ml were not further screened but advised to be re-screened four years later. This change was based on the observation that in 1451 men, 174 biopsies detected 4 cancers, resulting in a positive predictive value of 2% and a cancer detection rate of 0.3%.¹⁷ All centers adopted this change in protocol, saving a screening visit for 35% of the whole screening population and leading to an obvious reduction in costs of the study.

A second major change was omitting DRE and TRUS as screening tests and lowering the biopsy threshold to a PSA value of 3.0 ng/ml. This decision was based on a study from the Rotterdam center, where the value of DRE was investigated. Ideally, sensitivity and specificity of a screening test are calculated. However, the prevalence of the disease is unknown in the case of prostate cancer. Therefore, an estimate of the prevalence is set as the “gold standard” and used to calculate a relative sensitivity and specificity.¹⁹ Relative sensitivity and specificity of DRE was assessed in 10,523 consecutive men randomized to the screening arm of the Rotterdam section of the ERSPC, based on estimates of the predictive index (the number of cancers that would have been detected if all men had been biopsied, the “gold standard”). Of these men, 7055 were found to have PSA values < 4.0 ng/ml. In the PSA-ranges 0-0.9, 1-1.9, 2-2.9 and 3-3.9 ng/ml, positive predictive value (PPV) of a positive DRE was 4%, 10%, 11% and 33% respectively. In these PSA ranges, relative sensitivity and specificity levels were 21%, 24%, 14%, 39% and 94%, 92%, 91% and 98%. Overall PPV of DRE in the PSA-range =< 3.0 ng/ml was 8.8%. Omitting DRE as a screening test in men with a PSA <3.0 ng/ml, would have missed 57 of 473 cancers actually detected (12.1%) and saved 533 biopsies (23.5%). Biopsying every men with a PSA of 3-3.9 ng/ml, would have added 43 cancers and decreased the false-negative biopsy indication drastically. Furthermore, cancers detected in the PSA range < 4.0 ng/ml were classified as minimal, moderate and advanced in 42%, 42% and 16% in men screened during the first round in Rotterdam.^{20,21}

The new screening protocol, with only a PSA higher than or equal to 3.0 ng/ml as a biopsy indication, was implemented from February 1997 and a validation of this protocol was carried out on 7943 men consecutively randomized to the screening arm of the ERSPC Rotterdam.²² It was shown that the detection rate remained almost the same (5.0% vs. 4.7% in the new protocol) and the PPV of a PSA of 3.0 ng/ml or higher, predicted to be 12.3%, was actually 18.0%. The proportion of men with a biopsy indication decreased from 28.2% to 19.5%. Furthermore, the overall tumour characteristics found in the new protocol differed very little from those detected in the old regimen, based on PSA, DRE and TRUS. Therefore, the new protocol contributes to reaching a delicate balance between sensitivity and specificity. However, the search for an even better balance is continuing. More research on the evaluation of screening tests is still being performed and definite judgments on the validity of test regimens will only be possible after the final conclusions of the trial have been reached.

2.4.2 | Biopsy

Similar to the screening regimen policy, the biopsy regimen depends on the choice of the individual screening centers. At the time the ERSPC started, a systematic sextant needle biopsy was the general accepted biopsy regimen among urologists. In Rotterdam, a lateralized sextant biopsy scheme was chosen as the prostate cancer detection rate increases when the lateral peripheral zone is sampled.^{23,24} Nowadays, the trend is to obtain more than six biopsy cores, as this increases the detection rate of prostate cancer.²⁵ Some centers adopted this more extensive biopsy regimen to assure comparability with the control group. Others, including Rotterdam, continue performing sextant biopsies, aiming for maximum of data consistency during the study.

2.5 | Screening interval

2.5.1 | Lead-time

All centres, except for Sweden, have adopted a screening interval of four years. This was based on estimations of lead-time available at the beginning of the ERSPC. Lead-time was estimated to be 6-10 years, based on serum banks used for PSA-determinations and the subsequent diagnosis of clinical prostate cancer.^{26,27} Of course, data that could confirm the correctness of this relatively long interval were highly desirable.

A first evaluation of lead-time came from Finland.²⁸ Auvinen et al defined lead-time as the duration of follow-up needed to accrue the same expected number of incident prostate cancer cases in the absence of screening as detected in the initial screening round. Expected numbers were calculated using an age-cohort model. Based on findings among 10,000 men screened in 1996-1997 with 292 screen-detected cancers, lead-time was estimated as approximately 5-7

years. With the assumption that the cancers are detected on average at the midpoint of the detectable preclinical phase, this detectable preclinical phase was estimated to be 10-14 years.

In the Netherlands, a micro simulation model (MISCAN) was used to estimate lead-time.²⁹ Simulation models are based on results of the Rotterdam section, which enrolled 42,376 men and in which 1,498 cases of prostate cancer were identified, and on baseline prostate cancer incidence and stage distribution data. The models were used to predict mean lead times, over-detection rates, and ranges. Mean lead times and rates of over-detection depended on a man's age at screening. For a single screening test, the estimated mean lead-time was 12.3 years at age 55 and six years at age 75. For a screening program with a 4-year screening interval from age 55 to 67, the estimated mean lead-time was 11.2 years (range 10.8-12.1 years), and the over-detection rate was 48% (range 44%-55%). This screening program increased the lifetime risk of a prostate cancer diagnosis from 6.4% to 10.6%.

These studies seem to confirm the appropriateness of an inter-screening interval of at least 4 years.

2.5.2 | Distribution of prognostic factors at re-screening

For a screening interval to be appropriate, characteristics of tumours found at re-screen should be favourable for curative treatment.

The incidence of potentially advanced malignancies in the second screening round of the Rotterdam section was evaluated by Postma et al.³⁰ Potentially advanced malignancy was defined as a biopsy Gleason score of 7 or higher. During the second screening round, 503 prostate cancers were detected, of which 30 (6.0%) with features of potentially advanced malignancy, in 11,210 screened men. Curative treatment was offered to 26 men, 12 men were treated with radical prostatectomy. Of those 12 RP specimens, 11 showed organ-confined disease. This study showed that potentially advanced disease is a rare finding in the second screening round, and that prostate cancer was still potentially curable in most men.

In addition, other studies demonstrated a shift toward more favourable tumour characteristics in the second screening round, compared to the initial round.^{31,32} In Sweden, where screening is performed with a two-year interval, stage distribution showed a trend toward a lower stage at the second screening (an increase in T1 lesions from 60% to 74% in the second round) as well as lower PSA in men diagnosed with cancer.³¹

First and second round findings from the Rotterdam section were evaluated as well.³² In the second screening round, the mean prostate-specific antigen value was lower (5.6 versus 11.1 ng/mL), advanced clinical stage T3-T4 was 7.1-fold less common, and 76.4% versus 61.5% of the biopsy Gleason scores were less than 7. In the first screening round, 13 regional and 9 distant metastases were detected. In the second round, 2 cases with distant metastasis were found. Overall, a shift toward more favourable tumour characteristics was seen for the second round of screening. These results support the screening methods used and the inter-screening interval of 4 years.

2.5.3 | Interval cancers

An interval cancer is a cancer detected during the interval between 2 screening visits. The rate of interval cancers is an important parameter in determining the sensitivity of the screening procedure and screening interval. In the Swedish centre, 5,854 men participated in the first screening round and 145 prostate carcinomas were detected. During the second screening round, two years later, 5,267 men participated and 111 cancers were found. Nine interval carcinomas were diagnosed (10.6% of the control group prevalence).³¹ Of these, three men had metastatic disease, the others seemed confined to the prostate gland and were detected through opportunistic screening or because of urinary symptoms.

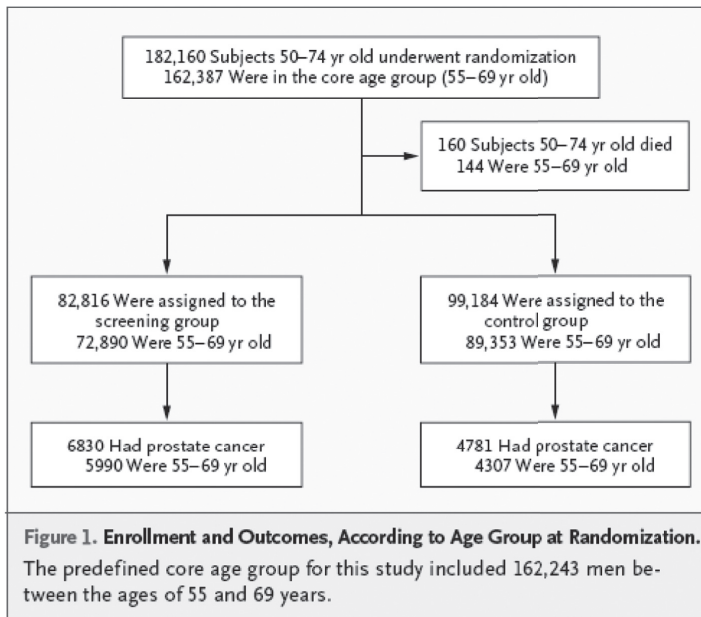
In the Rotterdam section, interval carcinomas were studied in a cohort of 17,226 men (8350 in the screening arm, 8876 in the control arm), enrolled consecutively to the ERSPC.³³ During the first screening round, 412 prostate cancers were detected. During the following four-year interval, 135 cancers were found in the control group, and 25 interval cancers were diagnosed in the screened arm (18.5% of the control group prevalence). Of these 25, seven men had refused a recommended biopsy in the initial screening round. The remaining 18 prostate cancers were all classified as T1 or T2A and none were poorly differentiated or metastatic. These data, which show a low interval carcinoma rate, suggest that the 4-year screening interval is reasonable.

This was supported by a comparative study of the Swedish and Dutch ERSPC study centres, comparing rates of overall interval cancers and high-grade interval cancers.³⁴ The cumulative incidence rates of interval cancers was 0.43% in Rotterdam with a 4-year interval versus 0.74% in Sweden with a 2-year interval ($p=0.51$). The cumulative incidence rates of high-grade interval cancer was 0.11% versus 0.12% ($p=0.72$).

2.6 | First end-point related results

2.6.1 | Recruitment and cancer detection

A total of 182,160 men aged 50-74 participated in the ERSPC, and part of this cohort is still participating in the ongoing study. Of those, 162,387 men were in the core age range group, i.e. 55-69 years old, of which 72,890 were randomized to the screening arm and 89,353 to the control arm (see figure 2 NEJM). Randomization is unequal because Finland did not randomize in a 1:1 ratio. Of the men randomized to the screening arm, 82.2% were actually screened at least once. A total of 5,990 PC cases were detected in the screening arm and 4,307 in the control arm, leading to cumulative incidence rates of 8.2% and 4.8%.¹

Figure 2 | Adapted from¹

2.6.2 | Distribution of prognostic features

For screening to be effective, a stage shift into the direction of a more favorable distribution of prognostic factors with respect to local tumour extent, grading (Gleason score) and presence of metastases is a prerequisite. An early report on this issue is given by Rietbergen et al³⁵ on the Rotterdam section of the ERSPC. The TNM classification of 459 screen-detected prostate cancers was compared to the TNM classification of a cohort of 4,708 men from the Amsterdam Cancer Registry. A stage-shift towards more favorable features was seen for the screen-detected cancers. Furthermore, the incidence of metastases was 24% in the cancer registry cohort, compared to 1.7% in the screen-detected series.

Another report on the issue of metastatic disease comes from the Swedish section.³⁶ Metastatic prostate cancer incidence at diagnosis in a screened cohort was compared with a control cohort, both 10,000 men. For the control group, diagnosis of metastatic prostate cancer was monitored by using the Swedish Cancer Registry. During a time period of 10 years, the risk of being diagnosed with prostate cancer with metastasis at the time of diagnosis differed by 48.9%, decreasing from 47 cases in the control group to 24 cases in the group randomized to PSA-based screening. However, the PC incidence in the screened cohort was 1.8-fold higher than in the control group.

A comparison of all cancers found in the screening and the control arm of the Rotterdam section was performed in 2003.³⁷ By January 1, 2003, 1,269 cancers were detected in the screening arm and 336 were detected in the control arm. A shift to more favourable clinical

stage was seen in the screening arm of the trial. T1C and T2 cancers were 5.8 and 6.2 times more often diagnosed, respectively, in the screening arm than in the control arm of the trial.

A grading shift towards lower Gleason scores in screen-detected PC was reported by Postma et al.³⁸ In radical prostatectomy specimens of the screening arm, 34.6% of the cancers had a Gleason score equal to or higher than 7, a significantly lower proportion as compared to the 53.5% of cancers in the control arm. Furthermore, the median tumour volume was significantly smaller in the screened population (1.0 ml versus 3.9ml).

These studies suggested that the prerequisite for effective screening, i.e. the shift towards more favourable prognostic features, is met in the ERSPC.

2.6.3 | Prostate cancer specific mortality

After a third interim analysis, a significant reduction of prostate cancer mortality was found in the screening arm compared to the control arm.¹ After a median follow-up of 9 years, there were 214 prostate cancer deaths in the screening arm and 326 in the control arm (figure 3). This resulted in an adjusted rate ratio of 0.80 (95% CI 0.65-0.98, p=0.04). Prostate cancer screening, based on PSA testing, thus reduced prostate cancer specific mortality by 20% in the intention-to-screen analysis.

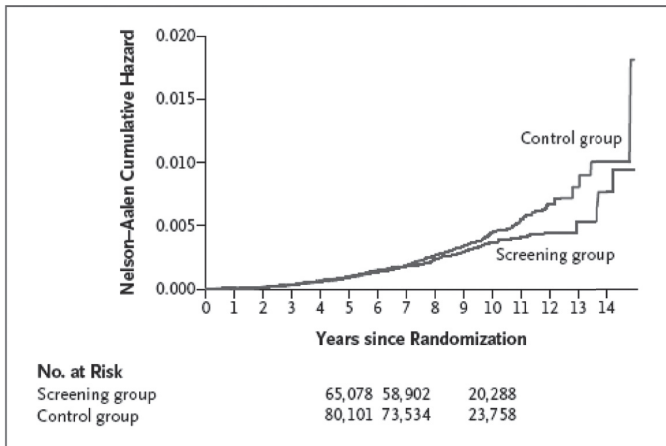


Figure 2. Cumulative Risk of Death from Prostate Cancer.
 As of December 31, 2006, with an average follow-up time of 8.8 years, there were 214 prostate-cancer deaths in the screening group and 326 in the control group. Deaths that were associated with interventions were categorized as being due to prostate cancer. The adjusted rate ratio for death from prostate cancer in the screening group was 0.80 (95% CI, 0.65 to 0.98; P=0.04). The Nelson-Aalen method was used for the calculation of cumulative hazard.

Figure 3 | Adapted from¹

2.7 | Contamination

2.7.1 | Contamination and effective contamination

PSA contamination, i.e. the opportunistic PSA based screening in the control arm of the study, can jeopardize the power of the ERSPC. With increasing contamination during the early years of the study, the power decreases, necessitating higher numbers of participants in both the screening and control arm. Therefore, the extent of contamination has been carefully studied in several participating centres.³⁹⁻⁴²

However, it is important to realize that “effective” contamination cannot be determined by assessing the number of PSA tests only. It is necessary to evaluate the number of men who were PSA tested and who subsequently had a biopsy if indicated by the PSA level.

Such an analysis of “effective contamination” in the Rotterdam area was reported by Otto et al.⁴² During a period of 2.9 years, 2895 of 14,349 men (20.2%) in the control arm were PSA tested and 1981 of 14,052 men (14.1%) in the screening arm were PSA-tested outside the screening protocol. The proportion of men in the control arm with a PSA ≥ 3.0 ng/ml followed by a biopsy and a prostate cancer diagnosis was 7-8% and 3% respectively (3% and 0.4-0.6% in the screening arm). Therefore, although the PSA testing rate in the control arm was high, this was not followed by a substantial increase in prostate biopsies and, more importantly, in detection of prostate cancer. The effective PSA contamination was relatively small and may not jeopardize the power of the trial. Furthermore, the rate of effective contamination was in line with the predicted contamination rate of 20% used in adjusting the initial sample size of the ERSPC.¹³

2.7.2 | Adjustment for contamination and non-compliance

Next to contamination, another process that may dilute the effects of the ERSPC is non-compliance. Non-compliers are participants randomized to the screening arm, who are not screened or do not participate in the whole screening protocol. When an adjustment is made for contamination and non-compliance during the final analysis of the ERSPC, the unbiased effect of screening can be determined for those who are willing to participate in a screening program. A method for this adjustment has been described by Cuzick et al.⁴³

Roemeling et al reported on a feasibility study and impact simulation of a secondary analysis, according to Cuzick et al, on the results of the Rotterdam center.⁴⁴ Endpoints in this analysis were simulated prostate cancer mortality reductions and contamination was defined as a PSA test only. This study concluded that the adjustment for contamination and non-compliance was feasible. A second analysis according to Cuzick was described by Kerkhof et al.⁴⁵ In this analysis data on the presence of metastasis at diagnosis, as a proxy for mortality, in both study arms from the Rotterdam section were evaluated. A non-compliance rate of 26% and a contamination rate of 12% were observed. Prostate cancer screening significantly reduced the occurrence of metastasis in the intention-to-screen analysis (risk ratio 0.75, 95% CI 0.59-0.95, $p=0.02$). After adjustment for both contamination and non-compliance, the risk was further reduced: RR 0.68,

95% CI 0.49-0.94, $p=0.02$. The authors concluded that the screening effect on those who are actually screened is approximately 28% greater than the overall effect seen without adjustment for contamination and non-compliance.

Finally, a secondary analysis on the final results of the ERSPC was carried out to provide accurate information for those men actually screened. Roobol et al⁴⁶ reported a relative risk reduction of 0.69 (95% CI 0.51-0.92) to 0.71 (95% CI 0.55-0.93) for men actually screened after adjustment for contamination and non-compliance using 2 different estimates for contamination. Thus, PSA screening reduces the risk of dying from prostate cancer by up to 31% in men who are actually screened.

2.8 | Overdiagnosis and indolent disease

2.8.1 | Overdiagnosis

Overdiagnosis is the detection of prostate cancer that would never have been diagnosed without screening. Those cases of prostate cancer would not have led to symptoms or death during life and therefore would not have been diagnosed clinically. This may relate to a relatively harmless tumour behaviour or to competing causes of death. However, if such a carcinoma is found through screening, adverse psychological and physical effects may arise. Especially, invasive treatment for prostate cancer, which would not have been clinically relevant, with possible side effects, may be harmful.

Using the micro simulation (MISCAN) model, the extent of overdiagnosis in the Rotterdam section of the ERSPC was estimated.²⁹ A 100% attendance rate for each screening program was assumed. Obviously, overdiagnosis leads to a rise in prostate cancer incidence. The first-round detection rate in the screened arm was almost 17 times as high as the prostate cancer detection rate in the control arm (54 versus 3.18 cases per 1000 man-years).

Screening men aged 55 to 67 years with a 4-year screening interval detects 41 irrelevant cancers in 1000 men. This corresponds with an overdiagnosis rate of 48% (range 44%-55%). Furthermore, the lifetime prostate cancer risk increases from 64 to 106 per 1000, a relative increase of 65%.

This amount of overdiagnosis may be unacceptable in population-based screening, for both health care providers and policy makers. Reduction of overdiagnosis by increasing specificity, i.e. through individualizing screening programs and finding new markers, will be of great importance during the years to come.⁴⁷

2.8.2 | Indolent prostate cancer

However, as long as screening can not be made more selective for aggressive or clinically significant prostate cancer, another approach was taken with the identification of potentially indolent cancers. Indolent prostate cancer is cancer that will not cause any symptoms or

mortality during life due to a relatively benign behaviour. If such a tumour is diagnosed and treated radically, side effects may occur, possibly reducing the quality of life. To avoid this unnecessary treatment or overtreatment, many attempts are made to identify indolent prostate cancer. For the screening situation, a nomogram was described by Steyerberg et al based on data from the ERSPC.⁴⁸ This nomogram can be used to identify 20-30% of screen-detected PC as “potentially indolent” depending on the probability level chosen (70-80%).

Instead of radical treatment, those cases of indolent cancer may be offered active surveillance, a strategy of closely monitoring the patient and offering radical treatment with curative intent if signs of progression occur. The safety and feasibility of such a treatment strategy is subject of investigation of the project Prostate Cancer Research International: Active Surveillance (PRIAS).⁴⁹ This international website-based study has been launched at the end of 2006 and has included around 850 prostate cancer patients up to June 2009. First reports of this study show favourable results on quality of life and feasibility. Furthermore, no adverse effects of active surveillance were shown in a retrospective study.⁵⁰ However, the safety of active surveillance remains to be proven.

2.9 | Quality of life and costs of screening

2.9.1 | Quality of life

The evaluation of quality of life in relation to screening and treatment aspects of prostate cancer has been considered essential from the start of the ERSPC. Therefore, next to mortality reduction, quality adjusted life years (QALYs) will be calculated. Unfortunately, only in one of the centers (Rotterdam) a truly systematic study of quality of life has been conducted and is still ongoing. So far, only short-term effects could be analyzed and final conclusions about the impact of prostate cancer screening on quality of life will only become clear after the final analysis of the ERSPC.

Up to now, it has been shown that prostate cancer screening induced no important short-term health status effects, with some exceptions of high levels of anxiety in subgroups.⁵¹ Health-related quality of life (HRQoL) was related to tumour stage and the detection method (screen vs. clinically detected PC).⁵² Furthermore, the type of post-treatment HRQoL impairments was dependent on treatment modality (prostatectomy or radiotherapy). Patients with screen-detected or clinically diagnosed PC reported similar post-treatment HRQoL.⁵³

Prostate cancer diagnosis was shown to worsen mental and self-rated overall health immediately after diagnosis in screened patients. However, six months later, health status scores improved and no longer differed significantly from pre-diagnosis scores.⁵⁴

2.9.2 | Costs of screening

Due to screening, some cases will be prevented from reaching the advanced stage. In order to evaluate a screening program thoroughly, it is important to quantify course, care, and accompanying costs of advanced disease. Data on these factors are reported in.⁵⁵ Together with the effects of advanced prostate cancer on quality of life, these data will be used for the evaluation of prostate cancer screening. A report on the costs of screening is in progress.

2.10 | Future prospects

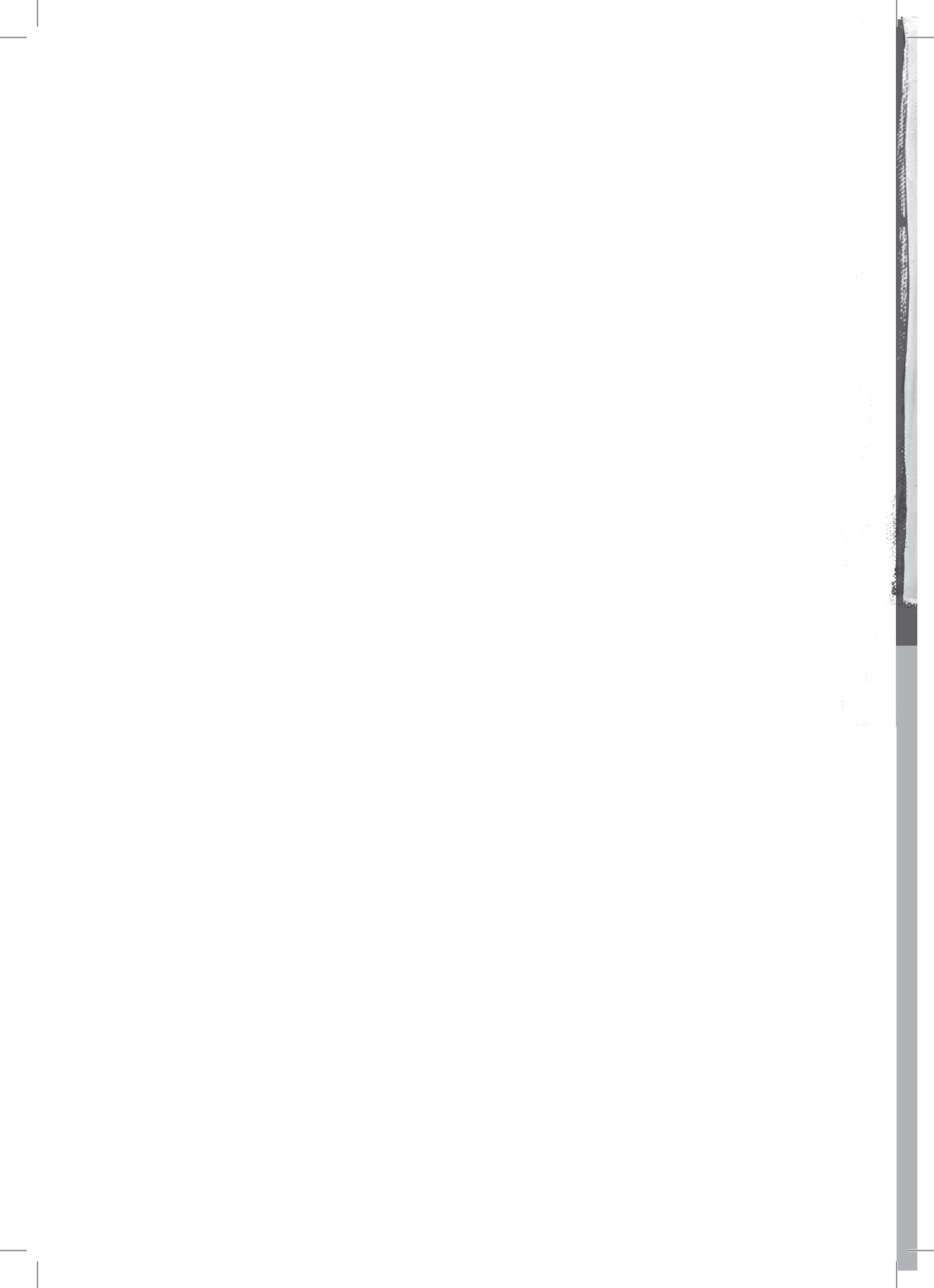
In the ERSPC, a prostate cancer specific mortality reduction has been shown due to PSA based screening. However, 1410 men need to be screened and 48 cases needed to be treated in addition to the control arm to avoid one prostate cancer death after 9 years of observation.¹ These numbers needed to screen and needed to treat are very high and could be lowered by increasing the screening specificity for aggressive disease and avoiding overdiagnosis and overtreatment. This is very important, as screening for prostate cancer is being performed on a growing scale by general practitioners and urologists, independently of the final recommendations of the ERSPC. Therefore, attention should be drawn to optimizing the screening regimen. For example, the use of individual characteristics of the screening participant, new biomarkers, and the use of nomograms may play a role in this search for optimal future screening programs.

References

1. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berebnguer A, Määttä L, Bangma CH, Aus G, Villers A, Rebillard X, Van der Kwast ThH, Blijenberg BG, Moss SM, De Koning HJ, Auvinen A. Screening and prostate-cancer mortality in a randomized European Study. *N Engl J Med* 2009;360:1320-8
2. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, Petros JA, Andriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991;324:1156-61
3. Schröder FH. (1993) Prostate cancer: to screen or not to screen? *BMJ* 1993;306:407-8.
4. Schröder FH, Boyle P. Screening for prostate cancer- necessity or nonsense? *Eur J Cancer* 1993;29A:656-61.
5. Schröder FH, Damhuis RAM, Kirkels WJ, De Koning HJ, Kranse R, Nijs HGT, Blijenberg BG. European randomized screening for prostate cancer- The Rotterdam pilots. *Intern J Cancer* 1996;65:145-51.
6. Schröder FH, Denis LJ, Kirkels WJ, De Koning HJ, Standaert B. European randomized study of screening for prostate cancer: Progress report of Antwerp and Rotterdam pilot studies. *Cancer* 1995;76:129-34.
7. Auvinen A, Tammela T, Stenman U-H, Uusi-Erkilä I, Schröder FH, Hakama M. Screening for prostate cancer using serum prostate-specific antigen; a randomized, population-based pilot study in Finland. *Br J Cancer* 1996;74:568-72
8. Adami HO, Baron JA, Rothman KJ. Ethics of a prostate cancer screening trial. *Lancet* 1994;343:958-60
9. Schröder FH. Screening for prostate cancer (letter to the editor). *Lancet* 1993;343:1438-9.
10. Schröder FH. Detection of prostate cancer, screening the whole population has not yet been shown to be worth while (letter to the editor). *BMJ* 1995;310:140-1.
11. Denis L, Standaert B. Contributors Chapter Prostate Cancer. (Döbrösy L ed.) Prevention in primary care. Recommendations for promoting good practice. Copenhagen, SHO regional office for Europe, 1995, 139-45.
12. Denis LJ, Murphy GP, Schröder FH. Report of the consensus workshop on screening and global strategy for prostate cancer. *Cancer* 1995;75:1187-207.
13. De Koning HJ, Liem MK, Baan CA, Boer R, Schröder FH, Alexander FE. Prostate cancer mortality reduction by screening: power and time frame with complete enrollment in the European Randomized Study of Screening for Prostate Cancer. *Int J Cancer* 2002;98:268-73.
14. Ciatto S, Bonardi R, Mazzotta A, Lombardi C, Santoni R, Cardini S, Zappa M. Comparing two modalities of screening for prostate cancer: digital rectal examination and transrectal ultrasonography vs. prostate specific antigen. *Tumori* 1995;81:225-9.
15. Bangma CH, Kranse R, Blijenberg BG, Schröder FH. The value of screening tests in the detection of prostate cancer. Part I: Results of a retrospective evaluation of 1726 men. *Urology* 1995;46:773-778.
16. Bangma CH, Kranse R, Blijenberg BG, Schröder FH. (1995) The value of screening tests in the detection of prostate cancer. Part II: Retrospective analysis of free/total prostate specific analysis ratio, age-specific reference ranges, and PSA density. *Urology* 1995;46:779-784.
17. Beemsterboer PMM, Kranse R, De Koning HJ, Habbema JDF, Schröder FH. Changing role of 3 screening modalities in the European Study of Screening for Prostate Cancer (Rotterdam). *Int J Cancer* 1999;84:437-41.
18. Kranse R, Beemsterboer PMM, Rietbergen JBW, Habbema D, Hugosson J, Schröder FH. (1999) Predictors for biopsy outcome in the European Study of Screening for Prostate Cancer (Rotterdam region). *The Prostate* 1999;39:316-22.
19. Schröder FH, Alexander FE, Bangma CH, Hugosson J, Smith DS. Screening and early detection of prostate cancer. *Prostate* 2000;44:255-63.
20. Schröder FH, van der Maas P, Beemsterboer P, Kruger AB, Hoedemaeker R, Rietbergen J, Kranse R. Evaluation of the digital rectal examination (DRE) as a screening test for prostate cancer. *JNCI* 1998;90:1817-23.
21. Schröder FH, Van der Crujisen-Koeter I, De Koning HJ, Vis AN, Hoedemaeker RF, Kranse R. Prostate cancer detection at low prostate specific antigen. *J Urol* 2000;163:806-12.

22. Schröder FH, Roobol-Bouts MJ, Vis AN, Van der Kwast ThH, Kranse R. PSA-based early detection of prostate cancer - validation of screening without rectal examination. *Urology* 2001;57:83-90.
23. Stamey TA. Making the most out of six systematic sextant biopsies. *Urology* 1995;45:2-12.
24. Eskew LA, Bare RL, McCullough DL. Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate. *J Urol* 1997;157:199-202.
25. Aus G, Ahlgren G, Hugosson J, Pedersen KV, Rensfeldt K, Söderberg R. Diagnosis of prostate cancer: optimal number of prostate-specific antigen and findings on digital rectal examination. *Scand J Urol Nephrol* 1997;31:541-4.
26. Stenman UH, Hakama M, Knekt P, Aromaa A, Teppo L, Leinonen J. Serum concentrations of prostate-specific antigen and its complex with α 1-antichymotrypsin before diagnosis of prostate cancer. *Lancet* 1994;334:1594-8.
27. Parkes C, Wald NJ, Murphy P, George L, Watt HC, Kirby R, Knekt P, Helzlsouer KJ, Tuomilehto J. Prospective observational study to assess value of prostate-specific antigen as a screening test for prostate cancer. *BMJ* 1995;311:1340-3.
28. Auvin A, Määttänen L, Stenman UH, Tammela T, Rannikko S, Aro J, Juusela H, Hakama M. Lead-time in prostate cancer screening (Finland). *Cancer Causes Control* 2002;13:279-85.
29. Draisma G, Boer R, Otto SJ, van der Crujisen IW, Damhuis RA, Schröder FH, de Koning HJ. Lead-time and overdiagnosis due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *JNCI* 2003;95:868-78.
30. Postma R, Roobol MJ, Schröder FH, Van der Kwast TH. Potentially advanced malignancies detected by screening for prostate carcinoma after an interval of 4 years. *Cancer* 2004;100:968-75.
31. Hugosson J, Aus G, Lilja H, Lodding P, Pihl C-G, Pileblad E. Prostate specific antigen based biennial screening is sufficient to detect almost all prostate cancers while still curable. *J Urol* 2003;169:1720-3.
32. Van der Crujisen-Koeter IW, Roobol MJ, Wildhagen MF, Van der Kwast TH, Kirkels WJ, Schröder FH. Tumour characteristics and prognostic factors in two subsequent screening rounds with four-year interval within prostate cancer screening trial, ERSPC Rotterdam. *Urology* 2006;68:615-20.
33. Van der Crujisen-Koeter IW, Van der Kwast TH, Schröder FH. Interval carcinomas in the European Randomized Trial of Screening for Prostate Cancer (ERSPC) – Rotterdam. *JNCI* 2003;95:1462-6.
34. Roobol MJ, Grenabo A, Schröder FH, Hugosson J. Interval cancers in prostate cancer screening: comparing 2- and 4-year screening intervals in the European Randomized Study of Screening for Prostate Cancer, Gothenburg and Rotterdam. *J Natl Cancer Inst* 2007;99:1296-303
35. Rietbergen JBW, Hoedmaeker RF, Boeken Kruger AE, Kirkels WJ, Schröder FH. The changing pattern of prostate cancer at the time of diagnosis: characteristics of screen detected prostate cancers in a population-based screening study. *J Urol* 1999;161:1191-8.
36. Aus G, Bergdahl S, Lodding P, Lilja H, Hugosson J. Prostate cancer screening decreases the absolute risk of being diagnosed with advanced prostate cancer- results from a prospective, population-based randomized controlled trial. *Eur Urol* 2007;51:659-64.
37. Van der Crujisen-Koeter IW, Vis AN, Roobol MJ, Wildhagen MF, De Koning HJ, Van der Kwast TH, Schröder FH. Comparison of screen-detected and clinically diagnosed prostate cancer in the European Randomized Study of Screening for Prostate Cancer, section Rotterdam. *J Urol* 2005;174:121-5.
38. Postma R, Van Leenders GJ, Roobol MJ, Schröder FH, Van der Kwast ThH. Tumour features in the control and screening arm of a randomized trial of prostate cancer. *Eur Urol* 2006;50:70-5.
39. Lujan M, Paez A, Pascual C, Angulo J, Miravalles E, Berenguer A. Extent of prostate-specific antigen contamination in the Spanish section of the European Randomized Trial of Screening for Prostate Cancer (ERSPC). *Eur Urol* 2006;50:1234-40.
40. Paez A, Lujan M, Llanes L, Romero I, de la Cal MA, Miravalles E, Berenguer A. PSA-use in a Spanish industrial area. *Eur Urol* 2002;41:162-6.
41. Beemsterboer PM, de Koning HJ, Kranse R, Trienekens PH, van der Maas PJ, Schröder FH. Prostate specific antigen testing and digital rectal examination before and during a randomized trial of screening for prostate cancer: European Randomized Study of Screening for Prostate Cancer, Rotterdam. *J Urol* 2000;164:1216-20.

42. Otto SJ, van der Crujisen IW, Liem MK, Korfage IJ, Lous JJ, Schröder FH, de Koning HJ. Effective PSA contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer. *Int J Cancer* 2003;20:394-9.
43. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med* 1997;16:1017-122.
44. Roemeling S, Roobol MJ, Otto SJ, Habbema DF, Gosselaar C, Lous JJ, Cuzick J, Schröder FH. Feasibility study of adjustment for contamination and non-compliance in a prostate cancer screening trial. *Prostate* 2007;67:1053-60.
45. Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schröder FH, Steyerberg EW. Effect of the correction for non-compliance and contamination on the estimated reduction of metastatic prostate cancer within a randomized screening trial (ERSPC section Rotterdam). *J Urol* 2009;181(suppl 1):233.
46. Roobol MJ, Kerkhof M, Schröder FH, Cuzick J, Sasieni P, Hakama M, Stenman UH, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis L, Recker F, Berenguer A, Ruutu M, Kujala P, Bangma CH, Aus G, Tammela TL, Villers A, Rebillard X, Moss SM, de Koning HJ, Hugosson J, Auvinen A. Prostate Cancer Mortality Reduction by Prostate-Specific Antigen-Based Screening Adjusted for Nonattendance and Contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 2009;56:584-591
47. Bangma CH, Roemeling S, Schröder FH. Overdiagnosis and overtreatment of early detected prostate cancer. *World J Urol* 2007;25:3-9.
48. Steyerberg EW, Roobol MJ, Kattan MW, Van der Kwast ThH, De Koning HJ, Schröder FH. Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. *J Urol* 2007;177:107-12.
49. Van den Bergh RC, Roemeling S, Roobol MJ, Roobol W, Schröder FH, Bangma CH. Prospective validation of active surveillance in prostate cancer: the PRIAS study. *Eur Urol* 2007;52:1560-3.
50. Van den Bergh RC, Roemeling S, Roobol MJ, Aus G, Hugosson J, Rannikko AS, Tammela TL, Bangma CH, Schröder FH. Outcomes of Men with Screen-Detected Prostate Cancer Eligible for Active Surveillance Who Were Managed Expectantly. *Eur Urol* 2009; 55:1-8.
51. Essink-Bot ML, De Koning HJ, Nijs HGT, Kirkels WJ, Van der Maas PJ, Schröder FH. Short-term effects of population-based screening for prostate cancer on health-related quality of life. *JNCI* 1998;12:925-31.
52. Madalinska JB, Essink-Bot ML, De Koning HJ, Kirkels WJ, Van der Maas PJ, Schröder FH. Health-related quality of life in patients with screen-detected versus clinically diagnosed prostate cancer preceding primary treatment. *Prostate* 2001;46:87-97.
53. Madalinska JB, Essink-Bot ML, De Koning HJ, Kirkels WJ, Van der Maas PJ, Schröder FH. Health-related quality-of-life effects of radical prostatectomy and primary radiotherapy for screen-detected or clinically diagnosed localized prostate cancer *J Clin Oncol* 2001;19:1619-28.
54. Korfage IJ, De Koning HJ, Roobol MJ, Schröder FH, Essink-Bot ML. Prostate cancer diagnosis: the impact on patient's mental health. *Eur J Cancer* 2006;42:165-70.
55. Beemsterboer PMM, De Koning HJ, Birnie E, Van der Maas PJ, Schröder FH. Advanced disease in prostate cancer, course, care and cost implications. *The Prostate* 1999;40:970-104.





CHAPTER 3

Scope and outline of the thesis

3.1 | Scope

PSA based screening for prostate cancer reduces prostate cancer specific mortality.¹ Due to screening the prostate cancer is diagnosed earlier in time and hopefully curable treatment can be offered and disease specific mortality can be reduced. Differences in treatment distribution between the screening and the control arm of the study and possible treatment bias were assessed in the second part of the thesis.

Although a mortality reduction has been shown due to screening, some important drawbacks prevent the introduction of a population based screening program at the moment. The first reason is the lack of specificity of PSA for prostate cancer in general, with a subsequent large number of unnecessary biopsies. The third part of this thesis contributes to improvement of the current screening protocol.

Another essential limitation of screening is the substantial overdiagnosis and subsequent overtreatment. This is reflected in a high number needed to treat ($n=48$, in excess to the control group). The overdiagnosis results from the lack of specificity of the PSA based screening protocol for clinically significant prostate cancer. However, as long as the protocol cannot be made more selective for aggressive disease, we will need to be able to differentiate indolent from aggressive prostate cancer and possibly prevent overtreatment. In the last part of the thesis an effort is made to improve this differentiation.

3.2 | Outline

A different distribution of treatment for prostate cancer in the screening and control arm seems logical, with the observed stage and grade shift between both study arms. However, if differences in age, stage and grade do not completely explain the discrepancy in treatment between both arms, the observed mortality reduction could be at least in part be due to differences in treatment. In Part II, **chapter 4** describes a study which evaluates treatment in both arms and the possible influence of differences in treatment on prostate cancer mortality.

In Part III, an attempt is made to improve the current screening protocol. First, the performance of PSA velocity, i.e. the rise in PSA per year, as a screening test was assessed in **Chapter 5**. Specifically, the detection of clinically significant prostate cancer was described. In **Chapter 6** it was investigated whether pathological characteristics of a biopsy could identify men with increased risk of prostate cancer diagnosis during a next screening visit. If such characteristics could be identified, the screening protocol could be intensified for men harbouring these high risk features or the protocol could be made less strict for men without these characteristics. Another study on prostate biopsies is described in **Chapter 7**. The number and characteristics of suspicious lesions and prostate cancer lesions are described that were

missed during the original pathological examination. Knowledge about these missed lesions could improve the diagnostic accuracy and prevent the delay of prostate cancer diagnosis.

In Part IV, **Chapter 8** contains an evaluation of the independent prognostic value of tumour volume. Although tumour volume alone unequivocally has been reported to be predictive for patient outcome as a sole predictor, its value after correction in multivariable analysis for other prognosticators like tumour stage and grade has been disputed. In **Chapter 9**, we reassessed the 0.5 ml tumour volume threshold commonly used to identify indolent prostate cancer on a radical prostatectomy series from the Rotterdam section of the ERSPC. Additionally, the prognostic value of the tumour volume threshold for indolent disease was evaluated. Finally, a study which aims to improve the differentiation of clinically significant and indolent cancer with the use of immunohistochemical staining on biopsy specimens is described in **Chapter 10**.

References

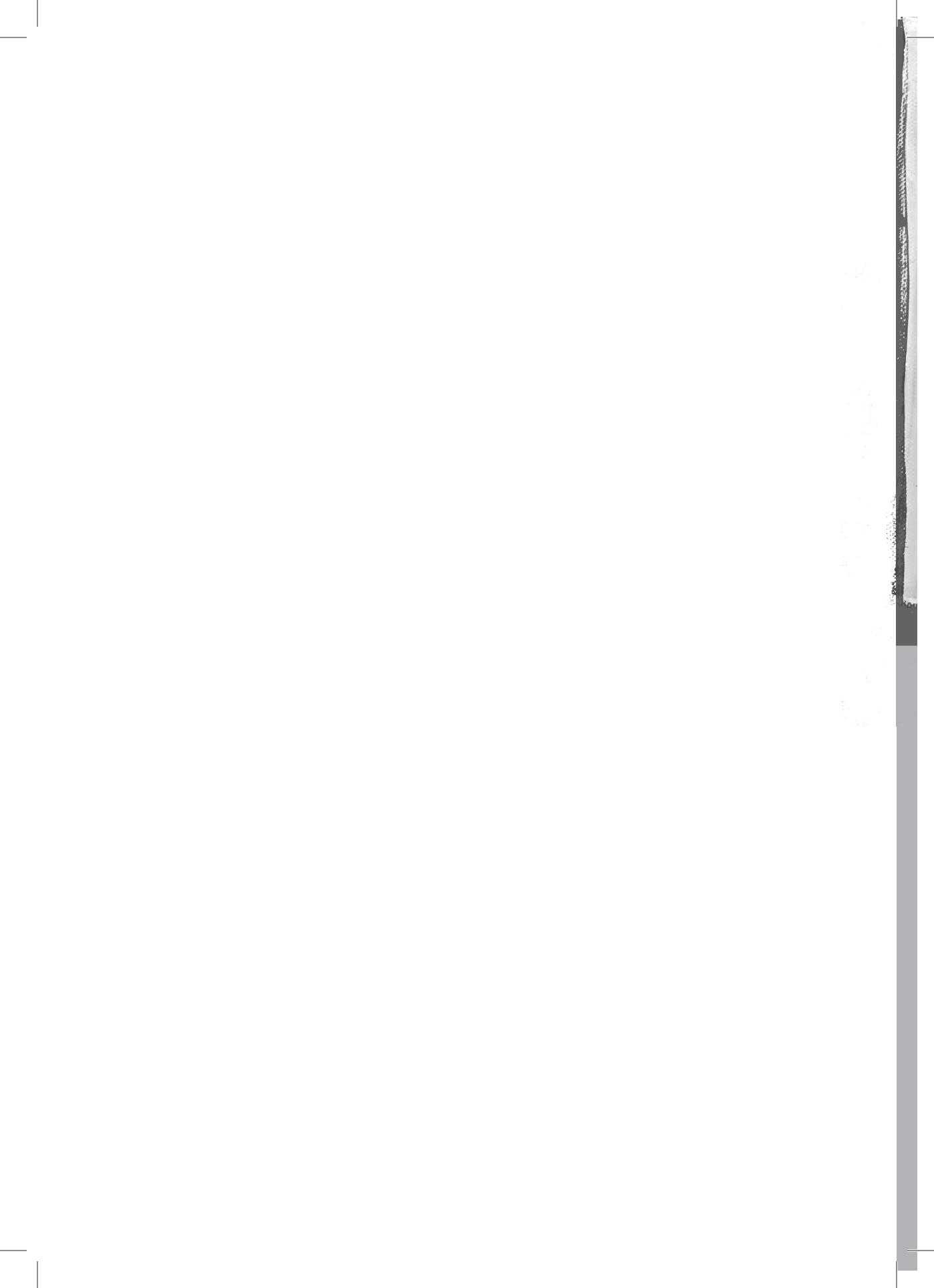
1. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Määttä L, Bangma CH, Aus G, Villers A, Rebillard X, Van der Kwast ThH, Blijenberg BG, Moss SM, De Koning HJ, Auvinen A. Screening and prostate-cancer mortality in a randomized European Study. *N Engl J Med* 2009;360:1320-8
2. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumour extent of nonpalpable (stage T1c) prostate cancer. *JAMA* 1994;271:368-74

PART II

THE DISEASE SPECIFIC MORTALITY REDUCTION DUE TO SCREENING FOR PROSTATE CANCER- THE ROLE OF TREATMENT

Chapter 4

The effect of study arm on prostate cancer treatment in a large screening trial (ERSPC)





CHAPTER 4

The effect of study arm on prostate cancer treatment in a large screening trial (ERSPC)

Int J Cancer. 2009, epub ahead of print, doi:10.1002/ijc.24870

T. Wolters, M.J. Roobol, E.W. Steyerberg, R.C.N. van den Bergh, C.H. Bangma, J. Hugosson, S. Ciatto, M. Kwiatkowski, A. Villers, M. Luján, V. Nelen, T.L.J. Tammela and F.H. Schröder

Abstract

Prostate cancer (PC) mortality is the most valid end-point in screening trials, but could be influenced by the choice of initial treatment if treatment has an effect on mortality. In this study, PC treatment was compared between the screening and control arms in a screening trial.

Data were collected from the European Randomized Study of Screening for Prostate Cancer (ERSPC). Characteristics and initial treatment of PC cases detected in the screening and the control arm were compared. Polytomous logistic regression analysis was used to assess the influence of study arm on treatment, adjusting for potential confounders and with statistical imputation of missing values.

A total of 8,389 PC cases were detected, 5,422 in the screening arm and 3,145 in the control arm. Polytomous regression showed that trial arm was associated with treatment choice after correction for missing values, especially in men with high risk PC. A control subject with high-risk PC was more likely than a screen subject to receive radiotherapy (OR 1.43, 95% CI 1.01-2.05, $p=0.047$), expectant management (OR 2.92, 95% CI 1.33-6.42, $p=0.007$) or hormonal treatment (OR 1.77, 95% CI 1.07-2.94, $p=0.026$) instead of radical prostatectomy. However, trial arm had only a minor role in treatment choice compared to other variables.

Concluding, a small effect of trial arm on treatment choice was seen, particularly in men with high-risk PC. Therefore, differences in treatment between arms are unlikely to play a major role in interpretation of the results of the ERSPC.

4.1 | Introduction

With the introduction of the serum prostate-specific antigen (PSA) test, early detection of prostate cancer has become possible. PSA is nowadays widely used as a screening test for prostate cancer. Recently, the European Randomized Study of Screening for Prostate Cancer (ERSPC) has reported a disease-specific mortality reduction due to screening¹.

The mortality reduction induced by screening for PC may be influenced by differences in treatment. Due to early detection, prostate cancer is diagnosed in an earlier stage and grade and may subsequently be more often suitable for radical treatment, like surgery. A different treatment distribution is thus expected between the intervention and the control arm of the ERSPC. However, similar PC cases are preferably treated similarly in both arms. Otherwise, the mortality reduction may not only be due to early detection and early treatment, but could possibly also be caused by different treatment choices for the same types of patients as treatment choice is intermediate between randomization to a study arm and mortality. Since study arm cannot be blinded, different treatment among diagnosed cancers could have an impact on study arm differences, i.e. mortality. We analyze treatment choice as an embedded observational study within the parent clinical trial, i.e. the ERSPC.

However, a complete separation between screening per se and treatment effect cannot be made since a screen detected tumour differs in various ways from a clinically detected tumour. Especially, correction may not be fully possible for the lead-time effect in screening.

In the current study, we aim to describe and compare the treatment modalities in the intervention and control arm of the ERSPC, with correction for available patient and tumour characteristics as far as possible.

4.2 | Methods

The main goal of the ERSPC trial was to show or exclude a 20% reduction in PC mortality due to screening. The trial started in 1993 in Belgium and the Netherlands, soon followed by Finland, Sweden, Spain, Italy, Switzerland and France.² The recruitment has been completed with 162,243 men randomized in the core age range of 55-69 years old, 72,890 in the intervention arm and 89,353 in the control arm¹. Informed consent was obtained from all subjects.

The ERSPC has established a Central Database for data monitoring and joint analyses.³ From this Central Database, numbers of cancers in the screening and control arm were obtained. Furthermore, age at diagnosis, PSA at diagnosis, Gleason score, TNM stage and initial treatment were retrieved. Data were complete up to December 31, 2006, with identical follow-up in the two study groups.¹

Patient and tumour characteristics, and the initial treatment were described for the whole cohort and compared between study arms. Treatment was classified as radical prostatectomy (RP), radiotherapy (RT), active surveillance (AS) or hormonal therapy (HT). The treatment which was chosen as original treatment was scored. For example, when AS was chosen but HT was applied after a year in case of symptoms, treatment was scored as AS. Men who declined treatment were classified in the AS group. To describe treatment per risk group, all PC cases were divided in three risk groups: low, intermediate and high risk PC, according to the criteria of d'Amico et al⁴. Low risk was defined as stage \leq T2a and PSA \leq 10 ng/ml and Gleason score \leq 6. High risk was defined as stage T2c or higher or a PSA-level $>$ 20 ng/ml or Gleason score $>$ 8. The remaining cases (stage T2b or a Gleason score of 7 or a PSA level $>$ 10 ng/ml and \leq 20 ng/ml) were defined as intermediate risk, unless data on PSA, Gleason score or T-stage were missing. Cases with known positive lymph nodes or distant metastases were defined as high risk.

The ERSPC was approved by the local institutional boards of all participating centers.

4.2.1 | Statistical analysis

Differences in treatment between trial arms, i.e. screening arm and control arm, were evaluated using polytomous multivariate logistic regression. In the multivariable analysis, adjustment was made for age (continuous), PSA (log transformation), Gleason score (3 categories: $<$ 7, =7, $>$ 7) and TNM stage. Statistical significance was assessed with the likelihood ratio test, which follows a chi-square distribution. Chi-square values were also used to indicate the relative importance of the variables in the model. The higher the overall total Chi-square value, the better the performance of the predictive model. In addition, the higher the Chi-square value for a particular variable, the more important this variable is in predicting treatment choice. We report odds ratios (OR) for a screen subject compared to a control subject for the 4 treatment modalities based on the polytomous analysis with RP as a reference, using SPSS software (v 15, SPSS Inc, Chicago, Ill).

This multivariable analysis was hampered by missing data for potential confounders. Therefore, an imputation procedure was followed for missing data. Missing values were filled in by a statistical method that accounted for correlations between the variables. We used the first imputation of a multiple imputation procedure with the Impute function in R software (v 2.7.2, R foundation for statistical computing), with inclusion of treatment, study arm, and all potential confounders as variables in the imputation model. A total of 3964 confounder values were missing, comprising 9.5% of all values. By filling in these values, the 3030 patients with any missing value (37.8% of all patients) could be included in the analysis. Apart from increasing sample size, imputation corrects for a possible selection bias due to selective missingness^{5,6}. A p-value less than 0.05 was considered statistically significant.

4.3 | Results

4.3.1 | PC characteristics in the screening and the control arm

A total of 10,297 PC cases were detected, 5,990 in the screening arm and 4,307 in the control arm. For 8,389 (81.5%) cases data on initial treatment were available and those men were included in the analyses (87.5% of all PC cases detected in the screening arm and 73.0% of all PC cases detected in the control arm).

Of 8,389 PC cases, 5,244 (62.5%) had been diagnosed in the screening arm and 3,145 (37.5%) in the control arm. Men diagnosed with PC in the control arm were on average 1 year older than those in the intervention arm (mean age 66.5 vs. 65.2 years, $p < 0.001$). The PSA level was much higher among PC cases in the control arm than in the screening arm (median 10.3 vs. 5.5 ng/ml, $p < 0.001$). Gleason scores and TNM stage were also significantly worse among cases in the control group (**table 1**). In the screened arm, 2766 (52.7%) cases were classified as low-risk PC, 1319 (25.2%) as intermediate risk PC and 1159 (22.1%) as high risk PC. For the control group, these numbers were 873 (27.8%), 976 (31.0%) and 1296 (41.2%) respectively.

Table 1 | General characteristics of the overall cohort and per study arm. All differences were significant at the $p < 0.001$ level.

	Total group N=8389	Screen N=5244	Control N=3145
Age in years Mean (median)	65.7 (66.0)	65.2 (66.0)	66.5 (67.0)
PSA in ng/ml Mean (median)	33.3 (6.8)	17.6 (5.5)	64.2 (10.3)
Gleason			
< 7	4815 (57.4)	3360 (64.1)	1455 (46.3)
= 7	1695 (20.2)	946 (18.0)	749 (23.8)
> 7	766 (9.1)	345 (6.6)	421 (13.4)
unknown	1113 (13.3)	593 (11.3)	520 (16.5)
T-stage			
T1 a-c	4581 (54.6)	3121 (59.5)	1460 (46.4)
T2	2453 (29.2)	1521 (29.0)	932 (29.6)
T3	1985 (11.7)	447 (8.5)	538 (17.1)
T4	161 (1.9)	55 (1.0)	106 (3.4)
unknown	209 (2.5)	100 (1.9)	109 (3.5)
N-stage			
N0	6627 (79.0)	4475 (85.3)	2152 (68.4)
N+	156 (1.9)	66 (1.3)	90 (2.9)
unknown	1606 (19.1)	703 (13.4)	903 (28.7)
M-stage			
M0	7353 (87.7)	4773 (91.0)	2580 (82.0)
M+	379 (4.5)	132 (2.5)	247 (7.9)
unknown	657 (7.8)	339 (6.5)	318 (10.1)

4.3.2 | Initial treatment and differences in treatment by trial arm

The proportions of the treatments were significantly different between the screen and control arms ($p < 0.001$, **table 2**). In the screening arm, surgery (40.3% vs. 30.3%) and AS (21.3% vs. 13.9%) were performed more often than in the control arm. Hormonal treatment (20.8% vs. 7.8%) was used more frequently in the control arm.

We excluded 379 men with distant metastases of whom 353 received HT, to improve the statistical analyses. The model estimation was compromised with inclusions of these cases due to small numbers of men with metastases. No difference in treatment distribution was found between study arms in this subgroup of cases with metastasis. This left a cohort of 8010 men for further analyses (see **table 3**). The treatment modalities in the screening arm and control arm were described stratified per risk group as well. This risk stratification showed that the differences in treatment between the study arms were largely caused by differences in the high-risk group: RP in 34.2% and 19.6% of screen and control cases, HT in 14.7% and 29.5% (Chi square test, $p < 0.001$).

Table 2 | Treatment modalities in the whole cohort and per study arm. Treatment distribution was significantly different between arms (Chi-square test, $p < 0.001$).

Treatment	Total group n=8389 (%)	Screen n=5244 (%)	Control n=3145 (%)
Radical prostatectomy	3067 (36.6)	2113 (40.3)	954 (30.3)
Radiotherapy	2704 (32.2)	1604 (30.6)	1100 (35.0)
Active surveillance	1553 (18.5)	1116 (21.3)	437 (13.9)
Hormonal therapy	1065(12.7)	411 (7.8)	654 (20.8)

The polytomous regression analysis showed that study arm was not a significant factor in treatment choice in the original dataset (**table 4**). In the analysis with completed data, however, study arm was a significant predictor in treatment choice, but with a very small impact compared to the other variables (chi-square 8 on a total chi-square of 2428, **table 5**). PSA and age were the most important factors for treatment choice, followed by T-stage.

A control subject was more likely to receive HT (OR 1.28, 95% CI 1.05-1.57, $p = 0.016$) or RT (OR 1.13, 95%CI 1.01-1.28, $p = 0.039$) than a screen subject using RP as reference treatment. Additional analyses with AS as reference treatment, showed that a control subject was more likely to receive HT instead of AS than a screen subject (OR 1.27 95% CI 1.02-1.59, $p = 0.036$). No differences were found between the other treatment modalities.

Table 3 | Treatment modalities in the cohort and per study arm, excluding men with distant metastases (n=379). Additionally, treatment per arm is described stratified by risk group according to the criteria by d'Amico. Differences in treatment distribution were statistically significant in all risk groups at the p<0.05 level.

Treatment	Total group (%)	Control (%)		Low risk PC		Intermediate risk PC		High risk PC	
		Screen (%)	Control (%)	Screen	Control	Screen	Control	Screen	Control
Radical prostatectomy	3064 (38.3)	2113 (41.3)	951 (32.8)	1099 (39.7)	342 (39.2)	663 (50.3)	403 (41.3)	351 (34.2)	206 (19.6)
Radiotherapy	2689 (33.6)	1597 (31.2)	1092 (37.7)	695 (25.1)	246 (28.2)	419 (31.8)	365 (37.4)	483 (47.0)	481 (45.9)
Active Surveillance	1545 (19.3)	1111 (21.7)	434 (15.0)	916 (33.1)	251 (28.8)	153 (11.6)	130 (13.3)	42 (4.1)	53 (5.1)
Hormonal therapy	712 (8.9)	291 (5.7)	421 (14.5)	56 (2.0)	34 (3.9)	84 (6.4)	78 (8.0)	151 (14.7)	309 (29.5)
Total	8010	5112	2898	2766	873	1319	976	1027	1049

Table 4 | Odds ratios (OR) for all included variables with radical prostatectomy as reference treatment, based on polytomous logistic regression analyses for the original data (n=5285) and the completed data (n=8010). * logarithmic transformation of PSA was used, effect per 10-fold increase.

Treatment	Radiotherapy		Active surveillance		Active surveillance		Hormonal treatment		Hormonal treatment	
	OR (95% CI) original data	OR (95% CI) completed data	OR (95% CI) original data	OR (95% CI) completed data	OR (95% CI) original data	OR (95% CI) completed data	OR (95% CI) original data	OR (95% CI) completed data	OR (95% CI) original data	OR (95% CI) completed data
age	1.13 (1.11-1.15)	1.12 (1.11-1.14)	1.20 (1.18-1.23)	1.19 (1.17-1.21)	1.28 (1.23-1.32)	1.27 (1.24-1.30)				
PSA*	1.91 (1.50-2.42)	2.02 (1.67-2.43)	0.22 (0.16-0.30)	0.30 (0.23-0.39)	16.90 (11.23-25.33)	16.01 (12.07-21.24)				
Tumour stage										
T1	1	1	1	1	1	1	1	1	1	1
T2	0.74 (0.64-0.86)	0.79 (0.71-0.90)	0.26 (0.21-0.32)	0.33 (0.28-0.39)	0.83 (0.61-1.12)	1.11 (0.89-1.39)				
T3	3.76 (2.65-5.32)	4.46 (3.57-5.58)	0.26 (0.11-0.64)	0.36 (0.21-0.61)	1.88 (1.08-3.28)	4.38 (3.18-6.03)				
T4	6.21 (1.27-30.34)	3.26 (1.29-8.21)	3.27 (0.24-43.99)	2.85 (0.68-12.01)	2.75 (0.43-17.47)	5.25 (1.93-14.29)				
Gleason score										
<=6	1	1	1	1	1	1	1	1	1	1
=7	0.95 (0.81-1.12)	1.03 (0.90-1.17)	0.27 (0.21-0.35)	0.30 (0.25-0.37)	0.99 (0.72-1.36)	0.95 (0.75-1.19)				
>=8	1.19 (0.91-1.56)	1.22 (0.99-1.50)	0.11 (0.06-0.23)	0.20 (0.13-0.32)	1.46 (0.93-2.27)	1.29 (0.96-1.75)				
Lymph node involvement										
Negative	1	1	1	1	1	1	1	1	1	1
Positive	0.23 (0.08-0.56)	0.72 (0.43-1.21)	5.48 (1.90-15.78)	1.75 (0.81-3.78)	5.11 (2.42-10.79)	6.24 (3.79-10.27)				
Study arm										
Screen	1	1	1	1	1	1	1	1	1	1
Control	1.11 (0.95-1.30)	1.13 (1.01-1.28)	1.06 (0.84-1.23)	1.01 (0.87-1.17)	1.21 (0.91-1.60)	1.28 (1.05-1.57)				

Subgroup analyses were performed for the risk groups, since treatment distribution seemed especially different between arms in the high risk group and less obvious in low and intermediate risk (table 3). Indeed, study arm remained a significant predictive factor in treatment choice in high risk PC: overall p-value 0.018, with a Chi square of 10 on a total of 401) (model not shown). Compared to a screening subject, a control subject was more likely to receive RT (OR 1.43, 95% CI 1.01-2.05, p=0.047), AS (OR 2.92, 95% CI 1.33-6.42, p=0.007) or HT (OR 1.77, 95% CI 1.07-2.94, p=0.026) instead of RP. No differences between other treatments were found. In low and intermediate risk PC, no significantly predictive value was observed for study arm (p=0.334 and p=0.701 respectively). These results remained unchanged in the completed data.

Table 5 | Overall effect of patient and tumour characteristics on treatment for the original data and the completed data after imputation. The importance of a variable for the model is expressed in the Chi-square value relative to the total Chi-square.

Effect	Original data		Completed data	
	Chi-square	p-value	Chi-square	p-value
Study arm	3	0.421	8	0.042
PSA	801	<0.001	647	<0.001
Age	574	<0.001	862	<0.001
Gleason score	197	<0.001	244	<0.001
T stage	293	<0.001	545	<0.001
N stage	78	<0.001	122	<0.001
Total Chi-square	1923		2428	

4.4 | Discussion

This study shows a major stage and a grade shift of prostate cancer between the screen and the control arm in the ERSPC, with more favorable characteristics in the screening arm. This is in line with earlier reports from individual centers of the ERSPC⁷⁻¹⁰, but has not previously been reported for the whole ERSPC.

Currently, there is a striking difference in the cumulative incidence of PC between the study arms: 5,990 (8.2%) PC cases were detected in the screening arm and 4,307 (4.8%) in the control arm¹. This markedly higher incidence in PSA-screened men has been reported by some individual centers, as well other large databases⁷⁻¹². Two main reasons should be mentioned. First, lead-time bias is inherent to screening: the disease is identified earlier in a screening setting than in the clinical setting. For the ERSPC, this lead-time is calculated to be around 10 years^{13,14}. Therefore, one would expect a higher detection rate in the screened arm and to detect more additional PC cases in the control arm in the future. Second, the high detection rate in the

screening arm is partly caused by the detection of indolent disease (overdiagnosis). Those PC cases would not have been diagnosed during a man's lifetime in the absence of screening, due to a non-aggressive natural course and competing causes of death. Therefore, screening currently leads to over-detection, which presents a challenge to optimize screening and treatment regimens in order to minimize the harms of screening.¹⁵

Our results indicate a dissimilar distribution of initial treatments in the two arms of the ERSPC. This is not surprising as stage and grade are important determinants in treatment choice and these were significantly lower in the screening arm. This is reflected in the distribution of treatment: AS and RP were more often used in the screening arm, whilst especially HT was more frequently chosen in the control group. This observation is in line with a report of the Swedish ERSPC group.⁹ It is worth noting that all expectant management strategies were categorized in one group (AS), while the intent of treatment may differ: curative, similar to active surveillance, or palliative as in watchful waiting.¹⁶ In the trial database, the intent was not recorded. Furthermore, all patients who received radiotherapy were categorized as a single group, including those receiving monotherapy and those who received RT in combination with HT. This approach was chosen because subdivision of the RT group resulted in too small numbers for reliable statistical modeling.

Arm was not significantly predictive for treatment in the total cohort (overall p-value < 0.001) after correction for important factors in PC treatment choice, i.e. age, PSA and tumour characteristics. A bias could be introduced in the analysis due to cases with missing values (3,030 of 8,010, 37.8%). Nowadays, advanced statistical procedures are readily available to fill in missing data.¹⁷ This increases efficiency and limits any selection bias. After correction for missing values using an imputation method some differences in treatment choice were found between both study arms: a control subject was more likely than a screen subject to receive HT or RT instead of RP, and was also more likely to receive HT instead of AS. This could be of importance if treatment affects the outcome within this group. This is especially true for men treated with HT, as those are often men with high-risk PC and thus more prone to die from their disease.

We note that study arm had statistically significant associations with treatment, but was far less important than age, PSA level, T-stage, or Gleason score. Therefore, no systematic discrepancy in treatment selection between arms could be shown and a mortality reduction solely caused by a treatment effect is very unlikely. Inclusion of ERSPC center, or exclusion of non-attendees and interval cancers in the screening arm did not change results (models not shown). The discrepancy was mainly observed in 2076 men with high-risk PC without metastases. Obviously, these patients are more likely to die from their disease. A total of 490 patients died from PC in the total cohort of men included in this report, and 187 (38.2%) of these were in the current selection of 2076 men. The remaining major part of PC deaths occurred in men with metastasis (207 or 42.2%) and in men with low or intermediate risk PC

(96 or 19.6%). In these groups, no treatment difference was observed. Therefore, only a minor part of the PC mortality could be influenced by a treatment difference between arms.

Although the role of study arm was small, it could have some influence on mortality if treatment indeed influences outcome in the high-risk PC group. A screen subject was more likely to receive RP than a control subject. Unfortunately, randomized studies comparing radical prostatectomy to a control group in high risk cases are not available. Available randomized studies in high-risk patients compare RT in combination with HT to RT alone¹⁸ or HT alone¹⁹ and show a mortality difference in favor of combined therapy. Further exploration of the treatment data in the 2076 high-risk patients showed that in the screening arm, 29.2% received RT alone, 14.7% HT alone and 17.8% received a combination of RT with HT. For the control arm these rates were 15.8%, 29.5% and 30.0% respectively. Thus, the only treatment which has shown to be superior is a combination of RT with HT and this treatment was given more often in the control arm. If an effect of different treatment on mortality would occur, this might be in favor of the control arm based on available randomized trials. Concluding, a very small overall effect on PC mortality, if any, is expected from the different treatment in the high-risk PC cases.

The most important reason for the uneven treatment distribution between arms in high risk disease, even when correction for patient and tumour characteristics is made, is that a screen-detected tumour per definition differs from a clinically detected tumour. Even after the correction for age, PSA, Gleason and TNM-stage, a screen-detected tumour is not similar to a clinically detected tumour due to a lead-time effect. For example a screen-detected T3 tumour most probably has a more favorable prognosis than a clinically detected T3 tumour. Clearly, this distinction plays a role in treatment selection. This should not be regarded as a bias, but a screening effect. However, no correction could be made for this factor and this could, at least partly, explain the fact that a control subject was more likely to receive HT. Moreover, men in the control arm were diagnosed clinically and at least some of them were thus presumably symptomatic, while the screen-detected cases were more likely to be non-symptomatic. The presence or absence of symptoms is important in choosing a particular treatment, but no data on symptoms were available.

Some other limitations of the analysis of treatment distribution between arms should be taken into account. Firstly, HT is especially selected in men with high-risk PC or patient-bound factors that make radical therapy less suitable, for instance men with locally advanced PC or men with high age or extensive co-morbidity. In the analysis, correction was made for tumour characteristics, PSA and age. However, extent of co-morbidity was not included in our data. Co-morbidity plays an important role in treatment choice²⁰ and its prevalence and degree may differ between study arms. It is likely that absence of this variable may explain part of the treatment difference: it has been shown that trial participants tend to be healthier than their general population counterparts (“healthy attendee bias”²¹). Therefore, the control subjects may have been less suitable for radical treatment.

Secondly, hospital of treatment was not included, although treatment choice is likely to differ between clinics. Moreover, choice of hospital possibly differs between study arms. Screening for PC within the ERSPC mostly takes place in large (university) hospitals and subsequently the screen subject is more likely to be treated in this university hospital than a control subject who is diagnosed in any hospital. Although participants were treated in similar hospitals in one center, in another center a discrepancy in hospital of treatment was seen: 36.3% of screened men were treated in the screening university hospital versus 7.9% of all control subjects (data not shown). Hypothetically, this could result in differences in treatment choice.

Finally, treatment was not available for all cases initially selected for the analysis, especially in the control arm. These cases were excluded from analysis. Explanations for the missing data include the effect of post randomization consent used in the Scandinavian countries and in Italy, which makes it more difficult to retrieve clinical information from the control group. Although unlikely, this may influence the outcomes of the current analysis. If the data become more complete during the coming years a second analysis can be considered .

4.5 | Conclusions

In conclusion, 8389 cases of prostate cancer with known treatment have been found so far in the core age range cohort of the ERSPC, of which 5,244 in the screening arm and 3,145 in the control arm. A stage and a grade shift were seen, with more favorable characteristics in the screening arm.

Study arm played a statistically significant but minor role in treatment selection in patients with high-risk PC: a control subject was more likely to receive RT, AS or HT instead of RP than a screening subject, but no major differences in other treatment choices were seen. This indicates that an effect of different treatment between arms on PC mortality may be possible but probably will be small. Some important factors could not be corrected for in the analyses, while they may differ by study arm and could explain treatment choice. However, even in absence of these factors, study arm played only a minor role. Therefore, these results show that a mortality reduction in the ERSPC based solely on unequal treatment in both arms is very unlikely.

References

1. Schröder FH, Hugosson J, Roobol MJ et al. Screening and prostate-cancer mortality in a randomized European study. *NEJM* 2009;360:1320-8
2. Schröder FH, Denis LJ, Roobol MJ and all participants of the ERSPC. The story of the European Randomized Study of Screening for Prostate Cancer. *BJU I* 2003;92 suppl 2:1-13
3. Smith P.H. The data monitoring committee-bridging the gap between urology and public health epidemiology. *BJUI* 2003;92 suppl 2:55-56
4. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, Tomaszewski JE, Renshaw AA, Kaplan I, Beard CJ, Wein A. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998;280:969-974
5. Little RJA, Rubin DB. *Statistical analysis with missing data*. John Wiley, New York. 2002
6. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*. Springer, New York. 2008
7. Postma R, Van Leenders GJLH, Roobol MJ, Schröder FH and Van der Kwast TH. Tumour features in the control and screening arm of a randomized trial of prostate cancer. *Eur Urol* 2006;50:70-75
8. Aus G, Bergdahl S, Lodding P, Lilja H and Hugosson J. Prostate cancer screening decreases the absolute risk of being diagnosed with advanced prostate-cancer – results from a prospective, population-based randomized controlled trial. *Eur Urol* 2007;51:659-664
9. Van der Crujisen-Koeter IW, Vis AN, Roobol MJ, Wildhagen MF, De Koning HJ, Van der Kwast TH and Schröder FH. Comparison of screen detected and clinically diagnosed prostate cancer in the European study of screening for prostate cancer, section Rotterdam. *J Urol* 2005;174:121-5
10. Mäkinen T, Tammela TLJ, Hakama M, Stenman U-H, Rannikko S, Aro J, Juusela H, Määttänen L and Auvinen A. Tumour characteristics in a population-based prostate cancer screening trial with prostate-specific antigen. *Clin Cancer Res* 2003;9:2435-9
11. Hugosson J, Aus G, Lilja H, Loddong P and Pihl C-G. Results of a randomized, population-based study of biennial screening using serum prostate-specific antigen measurement to detect prostate carcinoma. *Cancer* 2004;100:1397-1405
12. Potosky AL, Miller BA, Albertsen PC, Kramer BS. The role of increasing detection in the rising incidence of prostate cancer. *JAMA* 1995;273:548-52
13. Draisma G, Boer R, Otto SJ, Van der Crujisen IW, Damhuis RAM, Schröder FH and De Koning HJ. Lead times and overdiagnosis due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *JNCI* 2003;95:868-78
14. Auvinen A, Määttänen L, Stenman U-H, Tammela T, Rannikko S, Aro J, Juusela H and Hakama M. Lead-time in prostate cancer screening (Finland). *Cancer Causes Control* 2002;13:279-285
15. Bangma CH, Roemeling S, Schröder FH. Overdiagnosis and overtreatment of early detected prostate cancer. *World J Urol* 2007;25:3-9
16. Parker C. Active surveillance: towards a new paradigm in the management of early prostate cancer. *Lancet Oncol* 2004;5:101-06
17. Van Dijk MR, Steyerberg EW, Stenning SP, Habbema JDF. Survival estimates of a prognostic classification depended more on year of treatment than on imputation of missing values. *J Clin Epidemiol* 2006;59:246-53
18. Bolla M, Collette L, Blank L, Warde P, Dubois JB, Mirimanoff RO, Storme G, Bernier J, Kuten A, Sternberg C, Mattelaer J, Lopez Torecilla J, Pfeffer JR, Lino Cutajar C, Zurlo A, Pierart M. Long-term results with immediate androgen suppression and external irradiation in patients with locally advanced prostate cancer (an EORTC study): a phase III randomised trial. *Lancet* 2002;360:103-6.
19. Widmark A, Klepp O, Solberg A, Damber JE, Angelsen A, Fransson P, Lund JA, Tasdemir I, Hoyer M, Wiklund F, Fosså SD; Scandinavian Prostate Cancer Group Study 7; Swedish Association for Urological Oncology 3. Endocrine treatment, with or without radiotherapy, in locally advanced prostate cancer (SPCG-7/SFUO-3): an open randomised phase III trial. *Lancet* 2009;373:301-8

20. Desch CE, Penberthy L, Newschaffer CJ, Hillner BE, Whittemore M, McClish D, Smith T, Retchin S. Factors that determine the treatment for local and regional prostate cancer. *Med Care* 1996;34:152-162
21. Otto SJ, Schröder FH, and de Koning HJ. Low all-cause mortality in the volunteer-based Rotterdam section of the European randomised study of screening for prostate cancer: self-selection bias? *J Med Screen* 2004;11:89-92

PART III

HOW TO SCREEN FOR PROSTATE CANCER?

Chapter 5

Is prostate-specific antigen velocity selective for clinically significant prostate cancer in screening? ERSPC Rotterdam

Eur Urol 2009;55:385-92

Chapter 6

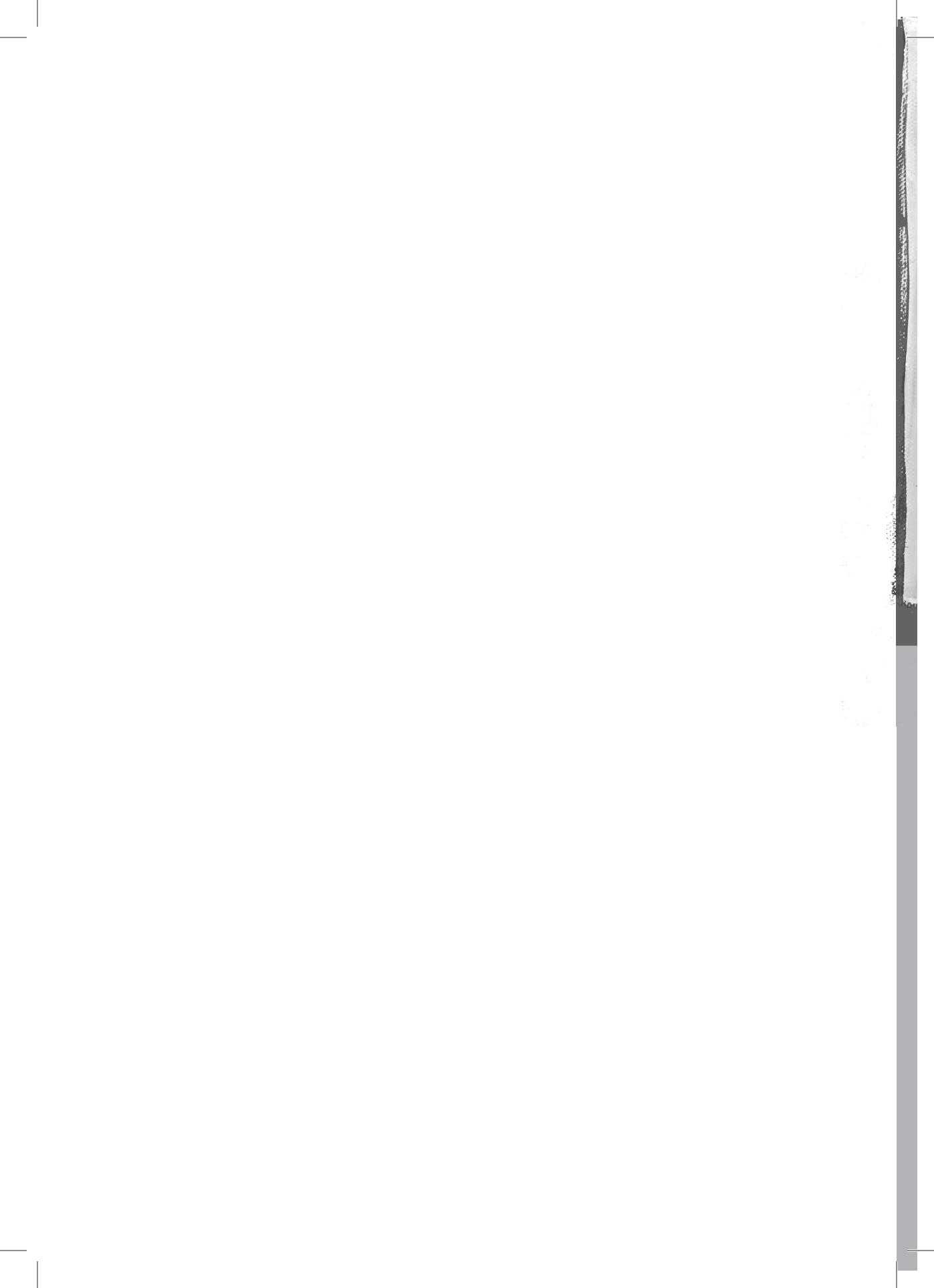
Can non-malignant biopsy features identify men at increased risk of biopsy-detectable prostate cancer at re-screening after four years? (ERSPC Rotterdam)

BJU Int 2008;101:283-8

Chapter 7

False-negative prostate needle-biopsies: frequency, histopathologic features and follow-up

Am J Surg Pathol 2010;34:35-43





CHAPTER 5

Is prostate-specific antigen velocity selective for clinically significant prostate cancer in screening? ERSPC Rotterdam

Eur Urol 2009;55:385-92

Tineke Wolters, Monique J. Roobol, Chris H. Bangma and Fritz H. Schröder

Abstract

Background: The value of prostate specific antigen velocity (PSAV) in screening for prostate cancer and especially for clinically significant PC is unclear.

Objective: To assess the value of PSAV in screening for PC. Specifically, the role of PSAV in lowering the number of unnecessary biopsies and reducing the detection rate of indolent PC was evaluated.

Design, setting and participants: All men included in the study cohort were participants in the European Randomized study of Screening for Prostate Cancer (ERSPC), Rotterdam section.

Intervention: During the first and second screening round, a PSA test was performed in 2217 men, and all underwent a biopsy during the second screening round 4 years later.

Measurements: PSAV was calculated and biopsy outcome was classified as benign, possibly indolent PC or clinically significant PC.

Results and limitations: A total of 441 cases of PC were detected, 333 were classified as clinically significant and 108 as possibly indolent. The use of PSAV cut-offs reduced the number of biopsies but led to important numbers of missed (indolent and significant) PC. PSAV was predictive for PC (OR 1.28, $p=0.000$) and specifically for significant PC (OR 1.46, $p=0.000$) in univariate analyses. However, multivariate analyses using age, PSA, prostate volume, DRE and TRUS outcome and previous biopsy (yes/no) showed that PSAV was not an independent predictor of PC (OR 1.01, $p=0.91$) or significant PC (OR 0.87, $p=0.30$).

Conclusions: The use of PSAV as a biopsy indicator would miss a large number of clinically significant PC cases with increasing PSAV cut-offs. In this study, PSAV was not an independent predictor of a positive biopsy in general or significant PC on biopsy. Therefore, PSAV does not improve the ERSPC screening algorithm.

5.1 | Introduction

Screening for prostate cancer (PC) by means of a serum prostate-specific antigen (PSA) test has become widespread practice. The optimal PSA cut-off level to indicate a prostate needle biopsy, however, is not easily identified. An optimal balance is to be found in detecting significant PC (sPC), while avoiding unnecessary biopsies and the detection of indolent disease (iPC) (PC which would not have been diagnosed in the absence of screening). Moreover, Thompson et al¹ showed that PC is biopsy-detectable throughout the whole PSA-range. Therefore, with the use of any PSA cut-off, cancers inevitably will be missed. However, lowering the biopsy threshold will lead to an increase in unnecessary biopsies and possibly an increased detection of iPC.

To improve specificity of PSA testing, PSA kinetic parameters, such as PSA-velocity (PSAV) has been studied extensively. PSAV is the change in PSA level during one year. Several studies have shown that PSAV is predictive of PC detection²⁻⁵ but others did not support this finding.^{6,7} In addition, d'Amico et al reported a significantly higher chance of PC death after radical prostatectomy (RP)⁸ or radiotherapy⁹ in men with a PSAV >2.0 ng/ml in the year before diagnosis. These findings were supported by another study that found a significantly higher median PSAV in men with relapse after RP than in men without relapse.¹⁰ Those studies suggest an association of PSAV with tumour aggressiveness and adverse outcome. If PSAV were able to distinguish sPC from iPC, this would be an important step forward in screening for PC. Some major concerns about screening, namely the large number of unnecessary biopsies and the high detection rate of iPC, could be decreased in part by a marker selective for sPC.

The aim of this study was to assess the value of PSAV as a predictor of biopsy outcome and tumour aggressiveness in a screened population. We evaluated the effects of applying a PSAV cut-off level as a biopsy indicator in terms of relative sensitivity, specificity and positive predictive value. In addition, the predictive value of PSAV for (significant) PC in the whole cohort was calculated.

5.2 | Methods

5.2.1 | Study cohort

All men included in this study were participants in the European Randomized Study of Screening for Prostate Cancer (ERSPC), Rotterdam section. The ERSPC was conducted to show or exclude a significant difference in PC mortality by screening for PC.

In the second round of the ERSPC Rotterdam (1997-2003), 12,529 men aged 55-74 yr were screened by means of a PSA serum test. PSA determinations were done with the Beckman-Coulter Hybritech Tandem E Assay (Hybritech Incorporated, San Diego, CA). After January 2000, this assay was replaced by the automatic version (Beckman-Access; Beckman-Coulter, Inc., Fullerton, CA). A PSA \geq 3.0 ng/ml prompted a systematic lateralized sextant prostate

needle biopsy. In 2502 men (20.0%) a $\text{PSA} \geq 3.0$ was found and 2217 men (88.6%) were actually biopsied. Our study cohort consisted of those 2217 men who underwent a biopsy in the second screening round (Figure 1). This choice allowed us to study the value of PSAV in this selected cohort without any verification bias, but it precluded an evaluation of the entire second-round population.

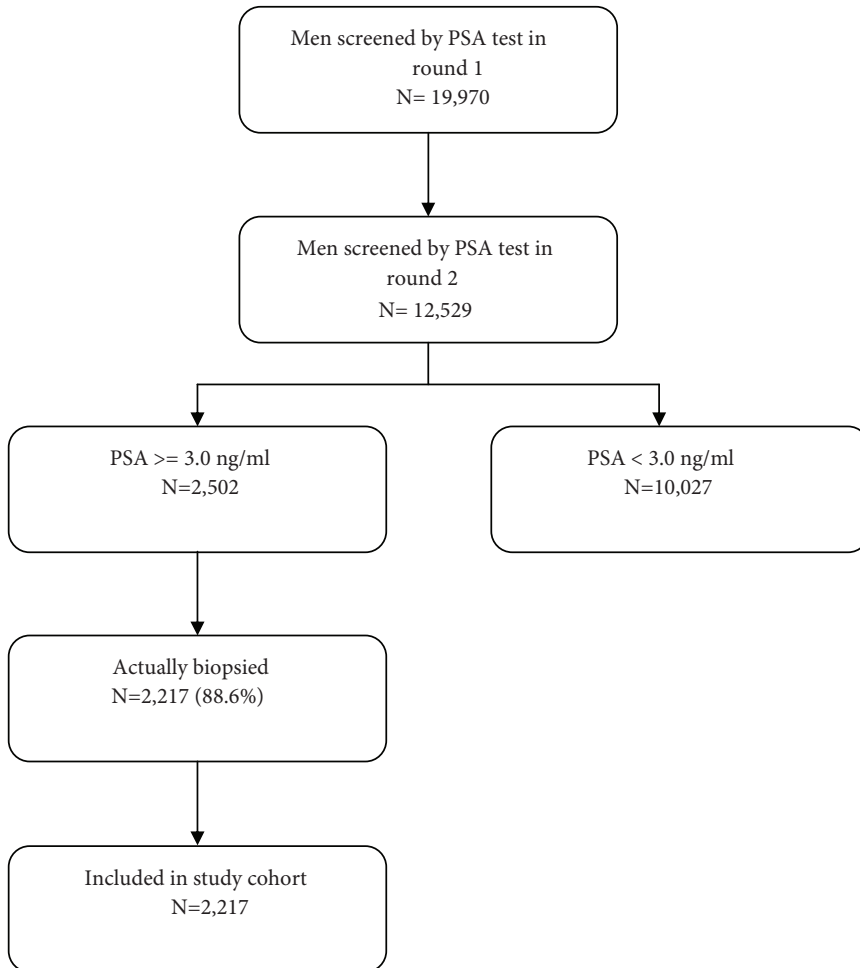


Figure 1 | Consort diagram: Method of patient selection. PSA = prostate-specific antigen.

* Significant PC was defined as PC with a probability of indolence less than 70% according to the nomogram described by Steyerberg et al¹¹

All men had been screened during the initial screening round four years earlier. Therefore, two subsequent PSA levels with a four-year interval were available for calculating the PSA-velocity (PSAV), which was calculated as the difference of the two PSA-levels divided by the exact time interval between the first and second screening visit.

In all cancers found, the probability of indolent disease was assessed. The probability of iPC was calculated using the nomogram described by Steyerberg et al.¹¹ For the development of this nomogram, iPC was defined as pathologically organ-confined disease, tumour volume <0.5 ml and no Gleason pattern 4/5 based on radical prostatectomy specimens. The probability of iPC was calculated with the nomogram based on pre-treatment information: PSA, prostate volume, Gleason patterns, total mm cancerous and non-cancerous biopsy core tissue. If the calculated probability of iPC was $\geq 70\%$, PC was prospectively defined as indolent. The remaining PC cases (including those not suitable for the nomogram based on the entry criteria¹¹) were classified as significant disease. PSAV cut-offs of 0.15 to 1.0 ng/ml/yr as a biopsy indicator were simulated in addition to the actual biopsy indicator of a PSA-level ≥ 3.0 ng/ml. Relative sensitivity and specificity and the positive predictive value (PPV) were calculated. Because overdiagnosed cancers cannot be reliably identified, all screen-detected cancers have to be included in the denominator for sensitivity. This reduces the value of sensitivity as a useful measure for clinically relevant disease. Therefore, we added another measure: the ratio benign biopsies/delayed PC diagnoses.

Theoretically, PSA and subsequently PSAV can rise quickly due to (subclinical) prostatitis. Because the PSAV is based on two PSA measurements, a period of PSA rise due to prostatitis is not being accounted for. Therefore, in addition to clinical data, the histological diagnosis of prostatitis was retrieved from the medical records.

Furthermore, follow-up and PC detection in the third screening round were evaluated. To account for PC cases missed by biopsy during the second screening round, PC detected during the third screening round were added to the cases detected during the second round and analyses were repeated.

5.2.2 | Statistical analyses

PSA, PSAV, age and prostate volume were assessed as continuous variables. Abnormal digital rectal examination (DRE) and transrectal ultrasonography (TRUS) outcome, having had a previous negative biopsy and prostatitis were assessed as binary variables.

Differences in proportions were evaluated with a Student's t-test (continuous variables, normal distribution), Mann-Whitney U-test (continuous variables, no normal distribution) or Chi-square test (binary variables). Univariate and multivariate logistic regression analyses were performed, the latter using a backward stepwise method. Parameters were rejected at a p-value >0.05 . In all multivariate analyses the following variables were included: PSA level, age, prostate volume, DRE and TRUS outcome, and previous biopsy. For PSA and volume, a logarithmic

transformation was used to optimize the model fit. For statistical analysis the statistical package for social sciences (SPSS 14.0; SPSS inc., Chicago, IL) was used.

5.3 | Results

In the 2217 men who underwent a biopsy during the second screening round, 441 (19.9%) carcinomas were found. General characteristics of the study cohort are listed in Table 1. Of the 441 PC diagnosed during the second round, 333 (75.5%) were classified as significant disease and 108 (24.5%) as possibly indolent.

Mean (median) PSAV was 0.47 (0.35) ng/ml/yr for the entire study population. PSAV-values are shown in Figure 2, depicted by the presence or absence of PC. Mean PSAV in men with PC was significantly higher than in men without PC (0.61 resp 0.44 ng/ml/yr, $p < 0.001$; Table 1). A negative PSAV was seen in 204 men (11.5%) with no PC, in 13 (12.0%) men with iPC, and in 10 (2.1%) men with sPC.

Table 1 | General descriptives of study cohort.

		Total group n=2217	PC n=441	No PC n=1776	p-value
Age (years, range 55-74)	Mean (median)	67.2 (67.3)	67.1 (67.1)	67.2 (67.3)	0.76*
PSA (ng/ml, ≥ 3.0 ng/ml)	Mean (median)	5.6 (4.5)	5.6 (4.4)	5.6 (4.5)	0.15**
Prostate volume (ml)	Mean (median)	52.6 (48.2)	44.0 (40.6)	54.7 (50.3)	0.000**
Abnormal DRE	N (%)	479 (21.6)	143 (32.4)	336 (18.9)	0.000***
Abnormal TRUS	N (%)	385 (17.4)	107 (24.3)	278 (15.7)	0.000***
Previous biopsy	N (%)	953 (43.0)	116 (26.3)	837 (47.1)	0.000***
PSAV (ng/ml/yr)	Mean (median)	0.47 (0.35)	0.61 (0.39)	0.44 (0.34)	0.000**

* Student's t-test

** Mann-Whitney U test

***Chi square test

5.3.1 | PSAV as a biopsy indicator

The effect of various PSAV cut-off levels as a biopsy indicator on PC detection was assessed (Table 2). With increasing PSAV cut-off levels, the PPV shows a tendency to increase (from 21.8% with a cut-off of 0.15 to 25.8% with a cut-off of 1.00 ng/ml/yr). Relative specificity increased significantly, at the expense of the relative sensitivity (84.4 and 24.9% with a PSAV cut-off of 0.15, 13.2 and 90.6% with a PSAV cut-off of 1.00 ng/ml/yr). The ratio spared benign

biopsies/delayed cancer diagnoses decreased with increasing PSAV cut-offs, indicating a less beneficial trade-off in terms of avoiding unnecessary biopsies and delaying PC diagnoses with higher cut-offs. The features of the carcinomas found and missed using the PSAV cut-offs in addition to the PSA cut-off of 3.0 ng/ml, are listed in Tables 3 and 4. The proportion of sPC rose with increasing PSAV cut-off levels (79.6% with a cut-off of 0.15 ng/ml/yr to 91.3% with a cut-off of 1.00 ng/ml/yr), showing the ability of PSAV to discriminate between sPC and iPC. However, although the larger proportion of missed PC is likely to be indolent, rapidly increasing proportions of all significant disease were missed with the use of a PSAV cut-off as a biopsy indicator (Table 4).

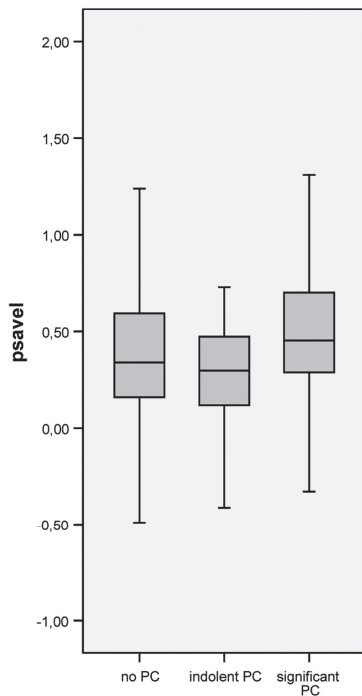


Figure 2 | Prostate-specific antigen velocity (ng/ml/yr) in men without prostate cancer (PC) (n=1776), men with iPC (n=108) and men with sPC (n=333).

The boxes represent the interquartile range, the line in the box the median value. The tails of the boxes represent the 95% interval.

5.3.2 | PSAV and overall PC detection

PSAV was a significant predictor of PC in univariate logistic regression analysis (OR 1.28, $p=0.000$). In a backward stepwise multivariate logistic regression analysis, PSA, prostate volume, DRE, TRUS and previous negative biopsy were included in the model. PSAV and age were omitted from the model, as they were not significant predictors of PC detection (PSAV OR=1.01, $p=0.913$).

Of the 1776 men with a non-malignant biopsy, 781 (44.0%) participated so far in the third screening round 4 years later, 598 biopsies were performed and 85 cases of PC were found (PPV 14.2%). Adding those PC cases to the PC detected during the second round and repeating the analyses described above did not change outcome: although PSAV was a significant predictor of PC detection in univariate analysis (OR 1.28, $p=0.000$), this parameter lost significance in multivariate analysis (OR 1.013, $p=0.907$).

Table 2 | Performance of PSAV as predictor of biopsy outcome with varying PSAV cut-off levels

	biopsied	PC (n)	PPV (%)	Rel. sens	Rel. spec	Ratio spared benign biopsy/delayed PC diagnosis
PSA ≥ 3	2217	441	19.9	100.0	0.0	
PSAV ≥ 0.15	1705	372	21.8	84.4	24.9	6.4
PSAV ≥ 0.25	1429	327	22.9	74.1	38.0	5.9
PSAV ≥ 0.35	1116	249	22.3	56.6	51.2	4.7
PSAV ≥ 0.50	750	168	22.4	38.1	67.2	4.4
PSAV ≥ 0.75	381	78	20.5	17.7	82.9	4.1
PSAV ≥ 1.00	225	58	25.8	13.2	90.6	4.2

Table 3 | Features of carcinomas found with varying PSAV cut-off levels

	Biopsied (n)	PC (n)	PC significant (% of PC found)	PC indolent (% of PC found)
PSA ≥ 3	2217	441	333 (75.5)	108 (24.8)
PSAV ≥ 0.15	1705	372	296 (79.6)	76 (20.4)
PSAV ≥ 0.25	1429	327	266 (81.3)	61 (18.7)
PSAV ≥ 0.35	1116	249	209 (83.9)	40 (16.1)
PSAV ≥ 0.50	750	168	143 (85.1)	25 (14.9)
PSAV ≥ 0.75	381	78	73 (93.6)	5 (6.4)
PSAV ≥ 1.00	225	58	53 (91.3)	5 (8.6)

5.3.3 | PSAV and sPC

All biopsy outcomes were divided in sPC ($n=333$) or iPC/no PC ($n=1884$) and the predictive ability of PSAV on significant disease in the entire study cohort was assessed. In univariate analysis, PSAV was a significant predictor of sPC (OR=1.46, $p=0.000$; Table 5). In multivariate analysis, PSA, age, prostate volume, DRE, TRUS and previous negative biopsy were included. In addition to age, PSAV was omitted from the model due to lack of significance (OR=0.87, $p=0.30$) (Table 5).

Table 4 | Features of carcinomas missed with varying PSAV cut-off levels

	biopsied	PC (n) missed	Missed PC significant (% of total cases of sPC)	Missed PC indolent (% of total cases of iPC)
PSA >=3	2217	0	0	0
PSAV>=0.15	1705	69	37 (11.1%)	32 (29.6%)
PSAV>=0.25	1429	114	67 (20.1%)	47 (43.5%)
PSAV>=0.35	1116	192	124 (37.2%)	68 (63.0%)
PSAV>=0.50	750	273	195 (58.6%)	78 (72.2%)
PSAV>=0.75	381	363	265 (79.6%)	98 (90.7%)
PSAV>=1.00	225	383	285 (85.6%)	98 (90.7%)

Table 5 | Predictive value of prostate-specific antigen velocity (PSAV) for detection of significant prostate cancer (PC)* in uni- and multivariate analyses. CI=confidence interval

Variable	Univariate analysis		Multivariate analysis	
	Odds ratio (95% CI)	p-value	Odds ratio (95% CI)	p-value
PSAV	1.54 (1.33-1.79)	0.000	1.01 (0.78-1.32)	0.939
Log PSA	1.96 (1.07-3.58)	0.029	5.66 (2.68-11.95)	0.000
Age	1.00 (0.97-1.02)	0.838	1.01 (0.98-1.04)	0.428
Log prostate volume	0.01 (0.01-0.03)	0.000	0.02 (0.01-0.04)	0.007
DRE	4.54 (3.51-5.87)	0.000	3.53 (2.64-4.72)	0.000
TRUS	3.43 (2.62-4.49)	0.000	2.20 (1.61-3.01)	0.000
Previous biopsy	0.41 (0.31-0.54)	0.000	0.43 (0.31-0.61)	0.000

5.3.4 | PSAV and prostatitis

Prostatitis was seen on biopsy in 177 men. In the total group, PSAV was not a predictor of prostatitis (OR1.14, p=0.125) in univariate analysis, but in men without PC (n=1776, n=171 with prostatitis) PSAV was a significant predictor (OR=1.28, p=0.015). The analyses on the predictive value of PSAV on (significant) PC as described above were repeated with exclusion of all cases of prostatitis. Small, non-significant changes were found for the odds ratios (data not shown).

5.4 | Discussion

Although some reports suggest that PSAV is a useful marker for aggressiveness in PC patients, this does not necessarily imply that PSAV is a useful marker for detecting aggressive PC in a screening setting.^{12,13} However, a recent report by Carter et al¹⁴ suggested that PSAV indeed was

a useful marker for identifying men at risk of deadly PC during the preclinical curable phase. We could not confirm those results.

The regression analyses make clear why the implementation of a PSAV cut-off as a biopsy indicator in screening for PC is not feasible in this cohort: PSAV was not an independent predictor of PC detection on biopsy. This finding is in line with previous reports from our study group.^{7,15,16} More importantly, PSAV was not a significant independent predictor of aggressive disease in the cohort (OR 1.10, $p=0.55$).

Two main factors may negatively influence the predictive value of PSAV for detection of PC: missing PC with biopsy and subclinical prostatitis. Subclinical prostatitis has been reported to elevate PSA levels¹⁷ and may cause an increased PSAV. The effect of these possible confounding factors is shown in Figure 2: the upper range of PSAV values in the men without PC is similar to the range in men with PC, especially men with sPC. On the other hand, the lower range shows more spread to negative values than in men with PC. Missed PC at biopsy or prostatitis could explain this upper range in the men without PC detected. For this reason, men with PC detected during the third screening round, were scored as having PC as well. However, PSAV still was not a significant predictor of PC detection in multivariate analysis (OR 1.013, $p=0.907$). Additionally, exclusion of all men with prostatitis in the subgroup without PC did not significantly change test results. False-negative biopsy or subclinical prostatitis as an explanation for the absence of the predictive value of PSAV for PC or sPC consequently seems unlikely.

It should be noted that our study differs in several aspects from the reports in which PSAV was described to be a predictor of PC aggressiveness and outcome. First of all, we studied a screened population, with cancers detected during the pre-clinical detectable phase. As previously described,^{7,13} it is likely that during this phase the PSAV does not yet show the significant increase as assessed in studies including clinically detected PC. Furthermore, our cohort consisted of men aged 55-75. As age was shown to be related to the predictive value of PSAV⁵, this may further explain the discordant results as some studies included men in the age range of 41 to 94.^{3,4}

A third and very important difference is that all men in our cohort were biopsied. This is in contrast with other reports, in which only a part of the population (mostly with an indication for biopsy) has been biopsied.^{2-5,18} The remaining part is assumed not to have PC, which is not true. Clearly, the verification bias that is inherent to this type of analysis will result in a higher predictive value of PSAV for the detection of PC.¹⁹ In our study every man was biopsied due to our patient selection method and verification bias could be kept to a minimum. Our results are concordant with a study by Thompson et al:⁶ in the control arm population of the PCPT trial all men were biopsied as well and PSAV lost its predictive value on PC detection in multivariate analyses.

The high mean PSAV in men without PC detected (0.44 ng/ml/yr) is a result of the patient-selection method: men with a PSA <3.0 ng/ml were excluded, as they were not biopsied. Earlier

reports from our study group showed that, when including those with low PSA levels, mean PSAV in men without PC was 0.09 ng/ml/yr.¹⁶

Nevertheless, there are some limitations. Most importantly, like every other retrospective analyses also this one is subject to verification bias. This results from the fact the PSAV was not used as a biopsy indication. Only a PSA ≥ 3.0 ng/ml indicated a biopsy and cancers with a high PSAV but a PSA < 3.0 ng/ml in the second round remain undetected (Figure 1). Furthermore, not all men with a PSA ≥ 3.0 ng/ml were biopsied (Figure 1), due to co-morbidity, medication use or refusal. This also adds to the verification bias.

The method of patient selection was based on the possibility to calculate a PSAV (at least 2 PSA levels available). Our population consisted of men screened during the second round and therefore they were all screened before (during the first round). This method of patient selection has two major consequences. First, this makes our cohort a pre-screened population, thereby restricting the applicability of our results. Secondly, the PSAV-calculations are only based on 2 PSA values. PSA change may not be linear, which cannot be accounted for using a two-point method for PSAV-calculation. Furthermore, the impact of biological variation in PSA-values cannot be accounted for²⁰. This limits our results. However, Connolly et al²¹ concluded that linear regression should be the method of choice, and that using two PSA values had a similar predictive value and may be adequate as long as measurements are separated by a sufficiently long time period. These findings are supported by a study of King et al.²² Therefore, the method used in our study may be adequate for these circumstances.

In this study, we explored whether PSAV could improve screening by identifying sPC and by reducing overdiagnosis. However overdiagnosis is not only based on PC features (and indolent disease), but also on patient related features (for example age and co-morbidity). Therefore, even a non-indolent cancer may be over-diagnosed. Moreover, even if a cancer is classified as indolent, this cancer may progress to more aggressive stages and become relevant.

The most ideal endpoint to define indolent and significant disease is PC-specific mortality. However, follow-up in this cohort is not long enough to consider this endpoint and a proxy had to be chosen. Clearly, using a proxy will influence the results. IPC was defined as a chance of indolence $\geq 70\%$ based on probabilities calculated using a nomogram that was developed for this cohort.¹¹ This 70% cut-off is arbitrary. Therefore, we assessed the predictive value of PSAV with varying probability cut-offs (range 0-80%). Little variation was seen in odds ratios and with all cut-offs PSAV lost significance in multivariate analyses (data not shown).

Finally, even if men are biopsied, the true incidence of PC remains unknown as a biopsy may miss PC. For this reason, “relative” sensitivity and specificity were calculated for the PSAV cut-off levels.²³

Despite its limitations, some points of strength of our study must be emphasized. Verification bias was kept to a minimum by selecting a cohort in whom every man was biopsied. In addition, all men were biopsied following identical biopsy protocols.

5.5 | Conclusion

In our screened cohort, PSAV was not an independent predictor of sPC. Using a PSAV cut-off as biopsy indicator would miss an important proportion of clinically sPC. Therefore, PSAV does not improve the ERSPC screening algorithm. Although PSAV seems a possible marker for tumour aggressiveness and outcome in PC patients, its value as a predictor of clinically sPC may not be applicable in a screening setting.

References

1. Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or = 4.0 ng per milliliter. *N Engl J Med* 2004;350:2239-46
2. Carter HB, Pearson JD, Metter EJ, Brant LJ, Chan DW, Andres R, Fozard JL, Walsh PC. Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. *JAMA* 1992;267:2215-20
3. Berger AP, Deibl M, Strasak A, Bektic J, Pelzer AE, Klocker H, Steiner H, Fritsche G, Bartsch G, Horninger W. Large-scale study of clinical impact of PSA velocity: long-term PSA kinetics as method of differentiating men with from those without prostate cancer. *Urology* 2007;69:134-8
4. Fang J, Metter EJ, Landis P, Carter HB. PSA Velocity for assessing prostate cancer risk in men with PSA levels between 2.0 and 4.0 ng/ml. *Urology* 2002;59:889-94
5. Smith DS, Catalona WJ. Rate of change in serum prostate specific antigen levels as a method for prostate cancer detection. *J Urol* 1994;152:1163-7
6. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA Jr. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *JNCI* 2006;98:529-34
7. Schröder FH, Roobol MJ, Van der Kwast ThH, Kranse R, Bangma CH. Does PSA velocity predict prostate cancer in pre-screened populations? *Eur Urol* 2006;49:460-5
8. D'Amico AV, Chen M-H, Roehl KA, Catalona WJ. Preoperative PSA velocity and the risk of death from prostate cancer after radical prostatectomy. *N Engl J Med* 2004;351:125-35
9. D'Amico AV, Renshaw AA, Sussman B, Chen M-H. Pretreatment PSA velocity and risk of death from prostate cancer following external beam radiation therapy. *JAMA* 2005;294:440-7
10. Berger AP, Deibl M, Strasak A, Bektic J, Pelzer AE, Steiner H, Spranger R, Fritsche G, Bartsch G, Horninger W. Relapse after radical prostatectomy correlates with preoperative PSA velocity and tumour volume: results from a screening population. *Urology* 2006;68:1067-71
11. Steyerberg EW, Roobol MJ, Kattan MW, Van der Kwast ThH, De Koning HJ, Schröder FH. Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. *J Urol* 2007;177:107-12
12. Pinsky PF, Andriole G, Crawford ED, Chia D, Kramer BS, Grubb R, Greenlee R, Gohagan JK. Prostate-specific antigen velocity and prostate cancer Gleason grade and stage. *Cancer* 2007;109:1689-95
13. Etzioni RD, Ankerst DP, Weiss NS, Inoue LYT and Thompson IM. Is prostate-specific antigen velocity useful in early detection of prostate cancer? A critical appraisal of the evidence. *J Natl Cancer Inst* 2007;99:1510-5
14. Carter HB, Ferrucci L, Ketterman A, Landis P, Wright EJ, Epstein JI, Trock BJ, Metter EJ. Detection of life-threatening prostate cancer with prostate-specific antigen velocity during a window of curability. *JNCI* 2006;98:1521-7
15. Roobol MJ, Kranse R, De Koning HJ, Schröder FH. Prostate-specific antigen velocity at low prostate-specific antigen levels as screening tool for prostate cancer: results of second screening round of ERSPC (Rotterdam). *Urology* 2004;63:309-15
16. Raaijmakers R, Wildhagen MF, Ito K, Pàez A, De Vries S, Roobol MJ, Schröder FH. Prostate-specific antigen change in the European Randomized Study of Screening for Prostate Cancer, section Rotterdam. *Urology* 2004;63:316-20
17. Nadler RB, Humphrey PA, Smith DS, Catalona WJ, Ratliff TL. Effect of inflammation and benign prostatic hyperplasia on elevated serum prostate specific antigen levels. *J Urol* 1995;154:407-13
18. Punglia RS, Cullen J, McLeod DG, Chen Y, D'Amico AV. Prostate-specific antigen velocity and the detection of Gleason score 7 to 10 prostate cancer. *Cancer* 2007;110:1973-8
19. Roobol MJ. Re: Longitudinal PSA changes in men with and without prostate cancer: assessment of prostate cancer risk. *Eur Urol* 2007;51:1441-2
20. Sölétormos G, Semjonow A, Sibley PEC, Lamerz R, Petersen PH, Albrecht W, Bialk P, Gion M, Junker F, Schmid H-P, Van Poppel H. Biological variation of total prostate-specific antigen: a survey of published estimates and consequences for clinical practice. *Clin Chem* 2005;51:1342-51

21. Connolly D, Black A, Murray LJ, Napolitano G, Gavin A, Keane PF. Methods of calculating prostate-specific antigen velocity. *Eur Urol* 2007;52:1044-50
22. King CR, Freedland SJ, Terris MK, Aronson WJ, Kane CJ, Amling CL, Presti JC Jr. Optimal timing, cutoff, and method of calculation of preoperative prostate-specific antigen velocity to predict relapse after prostatectomy: A report from SEARCH. *Urology* 2007;69:732-7
23. Schröder FH, Alexander FE, Bangma CH, Hugosson J, Smith DS. Screening and early detection of prostate cancer. *Prostate* 2000;44:255-63



CHAPTER 6

Can non-malignant biopsy features identify men at increased risk of biopsy-detectable prostate cancer at re-screening after four years? (ERSPC Rotterdam section)

BJU Int 2008;101:283-8

Tineke Wolters, Monique J. Roobol, Fritz H. Schröder, Theodorus H. van der Kwast,
Stijn Roemeling, Ingrid W. van der Crujisen-Koeter, Chris H. Bangma,
Geert J.L.H. van Leenders

Abstract

Objectives: To identify pathological features in non-malignant sextant prostate needle biopsies and assess their predictive value for detecting prostate cancer (PC) on biopsy 4 years later.

Patients and Methods: We selected and reviewed the biopsy specimens of 121 men that were diagnosed as non-malignant during the first screening round of the European Randomized Study of Screening for Prostate Cancer (ERSPC), Rotterdam section. Of these 61 (50.4%) were positive for PC during the second round (result of a matched random sample). The biopsies were indicated by prostate-specific antigen levels of ≥ 3.0 ng/ml. Specimens were scored for high-grade prostatic intraepithelial neoplasia (HG-PIN), active and chronic inflammation (AI, ChI), biopsy core length and glandular core length. The predictive value of the pathological features for detecting PC after 4 years was assessed.

Results: In the first-round biopsies the incidence of HG-PIN was 7.1%; there was AI in 22.4% and CI in 51.0%. The mean core length was 9.3 mm and mean glandular core length 7.4 mm. The mean total biopsy length (sum of core lengths) was 56.3 mm and mean total glandular length (sum of glandular core lengths) was 44.6 mm. None of the pathological features in the initial round was significantly related to PC detection in the second round.

Conclusions: In this study of non-malignant prostate biopsy specimens from a screened population, no pathological features could be identified that were predictive for PC detection on biopsy 4 years later.

6.1 | Introduction

With the introduction of serum prostate-specific antigen (PSA) measurement, screening for prostate cancer (PC) has become widespread practice. One of the important challenges in screening is the length of the screening interval. It should be short enough to prevent the occurrence of clinically important interval cancers and incurable tumours found at re-screening, but it should be long enough to avoid unnecessary biopsies and costs. Furthermore, men with features that are known to increase the risk of PC may benefit from more frequent screening than men without those features. For identifying men at increased risk, next to clinical variables, histopathological features might be important. Can features of a non-malignant biopsy define subgroups at increased risk of subsequent prostate cancer detection 4 years later?

Several histological predictors of subsequent prostate cancer detection have been described. First, high-grade prostatic intraepithelial neoplasia (HG-PIN) is considered a premalignant lesion¹. By contrast with earlier reports on its predictive value for adenocarcinoma,¹⁻³ recent studies indicate that men with limited HG-PIN are probably not at increased risk of PC.^{3,4}

Second, active (AI) or chronic inflammation (ChI) can be found in a biopsy. Although considered to have no direct clinical implications, these lesions may influence the risk of a prostate cancer diagnosis in the subsequent screening round. ChI is especially considered to be a risk factor for carcinoma,⁵ including prostatic carcinoma.⁶

Finally, the sextant biopsy core length is a variable of biopsy quality and a greater total core length raises the cancer detection rate.^{7,8} Therefore, it may be possible that a low total core length is not a good representation of the prostate and men with an initial benign biopsy with a low total core length have a higher risk of PC diagnosis during a subsequent screening round compared to men with a high total core length.

Thus the aim of the present study was to evaluate the predictive value of features of a non-malignant initial biopsy (PIN, inflammation, and biopsy length) on having biopsy detectable prostate cancer 4 years later in a screened population.

6.2 | Patients and methods

6.2.1 | ERSPC and patient selection

We retrospectively evaluated prostate needle biopsy specimens of a screened population. The European Randomized Study of Screening for Prostate Cancer (ERSPC) is investigating the impact of screening on mortality and quality of life in men aged 55-75 years.⁹ In the Rotterdam section of the ERSPC, men randomized to the screening arm are screened every 4 years, i.e. a systematic lateralized sextant prostate needle biopsy prompted by an elevated PSA level ($>=3.0$ ng/ml).

The biopsy was taken using longitudinal and cross-sectional transrectal ultrasound guidance. In our institution, a lateralized biopsy strategy is used, based on the observation that sampling the lateral peripheral zone increases the prostate cancer detection rate.^{10,11} A seventh biopsy was taken when a hypo-echoic lesion was visible at TRUS.

If there is an initial diagnosis of HG-PIN on biopsy, another biopsy is taken within 6 weeks. When no cancer is found on this immediate re-biopsy, men are enrolled in the regular screening programme (i.e. next screening visit after 4 years).

In the first screening round 1,850 men had a biopsy prompted by a PSA \geq 3.0 ng/ml (May 1997 to December 1999). Of the 1309 men without prostate cancer in the first round, 584 (44.6%) had a repeat biopsy during the second screening round 4 years later. In the second round, 61 prostate cancers were diagnosed and were included in the present study. Of the 523 men in whom no prostate cancer was found at repeat screening, 60 men were randomly chosen and formed the control group of the present study. Therefore, our study group comprised 121 men, all with a non-malignant biopsy in the first screening round (Figure 1a).

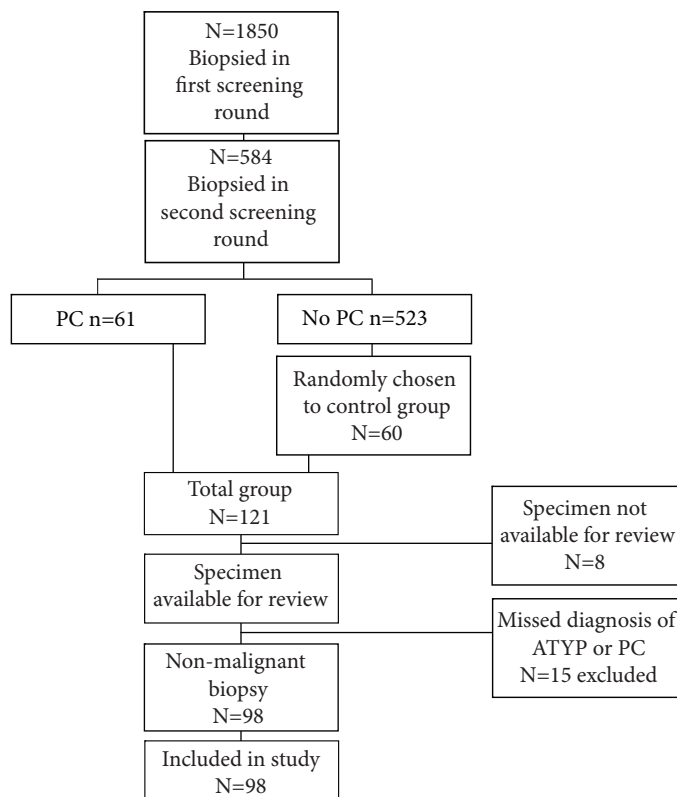


Figure 1 |Consort diagram of patient selection. PC= prostate cancer, ATYP= atypical lesion suspect for prostate cancer.

6.2.2 | Pathologic evaluation

All biopsy cores were labelled separately and processed as described previously⁸. At time of diagnosis, pathologists with no specialisation in uro-genital pathology had evaluated all samples. In the present study, the negative first- and second-round biopsies were blindly reviewed by an uro-pathologist (GvL). Within the scope of the ERSPC protocol, all positive second-round biopsies had been confirmed by another uro-pathologist (TvdK). Biopsy specimens from the first round were scored for HG-PIN (as described)². Furthermore, ChI and AI were scored; ChI was graded as negative (no or sporadic lymphocytes in stroma), mild (aggregates of lymphocytes, plasma cells and histiocytes in stroma, without influx in epithelial glands), severe (as mild, with influx in epithelial glands). AI was scored as negative (no neutrophilic granulocytes present), mild (sparse neutrophilic granulocytes in stroma, without influx in epithelial glands), severe (neutrophilic granulocytes in stroma, with influx in epithelial glands). The extent of inflammation was scored as the number of biopsy cores positive for inflammation. The total length of the needle-biopsies was measured manually with a ruler on the slides. In addition, the spatial distance between the most proximal and peripheral epithelial gland in the needle-biopsies was determined (glandular length) to exclude non-prostatic tissues, i.e. colonic mucosa and fat tissue, from our measurements. All pathological data were matched with clinical data (age, PSA, PSA-density, DRE and TRUS outcomes and prostate volume) as assessed at the initial screening round.

6.2.3 | Statistical analyses

For binary variables a chi-square test with continuity correction was used to assess differences (for small expected numbers, a Fisher's exact test was used). For ordinal and continuous variables t-tests (normally distributed variables), Mann-Whitney U tests (non-normally distributed variables) and logistic regression analyses were used.

The hypothesis that no difference or predictive value existed was tested using a two-sided assessment and rejected at a p-value <0.05.

6.3 | Results

Of the 121 men, both first- and second-round biopsy specimens could be retrieved from the pathology archive of the Erasmus Medical Center Rotterdam for 113. The review by the uro-pathologist showed 6 cases with a missed diagnosis of PC and 9 with missed atypical lesions suspect for adenocarcinoma (ATYP), all in first-round biopsy specimen. Those cases were excluded from further analyses (see figure 1b), leaving 98 men for the present study, of whom 44 (45%) were diagnosed with PC 4 years later. In men with cancer in this second screening round, advanced clinical stage (cT3-4) was seen in 1 man (2.3%). The biopsy Gleason sum score was 7 or more in 6 men (14%).

The general descriptive data of the first screening round are listed Table 1.

The incidence rates of the pathologic features and analyses of PC detection rates 4 years later are summarized in Table 2.

Table 1 | Descriptive parameters at the time of initial biopsy: median and interquartile range (IQR) of age, PSA, PSA density, prostate volume, transition zone volume, DRE results and hypo-echoic lesion at TRUS (total group values and values depicted by presence or absence of PC at repeat screening).

	Total group (n=98)	PC after 4 years (n=44)	No PC after 4 years (n=54)	p-value (two-sided)
Age (years)				
median	64.1	65.2	3.6	0.290*
IQR	61.0-67.2	60.9-67.9	61.1-65.9	
PSA (ng/ml)				
median	4.10	4.30	4.05	0.207**
IQR	3.30-5.10	3.63-5.75	3.30-4.90	
PSA density (nl/ml/gr)				
median	0.10	0.10	0.09	0.100**
IQR	0.07-0.13	0.08-.014	0.07-0.12	
Prostate volume (ml)				
median	44.6	43.8	46.2	0.126*
IQR	34.5-59.7	33.7-54.6	35.05-66.80	
Transition zone volume				
median	24.5	22.5	26.3	0.207**
IQR	18.4-35.5	17.7-33.8	18.65-44.65	
Abnormal DRE (%) mean	19.4	20.0	19.0	0.809 ***
Hypo-echoic lesion on TRUS (%) mean	15.3	9.0	21.0	0.123 ***

* Student's t-test

** Mann-Whitney U test

*** Chi-square test

Table 2 | Incidence of the first round pathologic lesions and the second round PC prevalence depicted by the presence or absence of this lesion. PC = prostate cancer, HG-PIN = high-grade prostatic intraepithelial neoplasia, ChI = chronic inflammation, AI = active inflammation.

	Prevalence in total group (n=98)	PC incidence when lesion present	PC incidence when lesion absent	P-value (two-sided)
HG-PIN	7.1 % (7/98)	42.9%	46.2%	1.00*
ChI	51.0% (50/98)	44.0%	48.0%	0.85**
AI	22.4 % (22/98)	50.0%	44.7%	0.85**

* Fisher's exact test

** Chi square test with continuity correction

6.3.1 | HG-PIN

All HG-PIN lesions found were focal lesions limited to one core; there was HG-PIN in 7 men (7.1%). As none of these cases was diagnosed during the original pathological examination, these were newly found HG-PIN cases. The PC-incidence after 4 years was the same for men with HG-PIN and men without HG-PIN (43% and 46%, $p=1.00$).

6.3.2 | Inflammation

Of the 51% of men with ChI, 47% had mild inflammation; there was a severe lymphoplasmacytic infiltrate in 4.1%. ChI was not predictive for detecting PC in the second screening round (Table 3). Clustering by degree (mild or severe) or extent of inflammation (number of cores with ChI) had no significant predictive value either ($p=0.981$ and $p=0.577$).

The incidence of AI was 22%, of which 21% was mild and 1.0% severe. AI was not predictive for detecting PC 4 years later (Table 3). There was no significant predictive value if AI was clustered by degree ($p=0.897$) or extent ($p=0.526$) of inflammation.

Table 3 | Biopsy length parameters (mean, median and interquartile range (IQR)) for the total group and depicted by presence or absence of prostate cancer (PC) in the second screening round.

	Total group (n=98)	PC after 4 years (n=44)	No PC after 4 years (n=54)	P-value (two-sided)
Core length mean	9.3 mm	9.2 mm	9.4 mm	0.47*
median	10.0 mm	10.0 mm	10.0 mm	
IQR	8.0-11.0 mm	8.0-11.0 mm	8.0-11.0 mm	
Glandular core length mean	7.4 mm	7.2 mm	7.5 mm	0.21*
median	7.0 mm	7.0 mm	7.0 mm	
IQR	6.0-9.0 mm	6.0-9.0 mm	6.0-10.0 mm	
Total biopsy length mean	56.3 mm	55.6 mm	56.9 mm	0.59*
median	57.0 mm	57.0 mm	57.0 mm	
IQR	50.0-63.0 mm	50.0-62.0 mm	51.0-64.0 mm	
Total glandular biopsy length mean	44.6 mm	43.6 mm	45.5 mm	0.34*
median	45.0 mm	45.0 mm	45.0 mm	
IQR	39.0-51.0 mm	38.5-49.5 mm	39.0-52.5 mm	

* Student's t-test

6.3.3 | Biopsy length

The effect of biopsy length is listed in Table 3; in 5 cores (0.8%) glandular structures were lacking. In 15 men, a seventh biopsy core was taken, prompted by a hypo-echoic lesion. The total biopsy length and total glandular length were calculated by adding the sextant biopsy core lengths, and values are for complete sextant biopsy cores only (96, thus with no possible seventh biopsy length or missing values). None of the biopsy length variables was predictive of subsequent PC.

Adding the length of the seventh biopsy to the sextant biopsy cores did not change the results ($p=0.655$ for total length and $p=0.421$ for total glandular length).

6.4 | Discussion

After a non-malignant initial biopsy, a screening participant in the ERSPC will be screened further, following the regular screening schedule (re-screen after 4 years). However, non-malignant initial biopsies might provide information that can be used to identify men at greater risk of prostate cancer detection in subsequent screening rounds. This information could possibly be used in optimizing future screening programmes. For example, men with pathological features at initial biopsy that are known to increase subsequent PC detection rate might benefit from a more intense screening protocol than men without these features.

The aim of the present study was to identify possible predictive features of a non-malignant initial biopsy. None of the features assessed (HG-PIN, active and chronic inflammation, biopsy length and glandular length) was a significant predictor of PC detection in the next screening round, using a 4-year interval.

The 7 cases of focal HG-PIN evaluated in the study were newly diagnosed at re-evaluation, and hence those men had not been re-biopsied at initial screening (as described by the screening algorithm) but were biopsied at the second screening round four years later. In all cases the PIN at initial screen was confined to 1 biopsy core. Nevertheless, those men were at no greater risk of PC after 4 years (43% in men with HG-PIN and 46% in men without). Although there were too few cases to draw conclusions, these findings are in agreement with recent review studies, which show that the risk of PC following a diagnosis of focal HG-PIN is only slightly higher than the risk of PC after a benign diagnosis, and questioning the need for repeat biopsy.^{3,4}

ChI is frequently found in prostate biopsies¹² and its role in prostatic carcinogenesis is currently being assessed.^{13,14} MacLennan et al⁶ recently published a study in which they evaluated the presence and degree of ChI in 177 prostate needle biopsies of men with clinical variables suspicious for malignancy. During a follow-up period of 5 years there was a higher incidence of cancer in men with ChI (20% vs. 6% in men without ChI). We could not confirm these findings, possibly because of the characteristics of the present study population, being all screen-detected prostate cancers, while that of MacLennan et al comprised men with a clinical suspicion of malignancy. Furthermore, all the present participants had had a repeat biopsy during the follow-up and had a chance of being diagnosed with PC, even when no clinical suspicion was present. In the study of MacLennan et al only 64% of men with ChI and 30% of men without ChI had had a repeat biopsy during the follow-up after an initial biopsy with no cancer (the reason for repeat biopsy was not described). Therefore, their chance of a diagnosis of prostate cancer was smaller than in the present study, especially in men with no ChI.

AI and its relation to PC has also been studied, especially when presenting clinically as prostatitis.^{13,1} However, the present goal was to evaluate the influence of AI found in biopsy specimens of men with no clinical signs of inflammation. Men with AI had no significantly greater risk of PC after four years than men without AI (50% vs. 45%, $p=0.85$). In 21 of 22 men with AI, ChI was also present; this finding is in agreement with the results of Anim et al,¹⁶ who found no AI without ChI during their review of prostate tissue specimens. As stated by Schatteman et al,¹² both forms of inflammation are probably dynamically related to each other. In multivariate analyses AI and ChI were not predictive of the PC detection rate (data not shown). In the present study, the presence of ChI or AI did not contribute to defining high-risk patients after four years, in this screening-based population.

Finally, we analysed the predictive value of the lateralized sextant biopsy core length. We hypothesized that a higher total biopsy length would decrease sampling artefacts, resulting in a lower incidence of PC after 4 years. However, no prognostic value of total biopsy length or total glandular length on cancer detection rate after 4 years could be identified in this dataset. When we corrected biopsy length for the total prostate volume it was also not predictive for having a biopsy-detectable PC at repeat screening 4 years later. An explanation could be the small variance in core lengths in the specimens studied (IQR 8.00-11.00 mm). Iczkowski et al⁷ and Van der Kwast et al⁸ reported that a higher total biopsy length is related to a higher cancer detection rate on that same prostate biopsy. Therefore, although not predictive for the PC detection rate four years later, biopsy length should be monitored by the urologist and the pathologist.

Atypical lesions suspicious but not diagnostic of malignancy (ATYP) (i.e. atypical small acinar proliferations, ASAP) are known to be predictive of concurrent or subsequent prostate cancer.¹⁷⁻¹⁹ An aim of the present study was to assess the predictive value of ATYP (with no PC at immediate repeat biopsy) on the detection of PC four years later. However, there were no cases of ATYP without PC at immediate re-biopsy in the study population and thus the predictive value could not be evaluated.

Previously, the relationship between atrophy and the incidence of PC in subsequent screening rounds in the ERSPC population (Rotterdam section) was assessed by Postma et al²⁰. They found no association of atrophy with the PC incidence, mainly because of its widespread occurrence. Therefore, atrophy was not studied as a possible predictor of PC detection in the present study.

The features of the tumours found after 4 years in the present population were in general favourable for curative treatment and comparable to the features found during the second screening round of the ERSPC²¹, in which there was significant down-grading and down-staging of PC. These findings support the view that there is no indication for a shorter screening interval in the present study population.

Our study has some general limitations. Although our data are based on lateralized sextant biopsies, the current trend is to obtain more than six biopsy cores, to obtain more adequate

information on cancer presence and its features. One of the side-effects is that more clinically insignificant cancers are diagnosed. The ERSPC study protocol was designed in the early 1990s, when sextant biopsies were the 'reference' standard. The Rotterdam study group has decided that it will follow this protocol for reasons of data consistency. Second, there were relatively few patients in the present study; although this affects the power of the study, the study is unique because patients are very well characterized by clinical variables, review of all slides and the exclusion of minimal PC and ATYP that had been missed on the first evaluation. The number of incidentally found minimal PC (n=6, 5%) and ATYP (n=9, 8%) at re-evaluation grossly overestimates the real prevalence of false-negative biopsies in the ERSPC, as the group was selected for patients with PC on the second screen. The prevalence of a false-negative diagnosis in the ERSPC was reported by Van der Kwast et al²², who found a false-negative biopsy outcome rate for adenocarcinoma of 4%. Combining ATYP and definite PC, in 6%-10% there was a missed lesion, the rate depending on the review pathologist. Therefore, our findings of missed lesions seem to be in agreement with earlier reports. However, the consequences of missing PC or ATYP are not yet entirely clear, and will be subject of further investigation. All the missed lesions in the present study were diagnosed during repeat screening rounds (Gleason scores =< 7, cT-stadium =< 2A). Those cases were excluded, because they fell beyond the scope of this study.

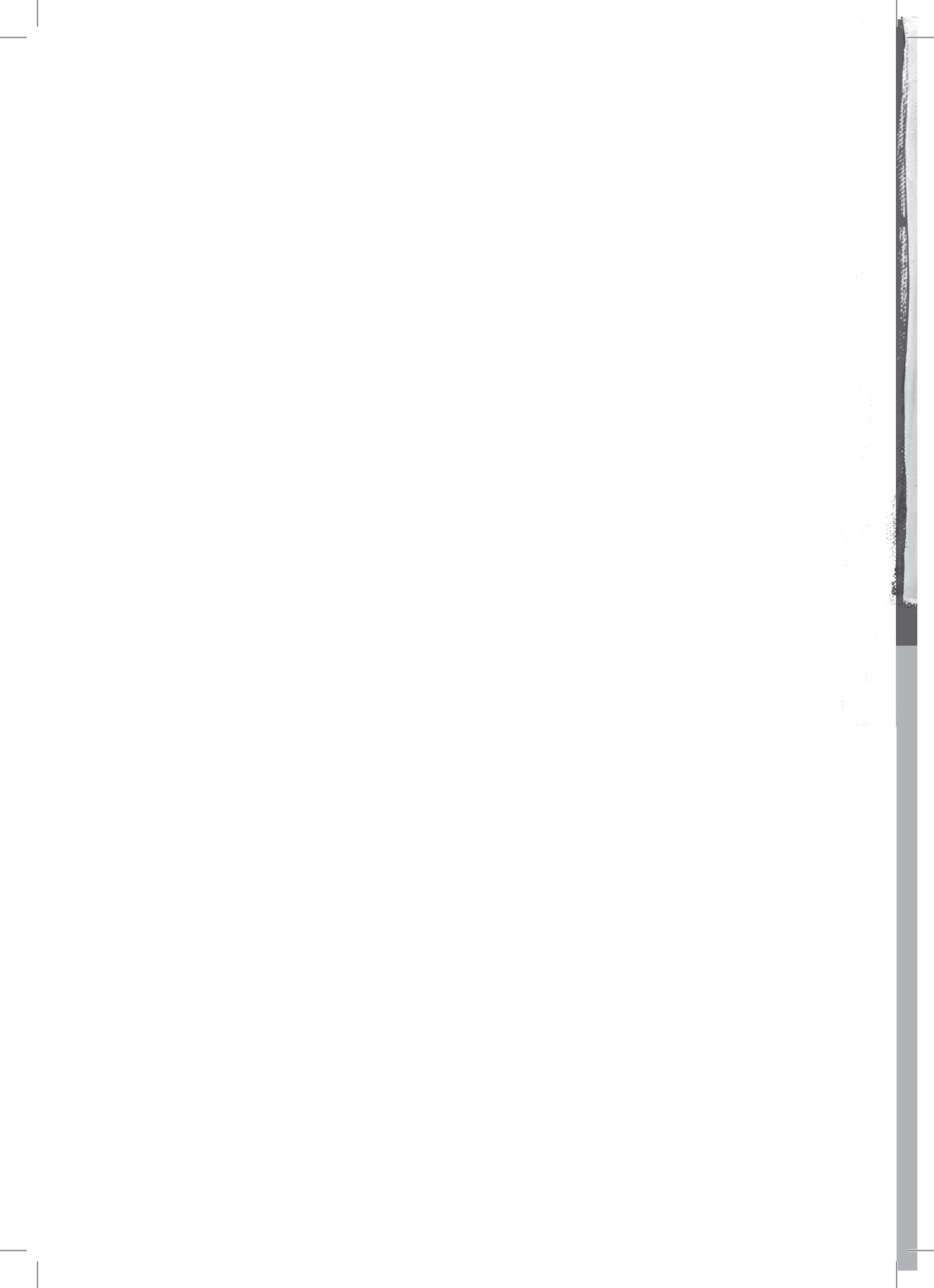
In this era of widespread screening for PC, the need for risk stratification of possible screening candidates is increasing; non-malignant biopsy features could not identify high-risk subgroups in the present study.

6.5 | Conclusions

In conclusion, we evaluated the predictive value of pathological features (HG-PIN, chronic and active inflammation and biopsy core length and glandular core length) in a non-malignant prostate needle biopsy for detecting prostate in a screened population on biopsy taken 4 years later. None of these features were significantly predictive of subsequent prostate cancer detection, and therefore no high-risk subgroups of PC-detection 4 years later could be identified.

References

1. Bostwick DG, and Qian J. High-grade prostatic intraepithelial neoplasia. *Mod Pathol* 2004;17:360-79
2. Keetch DW, P. Humphrey P, Stahl D, Smith DS and Catalona WJ. Morphometric analysis and clinical followup of isolated prostatic intraepithelial neoplasia in needle biopsy of the prostate. *J Urol* 1995;154(2):347-51
3. DeMarzo AM, Nelson WG, Isaacs WB and Epstein JI. Pathological and molecular aspects of prostate cancer. *Lancet* 2003;361:955-64
4. Epstein JI and Herawi M. Prostate needle biopsies containing prostatic intraepithelial neoplasia or atypical foci suspicious for carcinoma: Implications for patient care. *J Urol* 2006;175:820-34
5. Coussens LM and Werb Z. Inflammation and cancer. *Nature* 2002;420:860-7
6. MacLennan GT, Eisenberg R, Fleshman RL, Taylor JM, Fu P, Resnick MI and Gupta S. The influence of chronic inflammation in prostatic carcinogenesis: A 5-year followup study. *J Urol* 2006;176:1012-6
7. Iczkowski KA, Casella G, Seppala RJ, Jones GL, Mishler BA, Qian J and DG Bostwick. Needle core length in sextant biopsy influences prostate cancer detection rate. *Urology* 2002;59:698-703
8. Van der Kwast ThH, Lopes C, Santonja C, Pihl C-G, Neetens I, Martikainen P, Di Lollo S, Bubendorf L and Hoedemaeker RF. Guidelines for processing and reporting of prostatic needle biopsies. *J Clin Pathol* 2003;56:336-40
9. Roobol MJ, Kirkels WJ and Schröder FH. Features and preliminary results of the ERSPC (Rotterdam, the Netherlands). *BJU Int* 2003;92 suppl 2:48-54
10. Stamey TA. Making the most out of six systematic sextant biopsies. *Urology* 1995;45:2-12
11. Eskew LA, Bare RL, McCullough DL. Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate. *J Urol* 1997;157:199-202
12. Schatteman PHF, Hoekx L, Wyndaele JJ, Jeuris W and Van Marck E. Inflammation in prostate biopsies of men without prostatic malignancy or clinical prostatitis. *Eur Urol* 2000;37:404-12
13. Palapattu GS, Sutcliffe S, Bastian PJ, Platz EA, DeMarzo AM, Isaacs WB and Nelson WG. Prostate carcinogenesis and inflammation: emerging insights. *Carcinogenesis* 2004;26:1170-81
14. DeMarzo AM, DeWeese ThL, Platz EA, Meeker AK, Nakayama M, Epstein JI, Isaacs WB and Nelson WG. Pathological and molecular mechanisms of prostate carcinogenesis: implications for diagnosis, detection, prevention, and treatment. *J Cell Biochem* 2004;91:459-77
15. Roberts RO, Bergstralh EJ, Bass SE, Lieber MM and Jacobsen SJ. Prostatitis as a risk factor for prostate cancer. *Epidemiology* 2004;15:93-9
16. Anim JT, Kehinde EO, Prasad A, Sheikh M, Mojiminiyi OA, Ali Y and Al-Awadi K. Relationship between serum prostate specific antigen and the pattern of inflammation in both benign and malignant prostatic disease in Middle Eastern men. *Int Urol Nephrol* 2006;38:27-32
17. Montironi R, Scattoni V, Mazzucchelli R, Lopez-Beltran A, Bostwick DG and Montorsi F. Atypical foci suspicious but not diagnostic of malignancy in prostate needle biopsies (Also referred to as "Atypical small acinar proliferation suspicious for but not diagnostic of malignancy"). *Eur Urol* 2006;50:666-74
18. Bostwick DG and Meiers I. Atypical small acinar proliferation in the prostate. Clinical significance in 2006. *Arch Pathol Lab Med* 2006;130:952-7
19. Iczkowski KA. Current prostate biopsy interpretation. Criteria for cancer, atypical small acinar proliferation, high-grade prostatic intraepithelial neoplasia, and use of immunostains. *Arch Pathol Lab Med* 2006;130:835-43
20. Postma R, Schröder FH and Van der Kwast ThH. Atrophy in prostate needle biopsy cores and its relationship to prostate cancer incidence in screened men. *Urology* 2005;65:745-9
21. Cruikjens-Koeter IW, Roobol MJ, Wildhagen MF, Van der Kwast ThH, Kirkels WJ and Schröder FH. Tumour characteristics and prognostic factors in two subsequent screening rounds with four-year interval within prostate cancer screening trial, ERSPC Rotterdam. *Urology* 2006;68:615-20
22. Van der Kwast ThH, Lopes C, Martikainen PM, Pihl C-G, Santonja C, Neetens I, Di Lollo S, Hoedemaeker RF. Report of the pathology committee: false-positive and false-negative diagnoses of prostate cancer. *BJU I* 2003;92(suppl):62-5





CHAPTER 7

False-negative prostate needle-biopsies: frequency, histopathologic features and follow-up

Am J Surg Pathol 2010;34:35-43

Tineke Wolters, Theodorus H. van der Kwast, Cornelis J. Vissers, Chris H. Bangma,
Monique Roobol, Fritz H. Schröder, Geert J.L.H. van Leenders

Abstract

Little is known about the frequency, histopathological characteristics and clinical consequences of false-negative prostate biopsies, i.e. biopsies classified as benign but containing adenocarcinoma or atypical suspicious glands (ASAP). Objective of this study was to evaluate false-negative prostate biopsy in a prostate cancer screening setting. Prostate biopsy sets of 196 participants of a screening trial which had been reported as 'benign' at initial diagnosis, followed by a diagnosis of adenocarcinoma in a subsequent screening round were reviewed by two urologic pathologists. Adenocarcinoma was identified in 19 biopsy cores corresponding to 16 (8.2%) patients and ASAP in 24 cores, corresponding to 19 patients (9.7%). All missed prostate cancers were Gleason score 6 (3+3). After correction for patient selection, the overall false-negative biopsy rate was estimated to be 2.4%; 1.1% for prostate cancer and 1.3% for ASAP. Clinicopathological features at the time of initial biopsy and of subsequent prostate cancer diagnosis did not differ between patients with a false-negative or true benign biopsy. Relatively low number of atypical glands (<10 glands), intense intermingling with pre-existent glands or lack of architectural disorganization were the most prominent risk factors for a false-negative diagnosis. Another potential pitfall was the presence of prostate cancer variants, since one adenocarcinoma was of foamy gland type and three of pseudo-hyperplastic type. Routine examination of at least one level of prostate biopsy sets at high magnification and awareness of histologic prostate cancer variants might reduce the risk of missing or misinterpreting a relevant lesion at prostate biopsy evaluation.

7.1 | Introduction

Widespread PSA testing for prostate cancer has resulted in increased detection of small foci of adenocarcinoma on diagnostic needle-biopsies. Pathologists are therefore more often confronted with the presence of only a few atypical or malignant glands in diagnostic needle-biopsies. Small lesions of malignant glands or glands suspicious for malignancy, i.e. atypical small acinar proliferations (ASAP)¹ may be at increased risk to be overlooked, or to be misdiagnosed. Nevertheless, correct diagnosis of these lesions is required, because their presence may have clinical consequences.

The presence of ASAP is reported in about 5% of prostate biopsies² and warrant renewed and targeted needle-biopsies: the risk of cancer detection at re-biopsy is on average 40% after a diagnosis of ASAP in the first biopsy.^{2,3} The diagnosis of minimal prostate cancer might be clinically relevant as well because significant tumour may be found at radical prostatectomy after such a small prostate cancer lesion at biopsy.^{4,5} However, the recognition of a small focus of ASAP or minimal prostate cancer may be difficult.^{5,6}

Little is known about the frequency, histopathological characteristics and clinical consequences of false-negative prostate biopsies, i.e. biopsies in which an atypical or malignant lesion is missed. Nevertheless, study of missed lesions which have resulted in false-negative biopsies is important as this will lead to increased awareness and more accurate diagnosis of these challenging lesions. To provide more insight into false-negative prostate biopsy, a review study was performed of prostate biopsy specimens of participants in a large screening trial who were screened every four years. In this study, we assessed the frequency of false-negative prostate biopsies containing ASAP or carcinoma, their histopathological characteristics and potential clinical consequences.

7.2 | Materials and Methods

7.2.1 | Patient selection

All patients included in this study were participants in the screening arm of the ERSPC, Rotterdam section. First end-point results of the ERSPC have recently been published indicating a 20% mortality reduction due to screening⁷. Men in the age range of 55-74 were invited for screening every 4 years. From 1993 until May 1997, a PSA level ≥ 4.0 ng/ ml, an abnormal digital rectal examination (DRE) and/or transrectal ultrasound (TRUS) result prompted a lateralized sextant prostate biopsy. Hereafter, a biopsy was only indicated by a PSA level ≥ 3.0 ng/ ml. A lateralized sextant prostate needle biopsy was performed using TRUS guidance. If a hypo-echoic lesion was visible at TRUS, a seventh biopsy core was taken from this lesion. For all participants, written informed consent was obtained.⁸

To enhance the chance of finding missed lesions in the current study, we retrieved all patients from the Rotterdam ERSPC database who were diagnosed with prostate cancer in the second or third screening round (n=857) up to May 2008, and who had a previous biopsy diagnosis of benign prostate tissue (n=202, 23.6%). Clinical data and follow-up data were retrieved from the medical records.

7.2.2 | Histopathological processing of needle-biopsies

Prostate needle-biopsies were subjected to routine pathological processing with stretching of cores, inclusion of one core per cassette and H&E stainings on three levels.⁹ While all pathologists at the Department of Pathology participated in histological evaluation, the protocol required re-evaluation of all needle-biopsies signed-out as prostate adenocarcinoma, ASAP or prostate intraepithelial neoplasia (PIN) by a urologic pathologist (ThvdK, GvL) at the time of diagnosis. Needle-biopsies with a benign diagnosis including normal, hyperplasia, atrophy and inflammation were generally not reviewed at that time.

7.2.3 | Review of biopsy specimens

For this study, the slides of previous benign biopsy specimens were retrieved from the archives and evaluated by two urologic pathologists (GvL, ThvdK). If suspicious or malignant glands were identified, we manually counted the maximal number of atypical glands in the most affected level. For all false-negative biopsy cores the following histological criteria were recorded: number of atypical or malignant glands, architecture, cytoplasm, nucleus, nucleolus, intraluminal mucin, eosinophilic debris and crystalloids. In case of an uncertain diagnosis at re-evaluation, additional levels were cut and stained for H&E and immunohistochemically for basal cell keratins (34BE12) according to standard procedures.

7.2.4 | Statistical analysis

Differences between cases with and without a missed lesion were assessed using the Chi-square test for binomial variables and a Kruskal-Wallis test for continuous variables. The Fisher's exact test was used in case of small numbers. A two-sided p-value < 0.05 was considered statistically significant.

For statistical analysis the Statistical Package for Social Sciences (SPSS version 15.0; SPSS inc., Chicago, IL, USA) was used.

7.3 | Results

7.3.1 | Population characteristics

Up to May 2008, a total of 857 patients were diagnosed with prostate cancer in the second (n=550) or third (n=307) screening round, of whom n=96 (17.5%) and n=106 (34.5%) men had a benign biopsy in the prior screening. Of these 202 cases, 196 (97.5%) biopsy specimens could be retrieved from the archive and were available for review.

7.3.2 | Occurrence of relevant lesions in initially negative biopsy specimens

At initial review of the tissue slides, a total of 43 lesions consisting of atypical glands were identified in 35 patients (17.9%), 11 of which were considered diagnostic for adenocarcinoma and 32 suspicious for adenocarcinoma (ASAP). In six of 43 biopsy cores, the atypical glands were located adjacent to high-grade PIN. A total of 17 all-round pathologists with variable years of diagnostic experience had signed out the initial “benign” pathology reports, indicating that missing of atypical glands was a general phenomenon.

To establish a definitive diagnosis for lesions suspicious for adenocarcinoma or confirm a diagnosis of adenocarcinoma, we obtained deeper sections from the original paraffin blocks for H&E and immunostaining for basal cells. In 23 cases no staining was performed as definitive diagnosis did not warrant additional staining in case of adenocarcinoma (n=7), or the staining would not be informative because atypical glands were not present anymore in the deeper sections (n=16). For one biopsy the representative block could not be retrieved from the archive, leaving 19 biopsy cores for additional evaluation. Based on the outcome of deeper sections and lack of basal cell staining, 8 suspicious foci were now considered diagnostic for adenocarcinoma, including two of six lesions adjacent to high grade PIN. In 7 atypical lesions no definitive diagnosis could be reached due to a low number of atypical glands negative for basal cells. In 4 cases, the presence of adenocarcinoma was confirmed by absence of basal cells. Therefore, the final diagnosis was adenocarcinoma in 19 biopsy cores corresponding to 16 (8.2%) patients and ASAP in 24 biopsy cores, corresponding to 19 patients (9.7%). All missed prostate cancer cases were Gleason score 6 (3 + 3).

The rate of false-negative biopsy outcome was 17.9% in this subset of biopsied men diagnosed with prostate cancer during a subsequent screening round. This rate of false-negative biopsy outcome was extrapolated to the overall screening population. Of all screening participants in the ERSPC Rotterdam with a benign biopsy in the first or second screening round, and a second biopsy during the subsequent screening round, on average 13.4% were diagnosed with cancer during the following, i.e. second or third, screening round. Based on this mean cancer incidence of 13.4% in men with a previous biopsy, we estimate that 2.4% (i.e. 17.9% of 13.4%) of all biopsied screening participants had a false-negative biopsy outcome; for prostate cancer, this rate was 1.1% and for ASAP 1.3%.

7.3.3 | Histopathological characterisation of false-negative needle-biopsies

The histopathological characteristics of the missed cancers and ASAP are shown in Table 1. In 28 patients the missed atypical and malignant glands were identified in 1 core, in 6 patients in 2 cores and in 1 patient in 3 cores. In 16 patients, the missed lesion was observed in the same biopsy core as the final diagnostic lesion at the time of actual PC detection, in 14 patients the lesion was found in the same lobe, and in 4 patients the missed lesion and the final malignant lesion were both identified in the base or both in the apex of the prostate. Only in 1 patient a missed ASAP lesion was found in a core from the opposite side of the prostate compared to the location of actual PC detection during the subsequent screening round.

In four biopsy cores the missed lesion was easily identified and no specific aspect could be identified explaining why this lesion was missed. All other missed lesions showed at least one feature, which could explain why they were not diagnosed as such. Most common amongst these features were absence of architectural abnormality (n=21, 48.8%) and low number of atypical glands (less than 10 glands, n=22, 51.2%). Low number of glands was especially observed in the ASAP lesions, which consisted of median 5 glands compared to malignant lesions comprising a median of 17 glands (Table 1). Larger missed lesions consisting of more atypical glands generally resembled normal glands more closely at low magnification or they were intensely intermingled with pre-existent glands (Figure 1, 2).

Next to small size and lack of architectural abnormality, some other specific histopathological characteristics were observed that might explain why these lesions were missed or misinterpreted. One adenocarcinoma was of the foamy gland type and 3 lesions had atypical large-sized glands indicative for pseudo-hyperplastic prostate cancer, which can both be easily misinterpreted if one is not familiar with these variants (Figure 3). Four lesions were probably missed because they were small and were located at the border of a biopsy or in a small biopsy fragment. At low power, cases might have been less conspicuous due to normal gland size (n=16, 37.2%) or normal spacing between the glands (n=6, 14.0%). In one case, no definitive distinction between the atypical lesion and partial atrophy could be made. The most common findings, which triggered our suspicion for the presence of an atypical lesion were architectural abnormality, presence of enlarged nuclei, prominent or conspicuous nucleoli, amphophilic cytoplasm and presence of intra-luminal eosinophilic secretions or crystalloids (Table 1).

In 12 of the 24 cores with ASAP, the diagnosis might have been modified if immunohistochemistry could have been performed on the most relevant level.

Table 1 | Histopathological characterisation of false-negative needle-biopsies. PC = prostate cancer, ASAP = atypical small acinar proliferation, PIN = prostatic intraepithelial neoplasia.

	All false-negative lesions, number (%)	Missed ASAP lesion, number (%)	Missed PC lesion, number (%)
Number of glands	9	5	17
Median	13.3	6.1	22.5
Mean	13.3	6.1	22.5
Range	1-56	1-20	5-56
Architecture	21 (48.8)	14 (58.3)	7 (36.8)
Nodule, not invasive	22 (51.2)	10 (41.7)	12 (63.2)
Invasive between pre-existent glands	6 (14.0)	4 (16.7)	2 (10.5)
Spacing of the glands	37 (86.0)	20 (83.3)	17 (89.5)
Similar to pre-existent glands	16 (37.2)	12 (50.0)	4 (21.1)
Closely packed	27 (62.8)	12 (50.0)	15 (78.9)
Size of the glands	11 (25.6)	4 (16.7)	7 (36.8)
Normal (like pre-existent glands)	30 (69.8)	19 (79.2)	11 (57.9)
Smaller than pre-existent glands	1 (2.3)	1 (4.2)	-
Cytoplasm	1 (2.3)	-	1 (5.3)
Normal (like pre-existent glands)	1 (2.3)	-	1 (5.3)
Enlarged	42 (97.7)	24 (100)	18 (94.7)
Nucleus	25 (58.1)	15 (62.5)	10 (52.6)
Enlarged	18 (41.9)	9 (37.5)	9 (47.4)
Nucleolus	19 (44.2)	14 (58.3)	5 (26.3)
Conspicuous	24 (55.8)	10 (41.7)	14 (73.7)
Intraluminal eosinophilic secretions	33 (76.7)	19 (79.2)	14 (73.7)
No	10 (23.3)	5 (20.8)	5 (26.3)
Yes	40 (93.0)	23 (95.8)	17 (89.5)
Intraluminal crystalloids	3 (7.0)	1 (4.2)	2 (10.5)
No	37 (86.0)	20 (83.3)	17 (89.5)
Yes	6 (14.0)	4 (16.7)	2 (10.5)
Intraluminal mucin	43 (100)	24 (100)	19 (100)
Adjacent to high-grade PIN			
No			
Yes			
Total			

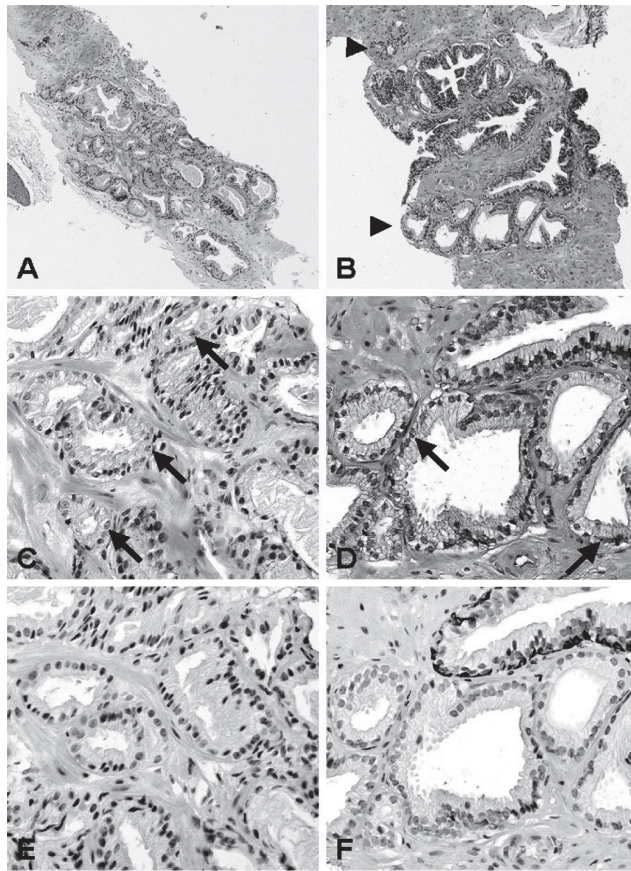


Figure 1 | False-negative prostate biopsies with adenocarcinoma. A, B: Prostate cancer glands are architecturally arranged in a nodule reminiscent of benign tissue. In B two foci (arrowheads) of atypical glands are present adjacent to normal pre-existent glands. C, D: At high magnification enlarged nuclei and conspicuous nucleoli (arrows) are visible, while cytoplasm is not conspicuous (D). E, F: Basal cells are absent (34BE12). In F the second focus of atypical glands also lacked basal cells (not shown); notice positive internal control (F). Both lesions were considered adenocarcinoma Gleason score 6 (3+3). The lesions were derived from two separate patients (A, C, E and B, D, F). Original magnifications: A, B H&E 40x; C, D H&E 200x; E 34BE12 100x; F 34BE12 200x. For color images, see appendix page 181.

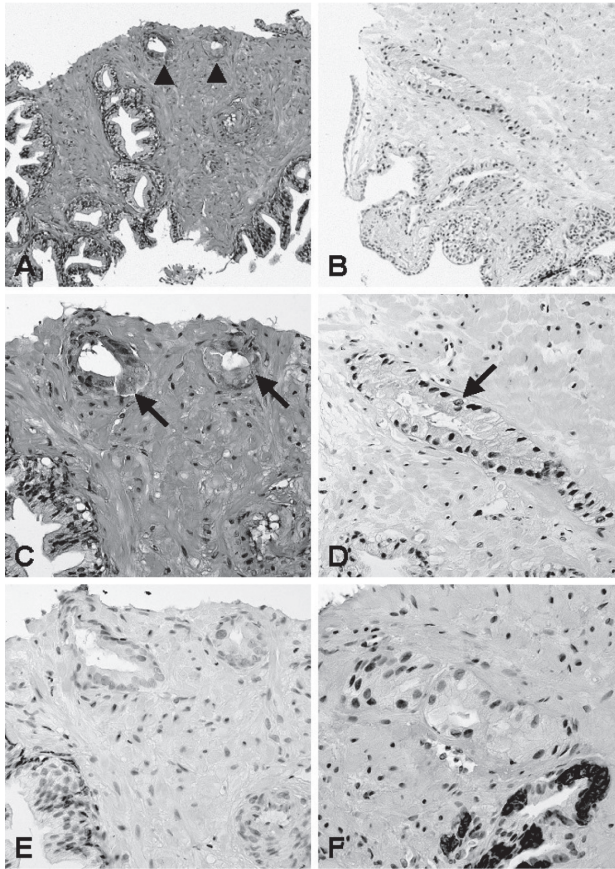


Figure 2 | False-negative prostate biopsies with atypical glands suspicious for adenocarcinoma. A, B: Two (A; arrowheads) and one (B) atypical glands were discovered at low magnification by their amphophilic cytoplasm and subtle architectural abnormality. C, D: At high magnification, the suspicious glands revealed enlarged nuclei and prominent nucleoli (arrows). E, F: The atypical glands showed lack of basal cells (34BE12). Both lesions were considered highly suspicious for malignancy. Due to a low number of atypical glands no definitive diagnosis for malignancy was given. The lesions were derived from two separate patients (A, C, E and B, D, F). Original magnifications: A H&E 40x; B H&E 100x; C, D H&E 200x; E, F 34BE12 200x. For color images, see appendix page 182.

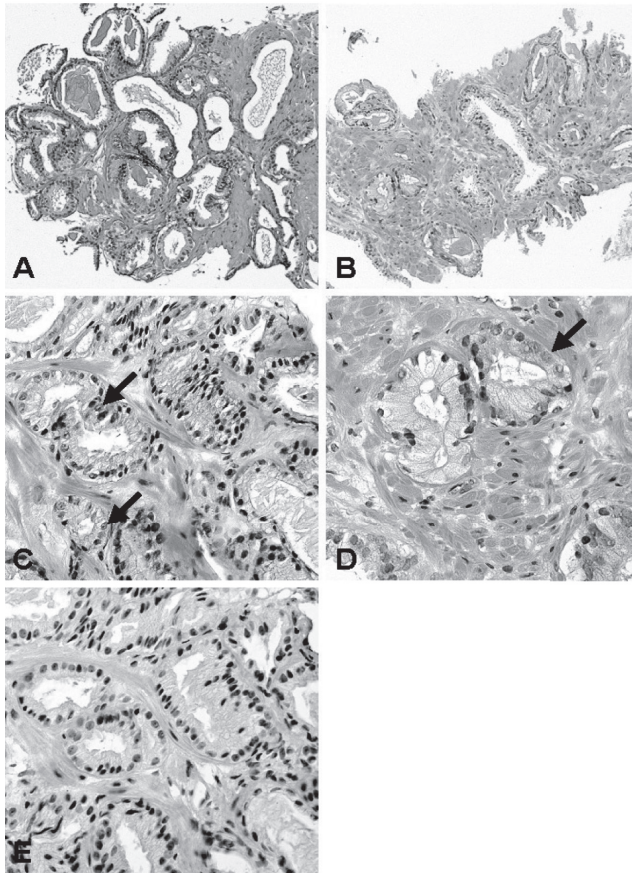


Figure 3 | False-negative prostate biopsy with pseudohyperplastic (A, C, E) and foamy gland (B, D) prostate adenocarcinoma. A, C: Dilated glands with cytoplasmic amphophilia, enlarged nuclei and conspicuous nucleoli characterize pseudohyperplastic cancer. E: Basal cells are absent in 34BE12 staining; notice positive staining in pre-existent atrophic glands. The lesion might be missed as large-sized glands can be interpreted as benign or hyperplastic glands at low magnification. B, D: Foamy gland cancer is characterized by architecturally disorganized glands with clear to foamy cytoplasm with some enlarged nuclei and sporadic nucleoli (arrow). The lesion might be missed as cytoplasmic amphophilia is not conspicuous at low magnification and only some of the nuclei are atypical with prominent nucleoli at high magnification. No immunohistochemical staining was performed on this lesion. Original magnifications: A, B H&E 40x; C, D H&E 200x; E 34BE12 200x. For color images, see appendix page 183.

7.3.4 | Clinico-pathological follow-up of false-negative prostate biopsies

At the time of the first biopsy, no differences were found between patients with a missed prostate cancer lesion, a missed ASAP lesion or “true benign” diagnosis (Table 2).

Patient and tumour characteristics at the time of prostate cancer detection are listed in Table 3. None of the cases were diagnosed with positive lymph nodes or metastasis at the time of prostate cancer diagnosis. When these characteristics were stratified by missed adenocarcinoma lesion, missed ASAP lesion and no missed lesion in the previous biopsy no significant differences in tumour characteristics were observed. Overall, men with a false-negative biopsy seemed to be in a curable stage at the time of PC diagnosis. Strikingly, the six prostate cancers with a Gleason score >7 detected in the next screening round were all found in patients with a previous “true benign” biopsy (no missed lesion).

Table 2 | Patient and tumour characteristics at the time of the previous benign biopsy. Characteristics are depicted for the total cohort and stratified by previous biopsy review outcome, i.e. missed adenocarcinoma lesion (n=161), missed suspicious lesion (n=19), no missed lesion (n=161).

	Total group n = 196	Missed PC lesion n= 16	Missed suspicious lesion n = 19	No missed lesion n = 161	P value
Age mean (median)	64.4 (64.5)	65.1 (66.4)	65.1 (64.9)	64.2 (64.3)	0.438
PSA ng/ml mean (median)	4.9 (4.3)	5.6 (4.9)	4.4 (4.0)	4.9 (4.3)	0.343
Volume ml mean (median)	46.1 (43.1)	51.4 (42.1)	44.3 (40.9)	45.8 (43.1)	0.667
Abnormal digital rectal examination	33 (16.8%)	4 (25%)	1 (5.3%)	28 (17.1%)	0.271
Visible lesion at transrectal ultrasound	20 (10.2%)	2 (12.5%)	1 (5.3%)	17 (10.6)	0.765

Table 3 | Patient and tumour characteristics at the time of prostate cancer detection. Characteristics are depicted for the total cohort and stratified by previous biopsy review outcome, i.e. missed adenocarcinoma lesion (n=16), missed suspicious lesion (n=19), no missed lesion (n=161).

	Total group n = 196	Missed PC lesion n= 16	Missed suspicious lesion n = 19	No missed lesion n = 161	P value
Age mean (median)	68.5 (68.7)	69.3 (70.5)	69.3 (69.1)	68.4 (68.6)	0.426
PSA ng/ml mean (median)	7.4 (5.7)	7.8 (6.8)	7.4 (5.4)	7.3 (5.5)	0.852
Volume ml mean (median)	55.8 (51.5)	57.8 (46.2)	52.3 (51.2)	56.0 (51.9)	0.698
Clinical tumour stage					
T1c	127 (64.8)	10 (62.5)	13 (68.4)	104 (64.6)	0.853
T2	64 (32.7)	5 (31.3)	6 (31.6)	53 (32.9)	
T3/4	5 (2.6)	1 (5.3)	0 (0.0)	4 (2.5)	
Gleason score					
<7	170 (86.7)	14 (87.5)	18 (94.7)	138 (85.7)	0.732
=7	20 (10.2)	2 (12.5)	1 (5.3)	17 (10.6)	
>7	6 (3.1)	0 (0.0)	0 (0.0)	6 (3.7)	
Number of positive cores					
1	118 (60.2)	9 (56.3)	13 (68.4)	96 (59.6)	0.923
2	51 (26.0)	4 (25.0)	4 (21.1)	43 (26.7)	
>2	27 (13.8)	3 (18.8)	2 (10.5)	22 (13.7)	
Maximum tumour invasion per core					
<50%	172 (87.8)	13 (81.3)	17 (89.5)	142 (88.2)	0.701
>=50%	24 (12.2)	3 (18.8)	2 (10.5)	19 (11.8)	

7.4 | Discussion

In this study, a false-negative biopsy rate of 17.9% was found in screening participants who were subsequently diagnosed with prostate cancer: in 8.2% a prostate cancer lesion was missed and in 9.7% a lesion suspicious for adenocarcinoma. After correction for the selection in this study, i.e. only men who were diagnosed with PC 4 years later during the subsequent screening round, the overall false-negative biopsy rate was estimated to be 2.4% in men who underwent a prostate biopsy. For this calculation it was assumed that in case of a second negative biopsy, the preceding biopsy was not false-negative, which is supported by a previous review study reporting no ASAP or carcinoma in 54 men with a second negative biopsy.¹⁰

The true incidence of false-negative prostate biopsy outcome is unknown and publications on this topic are scarce. Kronz et al¹¹ identified missed lesions at a consultation service, including high-grade PIN in 2.7% of the biopsies submitted for review. They considered this percentage an underestimation, since only selected biopsies were submitted for consultation. In a study by Van der Kwast et al,¹² two pathologists reviewed 141 prostate biopsy sets from a screened population, including 127 originally reported as benign and 14 reported as malignant. When

prostate cancer and ASAP were combined, one review pathologist reported missed lesions in 6.3% of the biopsies and the other in 10.2%. The opposite scenario, i.e. calling a benign lesion carcinoma, has been described as well. The rate of overcalling a lesion has been reported to be about 1.2%,^{13,14} but this phenomenon was not assessed in the current study.

Several histological explanations for the occurrence of false-negative needle-biopsy examination were identified. The majority of false-negative biopsies contained lesions with minimal architectural abnormality and were very small. Additionally, some lesions consisted of glands of normal size and spacing. To enhance the chance of finding small atypical foci we screened one level of a needle-biopsy at 10x objective magnification and the other levels on 4x. In this way, even one or two glands that may initially be overlooked at low magnification screening can be identified by their cytologic atypia or presence of abnormal luminal contents. It seems thus important to examine at least one level of prostate biopsies at high magnification even if no abnormalities are observed at low magnification. Furthermore, 4 of a total of 19 (21%) prostate cancer lesions showed specific characteristics of foamy gland adenocarcinoma¹⁵ or atypical large-sized glands typical for pseudo-hyperplastic prostate cancer,¹⁶ which may be difficult to recognize for a pathologist without special interest in urologic pathology. Missing such a specific entity at prostate biopsy is probably more difficult to prevent, but increasing the awareness of these histological variants could improve their identification.

Other explanations for the occurrence of false-negative prostate needle biopsies may be found in factors other than histopathological characteristics. Although the possibility could be entertained that during the earlier years of the screening study more atypical lesions would be missed than in later years, our analysis did not support such a “learning curve” (data not shown). Inter-observer variability may also explain the reporting of false-negative needle-biopsies. It is well known that pathologists with special interest in urologic pathology are more confident in interpretation of small atypical lesions. After review by a urologic pathologist, atypical lesions can be reclassified as malignant in 2.2%-45.1% and as benign in 5.2%-16.7%.^{2,6,9} In addition, the two review pathologists may have reported a false-negative atypical lesion as carcinoma or suspicious more frequently as they were aware of the final outcome due to the study design. Since consensus of the diagnosis was reached between the two pathologists for all cases, we think that this inter-observer variability and detection bias was minimal.

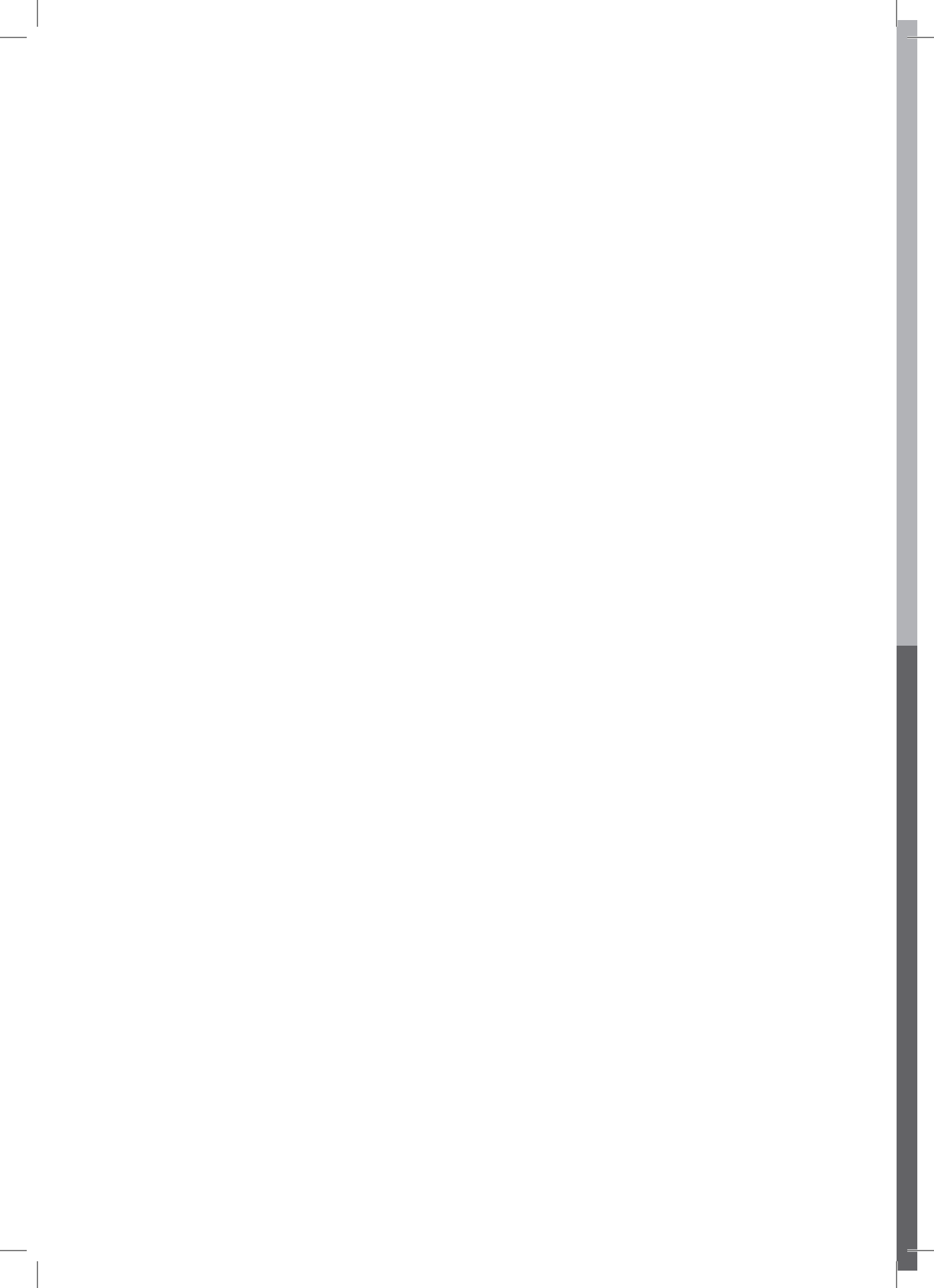
The majority of patients with a false-negative biopsy outcome seemed to be in a curable stage at the time of actual prostate cancer detection, even though 16 of 35 patients with a false-negative biopsy were classified as having a missed prostate cancer lesion 4 years prior to the eventual diagnosis. In only 2 of these 16 patients the diagnostic biopsy revealed a Gleason score 7 adenocarcinoma. This favorable outcome after a false-negative biopsy can be explained by two reasons. First, all men were enrolled in a screening program and were re-biopsied 4 years after the false-negative biopsy. Cancers in such a screening program are detected in a pre-clinical phase, with an estimated lead time of on average 10 years in our population³. Secondly, almost all missed lesions were small and all missed prostate cancer lesions were Gleason score 6 (3

+ 3). Consequently, missing a lesion in this particular setting may not necessarily negatively influence patient outcome.

Studies on false-negative or -positive diagnosis in pathology might be confronting to patients and doctors. Nevertheless, they reflect quality measures for daily clinical practice and offer opportunities to improve logistic and clinical expertise. In this study, the overall rate of false-negative biopsy was estimated to be 2.4%, of which 1.1% represented prostate cancer and 1.3% an atypical lesion suspicious for adenocarcinoma. Although missing a small suspicious or malignant lesion did not necessarily negatively influence patient outcome in this screening setting, the clinical effect of missing an atypical lesion in other settings is unclear. A relatively low number of atypical glands, intense intermingling with pre-existent glands or subtle architectural abnormality were potential risk factors for false-negative interpretation together with the occurrence of unusual histological variants such as pseudohyperplastic and foamy gland cancer. Routine examination at high magnification of at least one level of each prostate biopsy core, even in absence of abnormalities at low magnification, might reduce the number of false-negative prostate biopsies.

References

1. Iczkowski KA, MacLennan GT, Bostwick DG. Atypical small acinar proliferation suspicious for malignancy in prostate needle biopsies: clinical significance in 33 cases. *Am J Surg Pathol.* 1997;21:1489-95
2. Epstein JI, Herawi M. Prostate needle biopsies containing prostatic intraepithelial neoplasia or atypical foci suspicious for carcinoma: implications for patient care. *J Urol* 2006;175:820-834
3. Vis AN, Hoedemaeker RF, Roobol M, et al. The predictive value for prostate cancer of lesions that raise suspicion of concomitant carcinoma: an evaluation from a randomized, population-based study of screening for prostate cancer. *Cancer* 2001;92:524-34
4. Allan RW, Sanderson H, Epstein JI. Correlation of minute (0.5 MM or less) focus of prostate adenocarcinoma on needle biopsy with radical prostatectomy specimen: role of prostate specific antigen density. *J Urol* 2003;170:370-2
5. Thorson P, Vollmer RT, Arcangeli C, et al. Minimal carcinoma in prostate needle biopsy specimens: diagnostic features and radical prostatectomy follow-up. *Mod Pathol* 1998;11:543-51
6. Epstein JI. Diagnosis and reporting of limited adenocarcinoma of the prostate on needle biopsy. *Mod Pathol* 2004;17:307-15
7. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320-8
8. Roobol MJ, Kirkels WJ, Schröder FH. Features and preliminary results of the Dutch centre of the ERSPC (Rotterdam, the Netherlands). *BJU Int* 2003;92 suppl2:48-54
9. Van der Kwast ThH, Lopes C, Santonja C, et al. Guidelines for processing and reporting of prostatic needle biopsies. *J Clin Pathol* 2003;56:336-40
10. Wolters T, Roobol MJ, Schröder FH, et al; ERSPC Rotterdam section. Can non-malignant biopsy features identify men at increased risk of biopsy-detectable prostate cancer at re-screening after 4 years? *BJU Int* 2008;101:283-8.
11. Kronz JD, Milord R, Wilentz R, et al. Lesions missed on prostate biopsies in cases sent in for consultation. *Prostate* 2003;54:310-4
12. Van der Kwast ThH, Lopes C, Martikainen PM, et al. Report of the pathology committee: false-positive and false-negative diagnoses of prostate cancer. *BJU Int* 2003;92 suppl2:62-5
13. Kronz JD, Westra WH, Epstein JI. Mandatory second opinion surgical pathology at a large referral hospital. *Cancer* 1999;86:2426-35
14. DiGiuseppe JA, Sauvageot J, Epstein JI. Increasing incidence of minimal residual cancer in radical prostatectomy specimens. *Am J Surg Pathol* 1997;21:174-8
15. Nelson RS, Epstein JI. Prostatic carcinoma with abundant xanthomatous cytoplasm. Foamy gland carcinoma. *Am J Surg Pathol.* 1996;20:419-26
16. Levi AW, Epstein JI. Pseudohyperplastic prostatic adenocarcinoma on needle biopsy and simple prostatectomy. *Am J Surg Pathol.* 2000;24:1039-46



PART IV

PROGNOSIS OF PROSTATE CANCER: INSIGNIFICANT OR AGGRESSIVE?

Chapter 8

Should prostate tumour volume routinely be reported by the pathologist?: The prognostic value of tumour volume in prostate cancer

Eur Urol 2009, epub ahead of print, doi:10.1016/j.eururo.2009.07.027

Chapter 9

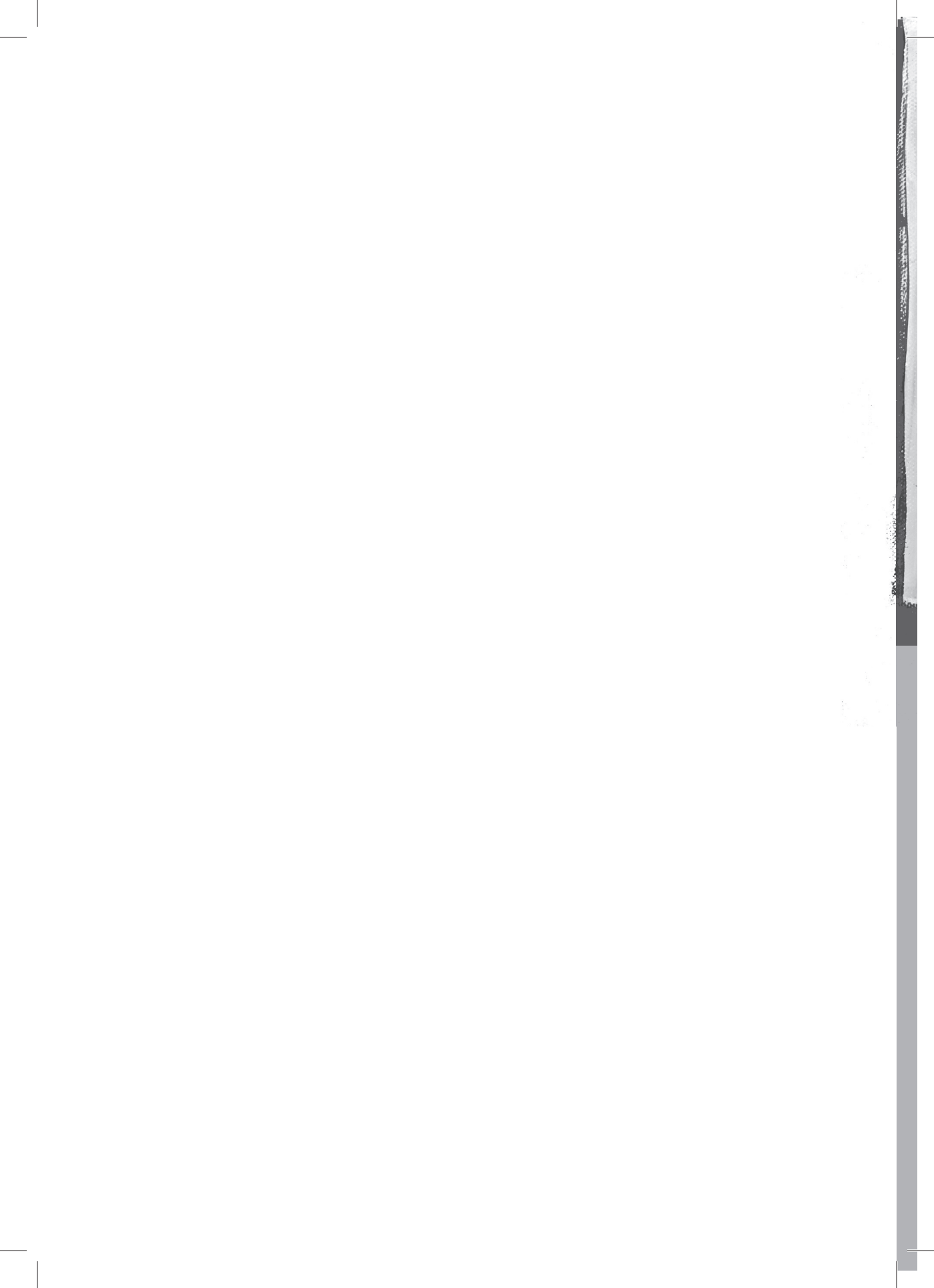
Reassessment of the tumour volume threshold defining clinically insignificant prostate cancer; results from a randomized trial

Submitted

Chapter 10

The value of EZH2, p27^{kip1}, BMI-1 and MIB-1 on biopsy specimens with low risk PC in selecting men with significant prostate cancer at prostatectomy

BJU Int 2009, epub ahead of print, doi:10.1111/j.1464-410x.2009.08998





CHAPTER 8



Should prostate tumour volume routinely be reported by the pathologist?: The prognostic value of tumour volume in prostate cancer

Eur Urol 2009, epub ahead of print, doi:10.1016/j.eururo.2009.07.027

Tineke Wolters, Monique J. Roobol, Pim J. van Leeuwen, Roderick C.N. van den Bergh,
Robert F. Hoedemaeker, Geert J.L.H. van Leenders, Fritz H. Schröder and
Theodorus H. van der Kwast

Abstract

Background: The independent prognostic value of tumour volume in radical prostatectomy (RP) specimens is controversial and it remains a matter of debate whether the pathologist should report a measure of tumour volume. In addition, tumour volume might be of value in substaging of pathological tumour stage (pT2) prostate cancer (PC).

Objective: To assess the prognostic value of PC tumour volume.

Design, setting and participants: The cohort consists of 344 participants of the European Randomized Study of Screening for Prostate Cancer (ERSPC), Rotterdam section whose PC was treated with RP. Mean time of follow-up was 96.2 months.

Measurements: Tumour volume was measured in totally embedded RP specimens with a morphometric, computer-assisted method and assessed as a continuous variable, as relative tumour volume (tumour volume divided by prostate volume) and in a binary fashion (\geq or $<$ than 0.5ml). These variables were related to PSA progression, local recurrence or distant metastasis and PC-related mortality using univariate and multivariable Cox proportional hazards analyses. The analyses were repeated in the subgroup with pT2 tumours.

Results and limitations: Tumour volume was related to tumour stage, Gleason score, seminal vesicle invasion and surgical margin status. In univariate analyses, tumour volume and relative tumour volume were predictive for all outcome variables. In multivariable analyses, including age, tumour stage, Gleason score, seminal vesicle invasion and surgical margin status, neither tumour volume nor relative volume were independent predictors of progression or mortality. Tumour volume ≥ 0.5 ml was predictive for PSA recurrence and local and/or distant progression in univariate analyses, but not in multivariable analyses. Tumour volume was not predictive for recurrence or mortality in univariate or multivariable analyses in the pT2 subgroup.

Conclusions: Tumour volume did not add prognostic value to routinely assessed pathological parameters. Therefore, there seems to be little reason to routinely measure tumour volume in RP specimens.

8.1 | Introduction

The relation of prostate cancer (PC) tumour volume with pathological features at radical prostatectomy specimens, PC progression and mortality has been reported by several independent study groups.^{1-5,8} However, whether tumour volume or a derivate of tumour volume should be included in the pathology report of radical prostatectomy specimens, remains subject of debate. This is due to the fact that the independent prognostic value of tumour volume, i.e. after correction for other pathological parameters like tumour stage, Gleason score, seminal vesicle invasion and surgical margin status, is controversial. Some studies report independent prognostic value of tumour volume after correction for other routinely used parameters,^{1,3,7,8} while others find opposite results^{2,5,6} (Table 1).

Therefore, although the prognostic value of tumour volume is not in doubt, the question is whether tumour volume adds value to the prognosticators tumour stage, Gleason score, seminal vesicle invasion and surgical margin status. This is an important discussion: if tumour volume does not add value to the more easily accessible data on tumour stage, Gleason score, seminal vesicle invasion and surgical margin status, there would be no need to report tumour volume.

One explanation for the discordant data on the independent prognostic value of tumour volume may be the differences between the cohorts in the reported studies. The studies that showed independent value, more often described a cohort with unfavorable tumour features^{1,8} (Table 1) compared to the more contemporary studies that failed to show similar results.^{2,6,9} Therefore, subgroup analyses in cases with high-risk PC will be done in this study to see whether the prognostic value of tumour volume is higher in this particular subset.

A second explanation may be the use of different tumour volume variables. Some use tumour volume as a continuous variable, log transformed or not^{1,2,5,6} while others use relative tumour volume, i.e. the percentage of the prostate volume invaded by tumour^{3,5,7} or use cut-off values, defining two or three categories.^{8,9} Two studies suggested that the relative tumour volume was more strongly associated with prognosis than tumour volume.^{10,11} In order to deal with these differences, we studied tumour volume as a continuous variable, as relative tumour volume and as a binary variable in the current study.

Finally, we considered that tumour volume might be of help in the substaging of pathological tumour stage 2 (pT2) prostate cancer. The current subclassification of pT2 PC,¹² distinguishes unilateral PC invading less than half a lobe (pT2a), more than half a lobe (pT2b) and bilateral tumour invasion (pT2c). This classification lacks uniform criteria for assigning a substage, especially in distinguishing pT2a from pT2b PC and lacks clinical relevance: the pathological substaging of pT2 tumours has no prognostic value.^{13,14} In order to assess whether tumour volume may be useful to differentiate separate prognostic groups within the set of pT2 tumours, we performed a subgroup analysis for pT2 PC.

Table 1 | Overview of publications on tumour volume. Only studies of independent study groups which included at least 100 men were selected. If one study group published more than one paper on this subject, the most recent paper was selected for this table.

Author, year of publication	Number of cases	Time period in which surgery was performed	Follow-up period in years mean (median)	Tumour volume parameter	Mean (median) tumour volume in ml	Outcome parameter	Independent predictive value
Cheng et al, 1993 (8)	894 (stage C prostate cancer)	1989-1993	4.4 (-)	Tumour volume, divided in 3 categories (<=3.0ml, 3.0-10.0ml, >10 ml)	-	Progression and disease-specific survival	yes
Dvorak et al, 2005 (26)	781	1986-2002	5.4 (-)	Maximum tumour diameter	-	Biochemical recurrence	no
Eichelberger et al, 2005 (27)	364	1993-2003	1.2 (1.0)	Maximum tumour diameter	2.5 (-)	Biochemical recurrence	yes
Epstein et al, 1993 (5)	185 (stage B prostate cancer)	-	At least 5 years	Tumour volume and relative tumour volume	- (-)	Progression (combining biochemical recurrence, local recurrence and distant metastasis)	no
Kikuchi et al, 2004 (2)	1,302	1983-2000	4.4 (3.9)	Tumour volume measured by computer planimetry, logarithmic transformation (after 1995)	2.16 (before 1995) and 1.25 (after 1995)	Biochemical recurrence	no
Marks et al., 2007 (28)	504	1990-1998	3.7 (-)	Positive block ratio	-	Biochemical recurrence	yes
Merrill et al., 2007 (9)	1,833	1998-2005	- (1.9)	3 categories (low, medium, extensive) based on visual estimation	-	Biochemical recurrence	Yes in Gleason score 7 PC, not in Gleason score <=6 or >= 8
Ramos et al., 2004 (7)	1,850	1988-2003	3.6 (-)	Visual estimation of relative tumour volume, using cut-off point	-	Biochemical recurrence, local recurrence or distant metastasis	yes

Table 1 | *Continued*

Author, year of publication	Number of cases	Time period in which surgery was performed	Follow-up period in years mean (median)	Tumour volume parameter	Mean (median) tumour volume in ml	Outcome parameter	Independent predictive value
Salomon et al., 2003 (6)	200	1992-1998	5.3 (-)	Area of the tumour was measured with unknown method. A shrinkage factor of 1.5 was applied.	1.35 (-)	Biochemical recurrence	no
Stamey et al., 1999 (1)	379	1983-1992	4.99 (4.95)	Tumour volume measured by computer planimetry, logarithmic transformation	4.66 (2.90)	Biochemical recurrence	yes
Van Oort et al., 2008 (31)	542	1992-2005	3.3 (-)	Maximum tumour diameter	-(2.0)	Biochemical recurrence	no
Vollmer, 2009 (3)	447	unknown	In patients alive at last follow-up (n=353): 6.2 (-)	Visual estimation of relative tumour volume	-	Overall survival	yes
Wolters et al., 2009 (current study)	344	1993-2000	8.0 (8.2)	Tumour volume measured by computer planimetry	1.05 (0.66)	Biochemical recurrence, local recurrence or metastasis, PC related mortality	no

8.2 | Methods

8.2.1 | Patient selection

All men included in the study were participants in the screening arm of the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC). Up to May 1997 a lateralized sextant prostate needle biopsy was indicated by a PSA ≥ 4.0 ng/ml, or a suspicious digital rectal examination or transrectal ultrasound. Hereafter, only a PSA ≥ 3.0 ng/ml prompted a biopsy. All patients diagnosed with PC during the first screening round or during the early re-screening round 1 year later and treated with radical prostatectomy were selected for the analysis, resulting in a total of 424 cases. For 344 (81.1%) cases the prostate cancer volume was known and those were included in the cohort. All patients were treated between November 1993 and June 2000 and none received adjuvant therapy.

8.2.2 | Pathologic examination

RP specimens were processed following the protocol described before.¹⁵ In short, after fixing the specimens, they were inked, cut at 4mm intervals perpendicular to the rectal surface. The apical slice is cut parasagittally at 2-3 mm intervals. The sections were divided in halves or quadrants to fit routine used cassettes for paraffin embedding. The whole prostate was sampled. Pathological tumour stage, Gleason score, surgical margin status (SM) and presence of seminal vesicle invasion (SVI) were assessed. Tumour areas were marked at each slide and measured using a computerized morphometric analysis.¹⁶ Subsequently the volumes were calculated by multiplying the area by the slice thickness. No separate calculations were made for tumour present in consecutive or non-consecutive slides. Since it was shown in a side-study (personal communication RFH) that fixation and processing of the RP specimens did not result in significant reduction of the tissue volume, shrinkage factor was not taken into account for the calculation of tumour volume.

Staging was done according to the 1992 TNM classification. The pT4 cases showed tumour invasion of the detrusor muscle. Seminal vesicle invasion was separately analyzed, because this feature might also be present in pT4 prostate cancer in addition to the pT3c cancers.

8.2.3 | Follow-up

Patients were followed after RP by serial PSA measurements every 3 months during the first year after RP, every 6 months in the second year and yearly after 2-3 years. Biochemical recurrence was defined as a PSA > 0.2 ng/ml.¹⁷ Local recurrence was defined as disease recurrence proven by biopsy or recorded as local recurrence in the medical records by the treating urologist. Metastasis was indicated by suspicious bone scintigraphy, MRI or CT scans. Total follow-up time was defined as the time from RP to death or last visit date. Time to recurrence was defined as the time of RP to the time of the first signs of recurrence. When no signs or recurrence were registered, cases were censored at the time of the last follow-up visit or date of death. The cause

of death was established by an independent committee¹⁸ and subsequently scored as PC-related or intercurrent cause of death.

8.2.4 | Statistical analysis

Univariate and multivariable Cox proportional hazards regression model analyses were performed to evaluate the value of the possible predictors on biochemical recurrence, local recurrence or metastasis and PC-related mortality. Tumour volume was assessed as a continuous variable. Additionally, relative tumour volume was calculated as tumour volume divided by prostate volume and multiplied by 100%. Tumour volume was also categorized in low and high volume groups using the cut-off of 0.5 ml, based on the cut-off suggested by Stamey for indolent PC.¹⁹ The base model resulting from the multivariable Cox analyses, using age, pT-stage, Gleason score, SVI and surgical margin status, was compared to the extended models, using one of the tumour volume related parameters in addition to the base model, with the likelihood ratio test. The analyses were repeated on 2 subgroups: high-risk PC (any of the following: pT3 or pT4, Gleason ≥ 8 , or positive surgical margins) and pT2 PC.

We analyzed the data using the statistical package for social sciences, version 15.0 (SPSS 15.0, Chicago, IL). Statistical significance was defined as a two-sided p-value < 0.05 .

8.3 | Results

8.3.1 | Pathologic characteristics

At the time of RP, mean (median) age was 63.6 (64.0) years. The pathologic characteristics found at RP are listed in Table 2. Advanced PC (pT3/4) was found in 82 (23.8%) patients and high-grade cancer (Gleason score ≥ 8) in 11 (3.2%) patients. Mean (median) tumour volume was 1.05 (0.66) ml, with a range from 0.001 to 13.48 ml. Tumour volume was related to pT-stage, Gleason score, seminal vesicle invasion and surgical margin status ($p < 0.001$, Table 2).

The mean (median) time of follow-up was 96.2 (98.0) months, with a range of 0 to 161 months. The number of events of biochemical and local progression, metastasis and PC specific mortality are described in Table 3.

8.3.2 | PSA recurrence

A total of 57 patients (16.6%) showed biochemical progression after a mean (median) period of 46.8 (43.0) months. Patients with biochemical progression had a significantly higher tumour volume (1.60 versus 0.95 ml, $p = 0.018$). All routinely assessed variables, i.e. tumour stage, Gleason score, seminal vesicle invasion and surgical margin status, were predictive for PSA recurrence in univariate analysis (Table 4 and Figure 1). In univariate Cox regression, tumour volume was a significant predictor of PSA recurrence: HR 1.26 (95% CI 1.12-1.42), $p < 0.001$.

Relative tumour volume (HR 1.10 (95% CI 1.05-1.16), $p < 0.001$) and tumour volume > 0.5 ml (HR 2.25 (95% CI 1.21-4.18) $p = 0.010$) were predictive as well.

In multivariable analyses, none of the tumour volume related variables remained a significant predictor of PSA recurrence: HR of 0.95 (95%CI 0.82-1.11), $p = 0.517$. The results for tumour volume are shown in table 4. Relative tumour volume and tumour volume ≥ 0.5 ml showed similar results: HR 1.00 (95%CI 0.94-1.06) $p = 0.971$ and HR 1.26 (95% CI 0.63-2.52) $p = 0.517$. Additionally, the base model did not significantly improve by adding one of the tumour volume variables ($p > 0.50$).

Table 2 | Pathologic characteristics at radical prostatectomy, prostate tumour volume and number of events per subgroup in 344 radical prostatectomy specimens. pTstage= pathologic tumour stage, SVI= seminal vesicle invasion, SM=surgical margins

	N (% of total)	Tumour volume mean (median) in ml	P-value
pTstage			
T2	262 (75.9)	0.80 (0.52)	0.000*
T3a/b	60 (17.4)	1.46 (1.17)	
T3c	4 (1.16)	2.48 (2.21)	
T4	18 (5.0)	2.97 (1.68)	
SVI			
Yes	6 (1.7)	2.11 (1.61)	0.006**
No	338 (98.3)	1.03 (0.64)	
Gleason score			
≤ 6	213 (61.9)	0.89 (0.54)	0.000*
3+4=7	102 (29.7)	1.04 (0.78)	
4+3=7	18 (5.2)	1.98 (1.25)	
≥ 8	11 (3.2)	2.80 (2.42)	
SM			
Positive	84 (23.8)	1.62 (1.20)	0.000**
Negative	260 (76.2)	0.87 (0.55)	

* Kruskal-Wallis test

** Mann-Whitney test

8.3.3 | Local recurrence and distant metastasis

A total of 15 patients (4.4%) showed local recurrence and/or distant metastasis. Seven patients (2.0%) showed local recurrence after a mean (median) period of 58.3 (57.0) months and six patients (1.7%) were diagnosed with distant metastasis after a mean (median) period of 80.9 (88.0) months in the screened group. Two (0.6%) patients showed both local recurrence and metastasis after 24 and 52 months. These men had a significantly higher tumour volume than patients without local recurrence or metastasis (2.40 ml versus 0.99 ml, $p = 0.004$).

Table 3 | Numbers of biochemical recurrence, local recurrence, metastasis and prostate cancer (PC) specific mortality stratified by pathological features in the total cohort, n=334. pTstage= pathological tumour stage, SVI= seminal vesicle invasion, SM=surgical margins.

	N (% of total)	Biochemical recurrence N (% of subgroup)	Local recurrence N (% of subgroup)	Metastasis N (% of subgroup)	PC specific mortality N (% of subgroup)
pTstage					
T2	262 (75.9)	31 (11.8)	5 (1.9)	2 (0.8)	2 (0.8)
T3a/b	60 (17.4)	14 (23.3)	0 (0.0)	2 (3.3)	0 (0.0)
T3c	4 (1.16)	3 (75.0)	0 (0.0)	0 (0.0)	0 (0.0)
T4	18 (5.0)	9 (50.0)	4 (22.2)	4 (22.2)	4 (22.2)
SVI					
Yes	6 (1.7)	5 (83.3)	1 (16.7)	1 (16.7)	1 (16.7)
No	338 (98.3)	52 (15.4)	8 (2.4)	7 (2.4)	5 (1.5)
Gleason score					
<=6	213 (62.5)	19 (8.9)	2 (0.9)	2 (0.9)	1 (0.5)
7 (3+4)	102 (29.7)	22 (21.6)	3 (2.9)	0 (0.0)	0 (0.0)
7 (4+3)	18 (5.2)	11 (61.1)	2 (11.1)	4 (22.2)	3 (16.7)
>=8	11 (3.1)	5 (45.5)	2 (18.2)	2 (18.2)	2 (18.2)
SM					
Positive	84 (23.8)	32 (38.1)	6 (7.1)	3 (3.6)	1 (1.2)
Negative	260 (76.2)	24 (9.2)	3 (2.6)	5 (1.9)	5 (1.9)
Total	344 (100)	57 (16.6)	9 (2.6)	8 (2.3)	6 (1.7)

Table 5 shows that in univariate analysis, tumour volume was a significant predictor of local recurrence and/or metastasis. In multivariable analysis however, tumour volume was no longer significant. Similar results were found for relative tumour volume and tumour volume ≥ 0.5 ml: in univariate analysis, relative tumour volume and tumour volume ≥ 0.5 ml were both significant predictors of local recurrence and/or metastasis (HR 1.15 (95% CI 1.07-1.24), $p=0.000$ and HR 8.34 (95% CI 1.10-63.49), $p=0.041$ respectively). In multivariable analysis, both relative tumour volume and tumour volume ≥ 0.5 ml were no longer significantly predictive (HR 1.04 (95% CI 0.94-1.15) $p=0.457$ and HR 4.50 (95% CI 0.52-38.60), $p=0.171$ respectively). The predictive base model did not significantly improve after including any tumour variable ($p>0.10$).

8.3.4 | PC related mortality

A total of 58 patients died during follow-up. Six deaths (10.3%) were PC-related, 47 (81.0%) due to intercurrent disease and for 5 (8.6%) cases no cause of death was established yet. Men who died from PC had a significantly higher tumour volume than those who died of intercurrent disease or were still alive at the time of last follow-up visit (2.32 ml versus 1.03 ml, $p=0.009$). The number of events was very small, leading to statistical difficulties. In order to get a rough

estimation of the value of tumour volume, we performed the regression analyses nonetheless. In univariate analysis, tumour volume and relative tumour volume were predictive for PC. In multivariable analyses, both variables lost statistical significance. Tumour volume ≥ 0.5 ml was not significantly predictive for PC mortality in univariate analysis nor in multivariable analysis. No improvement of the model was found by adding one of the tumour volume variables to the model ($p > 0.50$).

Table 4 | Univariate and multivariable Cox proportional hazards model for PSA recurrence. RP= radical prostatectomy, SVI= seminal vesicle invasion, HR= hazard ratio, CI= confidence interval.

	Univariate analysis HR (95% CI)	p-value	Multivariable analysis HR (95% CI)	p-value
Age at RP	1.14 (1.07-1.22)	0.000	1.10 (1.03-1.17)	0.008
Pathologic stage				
T2	baseline		baseline	
T3a/b	1.83 (0.97-3.46)	0.061	1.03 (0.53-2.00)	0.928
T3c	9.61 (2.93-32.57)	0.000	1.17 (0.12-11.37)	0.891
T4	5.50 (2.61-11.57)	0.000	2.24 (0.85-5.92)	0.103
SVI				
No	baseline		baseline	
Yes	9.38 (3.73-23.59)	0.000	2.32 (0.43-12.46)	0.327
Gleason score				
≤ 6	baseline		baseline	
7 (3+4)	2.43 (1.31-4.49)	0.005	1.65 (0.86-3.19)	0.135
7 (4+3)	9.20 (4.35-19.45)	0.000	5.95 (2.60-13.64)	0.000
≥ 8	6.56 (2.44-17.63)	0.000	2.66 (0.79-8.93)	0.14
Surgical margins				
Negative	baseline		baseline	
Positive	4.69 (2.76-7.96)	0.000	3.84 (2.15-6.85)	0.000
Tumour volume	1.26 (1.12-1.42)	0.000	0.95 (0.82-1.11)	0.517

8.3.5 | Value of tumour volume in high-risk PC and pT2 PC

In patients with high-risk PC ($n=130$), 41 (31.5%) showed PSA recurrence, 7 (5.4%) showed local recurrence, 6 (4.6%) had distant metastasis and 4 (3.1%) men died from PC. The results of the univariate and multivariable analyses are shown in Table 6: although tumour volume was predictive for PSA recurrence, local recurrence and the combination of local recurrence and/or distant metastasis, no independent statistical significance was reached for any of the endpoint variables in the multivariable analyses. Similar results were shown for relative tumour volume and tumour volume ≥ 0.5 ml and the predictive models were not improved.

Finally, the analyses were repeated in patients with pT2 PC ($n=262$), in which 67 (25.6%) men had pT2a, 7 (2.7%) men pT2b and 188 (71.8%) pT2c. In this group, 3 (1.1%) patients

had a Gleason score ≥ 8 and 46 (17.6%) had positive surgical margins. Thirty-one (11.8%) patients showed biochemical progression, 5 (1.9%) patients local recurrence, 2 (0.8%) patients showed distant metastasis and 2 (0.8%) died from PC. Tumour volume was not predictive for any of the outcome variables in univariate or multivariable analyses and the predictive model did not improve ($p > 0.50$). The same results were found for relative tumour volume and tumour volume ≥ 0.5 ml (data not shown). In multivariate analysis, surgical margin status was the only significantly predictive factor for PSA recurrence (OR 5.2 (95% CI 2.4-11.2), $p < 0.001$). Tumour substage (pT2a, pT2b or pT2c), Gleason score and age were not predictive in multivariable analysis (data not shown).

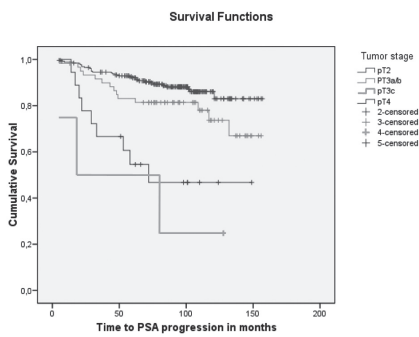


Figure 1a

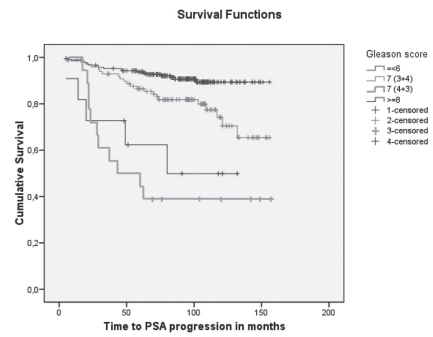


Figure 1b

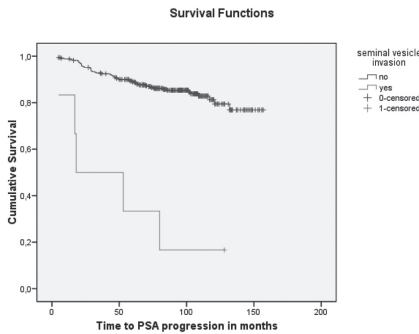


Figure 1c

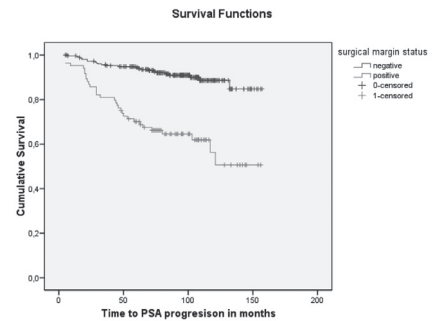


Figure 1d

Figure 1 | Survival curves of routinely assessed pathological characteristics, with time to PSA progression in months. 1a: Time to PSA progression stratified by tumour stage. 1b: Time to PSA progression stratified by Gleason score. 1c: Time to PSA progression stratified by seminal vesicle invasion. 1d: Time to PSA progression stratified by surgical margin status.

Table 5 | Univariate and multivariable Cox proportional hazards model for local and/or distant recurrence. RP= radical prostatectomy, SVI= seminal vesicle invasion, HR= hazard ratio, CI= confidence interval.

	Univariate analysis HR (95% CI)	p-value	Multivariable analysis HR (95% CI)	p-value
Age at RP	1.09 (0.97-1.23)	0.130	0.96 (0.84-1.11)	0.595
Pathologic stage				
T2	baseline		baseline	
T3a/b	0.89 (0.18-4.34)	0.887	0.37 (0.07-1.99)	0.248
T3c	0.00 (0.00- .)	0.987	0.00 (0.00- .)	0.984
T4	12.97 (4.35-38.67)	0.000	3.82 (0.81-18.07)	0.091
SVI				
No	baseline		baseline	
Yes	6.39 (1.43-28.49)	0.015	4.20 (0.51-34.80)	0.183
Gleason score				
<=6	baseline		baseline	
7 (3+4)	1.33 (0.30-5.97)	0.712	0.81 (0.16-4.16)	0.799
7 (4+3)	13.98 (3.90-50.06)	0.000	8.55 (1.87-39.14)	0.006
>=8	8.90 (1.62-49.00)	0.012	8.16 (1.28-52.01)	0.026
Surgical margins				
Negative	baseline		baseline	
Positive	3.26 (1.18-9.00)	0.023	4.13 (1.25-13.59)	0.020
Tumour volume	1.40 (1.20-1.63)	0.000	1.06 (0.85-1.33)	0.609

Table 6 | Hazard ratios (HR) resulting from univariate and multivariable Cox proportional hazards analyses of tumour volume in cases with high-risk PC (defined as any of the following: pT3 or pT4, Gleason >=8, or positive surgical margins) per endpoint variable. The HR's resulting from the multivariable analyses are corrected for age, T-stage, presence of seminal vesicle invasion, Gleason score, and surgical margin status.

Tumour volume	Univariate analysis HR (95% CI)	p-value	Multivariable analysis HR (95% CI)	p-value
PSA recurrence	1.16 (1.01-1.33)	0.032	1.06 (0.90-1.24)	0.496
Local recurrence	1.31 (1.09-1.59)	0.004	1.20 (0.91-1.59)	0.206
Distant metastasis	1.14 (0.89-1.46)	0.286	0.90 (0.60-1.35)	0.617
Local recurrence and/or distant metastasis	1.31 (1.11-1.56)	0.002	1.18 (0.94-1.47)	0.150
PC related mortality	1.21 (0.93-1.59)	0.165	0.97 (0.59-1.59)	0.907

8.4 | Discussion

Tumour volume is an important prognostic factor for prostate cancer progression and mortality. However, in this study tumour volume did not add predictive value to routinely assessed pathological parameters, like tumour-stage, Gleason score, seminal vesicle invasion and surgical margin status.

A possible explanation for the discrepant results in literature concerning the independent predictive value of tumour volume is the cohort selection. For example, the study by Stamey et al reports on a pre-PSA era cohort with Gleason pattern 4/5 in 85%, seminal vesicle invasion in 18% and a mean tumour volume of 4.66 ml. In the current study, those rates were 37.5%, 1.7% and the mean (median) tumour volume was 1.05 (0.66) ml. However, even in our subset analysis of patients with high-risk PC, tumour volume did also not show independent prognostic value in multivariable analyses. Additionally, we excluded all insignificant tumours (organ-confined without Gleason pattern 4/5 and tumour volume < 0.5ml, n=98, 28.5%) and repeated all analyses (data not shown). In this subgroup of significant PC (n=246, 71.5%), no independent predictive value of tumour volume was found for any of the outcome parameters.

Furthermore none of the different methods of tumour volume assessment, i.e. continuous tumour volume, relative tumour volume or tumour volume \geq 0.5ml, resulted in independent prognostic value in this study. Tumour volume was also assessed on a log-transformed scale (data not shown), which produced similar results. In conclusion, none of the tumour volume based variables showed additional value to the routinely used parameters of tumour stage, Gleason score and surgical margin status in the RP specimen.

This lack of prognostic significance of tumour volume also held true for pT2 PC, in which surgical margin status was the only parameter with independent prognostic value for PSA recurrence. Tumour substage was not prognostic either (data not shown). This is in line with previous reports demonstrating the lack of prognostic significance of pT2 subclassification.^{13,14}

Although tumour volume did not add prognostic value to other pathological parameters, this does not necessarily mean that tumour volume has no value at all. Since tumour volume correlates well with pathological tumour stage and RP Gleason score, this may be of help in pretreatment planning. In this situation, only clinical and biopsy specimen parameters are available for decision making and pathological tumour stage and RP Gleason score are unknown. Magnetic resonance imaging (MRI) is showing promising results in measurement of prostate tumour volume,^{20,21} although it is not yet optimal.^{22,23} If these MRI-based tumour volume measurements would correlate to the pathological characteristics at RP specimens to the same extent as the planimetric tumour volume measurements reported in this study, these tumour volumes may be of help in the pretreatment risk assessment and subsequent treatment choice. In addition, it is important for the surgeon to receive some feedback information on the size of the tumour. Furthermore, we have previously shown that the estimated proportion of high-grade cancer has independent prognostic value, contrary to overall tumour volume.²⁴

However, a detailed calculation of tumour volume does not seem necessary and an estimation of the percentage of the prostate volume involved by carcinoma or a maximum tumour diameter will be sufficient.

Some strong points of this study are worth mentioning. Firstly, tumour volume was assessed in a very accurate manner, by computer-assisted morphometric analysis. This method is laborious and therefore not a valid option for routine use. Several surrogate parameters have been described with a good correlation with tumour volume, for example visual estimation, positive-block ratio, point-count or grid method or measuring maximal tumour diameter.²⁵⁻³² Secondly, follow-up time was relatively long, with a mean (median) time of follow-up of 96.2 (98.0) months and clinical outcome data on local recurrence, metastasis and mortality were available. This is a major advantage compared to other studies with much shorter follow-up (Table 1).

Simultaneously, some limitations must be noted. No data were available on multifocality and tumour volume of the index tumour. It seems reasonable to assume that one tumour focus of 1.0 ml shows a different behavior than 2 foci of 0.5 ml. This may also be important in pT2 substaging: a large unilateral tumour focus can be assigned a lower pathological stage than two small bilateral foci in the current staging system. Data of volume of the separate tumour foci may be helpful in improving the pT2 substaging.

A second limitation is the small number of events of local recurrence, metastasis and PC related mortality. This may cause statistical difficulty in reaching the significance level and over fitting. However, for PSA recurrence, which reached an occurrence rate of 16.6% (57 cases), no statistical significance could be reached either. Finally, it should be noted that all PC cases in the current cohort were detected by screening and were thus subject to a lead-time. Therefore, these cases may not be totally comparable to clinically detected cases. On the other hand, their pathological features may resemble more closely the cancers detected by PSA testing, a widespread practice in most western countries. For this reason we consider our findings of relevance for the ongoing discussion on the inclusion of a tumour volume parameter in standard pathology reporting of radical prostatectomy specimens.

8.5 | Conclusions

In this study tumour volume did not add prognostic value to routinely assessed pathological parameters, like tumour-stage, Gleason score, seminal vesicle invasion and surgical margin status. Therefore, there seems to be little reason to include tumour volume in RP specimens routinely in the pathology report.

References

1. Stamey TA, McNeal JE, Yemoto CM, Sigal BM, Johnstone IM. Biological determinants of cancer progression in men with prostate cancer. *JAMA* 1999;215:1395-1400
2. Kikuchi E, Scardino PT, Wheeler TM, Slawin KM, Ohori M. Is tumour volume an independent prognostic factor in clinically localized prostate cancer? *J Urol* 2004;172:508-511
3. Vollmer RT. Percentage of tumour in prostatectomy specimens. *Am J Clin Pathol* 2009;131:86-91
4. McNeal JE, Bostwick DG, Kindrachuk RA, Redwine EA, Freiha FS, Stamey TA. Patterns of progression in prostate cancer. *Lancet* 1986;1:60-3
5. Epstein JI, Carmichael M, Partin AW, Walsh PC. Is tumour volume an independent predictor of progression following radical prostatectomy? A multivariate analysis of 185 clinical stage B adenocarcinomas of the prostate with 5 years of follow-up. *J Urol* 1993;149:1478-81
6. Salomon L, Levrel O, Anastasiadis AG, Irani J, De la Taille A, Vordos D, Cicco A, Hoznek A, Chopin D, Abbou C-C. Prognostic significance of tumour volume after radical prostatectomy : a multivariate analysis of pathological prognostic factors. *Eur Urol* 2003;43:39-44
7. Ramos CG, Roehl KA, Antenor JAV, Humphrey PA, Catalona WJ. Percent carcinoma in prostatectomy specimens is associated with risk of recurrence after radical prostatectomy in patients with pathologically organ confined prostate cancer. *J Urol* 2004;172:137-140
8. Cheng WS, Frydenberg M, Bergstralh EJ, Lrson-Keller JJ, Zincke H. Radical prostatectomy for pathologic stage C prostate cancer: influence of pathologic variables and adjuvant treatment on disease outcome. *Urology* 1993;43:283-291
9. Merrill MM, Lane BR, Reuther AM, Zhou M, Magi-Galluzzi C, Klein EA. Tumour volume does not predict for biochemical recurrence after radical prostatectomy in patients with surgical Gleason score 6 or less prostate cancer. *Urology* 2007;70:294-298
10. Epstein JI, Oesterling JE, Walsh PC. Tumour volume versus percentage of specimen involved by tumour correlated with progression in stage A prostatic carcinoma. *J Urol* 1988;139:980-4
11. Partin AW, Epstein JI, Cho KR, Gittelsohn AM, Walsh PC. Morphometric measurement of tumour volume and percent of gland involvement as predictors of pathological stage in clinical stage B prostate cancer. *J Urol* 1989;141:341-5
12. International Union against Cancer (UICC): TNM classification of malignant tumours, 6th edition, Sobin LH, Wittekind Ch (eds). New York: Wiley-Liss, 2002, pp.184-7
13. Freedland SJ, Partin AW, Epstein JI, Walsh PC. Biochemical failure after radical prostatectomy in men with pathologic organ-confined disease: pT2a versus pT2b. *Cancer* 2004;100:1646-9
14. Hong SK, Han BK, Chung JS, Park DS, Jeong SJ, Byun SS, Choe G, Lee SE. Evaluation of pT2 subdivisions in the TNM staging system for prostate cancer. *BJU Int* 2008;102:1092-6
15. Hoedemaeker RF, Ruijter ETG, Ruizeveld-de Winter JA, van der Kaa CA and van der Kwast ThH. The BIOMED II MPC study group. Processing radical prostatectomy specimens. A comprehensive and standardized protocol. *J Urol Pathol* 1998;9:211-22
16. Hoedemaeker RF, Rietbergen JBW, Kranse R, Van der Kwast ThH, Schröder FH. Comparison of pathological characteristics of T1c and non-T1c cancers detected in a population-based screening study, the European Randomized Study of Screening for Prostate Cancer. *World J Urol* 1997;15:339-345
17. Freedland SJ, Sutter ME, Dorey F, Aronson WJ. Defining the ideal cutpoint for determining PSA recurrence after radical prostatectomy. *Urology* 2003;61:365-9
18. De Koning HJ, Blom J, Merkelbach JW, Raaijmakers R, Verhaegen H, Van vliet P, Nelen V, Coebergh JWW, Hermans A, Ciatto S, Mäkinen T. Determining the cause of death in randomized screening trial(s) for prostate cancer. *BJU Int* 2003;92 (suppl2) :71-78
19. Stamey TA, Freiha FS, McNeal JE, Redwine EA, Whittemore AS, Schmid HP. Localized prostate cancer. Relationship of tumour volume to clinical significance for treatment of prostate cancer. *Cancer* 1993;71:933-8

20. Villers A, Puech P, Mouton D, Leroy X, Ballereau C, Lemaitre L. Dynamic contrast enhanced, pelvic phased array magnetic resonance imaging of localized prostate cancer for predicting tumour volume: correlation with radical prostatectomy findings. *J Urol* 2006;176:2432-7
21. Lemaitre L, Puech P, Poncelet E, Bouyé S, Leroy X, Biserte J, Villers A. Dynamic contrast-enhanced MRI of anterior prostate cancer : Morphometric assessment and correlation with radical prostatectomy findings. *Eur Radiol* 2009;19:470-480
22. Kirkham APS, Emberton M, Allen C. How good is MRI at detecting and characterizing cancer within the prostate? *Eur Urol* 2006;50:1163-75
23. Coakley FV, Kurhanewicz J, Lu Y, Jones KD, Swanson MG, Chang SD, Carroll PR, Hricak H. Prostate cancer tumour volume: measurement with endorectal MR and MR spectroscopic imaging. *Radiology* 2002;223:91-7
24. Vis AN, Roemeling S, Kranse R, Schröder FH, van der Kwast TH. Should we replace the Gleason score with the amount of high-grade prostate cancer? *Eur Urol* 2007;51:931-9
25. Chen ME, Johnston D, Reyes AO, Soto CP, Babaian RJ, Troncso P. A streamlined three-dimensional volume estimation method accurately classifies prostate tumours by volume. *Am J Surg Pathol* 2003;27:1291-301
26. Dvorak T, Chen M-H, Renshaw AA, Loffredo M, Richie JP, d'Amico AV. Maximal tumour diameter and the risk of PSA failure in en with specimen-confined prostate cancer. *J Urol* 2005;66:1024-8
27. Eichelberger LE, Koch MO, Eble JN, Ulbright TM, Juliar BE, Cheng L. Maximum tumour diameter is an independent predictor of prostate specific antigen recurrence in prostate cancer. *Mod Pathol* 2005;18:886-90
28. Marks RA, Lin H, Koch MO, Cheng. Positive-block ratio in radical prostatectomy specimens is an independent predictor of prostate-specific antigen recurrence. *Am J Surg Pathol* 2007;31:877-81
29. Billis A, Magna LA, Ferreira U. Correlation between tumour extent in radical prostatectomies and preoperative PSA, histological grade, surgical margins and extraprostatic extension; application of a new practical method for tumour extent evaluation. *Int Braz J Urol* 2003;29:113-20
30. Carvalhal GF, Humphrey PA, Thorson P, Yan Y, Ramos CG, Catalona WJ. Visual estimate of the percentage of carcinoma is an independent predictor of prostate carcinoma recurrence after radical prostatectomy. *Cancer* 2000;89:1308-14
31. Van Oort IM, Witjes JA, Kok DEG, Kiemeney LALM, Hulsbergen-vandeKaa. Maximum tumour diameter is not an independent prognostic factor in high-risk localized prostate cancer. *World J Urol* 2008;26:237-241
32. Scattoni V, Montironi R, Mazzucchelli R, Freschi M, Nava L, Losa A, Terrone C, Scarpa RM, Montorsi F, Pappagallo G, Rigatti P. Pathological changes if high-grade prostatic intraepithelial neoplasia ad prostate cancer after monotherapy with bicalutamide 150 mg. *BJU Int* 2006;98:54-8



CHAPTER 9

Reassessment of the tumour volume threshold defining clinically insignificant prostate cancer; results from a randomized trial

Submitted

Tineke Wolters, Monique J. Roobol, Pim J. van Leeuwen, Roderick C.N. van den Bergh, Robert F. Hoedemaeker, Geert J.L.H. van Leenders, Fritz H. Schröder and Theodorus H. van der Kwast

Abstract

Background: The identification of clinically insignificant prostate cancer (PC) could help avoid overtreatment. Currently, the most frequently used criteria for insignificant PC employ a tumour volume (TV) threshold of <0.5 ml.

Objectives: To reassess the TV threshold for clinically insignificant PC using data of the prevalence screening round of a randomized trial and to critically assess its prognostic value.

Design, setting and participants: The rate of insignificant PC was calculated by modeling life time risk estimates of PC diagnosis in screened and non-screened participants of a randomized PC screening trial.

Measurements: Using life time risk estimates, 50.8% of screen-detected PC were calculated to be clinically insignificant and the 49.2% largest TV of 325 prostatectomy specimens were used to determine the threshold TV for clinically insignificant prostate cancer. The predictive value of various TV thresholds for biochemical progression, in addition to stage and grade, were analyzed using Cox proportional hazards analyses.

Results and limitations: The TV threshold for clinically insignificant PC was found to be at least 0.70 ml in our screened population of men aged 55-75 years. Although patients with a TV>0.70ml were at increased risk of biochemical progression (HR 2.89, p<0.001), in the selection of men with possible insignificant PC, i.e. organ-confined PC without Gleason pattern 4/5, TV was not predictive for biochemical recurrence. This study was limited firstly by a more favourable pre-treatment risk profile of the study cohort compared to the entire population, resulting in an underestimation of the TV threshold. Secondly, all patients were surgically treated and true significance of the PC cannot be established.

Conclusions: Clinically insignificant prostate cancers likely includes organ confined PC without Gleason pattern 4/5 with a volume up to at least 0.70 ml. The lack of prognostic impact of TV in addition to stage and grade raises doubt about the inclusion of this parameter as a criterion for insignificant prostate cancer.

9.1 | Introduction

Population-based screening for prostate cancer (PC), using a prostate-specific antigen (PSA) test, has led to a dramatic increase in incidence of insignificant PC.¹ Insignificant PC is PC diagnosed in the absence of symptoms, which would not have caused disease-specific morbidity or mortality during the patient's life. Given the sharp increase in PC incidence and the harmful effects of overtreatment, the identification of insignificant cancers is now a high priority in PC screening practice.²

Currently, the most frequently used criteria for insignificant PC include organ-confined PC, absence of Gleason pattern 4/5 and a maximum tumour volume (TV) of 0.5 ml in radical prostatectomy (RP) specimens.³ It is well established that significantly worse patient outcome is observed when a PC displays extracapsular extension or high grade PC, independent of other prognostic factors.^{4,5} The <0.5 ml TV threshold is based on incidentally detected PC in a radical cystoprostatectomy series, published by Stamey et al⁶ based on a 8% life time risk to be diagnosed with clinically significant PC. The largest tumour foci in his cystoprostatectomy series were considered as the most aggressive and these largest foci all had tumour volumes of 0.5 ml or higher. Whether the same threshold would apply to a screen-detected prostate cancer patient cohort is unclear. The assumption that TV is an important prognosticator for PC is questionable, especially since the largest tumour foci are not necessarily the most significant tumours. Possibly, larger tumours could be considered clinically insignificant as well⁷. To our knowledge, no other study on an independent population has been performed to confirm the 0.5 ml TV threshold.

In this study, we applied at first the approach used by Stamey et al to our cohort of participants of a randomized screening trial (European Randomized Screening study of Prostate Cancer – Section Rotterdam) to estimate the TV threshold for clinically insignificant PC. Secondly, we made an attempt to evaluate the validity of the incorporation of a TV threshold in the definition of insignificant PC. This was done by assessing the prognostic value for biochemical recurrence after RP in patients with organ-confined prostate cancer without Gleason pattern 4/5 of the conventional 0.5 ml and the newly found TV threshold.

9.2 | Methods

9.2.1 | Study population

A total of 1014 patients were diagnosed with PC during the first screening round of the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC). According to the screening protocol, a lateralized sextant prostate biopsy was induced by a prostate-specific antigen (PSA) serum level ≥ 4.0 ng/ml, an abnormal digital rectal examination or a abnormal transrectal ultrasound (TRUS) up to May 1997. From May 1997, only a PSA ≥ 3.0 ng/ml prompted a biopsy.

Of these 1014 patients, 400 (39.5%) were primarily treated with RP. For 75 (18.8%) patients the data on pathological tumour stage, Gleason score and prostate tumour volume were incomplete and these patients were excluded from the analyses. Therefore, the study cohort consisted of 325 (81.3%) patients who were all treated between November 1993 and June 2000.

By assessing only patients treated with RP, a selection of the total screened population is made. To examine the validity of the use of cancer characteristics in the selection of patients treated with RP the proportion of low, intermediate and high risk PC according to the d'Amico criteria⁸ of all cancers of the screening arm was compared to those treated by RP.

9.2.2 | Radical prostatectomy specimens and pathologic examination

RP specimens were processed in toto following the protocol described before.⁹ In short, after fixing the specimens, they were inked, cut at 4mm intervals perpendicular to the rectal surface. The apical slice is cut parasagittally at 4mm intervals. The sections were divided in halves or quadrants to fit routine used cassettes for paraffin embedding. Pathological tumour stage, Gleason score, and surgical margin status were assessed. Tumour areas were marked at each slide and measured using a computerized morphometry.¹⁰ Subsequently the volumes were calculated by multiplying the area by the slice thickness. In a side-study we showed (personal communication RFH) that fixation and processing of the RP specimens did not result in significant reduction of the tissue volume. Therefore, a shrinkage factor was not taken into account for the calculation of TV. Staging was done according to the 2002 TNM classification. The pT4 cases showed tumour invasion of the detrusor muscle.

9.2.3 | TV thresholds

Stamey assessed a TV threshold for clinically insignificant PC using the life time risk of 8% for a man to be clinically diagnosed with prostate cancer and observed a threshold of 0.5 ml.⁶ We employed essentially the same approach to assess a TV threshold for insignificant PC using predicted life time risks obtained in the micro simulation model MISCAN, based on data of the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC).¹¹ The predicted life time risk of prostate cancer in a screening situation similar to our screening protocol (i.e. every 4 years, men aged 55-75 years) was 130 per 1000 men (13.0%), compared to a predicted life time risk of 64 per 1000 men (6.4%) of a clinically detected prostate cancer. Thus, of all screen-detected cancers only 49.2% (64 of 130) would have been clinically detected. The new TV threshold was established by considering the 49.2% largest TV in our cohort of screen-detected cancers clinically relevant. The predictive value for biochemical recurrence of the TV threshold was assessed in the entire study population and in a selection of patients with possibly insignificant PC, i.e. patients with organ-confined disease without Gleason pattern 4/5.

9.2.4 | Follow-up

Patients were followed after RP by serial PSA measurements every 3 months during the first year after RP, every 6 months in the second year and yearly after 2-3 years. Biochemical recurrence was defined as a PSA > 0.2 ng/ml.¹² Metastasis was indicated by suspicious bone scintigraphy, MRI or CT scans. Total follow-up time was defined as the time from RP to death or last visit date. Time to recurrence was defined as the time of RP to the time of the first signs of recurrence. When no signs of recurrence were registered, cases were censored at the time of the last follow-up visit or date of death. The cause of death was established by an independent committee¹³ and subsequently scored as PC-related or intercurrent cause of death.

9.2.5 | Statistical analysis

Proportions were compared with a Fisher's exact test. Univariate Cox proportional hazards regression model analyses were performed to evaluate the predictive value for biochemical recurrence of the traditional TV threshold, the new TV threshold and TV as a continuous variable in addition to stage and grade related criteria (i.e. organ-confined and no Gleason pattern 4/5) for insignificant PC. The data were analyzed using the statistical package for social sciences, version 15.0 (SPSS 15.0, Chicago, IL). Statistical significance was defined as a two-sided p-value < 0.05.

9.3 | Results

9.3.1 | Overall characteristics of the cohort

The distribution of low, intermediate and high risk cancer in the screening arm and the 325 cases treated with RP was significantly different: the surgically treated cases were skewed to a lower risk profile (Table 1). The overall characteristics of the 325 cases who underwent RP are listed in Table 2. TV ranged from 0.001 ml to 13.48 ml. Among these 325 men, 128 (39.4%) had a TV < 0.5 ml. Using the rate of 49.2% for clinically significant disease, a new TV threshold of at least 0.70 ml was found. A total of 160 (50.8%) patients had a TV < 0.70 ml.

Follow-up data were available for 322 (99.1%) patients and the median (mean) period of follow up after the prostate cancer diagnosis was 98 (96.1) months. A total of 54 (16.8%) patients showed PSA progression after a median (mean) period of 35 (43.9) months.

A total of 174 (53.5%) patients had organ-confined disease without Gleason pattern 4 or 5 (table 3) and 16 (9.2%) patients showed biochemical progression. Cox proportional hazards models showed that patients with organ-confined disease without Gleason pattern 4/5 were at significantly lower risk of biochemical progression (HR 0.35, 95% CI 0.20-0.63, p < 0.001) than patients with extra capsular tumour growth and/or Gleason pattern 4/5. The numbers of metastasis and PC specific mortality were too small to perform Cox regression analyses: 1 patient presented with metastasis during follow-up and no PC specific mortality was observed in men with organ-confined disease without Gleason pattern 4/5.

Table 1 | Comparison of the distribution of low, intermediate and high risk prostate cancer (PC) in the selected 325 cases treated with radical prostatectomy and the unselected men of the screening arm (n=1014). Low risk was defined as clinical tumour (cT) stage =< T2a and PSA=< 10 ng/ml and Gleason score = 6. High risk was defined as cT stage T2c or higher or a PSA-level >20 ng/ml or Gleason score >=8. The remaining cases (cT stage T2b or a Gleason score of 7 or a PSA level >10 ng/ml and <= 20 ng/ml) were defined as intermediate risk.

	All PC cases n=1014	RP selection n=325	Not included in the selection n=689	p-value
Low risk	455 (44.9%)	187 (57.5%)	268 (38.9%)	<0.001
Intermediate risk	223 (22.0%)	73 (22.5%)	150 (21.8%)	
High risk	336 (33.1%)	65 (20.0%)	271 (39.3%)	
Total	1014 (100%)	325 (100%)	689 (100%)	

Table 2 | Characteristics of the total cohort of 325 patients treated with RP and of the 174 cases with organ confined disease without Gleason pattern 4/5. PSA = prostate-specific antigen, pT=pathological tumour stage.

Features	Screen detected PC: total cohort No. (%) of cases	Organ confined prostate cancer without Gleason pattern 4/5 No. (%) of cases
No of cases	325	174
Age in years, mean (range)	63.6 (55.2-75.5)	63.2 (55.2-73.1)
PSA (ng/ml), mean (range)	6.2 (0.8-43.0)	5.5 (0.8-37.0)
PSA density (ng/ml/ml), mean (range)	0.17 (0.02-1.01)	0.14 (0.02-0.93)
Focality		
Monofocal	91 (28.0)	45 (25.9)
Multifocal	222 (68.3)	121 (69.5)
unknown	12 (3.7)	8 (4.6)
Gleason score		
≤ 6	202 (62.2)	174 (100)
7 (3+4)	96 (29.5)	
7 (4+3)	16 (4.9)	
8 to 10	11 (3.4)	

Table 2 | *Continued*

Features	Screen detected PC: total cohort No. (%) of cases	Organ confined prostate cancer without Gleason pattern 4/5 No. (%) of cases
pT (TNM system 2002)		
pT2a	63 (19.4)	47 (27.0)
pT2b	7 (2.2)	3 (1.7)
pT2c	180 (55.4)	124 (71.3)
pT3a	55 (16.9)	
pT3b	4 (1.2)	
pT4	16 (4.9)	
Surgical margin status		
Negative	243 (74.8)	148 (85.1)
Positive	81 (24.9)	25 (14.4)
unknown	1 (0.3)	1 (0.6)
Tumour volume (cc)		
Mean (median)	1.07 (0.66)	0.76 (0.47)
Range	0.001-13.48	0.001-6.34

Table 3 | Progression and mortality rates for the total study cohort and for the selection of patients with organ-confined without Gleason pattern 4/5. PC=prostate cancer

	Total cohort N=322	Organ-confined PC and no Gleason pattern 4/5 N=172
Biochemical progression	54 (16.8%)	16 (9.3%)
Metastasis	8 (2.5%)	1 (0.6%)
Overall mortality	54 (16.8%)	29 (16.9%)
PC specific mortality	5 (1.6%)	0 (0%)

9.3.2 | Tumour volume as a criterion of clinically insignificant PC

The 0.7 ml TV threshold was a significant predictor for biochemical progression in the entire study population: HR 2.88, 95% CI 1.59-5.24, $p < 0.001$. In the selection of 174 patients with organ-confined prostate cancer and no Gleason pattern 4/5, TV ranged from 0.001 ml to 6.34 ml, 92 (52.9%) had a TV < 0.5ml and 111 (63.8%) had a TV < 0.70ml. None of the thresholds showed prognostic value in any of the analyses (Table 4).

Table 4 | The prognostic value of a tumour volume threshold: results of Cox proportional hazards analyses in patients with organ-confined PC and without Gleason pattern 4/5. PC=prostate cancer, TV=tumour volume

	Number of patients with organ confined PC without Gleason pattern 4/5	Biochemical progression	Hazard ratio (95% confidence interval)	p-value
TV<0.5ml	91	9 (9.9%)	1 (reference)	0.818
TV>=0.5 ml	81	7 (8.6%)	0.89 (0.33-2.39)	
TV<0.7ml	110	11 (10.0%)	1 (reference)	0.719
TV>=0.7ml	62	5 (8.1%)	0.82 (0.29-2.37)	
TV (continuous variable)	172	16 (9.3%)	0.55 (0.22-1.37)	0.198

9.4 | Discussion

In this first study attempting to confirm the previously reported TV threshold of 0.5 ml for clinically insignificant PC, we found a TV threshold for insignificant PC of 0.70 ml. The threshold was assessed using life time risks of prostate cancer diagnosis that were specifically modelled for the studied population. This method was similar to the design of Stamey⁶ who observed a lower threshold of 0.5 ml, the threshold which is nowadays widely used as a criterion in the definition of insignificant PC. The difference in thresholds may very well be explained by the different study cohorts: screen-detected PC in patients, aged 55-75 years with a PSA>=3.0 ng/ml in our cohort compared to patients who had a cystoprostatectomy for bladder carcinoma without prior history of prostate cancer. Our results suggest that tumours larger than 0.5 ml could be insignificant as well, with volumes at least up to 0.70 ml.

It is very likely that our threshold of 0.70 ml as determined by the approach taken by Stamey⁶ is an underestimation. First, it should be noted that patients of our cohort treated with RP are a selection of all patients with prostate cancer detected in the screening program. A different distribution of low and high risk PC was observed in the selection of patients treated by RP compared to the total cohort of patients diagnosed with PC during the first screening round. Since they represented a selection of more favorable tumours as compared to the entire series of screen detected cancers, the TV threshold of 0.70 ml may be too restrictive and much on the safe side.

Secondly, it is commonly accepted that PC with extracapsular extension and/or Gleason scores > 6 should be considered clinically significant. Most studies agree that Gleason score and pathological stage are stronger determinants of biological behaviour of PC than TV. As a consequence, one could argue that the threshold for insignificant PC should be extracted from the selection of patients with organ-confined PC and no Gleason pattern 4/5 only. IN the

present population, a TV threshold of 2.50 ml was observed after repeating the analysis in this selection of patients.

This view is corroborated by our finding that TV was not useful for the differentiation between men with or without biochemical progression, in patients with organ confined disease without Gleason pattern 4/5. Patients with a low TV were at a similar risk of biochemical progression as patients with larger TV. This is not surprising: although TV is related to Gleason score and tumour stage,¹⁴⁻¹⁶ several studies, including one on the same set of patients used for this report, demonstrated a lack of independent prognostic value for biochemical and clinical progression.¹⁷⁻²¹ This suggests that a TV threshold might not be effective for the identification of insignificant PC. The use of a TV threshold might unnecessarily exclude patients with higher TV from active surveillance strategies. On the other hand, patients with small TV might mistakenly be reassured.

Overall, the selection of organ-confined PC with no Gleason pattern 4/5 seems to represent the most important step in the identification of patients with favourable prognosis after RP. This was illustrated by the fact that only 1 patient showed metastasis during follow-up and none of the patients died of prostate cancer after selecting for organ-confined disease without Gleason pattern 4/5. This selection seems therefore very useful for the identification of insignificant PC. Which factors will be valuable in further differentiation between insignificant and clinically relevant PC in this low-risk group remains unknown, but TV did not seem valuable in this differentiation.

Some limitations are worth mentioning. Since all patients in the current study were treated with RP, we cannot possibly know the natural course of the disease in this patient series. Although biochemical progression after RP would suggest that the tumour probably is not insignificant, a rising PSA does not necessarily indicate its clinical significance as it may occur in the absence of any symptoms.^{23,24} Additionally, in absence of any progression after treatment, the true significance of the tumour is unknown. The lack of prognostic value of TV in patients with organ-confined PC without Gleason pattern 4/5 and treated with RP might also be interpreted as RP being a very effective therapy in this selection of patients regardless of TV, as stated recently by Lee et al.²⁵

A second limitation is that in the present study cohort total TV was assessed, while the report by Stamey was based on the index TV. For a selection of 100 consecutive men in the present study, index TV was known and repeating the analyses showed a TV threshold of 0.55 ml (data not shown). Index TV was available for 37 men in the selection of organ-confined PC with no Gleason pattern 4/5 and ranged up to 4.3 ml. Although numbers are too small to draw valid conclusions, no prognostic value of TV was observed in this selection either. Finally, the same arguments as mentioned above for the 0.70 ml threshold apply to consider this 0.55 ml threshold an underestimation.

9.5 | Conclusions

Using life time risk rates for our screened cohort, a minimum total tumour volume threshold for insignificant PC of 0.70 ml was observed. This suggests that insignificant PC includes tumours larger than 0.5 ml, especially since the threshold of 0.70 ml is an underestimation, because it was based on a study cohort with a more favourable risk profile than the overall screening cohort. Additionally, tumour volume could not predict adverse outcome after prostatectomy in men with organ confined prostate cancer without Gleason pattern 4/5. Therefore, we question the validity of the incorporation of TV as a criterion for insignificant prostate cancer.

References

1. Gosselaar C, Roobol MJ, Schröder FH. Prevalence and characteristics of screen-detected prostate carcinomas at low prostate-specific antigen levels: aggressive or insignificant? *BJU Int.* 2005;95:231-7
2. Bangma CH, Roemeling R, Schröder FH. Overdiagnosis and overtreatment of early detected prostate cancer. *World J Urol* 2007;25:3-9
3. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumour extent of nonpalpable (stage T1c) prostate cancer. *JAMA* 1994;271:368-74
4. Han M, Partin AW, Pound CR, Epstein JI, Walsh PC. Long-term biochemical disease-free and cancer-specific survival following anatomic radical retropubic prostatectomy. The 15-year Johns Hopkins experience. *Urol Clin North Am.* 2001;28:555-65
5. Kupelian P, Katcher J, Levin H, Zippe C, Klein E. Correlation of clinical and pathological factors with rising prostate-specific antigen profiles after radical prostatectomy alone for clinically localized prostate cancer. *Urology* 1996;48:249-60
6. Stamey TA, Freiha FS, McNeal JE, Redwine EA, Whittemore AS, Schmid HP. Localized prostate cancer. Relationship of tumour volume to clinical significance for treatment of prostate cancer. *Cancer* 1993;71(3 Suppl):933-8
7. Stamey TA. Editorial: More information on prostate specific antigen and prostate cancer. *J Urol* 2003;170:457-8
8. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, Tomaszewski JE, Renshaw AA, Kaplan I, Beard CJ, Wein A. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998;280:969-974
9. Hoedemaeker RF, Ruijter ETG, Ruizeveld-de Winter JA, Van der Kaa CA and Van der Kwast ThH. The BIOMED II MPC study group. Processing radical prostatectomy specimens. A comprehensive and standardized protocol. *J Urol Pathol* 1998;9:211-22
10. Hoedemaeker RF, Rietbergen JBW, Kranse R, Van der Kwast ThH, Schröder FH. Comparison of pathological characteristics of T1c and non-T1c cancers detected in a population-based screening study, the European Randomized Study of Screening for Prostate Cancer. *World J Urol* 1997;15:339-345
11. Draisma G, Boer R, Otto SJ, van der Crujisen IW, Damhuis RA, Schröder FH, de Koning HJ. Lead times and over-detection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003;95:868-78
12. Freedland SJ, Sutter ME, Dorey F, Aronson WJ. Defining the ideal cutpoint for determining PSA recurrence after radical prostatectomy. *Urology* 2003;61:365-9
13. De Koning HJ, Blom J, Merkelbach JW, Raaijmakers R, Verhaegen H, Van vliet P, Nelen V, Coebergh JWW, Hermans A, Ciatto S, Mäkinen T. Determining the cause of death in randomized screening trial(s) for prostate cancer. *BJU Int* 2003;92 (suppl2) :71-78
14. Stamey TA, McNeal JE, Yemoto CM, Sigal BM, Johnstone IM. Biological determinants of cancer progression in men with prostate cancer. *JAMA* 1999;281:1395-1400
15. McNeal JE, Bostwick DG, Kindrachuk RA, Redwine EA, Freiha FS, Stamey TA. Patterns of progression in prostate cancer. *Lancet* 1986;1:60-3
16. Vollmer RT. Percentage of tumour in prostatectomy specimens. *Am J Clin Pathol* 2009;131:86-91
17. Sengupta S, Blute ML, Bagniewski SM, Inman B, Leibovich BC, Slezak JM, Myers RP, Zincke H. After radical retropubic prostatectomy 'insignificant' prostate cancer has a risk of progression similar to low-risk 'significant' cancer. *BJU Int.* 2008 Jan;101(2):170-4.
18. Merrill MM, Lane BR, Reuther AM, Zhou M, Magi-Galluzzi C, Klein EA. Tumour volume does not predict for biochemical recurrence after radical prostatectomy in patients with surgical Gleason score 6 or less prostate cancer. *Urology.* 2007 Aug;70(2):294-8.
19. Kikuchi E, Scardino PT, Wheeler TM, Slawin KM, Ohori M. Is tumour volume an independent prognostic factor in clinically localized prostate cancer? *J Urol.* 2004 Aug;172(2):508-11.

20. Porten SP, Cooperberg MR, Carroll PR. The independent value of tumour volume in a contemporary cohort of men treated with radical prostatectomy for clinically localized disease. *BJU Int.* 2009, doi: 10.1111/j.1464-410X.2009.08774.x
21. Wolters T, Roobol MJ, van Leeuwen P, van den Bergh RCN, Hoedemaeker RF, van Leenders GJLH, Schröder FH and van der Kwast ThH. Should prostate tumour volume routinely be reported by the pathologist?: The prognostic value of tumour volume in prostate cancer. *Eur Urol* 2009, doi:10.1016/j.eururo.2009.07.027
22. Han M, Partin AW, Zahurak M, Piantadosi S, Epstein JI, Walsh PC. Biochemical (prostate specific antigen) recurrence probability following radical prostatectomy for clinically localized prostate cancer. *J Urol.* 2003 Feb;169(2):517-23.
23. Roehl KA, Han M, Ramos CG, Antenor JA, Catalona WJ. Cancer progression and survival rates following anatomical radical retropubic prostatectomy in 3,478 consecutive patients: long-term results. *J Urol.* 2004;172:910-4.
24. Haukaas SA, Halvorsen OJ, Daehlin L, Hostmark J, Akslen LA. Is preoperative serum prostate-specific antigen level significantly related to clinical recurrence after radical retropubic prostatectomy for localized prostate cancer? *BJU Int.* 2006 Jan;97(1):51-5.
25. Lee MC, Dong F, Stephenson AJ, Jones JS, Magi-Galluzzi C and Klein EA. The Epstein criteria predict for organ-confined but not insignificant disease and a high likelihood of cure at radical prostatectomy. *Eur Urol* (2009), doi:10.1016/j.eururo.2009.10.025



CHAPTER 10

The value of EZH2, p27^{kip1}, BMI-1 and MIB-1 on biopsy specimens with low risk PC in selecting men with significant prostate cancer at prostatectomy

BJU Int 2009, epub ahead of print, doi:10.1111/j.1464-410x.2009.08998

Tineke Wolters, Kees J. Vissers, Chris H. Bangma, Fritz H. Schröder and Geert J.L.H. van Leenders

Summary

Objectives: To assess the additional prognostic value of EZH2, MIB-1, p27^{kip1} and BMI-1 on needle-biopsies from men with low-risk prostate cancer, as this disease in needle biopsies shows a heterogeneous clinical outcome, and while it is known that the expression of molecular tissue markers EZH2, BMI-1, MIB-1, and p27^{kip1} are predictive for clinical outcome after radical prostatectomy (RP), their value in prostate biopsies is largely unknown.

Patients and Methods: The study included men participating in a screening study, diagnosed with low-risk prostate cancer and subsequently treated with RP. Immunohistochemical staining for EZH2, MIB-1, p27^{kip1} and BMI-1 on the needle-biopsies were (semi)quantitatively scored and expression levels were related to significant disease at RP. Clinical low-risk prostate cancer was defined as prostate specific antigen (PSA) ≤ 10 ng/ml, clinical T-stage ≤ 2, biopsy Gleason score ≤ 6, a PSA density < 0.20 ng/ml/g and two or fewer positive cores. Significant PC at RP was defined as presence of any of extracapsular extension, Gleason pattern 4/5, or tumour volume ≥ 0.5 ml.

Results: In all, 86 biopsy specimens were included; there was high EZH2 expression (>1.0%) in 42% and a low p27^{kip1} expression (<90%) in 63%. Significant disease was present in 44 (51.2%) RP specimens. A high EZH2 (odds ratio 3.19, p=0.043) and a low p27^{kip1} (odds ratio 4.69, p=0.036) were independent predictors for significant prostate cancer at RP.

Conclusions: The determination of EZH2 and p27^{kip1} on diagnostic needle-biopsies supports the selection of men with indolent prostate cancer at RP. Especially p27^{kip1} could improve the pretreatment risk assessment of patients with low-risk prostate cancer.

10.1 | Introduction

The pretreatment risk assessment of prostate cancer (PC) is important to guide clinical decision making, and is based on clinical and pathological characteristics, e.g. PSA level, clinical tumour stage, number of positive biopsy cores and biopsy Gleason score. PC classified as low-risk disease has a high probability of being clinically insignificant, i.e. not causing any symptoms, or death. However, some cases classified as low-risk show more aggressive behaviour, urging radical treatment.¹ Unfortunately, the identification of patients at risk of harbouring significant PC is difficult.

To better assess the aggressiveness of prostate cancer before treatment, the use of molecular biomarkers on needle biopsy specimens might offer additional predictive value. Various immunohistochemical markers at radical prostatectomy (RP) had additional value in determining tumour aggressiveness and PSA relapse after RP. Several independent groups reported prognostic value for EZH2, BMI-1, p27^{kip1} and MIB-1 at radical prostatectomy specimens.²⁻¹⁴ While these markers have predictive value in this setting, their most important clinical value could be patient stratification for therapeutic decision-making.

Therefore, in the present study, we investigated the expression of EZH2, BMI-1, p27^{kip1} and MIB-1 in diagnostic needle-biopsies from men with low-risk PC, and the predictive value of these markers for significant disease in RP specimens.

10.2 | Patients and Methods

10.2.1 | Patient selection

All patients included in this study were participants in the screening arm of the European Randomized Study of Screening for Prostate Cancer (ERSPC), Rotterdam section. Men aged 55-74 years were invited for a screening visit every 4 years. Until May 1997, a PSA level of ≥ 4.0 ng/ml or an abnormal DRE and/or TRUS result prompted a lateralized sextant prostate biopsy. From May 1997, a biopsy was only indicated by a PSA level ≥ 3.0 ng/ml.

Patients were defined as having low-risk PC according to the criteria in the international active surveillance study PRIAS¹⁵, i.e. a PSA level of ≤ 10 ng/ml, clinical T-stage ≤ 2 , a PSA density of < 0.20 ng/ml/g, a Gleason score of < 7 and two or fewer 2 positive cores.

All 151 men with low-risk PC who had a RP between 1993 and 2008 at Erasmus Medical Center Rotterdam were selected. To enhance the chance of finding residual tumour in paraffin blocks, biopsy tissues had to contain at least 2 mm of PC (100 samples). For 98 of 100 cases the tissue blocks could be retrieved from the archive. The core containing the largest amount of PC was selected. All original biopsy specimens were reviewed by a urogenital pathologist (GvL) and in 12 (12.2%) cases the Gleason score was upgraded, due to interobserver variability and updated criteria for Gleason grading.¹⁷ Those cases were excluded, leaving 86 cases suitable for

the present study. Preoperative clinical and pathological data on the diagnostic biopsies and RP specimens were obtained from the ERSPC database., in which the number of positive cores and estimated maximum percentage tumour invasion per biopsy were recorded.

10.2.2 | Pathological assessment of the diagnostic biopsies and RP specimens

All diagnostic biopsy cores were labelled and processed separately, as described previously¹⁷. RP specimens were processed following the ERSPC procedure¹⁸. Briefly, after fixing the specimens, they were inked, cut at 4 mm intervals and totally embedded. Pathological stage, Gleason score and surgical margin status were assessed, and tumour volume was measured using a computerized morphometric analysis.¹⁹ PC was classified as significant if there were any of extracapsular extension (EPE, which includes seminal vesicle invasion), Gleason pattern 4/5, or a tumour volume ≥ 0.5 cc.²⁰

10.2.3 | Immunostaining

For immunostaining, sections (4 μ m) were cut and deparaffinized through xylene and 100% ethanol. Sections were immunostained using the Envision system (DAKO, Glostrup, Denmark). Briefly, endogenous peroxidase activity was quenched for 20 minutes in 0.3% H₂O₂ /PBS. Antigen was retrieved by placing the slides in a microwave (600 W) for 15 minutes in a pH9 Tris-EDTA buffer (Klinipath, Duiven, Netherlands). After cooling, the slides were rinsed with Tris-HCL pH 7.5 (Klinipath, Duiven, Netherlands) and incubated with the primary antibody at an optimal dilution. Incubation was for 1 hour at room temperature for EZH2 (clone 11/EZH2, DB transductions laboratories; 1:250), MIB-1 (clone M7240, DAKO; 1:50) and p27 (clone NCLp27, Novocastra; 1:40). For BMI-1 (clone F6, Millipore; 1:50), incubation was performed overnight at 4°C. The slides were rinsed with Tris-HCL followed by a 30-minute incubation with goat anti-rabbit (EZH2) or goat anti-mouse (p27, MIB-1, BMI-1) antibody from the Envision kit. The antigen antibody complexes were visualized with 3,3'-diaminobenzidine tetrahydrochloride (DAB+) chromogen diluted in substrate buffer (Dako REAL Envision Detecting system), during two incubations of 5 minutes. All slides were counterstained with haematoxylin (Klinipath 4085.9005, Duiven, Netherlands).

10.2.4 | Quantification process

For p27^{kip1} and for BMI-1 the positive nuclei were assessed relative to the total number of tumour cells at a scale of 0-100%. For EZH2 and MIB-1, the percentage of positive tumour nuclei of the total number of tumour cells in the whole core was calculated after actual visual counting of all positive and all negative cells. This very accurate quantification method was applied because of low expression of these immunostains.

All biopsy cores were scored independently by two investigators (TW and GvL) in a blinded fashion. When scores were discrepant, consensus was reached in a combined session. The staining intensity was scored for EZH2 and BMI-1 as weak (only visible at high magnification),

moderate (visible at low magnification) or strong (strikingly positive at low magnification). However, as for both EZH2 and BMI-1 intensity was weak in more than 95%, intensity was not included in further analyses.

10.2.5 | Statistical analysis

Logistic regression analyses were used to assess the predictive value of the immunostainings for significant PC. Analyses were used in univariate and multivariable settings, the latter using a backward stepwise method; variables were rejected at a p-value >0.05. In all multivariate analyses the following variables were included: PSA level, age, prostate volume, PSA density, clinical tumour (cT) stage (T1c or T2), number of positive cores (one or two) and maximum tumour invasion in the biopsy core.

For the immunostaining scores an optimum threshold was chosen based on optimum predictive value, because thresholds previously might not apply to the current cohort due to sample selection of low-risk PC. For independent predictive immunomarkers, the actual additional value was assessed by comparing the base model including PSA, age, prostate volume, PSA density, cT-stage, number of positive cores and maximum tumour invasion per core, with the extended model also including the respective immunomarker, using the likelihood ratio test. The area under the receiver-operator curve (AUC) was used to express the performance of the model. All statistical analyses were done using SPSS version 15.0 (SPSS Inc, IL, USA). In all tests, a two-sided $p < 0.05$ was considered to indicate significance.

10.3 | Results

10.3.1 | Clinical and pathological characteristics

Table 1 shows the clinical and needle-biopsy characteristics; in the RP specimens, EPE was found in 9 (10.4%) cases and the Gleason score was upgraded from 6 to 7 in 16 (18.6%) cases (Table 2). In all, 44 (51.2%) cases fulfilled the criteria of significant disease on RP.

The PSA density and the maximum PC invasion per core were significant univariate predictors of significant PC at RP (Table 3). In multivariable analysis, both factors remained independent predictors, while all other variables were excluded from the model due to lack of significance (Table 3). The area under the curve was 0.744 (95% confidence interval (CI) 0.616-0.871) for this model.

10.3.2 | Immunostaining results

Analysing EZH2 staining, 12 cases did not contain residual tumour and in 2 the immunostaining was considered false-negative based on complete absence of staining, leaving 72 (83.7%) cases for further analyses. EZH2 was expressed in nuclei of both normal and malignant cells. Its expression level was generally low, with a median(range) value of 1.1% (0-6.0%) in malignant

cells. An EZH2 level of >1.0% was assessed as the optimum threshold and was present in 30 (41.7%) cases.

Table 1 | Preoperative clinicopathologic characteristics.

Characteristic	Low-risk PC
Total cohort	86 (100%)
Age at diagnosis (yrs)	
Mean	63.5
Median	63.8
Range	55.5-71.6
PSA value at diagnosis (ng/ml)	
Mean	4.5
Median	4.2
Range	0.8-10.0
Prostate volume (ml)	
Mean	42.8
Median	36.3
Range	17.6-144.7
Clinical T-stage	
T1c	45 (52.3%)
T2a	31 (36.0%)
T2b	6 (7.0%)
T2c	4 (4.7%)
T3/4	-
Number of positive biopsy cores	
1	33 (38.4%)
2	53 (61.6%)
>2	-
Maximum tumour invasion	
=<10%	16 (18.6%)
10-50%	58 (67.4%)
>50%	12 (14.0%)

In MIB-1 stained biopsies, 13 cases showed no remaining tumour and in 2 the staining was false-negative, leaving 71 cases for the analyses. MIB-1 was expressed at low levels in benign and malignant nuclei. The median (range) expression was 2.4 (0.2-24.8) %. For MIB-1 the threshold was 3.0%, and 25 (35.3%) cases expressed >3.0% MIB-1 positive nuclei.

In p27^{kip1}-stained specimens, 17 (19.8%) cases had no tumour and in 9 (10.5%) the staining was false-negative, leaving 60 (69.8%) for analyses. The median p27^{kip1} expression was 87.5% (range 20%-100%). The threshold for p27^{kip1} was set at <90%, which was present in 38 (63.3%) cases.

BMI-1 staining could be scored in 69 (80.2%) specimens; there was no tumour in the remaining 17. The median BMI-1 expression was 80% (range 40%-93%), and the threshold for BMI-1 was set at >70% positive cells, which was present for 50 (72.5%) cases. Examples of all markers are shown in figure 1.

Table 2 | Radical prostatectomy characteristics. *Significant disease was defined as extracapsular extension, or Gleason pattern 4/5, or tumour volume \geq 0.5ml.

Characteristic	
Total cohort	86 (100%)
Pathological T-stage	
T2a	22 (25.6%)
T2b	4 (4.7%)
T2c	51 (59.3%)
T3	5 (5.8%)
T4	4 (4.7%)
Gleason score	
<7	71 (81.4%)
=7	15 (18.6%)
>7	-
Seminal vesicle invasion	
Yes	1 (1.2%)
No	85 (98.8%)
Extracapsular extension	
Yes	5 (5.8%)
No	81 (94.2%)
Surgical margins	
Negative	77 (89.5%)
Positive	9 (10.5%)
Tumour volume (ml)	
Mean	0.63
Median	0.43
Range	0.003-5.04
Significant disease*	
No	42 (48.8%)
Yes	44 (51.2%)

10.3.3 | The predictive value of immunostaining in low-risk PC

The distributions of Gleason score 7, EPE, tumour volume of \geq 0.5 ml and overall significant PC are shown for each immunomarker in Table 4. In univariate analysis, low expression of p27^{kip1} was significantly related to a higher probability of significant PC (OR 3.29, 95% CI 1.09-9.95, $p=0.035$). For high EZH2 expression there was a trend to a higher probability of significant PC in univariate analysis, although this was not statistically significant (OR 2.30, 95% CI 0.88-

6.03, $p=0.089$. When EZH2 or p27^{kip1} were included in the multivariable model with backward exclusion, both remained statically significant predictors of significant PC at RP. The area under the curve increased to 0.828 (95% CI 0.714-0.941) and 0.821 (95% CI 0.708-0.935) respectively. However, the predictive accuracy did not improve with the addition of either EZH2 or p27^{kip1}. Including both immunomarkers simultaneously in the model showed that EZH2 and p27^{kip1} had independent predictive value (Table 5). The predictive model significantly improved ($p<0.05$) with the inclusion of both markers. The area under the curve for this model including both immunomarkers was 0.863 (95% CI 0.763-0.964). BMI-1 and MIB-1 expression were not predictive for significant at RP.

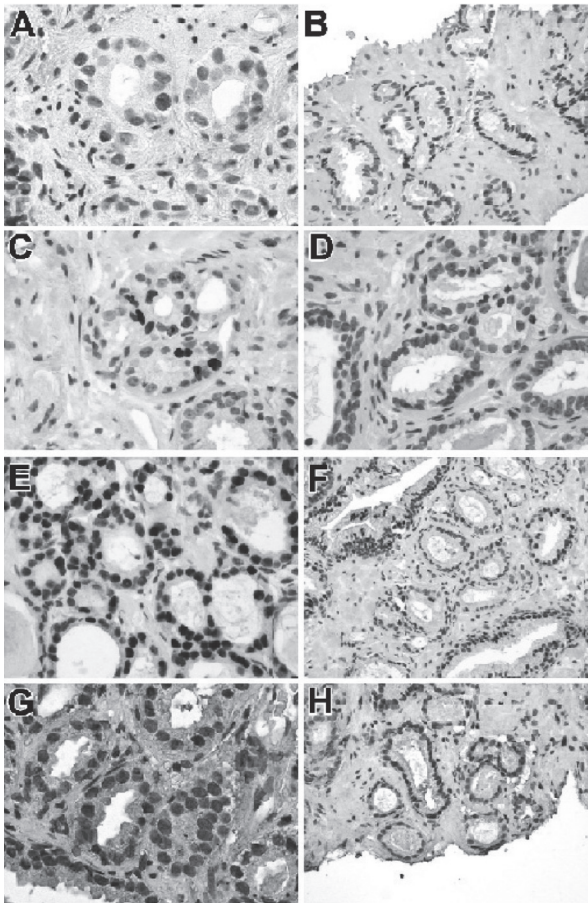


Figure 1 | Various expression levels of the immunomarkers. A: High EZH2 expression, 400x. B: Very low EZH2 expression, 200x. C: High MIB-1 expression, 400x. D: Low MIB-1 expression, 400x. E: High p27^{kip1} expression, 400x. F: Low p27^{kip1} expression, 200x. G: High BMI-1 expression, 400x. H: Low BMI-1 expression, 200x. For color images, see appendix page 184.

Table 3 | Predictive value of the preoperative variables, excluding immunostaining scores, for significant PC at RP (n=86). OR= odds ratio, CI= confidence interval, n.s.= not significant

Variable	Univariate OR (95% CI) significant PC	Multivariable OR (95% CI) significant PC
Age	1.03 (0.92-1.15) p=0.578	n.s.
PSA	1.19 (0.95-1.50), p=0.124	n.s.
PSA density	2.21 (1.26-3.86), p=0.006	2.11 (1.18-3.79), p=0.012
Prostate volume	1.00 (0.98-1.02), p=0.989	n.s.
Tumour stage	1.00 (0.43-2.34), p=0.992	n.s.
Number of positive cores	1.77 (0.74-4.26), p=0.203	n.s.
Maximum PC invasion	1.03 (1.01-1.06), p=0.012	1.03 (1.004-1.06), p=0.025

Table 4 | Distribution of Gleason score 7 at RP, extraprostatic extension (EPE), tumour volume ≥ 0.5 ml and significant prostate cancer (PC) per immunomarker in patients with low-risk prostate cancer.

	Total	Gleason 7 at RP	EPE	Tumour volume ≥ 0.5 ml	Significant PC
EZH2					
<1%	42	5 (11.9%)	2 (4.8%)	17 (40.5%)	18 (42.9%)
$\geq 1\%$	30	7 (23.3%)	4 (13.3%)	13 (43.3%)	19 (63.3%)
MIB-1					
$\leq 3\%$	46	5 (10.9%)	4 (8.7%)	18 (39.1%)	22 (47.8%)
$> 3\%$	25	7 (28.0%)	2 (8.0%)	11 (44.0%)	14 (56.0%)
p27^{kip1}					
$\geq 90\%$	22	5 (22.7%)	1 (4.5%)	5 (22.7%)	7 (31.8%)
$< 90\%$	38	4 (10.5%)	5 (13.2%)	19 (50.0%)	23 (60.5%)
BMI-1					
$\leq 70\%$	19	4 (21.1%)	4 (21.1%)	10 (52.6%)	10 (52.6%)
$> 70\%$	50	7 (14.0%)	3 (6.0%)	16 (32.0%)	22 (44.0%)

Table 5 | Predictive value of the immunostainings for significant PC at RP. Odds ratio's (OR) and 95% confidence interval (95%CI), resulting from the univariate and multivariable regression analyses. OR= odds ratio, CI= confidence interval, n.s.= not significant

Variable	Univariate OR (95% CI) P=	Multivariable OR (95% CI) including 1 immunomarker P=	Multivariable OR (95% CI) including all immunomarkers P=
EZH2 >1%	2.30 (0.88-6.03) P=0.089	3.19 (1.04-9.78) P=0.043	11.52 (1.91-69.37) P=0.008
MIB-1 >3.0%	1.39 (0.52-3.70) P=0.511	n.s.	n.s.
P27 <90%	3.29 (1.09-9.95) P=0.035	4.69 (1.11-19.88) P=0.036	5.78 (1.14-29.29) P=0.034
BMI-1 >70%	0.71 (0.25-2.04) P=0.522	n.s.	n.s.

10.4 | Discussion

One of the major current challenges in prostate cancer care is to find a reliable method to differentiate significant PC from indolent disease before RP. The present study showed that men with low-risk PC before treatment and high EZH2 expression (>1.0 %; OR 3.19; p=0.043) or low p27^{kip1} expression (<90%; OR 4.69; p=0.036) at biopsy had a significant higher chance of harboring significant PC at RP. To the best knowledge, the value of these immunostains in biopsy specimens of men with low-risk PC has not been reported before, and can possibly help in treatment stratification.

The Polycomb-group of genes are important in maintaining cell identity and regulation of the cell cycle.²¹ Overexpression of the Polycomb-group genes EZH2 and BMI-1 is associated with poor prognosis in several human tumours, including prostate cancer.^{3,5,21-23} EZH2 is overexpressed in hormone-refractory and/or metastatic disease compared to localized PC^{1,22} and has a predictive value for progression and survival after prostatectomy.^{1,3-6} EZH2 was found to be associated with adverse pathological characteristics by some investigators,^{4,23} although others could not confirm these results²². In our study of men with low-risk PC, a high EZH2 was predictive for significant PC at RP.

p27^{kip1} belongs to the Cip/Kip family of cyclin dependent kinase inhibitors (CDKIs), which are important in regulating cell-cycle processes. Loss of these inhibitors leads to deregulation of the cell cycle and uncontrolled proliferation.²⁴ Loss of p27^{kip1} expression in PC at RP has been associated with a higher risk of biochemical recurrence and decreased survival,⁷⁻⁹ and a high Gleason score.^{7,9,25,26} Thomas et al²⁵ reported similar findings for loss of p27^{kip1} expression on biopsy specimens: low expression (<30%) was related to a high Gleason score and more advanced pathologic stage, and there was a trend towards a higher probability of biochemical recurrence. Furthermore, low biopsy p27^{kip1} expression (<50%) was an independent predictor

of clinically significant PC at RP in a study by Vis et al.²⁷ The present study yielded parallel results: low expression of p27^{kip1} was independently predictive of significant PC, even though only low-risk patients were included.

MIB-1 is an antibody targeting Ki-67 antigen, which is a proliferation marker of cells in the G1 phase. A high Ki-67 expression at RP has been associated with a high GS and advanced stage,^{6,14,28,29} biochemical progression after treatment^{8,11-13} and PC related death.¹⁴ For Ki-67, there are several studies also on biopsy tissue, with similar results.^{8,30-33} In the present study, MIB-1 expression was not predictive of significant PC.

An important difference between immunohistochemical studies are the threshold values used to define high or low expression. For example, the definition of low p27^{kip1} ranges from 10% to 50%. The threshold of 90% in the present study is much higher, which can be explained by tissue selection: only low-risk PC was included, instead of RP tissue with all Gleason scores. Therefore, the most unfavourable tumours were excluded from this study, which most probably are those with the lowest p27^{kip1} expression. In the present patients, only 5 had a p27^{kip1} <50%. Interestingly, the reported threshold of 1% for EZH2 is in concordance with the value used in a RP study by one of the authors.⁴

While cut-offs were applied in most immunohistochemical studies, we do not suggest using strict threshold values in clinical practice. Although there is a significant difference in the tumour characteristics of patients below or above the threshold, using them will lead to missing relevant men with indolent disease. This is shown in table 4: e.g. for EZH2, applying the threshold and excluding men with high expression from active surveillance would exclude 11 patients with insignificant PC. A more promising application of immunomarkers might be in nomograms predicting the individual chance of significant PC at RP. These markers might be of help in differentiating between patients eligible for active surveillance and patients more prone to progression and subsequent need for radical treatment. The use of MIB-1 in this particular setting was recently described, with a significant predictive value of this marker in biopsy tissue.^{31,32} However, before EZH2 and p27^{kip1} can be implemented in routine clinical care, more evidence will be needed to confirm their value in biopsy specimens, preferably in a prospective study.

The present study has several limitations. First, an important problem arises with the threshold of 1.0% for EZH2. Deciding whether the expression is 0.5% or 1.5% is not easy. Counting of positive and negative cells is warranted, but time-consuming. Furthermore, most EZH2 staining was weak, making it more susceptible to a subjective interpretation. Also, at least 2 mm of prostate cancer tissue was studied to avoid a very high discard rate. However, in smaller PC lesions, establishing a 1% expression might be impossible because at least 100 cells seem to be required to assess such a low expression. Although EZH2 was an independent predictor of significant PC, further evaluation on the feasibility of using EZH2 in routine clinical practice is needed.

The present study was also limited because it was retrospective, using archived tissue blocks. This had several consequences. First, due to the updated Gleason scoring consensus¹⁶ and interobserver variability, 12 cases initially assigned a Gleason score 6 were now judged as Gleason score 7. Because our aim was to differentiate between possibly indolent and significant PC, we excluded specimens with Gleason score 7, as those are by definition not classified as low-risk disease. This could cause a bias, from excluding a part of the cohort. Second, by selecting patients with at least 2 mm of tumour we improved the chance of finding residual tumour for immunostaining and excluded patients with a high a priori chance of having indolent disease. Despite of our selection, 12-17 cases per marker no longer had tumour in the biopsy. However, in routine diagnostic tests, tissue from all 3 biopsy levels initially cut, are directly available for additional immunohistochemistry, solving this problem. Third, the numbers of cores available for additional immunostaining are relevant; only one core per sextant biopsy was used for immunostaining, to preserve as much as possible of the valuable tissue database. Staining of all positive cores possibly leads to a more reliable staining score and a higher predictive value of the marker expression scores. Finally, the relatively few samples assessed might have impaired the statistical power to assess the predictive value of the immunomarkers. We also intended to assess the predictive value of the markers for biochemical recurrence after RP, defined as PSA >0.2 ng/ml after RP, but only 7 (8.1%) patients had a biochemical recurrence, and no valid statistical testing was possible (data not shown).

In conclusion, the expressions of EZH2 and p27^{kip1} in biopsy specimens were independent predictors of significant PC at RP in men with low-risk disease. The determination of p27^{kip1} on diagnostic needle-biopsies might prospectively support therapeutic decision making.

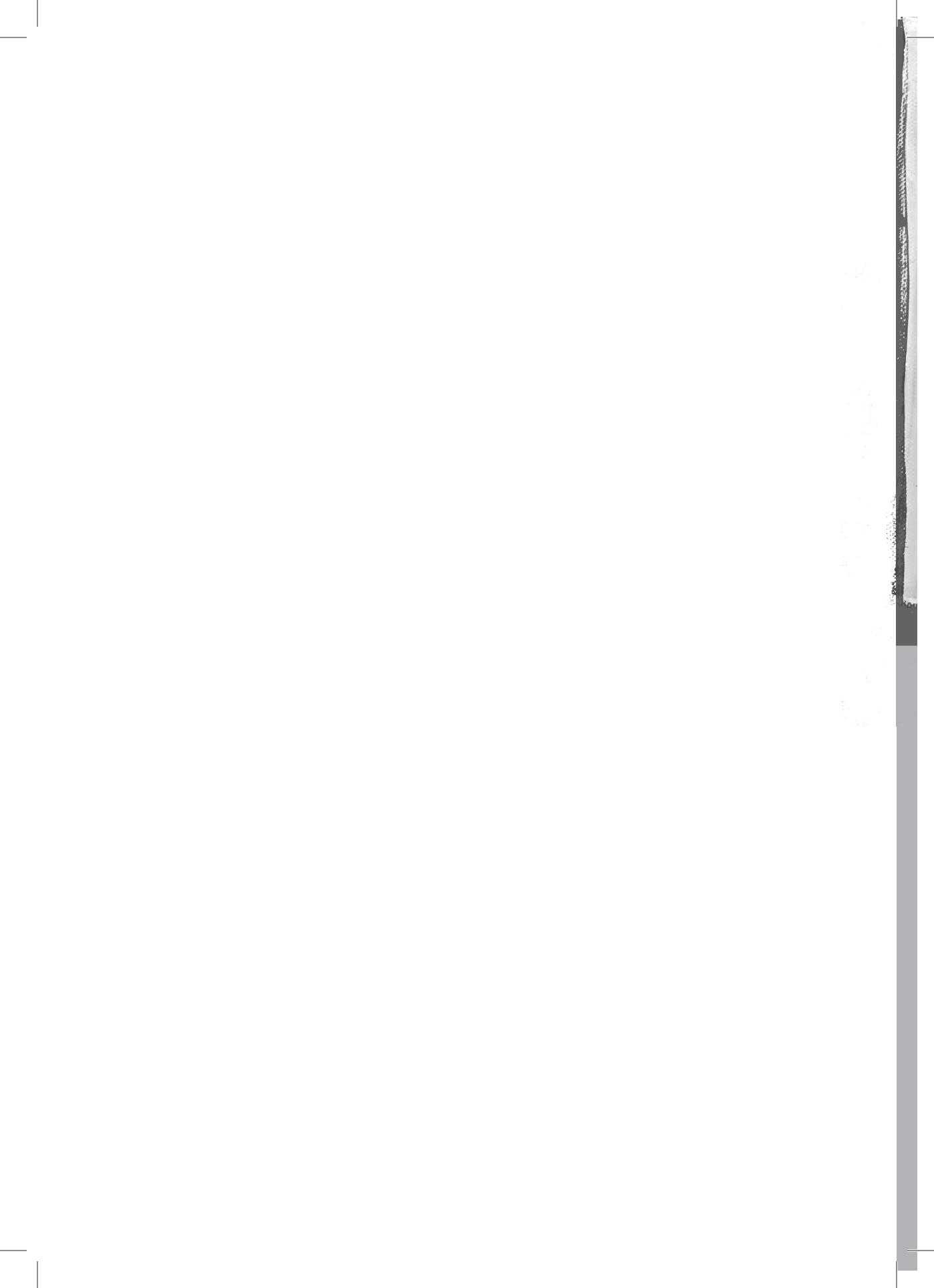
References

1. Noguchi M, Stamey TA, McNeal JE, Yemoto CM. Relationship between systematic biopsies and histological features of 222 radical prostatectomy specimens: lack of prediction of tumour significance for men with nonpalpable prostate cancer. *J Urol* 2001;166:104-9
2. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RGAB, Otte AP, Rubin MA, Cinnaiyan AM. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* 2002;419:624-9
3. Berezovska OP, Glinskii AB, Yang Z, Li X-M, Hoffman RM and Glinsky GV. Essential role for activation of the polycomb-group (PcG) protein chromatin silencing pathway in metastatic prostate cancer. *Cell Cycle* 2006;5:1886-1901
4. Van Leenders GJLH, Dukers D, Hessels D, Van den Kieboom SWM, Hulsbergen CA, Witjes JA, Otte AP, Meijer CJ and Raaphorst FM. Polycomb-group oncogenes EZH2, BMI1, and RING1 are overexpressed in prostate cancer with adverse pathologic and clinical features. *Eur Urol* 2007;52:455-463
5. Bryant RJ, Cross NA, Eaton CL, Hamdy FC and Cunliffe VT. EZH2 promotes proliferation and invasiveness of prostate cancer cells. *Prostate* 2007;67:547-556
6. Laitinen S, Martikainen PM, Tolonen T, Isola J, Tammela TLJ and Visakorpi T. EZH2, Ki-67 and MCM7 are prognostic markers in prostatectomy treated patients. *Int J Cancer* 2008;122:595-602
7. Cote RJ, Shi Y, Groshen S, Feng A-C, Cordon-Cardo C, Skinner D, Lieskovosky G. Association of p27^{kip1} levels with recurrence and survival in patients with stage C prostate carcinoma. *J Natl Cancer Inst* 1998;90:916-20
8. Kuczyk M, Machtens S, Hradil K, Schubach J, Christian W, Knüchel R, Hartmann J, Bokemeyer C, Jonas U and Seth J. Predictive value of decreased p27^{kip1} expression for the recurrence-free and long-term survival of prostate cancer patients. *Br J Cancer* 1999;81:1052-8
9. Vis AN, Noordzij MA, Fitoz K, Wildhagen MF, Schröder FH and van der Kwast ThH. Prognostic value of cell cycle proteins p27^{kip1} and MIB-1, and the cell adhesion protein CD44s in surgically treated patients with prostate cancer. *J Urol* 2000;164:2156-61
10. Rubio J, Ramos D, López-Guerrero JA, Iborra I, Collado A, Solsona E, Almenar S, Llombart-Bosch A. Immunohistochemical expression of Ki-67 antigen, Cox-2 and Bax/Bcl-2 in prostate cancer; prognostic value in biopsies and radical prostatectomy specimens. *Eur Urol* 2005;48:745-51
11. Sadi MV, Barrack ER. Determination of growth factor fraction in advanced prostate cancer by Ki-67 immunostaining and its relationship to the time to tumour progression after hormonal therapy. *Cancer* 1991;67:3065-71
12. Revelos K, Petraki C, Gregorakis A, Scorilas A, Papanastasiou P, Tenta R, Koutsilieris M. p27 (kip1) and ki-67 (MIB1) immunohistochemical expression in radical prostatectomy specimens of patients with clinically localized prostate cancer. *In Vivo* 2005;19:911-20
13. Miyake H, Muramaki M, Kurahashi T, Takenaka A, Fujisawa M. Expression of potential molecular markers in prostate cancer: Correlation with clinicopathological outcomes in patients undergoing radical prostatectomy. *Urol Oncol* 2008; Epub ahead of print. doi:10.1016/j.urolonc.2008.08.001
14. Aaltomaa S, Kärjä V, Lipponen P, Isotalo T, Kankkunen JP, Talja M, Mokka R. Expression of Ki-67, cyclin D1 and apoptosis markers correlated with survival in prostate cancer patients treated by radical prostatectomy. *Anticancer Res* 2006;26:4873-8
15. 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;52:1560-3
16. Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL; ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol*. 2005;29:1228-42
17. Van der Kwast ThH, Lopes C, Santonja C, Pihl C-G, Neetens I, Martikainen P, Di Lollo S, Bubendorf L and Hoedemaeker RF. Guidelines for processing and reporting of prostatic needle biopsies. *J Clin Pathol* 2003;56:336-40

18. Hoedemaeker RF, Ruijter ETG, Ruizeveld-de Winter JA, van der Kaa CA and van der Kwast ThH. The BIOMED II MPC study group. Processing radical prostatectomy specimens. A comprehensive and standardized protocol. *J Urol Pathol* 1998;9:211-22
19. Hoedemaeker RF, Rietbergen JBW, Kranse R, Van der Kwast ThH, Schröder FH. Comparison of pathological characteristics of T1c and non-T1c cancers detected in a population-based screening study, the European Randomized Study of Screening for Prostate Cancer. *World J Urol* 1997;15:339-345
20. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumour extent of nonpalpable (stage T1c) prostate cancer. *JAMA*. 1994;271:368-74
21. Pasini D, Bracken AP, Helin K. Polycomb group proteins in cell cycle progression and cancer. *Cell cycle* 2004;3:396-400
22. Saramäki OR, Tammela TLJ, Martikainen PM, Vessella RL and Visakorpi T. The gene for polycomb group protein enhancer of zeste homolog 2 (EZH2) is amplified in late-stage prostate cancer. *Genes Chromosomes Cancer* 2006;45:639-645
23. Bachmann IM, Halvorsen OJ, Collett K, Stefansson IM, Straume O, Haukaas SA, Salvesen HB, Otte AP, Akslen LA. EZH2 expression is associated with high proliferation rate and aggressive tumour subgroups in cutaneous melanoma and cancers of the endometrium, prostate and breast. *JCO* 2006;24:268-273
24. MacLachlan TK, Sang N, Giordano A. Cyclins, cyclin-dependent kinases and cdk inhibitors: implications in cell cycle control and cancer. *Crit Rev Eukaryot Gene Expr* 1995;5:127-56
25. Thomas GV, Schrage MI, Rosenfelt L, Kim JL, Salur G, deKernion JB, Dorey F, Said J and Reiter RE. Preoperative prostate needle biopsy p27^{kip1} correlates with subsequent radical prostatectomy p27^{kip1}, Gleason grade and pathological stage. *J Urol* 2000;164:1987-91
26. Chevillet JC, Lloyd RV, Sebo TJ, Cheng L, Erickson L, Bostwick DG, Lohse CM and Wollan P. Expression of p27^{kip1} in prostatic adenocarcinoma. *Mod Pathol* 1998;11:324-8
27. Vis AN, van Rhijn BWG, Noordzij MA, Schröder FH and van der Kwast ThH. Value of tissue markers , MIB-1, and CD44s for the pre-operative prediction of tumour features in screen-detected prostate cancer. *J Pathol* 2002;197:148-54
28. Moul JW. Angiogenesis, p53, bcl-2 and Ki-67 in the progression of prostate cancer after radical prostatectomy. *Eur Urol* 1999;35:399-407
29. Khatami A, Hugosson J, Wang W, Damber JE. Ki-67 in screen-detected, low-grade prostate cancer, relation to prostate-specific antigen doubling time, Gleason score and prostate-specific antigen relapse after radical prostatectomy. *Scand J Urol Nephrol* 2008;23:1-7
30. Bubendorf L, Tapia C, Gasser TC, Casella R, Grunder B, Moch H, Mihatsch MJ, Sauter G. Ki67 labeling index in core needle biopsies independently predicts tumour-specific survival in prostate cancer. *Hum Pathol* 1998;29:949-54
31. Jhavar S, Bartlett J, Kovacs G, Corbishley C, Dearnaley D, Eeles R, Khoo V, Huddart R, Horwich A, Thompson A, Norman A, Brewer D, Cooper CS and Parker C. Biopsy tissue micro array study of Ki-67 expression in untreated, localized prostate cancer managed by active surveillance. *Prostate Cancer Prostatic Dis.* 2009;12:143-7
32. Zellweger T, Günther S, Zlobec I, Savic S, Sauter G, Moch H, Mattarelli G, Eichenberger T, Curschellas E, Rüfenacht H, Bahcman A, Gasser TC, Mihatsch MJ, Bubendorf L. Tumour growth fraction measured by immunohistochemical staining of Ki67 is an independent prognostic factor in preoperative prostate biopsies with small-volume or low-grade prostate cancer. *Int J Cancer* 2009;124:2116-23
33. Cowen D, Troncso P, Khoo VS, Zagars GK, von Eschenbach AC, Meistrich ML, Pollack A. Ki-67 staining is an independent correlate of biochemical failure in prostate cancer treated with radiotherapy. *Clin Cancer Res* 2002;8:1148-54

PART V

DISCUSSION AND
SUMMARY





CHAPTER 11

General discussion

General discussion

Prostate cancer shows a very diverse clinical course. Some men present with extensive disease, rapid progression and mortality. Cure is out of reach already at the moment of diagnosis. On the other hand, some cases do not induce symptoms or mortality and do not surface during a mans life. The existence of these latter cases is known through autopsy studies and cystoprostatectomy studies. Autopsy studies have shown an overall incidence of latent prostate cancer of 19-39% in entirely processed prostate specimens¹⁻⁴ with the highest incidence rates in the highest age ranges: up to 86.6% in men aged 81-95 years.³ Similar high rates of incidentally detected prostate carcinoma of 14.2%-54.0% are described in radical cystoprostatectomy series of patients treated for muscle invasive bladder carcinoma and no evidence of prostate cancer at the time of surgery.⁵⁻¹⁰ This illustrates the heterogeneity of prostate cancer: a high prevalence of indolent prostate cancer which often remains undetected and in other men the occurrence of aggressive disease causing serious symptoms and even mortality.

In the pre PSA era, prostate cancer was mostly diagnosed after a suspicious rectal examination, which was performed because a patient presented with complaints, or due to symptoms caused by metastasized disease. Additionally, prostate cancer cases were detected incidentally in transurethral resection tissue or in cystoprostatectomy specimens. The majority of prostate cancer cases were thus clinically significant and a large proportion of these men died of their disease. With the introduction of PSA as a serum test for detection of prostate cancer, diagnosis was shifted to an earlier stage of disease. This was even further enhanced by reducing the PSA threshold to prompt a prostate biopsy to a total level of 4.0 ng/ml and subsequently to 2.5 ng/ml.^{11,12} Additionally, systematic biopsies were introduced, improving the overall prostate cancer detection.¹³ As a consequence, prostate cancers were now diagnosed in absence of symptoms or suspicious findings, except for an elevated serum PSA value. Thereby, population-based screening became feasible and the effect of screening is subject of large trials.^{14,15}

In 2009, the European Randomized Study of Screening for Prostate Cancer showed a significant mortality reduction of 20% in the screening arm.¹⁴ In this general discussion, the studies addressed in this thesis will be discussed in the perspective of the question 'how to screen for prostate cancer'. First, the role of prostate cancer treatment in both study arms will be discussed. The second part focuses on improvement of screening strategies, especially screening markers, the screening interval, and future directions. Next, the characterization and identification of clinically insignificant prostate cancer will be addressed and finally the possible role of prevention and imaging in screening for prostate cancer is discussed.

The reduction of disease specific mortality due to screening for prostate cancer

The role of prostate cancer treatment

Screening advances the moment of diagnosis, increasing the probability of detecting the cancer in a curable phase. Consequently, treatment will be advanced as well, causing a shift to more radical treatment modalities like radical prostatectomy or radiotherapy, instead of hormonal treatment or watchful waiting.¹⁶ However, if a treatment shift towards more aggressive treatment modalities would occur between the screening and control arm independent of the expected shift due to stage and grade migration, a mortality reduction in favour of the screening arm could be caused by this treatment shift alone. The mortality reduction in the screening arm could then no longer be related to effect of screening but to the effect of treatment. This possible treatment bias was addressed in **Chapter 4**. A small systematic difference in treatment between both trial arms was observed in patients with high risk PC, after correction for several important factors like age, PSA, tumour stage and grade: a participant in the screening arm was more likely to be treated with surgery than with any other treatment modality compared to a participant in the control arm. Whether this could result in a mortality difference is uncertain, but if a difference would occur it is likely to be small. Moreover, if this indeed would result in a mortality reduction, it is yet unclear which arm would be most likely to benefit. Although a benefit in favour of the control arm might seem contra intuitive, this indeed may be possible due to changing treatment patterns over time. For high-risk prostate cancer, important insights have been provided by studies reported during the last years. For example, a combination therapy of radiotherapy and hormonal treatment has shown to be superior to radiation or hormones alone.^{17,18} Prostate cancer cases are detected in the control arm later in time than in the screening arm, due to lead time produced by screening. It was observed in the ERSPC cohort, after correction for age, stage and grade, that patients with high risk disease in the control arm were significantly more often treated with combination therapy than patients with high risk disease in the screening arm. Therefore, patients in the control arm might theoretically benefit more from new insights on optimal treatment which reduces prostate cancer specific mortality because they are detected at a later point in time.

If future new prostate cancer treatments improve disease specific survival, the control arm may benefit more than the screening arm and the effect of screening on mortality might seem smaller than it actually is. Future studies within the ERSPC should therefore continue to monitor treatment patterns in both arms over time and their possible effect on prostate cancer mortality.

Overall, the mortality reduction observed due to screening is not caused by a systematic treatment bias. But if a treatment bias indeed were responsible for most of the mortality reduction due to more aggressive treatment in the screening arm, independent of tumour stage and grade, this could possibly indicate that treatment in general should be more aggressive and

screening would not necessarily be effective. However, this was not observed and screening indeed reduces prostate cancer mortality. Therefore, the following part of the thesis will focus on the question: how to screen for prostate cancer?

Improving screening strategies

Part II of this thesis aims at the improvement of the currently used screening protocol. Especially, the prevention of unnecessary biopsies and the selective detection of significant prostate cancer while avoiding the detection of clinically insignificant prostate cancer is important for the improvement of screening protocols.

The screening marker

At present, screening is PSA based and if the PSA level exceeds the threshold of 3.0 ng/ml a prostate biopsy is performed. However, a lot of unnecessary biopsies are performed due to lack of specificity of PSA for prostate cancer. It has been previously shown that in repeat screening the association between rising PSA levels and rising prostate cancer risk is lost.¹⁹ This is probably due to the presence of benign enlargement of the prostate (BPH), which induces higher levels of PSA. After a PSA based screen the PSA level might be stronger related to prostate volume than presence of cancer during a next screen²⁰. This is confirmed in the ERSPC data: prostate volume is an important negative predictor for prostate cancer detection particularly in repeat screening²¹.

The change in PSA level during time, i.e. PSA velocity, has been suggested to be a better screening marker than PSA. PSA velocity was suggested to differentiate between BPH and prostate cancer in a large longitudinal study²² and showed to have prognostic value in prostate cancer patients treated surgically or with radiotherapy.^{23,24} However, PSA velocity did not improve the performance of PSA alone in our screening program as a biopsy indicator (**Chapter 5**). This was observed for prostate cancer detection in general as well as for clinically significant prostate cancer. For the definition of clinically significant disease, the risk of clinically significant disease according to the Epstein criteria was used²⁵. Although the tumour volume threshold used in these criteria was not a useful criterion for differentiating between indolent and aggressive prostate cancer in a later report published by our study group, using the traditional criteria does improve comparability with other studies. Additionally, a more recent report on the use of PSA velocity as a screening test by Vickers et al²⁶ used a biopsy Gleason score ≥ 7 as a single criterion for aggressive disease and showed results similar to our data: PSA velocity was not selective for aggressive prostate cancer.

This report by Vickers²⁶ included Rotterdam ERSPC data, with PSA measured every 4 years and Swedish ERSPC data, with PSA measurements at a 2-year interval, but the interval did not

influence the predictive value of PSA velocity. Additionally, PSA velocity did not add prognostic information on prostate cancer detection in the Prostate Cancer Prevention Trial (PCPT) study, in which the PSA was measured yearly.²⁷ Consequently, the time interval between PSA measurements to calculate the PSA velocity did not explain the absence of additional predictive value of PSA velocity in addition to PSA. In conclusion, PSA velocity will probably not improve the screening protocol for prostate cancer detection overall nor for selectively detecting aggressive disease.

It should however be emphasized that the results described in Chapter 5 holds true for men with PSA levels equal to or higher than 3.0 ng/ml. Only those men were biopsied, so only for those men the predictive value of PSA velocity for biopsy detectable prostate cancer could be evaluated. Hypothetically, PSA velocity might improve screening protocols in the PSA ranges below 3.0 ng/ml. In the Prostate Cancer Prevention Trial (PCPT) participants were biopsied irrespective of PSA value at the end of the study, which creates a good opportunity to assess the value of PSA velocity on the whole PSA range: PSA velocity did not add prognostic information for prostate cancer detection.²⁷ This was confirmed in a subset of ERSPC participants enrolled in side studies within the ERSPC in which biopsies were performed in PSA ranges below 3.0 ng/ml; no additional value of PSA velocity was observed.²⁸ Therefore, PSA velocity as a biopsy indicator in the lower PSA ranges will probably not improve specificity of the screening protocol.

Other PSA related markers for prostate cancer detection have been extensively studied. PSA kinetics, which for example include PSA slope and PSA doubling time in addition to PSA velocity²⁹ and PSA isoforms, for instance total and free PSA, human kallikrein, pro-PSA and nicked PSA,³⁰ have been reviewed for their use as possible screening tests. Although some markers like proPSA and human kallikrein 2 did show some additional value to PSA particularly in the PSA range 4.0-10.0 ng/ml, none could replace PSA and improve sensitivity and specificity. PSA based markers are still under investigation. It is likely that future screening will continue to be PSA based, probably using PSA in combination with other markers, due to low costs, widely available valid assays and a large experience with PSA-based prostate cancer care.

Extensive research on new markers that are not PSA related is done and includes for example the gene-fusion TMRSS2:ERG,³¹ prostate cancer antigen 3 (PCA3),^{32,33} circulating tumour cells,³⁴ and exosomes³⁵ amongst many others. An important issue for all markers under investigation, is the verification bias (or attribution bias). Generally, new markers are tested in cohorts of men who are biopsied based on a elevated PSA or other abnormal tests, for example an suspicious digital rectal examination. Of course, these cohorts are most easily accessible and will therefore often be the first step taken in a evaluation of a new marker. It should be kept in mind though, that these studies are subject to attribution bias and prospective evaluation of promising markers in unbiased cohorts is needed before general statements can be made.

An ideal marker would have high sensitivity and specificity for clinically significant disease. However, no such marker has been discovered yet: neither TMRPSS2:ERG nor PCA3, or any other marker, seem to be able to replace PSA while improving its performance. Moreover, all of the current knowledge on prostate cancer diagnosis mainly relies on PSA based cohorts and subsequent treatment decisions. Most likely, future screening strategies will rely on a set of markers instead only on PSA level. For example, PCA3 seems useful in avoiding unnecessary repeat biopsies, which also has been observed in the Rotterdam ERSPC screening protocol (personal communication M.J. Roobol). In addition, the combined use of TMRPSS2:ERG with PCA3 showed higher predictive accuracy for prostate cancer detection than PCA3 alone.³⁶ To be suitable as a screening test, a new marker has to be easily detectable, mostly in body fluids like blood or urine, a valid and robust test should be available and costs should be limited. If a marker meets these requirements and improves screening with PSA alone, it could be added to screening protocols.

The screening interval

Next to offering screening with a fixed PSA threshold to indicate a prostate biopsy, a fixed screening interval of 4 years is applied in the Rotterdam section of the ERSPC. This may be too short for some, but too long for other cancers. It might be possible to establish an individual screening interval based on individual patient information. As an example, it has been shown by Roobol et al that men with an initial PSA below 1.0 ng/ml can safely be re-screened after an interval of 8 years instead of 4 years.³⁷ **Chapter 6** also addresses the possibility of individualizing screening intervals, based on biopsy information. None of the addressed features in non-malignant biopsies, including the presence of high-grade prostatic intraepithelial neoplasia (PIN), were predictive for prostate cancer detection in the next screening round. Thus, no recommendations for the adjustment of the screening interval based on these characteristics could be made based on these data. However, one specific result needs further discussion. The screening protocol of the ERSPC Rotterdam advises an immediate re-biopsy within six weeks if high grade PIN is observed in the prostate biopsy. However, the cancer detection rate at immediate rebiopsy is not increased in patients with HG-PIN.^{38,39} No increased risk at subsequent screening rounds after 1 or 4 years was observed either after a previous diagnosis of HG-PIN in Chapter 6 and by Vis et al.³⁸ It is worth noting that the cohort of participants described in Chapter 6 was not immediately rebiopsied after this detection of HG-PIN (these cases were newly found at review for the study), but only after 4 years. Although numbers are small and only focal presence of HG-PIN was observed, these findings question the need of the immediate re-biopsy within 6 weeks. Some studies reported that an increased risk of prostate cancer detection was only observed in wide-spread presence of high-grade PIN, not in focal high grade PIN^{40,41} while most report no effect of extensiveness of HG-PIN on subsequent cancer detection.⁴² No difference in cancer detection rate was found in men with HG-PIN in only one biopsy core, or more extensive presence of HG-PIN in the Rotterdam cohort either.⁴³ This held

also true for participants who refused the immediate re-biopsy. Furthermore, the cancer cases detected after a previous diagnosis of HG-PIN showed favourable features, which was described in the study by Vis and co-workers as well.³⁸ Concluding, the need of the immediate re-biopsy in the screening protocol needs to be evaluated. An immediate re-biopsy may not be required and probably can safely be postponed in a screening setting. Further evaluation is needed, possibly first by comparing the Rotterdam data with those from Sweden, where the immediate re-biopsy has already been omitted (personal communication prof. Dr. Hugosson).

Contrary to a diagnosis of HG-PIN, the presence of atypical lesions suspicious for prostate cancer, also known as atypical small acinar proliferations or ASAP, does warrant a re-biopsy due to a high cancer detection rate of about 40%,⁴² also in a screening program.^{38,39} The consequences of omitting the immediate re-biopsy after ASAP at initial biopsy, although not deliberately, are described in **Chapter 7**. No negative impact was observed, but these results cannot be generalized to all ASAP lesions: the ASAP lesions described in Chapter 7 were initially missed because they were mostly small and less suspicious. Therefore, they probably are a favourable selection of ASAP lesions in general. Thus, these results should not question the need for immediate re-biopsy and should not lead to reconsideration of the current guideline to perform an immediate rebiopsy, especially since prostate cancer at immediate re-biopsy may show aggressive features. For example, in participants of the ERSPC with an ASAP lesion at initial biopsy, 40% were diagnosed with prostate cancer of which 20% had a Gleason score of 7 or higher.³⁸

In addition to the effect of missing an ASAP lesion, Chapter 7 describes the occurrence of false-negative biopsies for prostate cancer. In 16 patients, around 1.1% of all men biopsied, malignant glands remained undetected in the biopsy cores. After a delay of four years these men were re-biopsied during the next screening round and diagnosed with prostate cancer. Overall, no negative effects in terms of a higher rate advanced tumour stages or Gleason scores ≥ 7 were observed in men with a missed prostate cancer lesion compared to participants with a previous benign diagnosis. Similar to the undiagnosed ASAP lesions, these prostate cancer foci are a selection of small cancers and are not representative for all prostate cancer cases. Most were small, consisting of only a few glands, confined to 1 or 2 cores and well differentiated with a Gleason score of 6. These results might be supportive of active surveillance for selected prostate cancer cases. Four years of natural history were observed, without interventions, and all cases seemed to be in a curable stage at the time of detection.

Furthermore, this study provided insights in the type of ASAP and malignant lesions that were at risk of being missed during pathological review of biopsy specimens. This is important educational information, which may reduce the rate of false-negative prostate biopsies.

Future screening strategies

The results of **Chapter 5** and **Chapter 6** showed that neither PSA velocity nor histopathological findings at a previous biopsy seemed to improve the screening protocol. However, it is very likely that results from previous screening visits will be incorporated in a future screening strategy. For

example, PSA might also be used as a marker for risk assessment of developing prostate cancer in the future. The PSA level has been reported to be predictive for prostate cancer detection, as much as 25 years before diagnosis.⁴⁴ An increase in the total PSA level of 1 ng/ml was associated with an increase in odds of detecting cancer of 3.69. This was also observed when only advanced cases were selected for analysis: 66% of advanced prostate cancers occurred in men with a total PSA level in the highest quintile (0.9 ng/ml or higher) measured before the age of 50 and total PSA was a strong predictor of advanced cancer (AUC 0.791, $p < 0.0001$)⁴⁵. In future screening strategies, this may imply a baseline PSA measurement, for example at the age of 40, which then could guide further individual screening intensity.

In addition to age, other patient related factors like co-morbidity,⁴⁶ family history,^{47,48} and race⁴⁹ will possibly be incorporated in future screening strategies as well because all of them influence the risk of dying from prostate cancer. Based on the combined evaluation of data obtained at previous screening visits, patient related factors and a set of markers for prostate cancer, a decision whether to perform a prostate biopsy will probably be made in the future. Furthermore, an individualized interval can be established at which re-screening would be reasonable. Such an individualized protocol has been described by Roobol et al,⁵⁰ using a risk calculator based on the ERSPC data.²¹ This calculator included prostate volume, digital rectal examination outcome and transrectal ultrasound in addition to PSA. Individualization of the screening protocol using this calculator resulted in a considerable reduction of biopsies of 33% at the cost of missing 13% of all detected prostate cancer cases of which 79-81% could be considered as possibly indolent. Another example of such a risk calculator is provided based on data from the Prostate Cancer Prevention Trial.⁵¹ Further exploration of other screening markers, prostate cancer risk factors and intervals will help to improve individual-based screening strategies. In these future screening programs, prostate cancer risk reduction,^{52,53} advanced imaging techniques⁵⁴⁻⁵⁶ or other aspects that are currently not implemented in the ERSPC screening protocol, may have a role. These aspects all may decrease the number of unnecessary biopsies. Most importantly, future screening protocols should not only aim at early diagnosis of prostate cancer, but aim selectively at early diagnosis of clinically significant prostate cancer.

Definition and identification of indolent prostate cancer

With the introduction of the PSA test, and especially with the introduction of screening using this PSA test, the detection of prostate cancer has increased. This increase is particularly observed in detection of small, well-differentiated prostate cancers with a low risk for morbidity or mortality.⁵⁷ A considerable part of these prostate cancer cases will probably cause no morbidity or mortality and may have gone undetected in absence of screening. The terminology used for these cases is diverse and can be confusing. First a short description of the terms

“clinically insignificant prostate cancer” and “indolent prostate cancer” will be given to improve understanding of the differences and to clarify the use of these terms in the following sections.

The term “clinically insignificant prostate cancer” is often used and defines prostate cancer that will not cause morbidity or mortality during life. The fact that these cases do not cause harm results from favourable tumour characteristics or from other, patient-related factors like age, or a combination. The subgroup of clinically insignificant cases that do not cause morbidity or mortality because of very favourable tumour characteristics, even in men with long life expectancy, is sometimes called “indolent prostate cancer”. Indolent prostate cancer thus refers to prostate cancer that will not cause any morbidity or mortality based on favourable tumour characteristics, irrespective of patient related factors. However, some refer to these cases using “clinically insignificant prostate cancer”. Especially in older reports, clinically insignificant is mostly used^{25,58} and more recent studies often use the terms intermixed. The differentiation between tumour-related factors and patient-related factors is a recent trend, and the term indolent disease is mostly used in more recent studies.

In this discussion, “indolent prostate cancer” is used for cases that do not cause morbidity or mortality because of very favourable tumour characteristics and subsequent low risk of progression, even in young men with long life expectancy and no co-morbidity. Clinically insignificant prostate cancer will be used for all cases that cause no morbidity or mortality. This includes indolent prostate cancer, but also prostate cancer that harbour more aggressive features but do not cause harm due to competing causes of death, like high age or severe co-morbidity. For example, a prostate cancer that contains Gleason pattern 4 disease is not indolent, but may very well be clinically insignificant in an 80-year old man. The third part of this thesis focuses especially on the identification of indolent disease using pathological criteria based on radical prostatectomy specimens, and the identification of indolent prostate cancer prior to treatment.

What are the criteria for indolent prostate cancer?

The most accurate tumour characteristics available are those observed in radical prostatectomy specimens, which are known to be prognostic for progression and mortality.^{59,60} The criteria for indolent prostate cancer most often used are subsequently based on radical prostatectomy specimens and include tumour stage, Gleason score and tumour volume. Inclusion of tumour volume as a criterion for indolent prostate cancer was suggested by Stamey⁵⁸ and Epstein.²⁵ Although they both used the term “clinically insignificant”, we will refer to these favourable cancers as “indolent”, in line with the previously described definitions. It seems logical to include a tumour volume threshold in the definition of indolent prostate cancer: tumour volume is an important prognostic factor for progression and mortality when assessed as a single prognosticator.⁶¹⁻⁶³ Additionally, one may assume that larger tumours are more easily identified and more likely to produce symptoms compared to small tumours.

However, whether the measurement of tumour volume actually adds prognostic value in multivariable analyses to more routinely assessed parameters in radical prostatectomy specimens

like tumour stage and Gleason score is subject of debate.^{61,63} Moreover, determination of tumour volume is very time-consuming. In **Chapter 8** was shown that tumour volume did not have any additional predictive value for patient outcome after correction for tumour stage and grade in our study of 344 radical prostatectomy specimens with a mean follow-up of 96 months. Obviously, the determination of tumour volume may be valuable for research purposes, but for clinical care settings, we suggest not to routinely measure tumour volume.

These results of Chapter 8 triggered us to further explore the prognostic value of tumour volume in the identification of indolent disease. The method that Stamey used to calculate the tumour volume threshold for indolent prostate cancer was used in **Chapter 9** to reassess this threshold value. Based on life-time risk estimates, a tumour volume threshold of 0.7 ml was found in the screened population in Rotterdam. Although the difference with the threshold of 0.5 ml might seem small, the proportion of men with a tumour volume ≤ 0.7 ml was considerably higher than the proportion of patients with a volume ≤ 0.5 ml: 50.8% versus 39.4%. Unfortunately, for the majority of the study population, only a total tumour volume was available. Although many studies on tumour volume did not explicitly describe whether total or index tumour volumes are studied, the differentiation might be important. In Chapter 9, the tumour volume threshold was also assessed specifically for index tumour volumes, which were available for a selection of 100 men, and a threshold of 0.55 ml was observed. This is remarkably similar to the traditional threshold of 0.5 ml, although the proportion of men with a tumour volume below the threshold increased from 47% for the 0.5 ml threshold to 52% for the 0.55 ml threshold.

It is important to understand that the reported thresholds in Chapter 9 are underestimating the true tumour volume threshold. Two arguments are shortly described in the discussion of the Chapter: first, the surgically treated patients were a favourable selection of prostate cancer patients in the screening arm of the ERSPC. Therefore, it is likely that overall tumour volume in the entire screened cohort was higher than in this selection of surgically treated patients, and the assessed threshold was an underestimation. Second, one might argue that Gleason score and tumour stage are more important parameters to determine patient outcome. It is commonly accepted, as described in the criteria for indolent prostate cancer by Epstein,²⁵ that only organ confined disease without Gleason pattern 4/5 can be considered possibly indolent. In the study population, 53.5% had organ confined prostate cancer with no Gleason pattern 4 or 5. We calculated that 50.8% of the study population had indolent prostate cancer, according to the life time risk estimates for this population. Therefore, almost patients with organ-confined prostate cancer and no Gleason pattern 4/5 had possibly indolent disease and a tumour volume threshold of 2.5 ml was assessed.

In addition to a reassessment of the tumour volume threshold, the prognostic value of the threshold was evaluated to assess the effectiveness of such a threshold in addition to the tumour stage and grade criteria for indolent prostate cancer. Biochemical progression was assumed to be a sign of significant prostate cancer in Chapter 9. It should be realized however,

that this is only a surrogate endpoint. Biochemical progression does not necessarily lead to symptoms or mortality. However, biochemical progression can be seen as a surrogate for clinically significant disease, because symptoms and mortality are most usually preceded by biochemical progression. This was also observed in our study (data not shown): not all patients with biochemical progression showed signs of metastasis or died of prostate cancer which may be caused by short follow-up, but all cases with metastasized disease or disease specific death, were observed to have biochemical progression as well.

It was shown that in organ-confined disease without Gleason pattern 4/5, tumour volume was not predictive of biochemical progression. These results are in line with the results described in Chapter 8. If we assume that biochemical progression indicated clinically significant disease, the results in Chapter 9 suggest that a tumour volume threshold does not differentiate between clinically insignificant or clinically significant disease, and tumour foci larger than 0.5 ml or 0.7 ml could be insignificant as well. This is an important result, because a considerable number of patients with well-differentiated, organ-confined disease have a tumour larger than 0.5 ml (in our study 47% of all patients with organ-confined prostate cancer without Gleason pattern 4/5) and are possibly unnecessarily treated. On the other hand, not all organ-confined prostate cancer with a Gleason score below 7 will be insignificant: in our study 16 (9.2%) patients showed biochemical progression. Therefore, other criteria are needed to further differentiate between insignificant and relevant disease within this selection of organ-confined prostate cancer without Gleason pattern 4/5.

The absence of additional prognostic value of tumour volume after correction for tumour stage and grade does not necessarily mean that tumour volume has no role in prostate cancer care at all. As was mentioned in the discussion of Chapter 8 as well, tumour volume may improve pre-treatment risk stratification due to the association of tumour volume with tumour stage and tumour grade at radical prostatectomy. If tumour volume could reliably be measured prior to treatment using imaging techniques, it would be valuable in future studies to assess the independent prognostic value of tumour volume after correction for other pre-treatment characteristics. Clinical tumour stage and clinical Gleason score have less prognostic value than stage and grade assessed in radical prostatectomy specimens. Therefore, hypothetically, tumour volume may indeed add prognostic value to these pre-treatment factors.

Two limitations of the study described in Chapter 9 are important to mention, even though they have already been briefly discussed in the discussion of this article as well. Firstly, the natural course of the disease has been altered due to treatment of all patients and the true clinical relevance of the tumour cannot be established anymore. Whether a tumour is clinically insignificant can only be defined at the time the untreated patient has died. At that moment it is known whether the tumour has caused symptoms or death and could have safely been left undiagnosed. This limitation is inherent to the use of radical prostatectomy specimens for the identification of indolent prostate cancer.

Secondly, another disadvantage of using prostatectomy-based criteria to identify indolent prostate cancer is that the patient needs to be treated before the tumour can be classified as possibly indolent or not. Assessing tumour aggressiveness is however most important in the pre-treatment phase when the choice for a particular treatment is made. At this point, overtreatment could be avoided by identification of indolent disease, and preferably all clinically insignificant disease, and by offering non-invasive active surveillance strategies to those cases, although this is a second choice approach. Research has aimed at predicting the presence of indolent disease at radical prostatectomy specimens with the use of pre-treatment data, like PSA, clinical tumour stage and histopathological features.^{64,65} Unfortunately, this pre-treatment risk assessment is rather inaccurate.^{66,67}

The pre-treatment identification of indolent prostate cancer

A possible method to improve the pre-treatment identification of indolent prostate cancer is the use of immunohistochemistry on prostate biopsy specimens. The use of immunohistochemistry as a marker for prostate cancer prognosis has been described for a large number of different markers in radical prostatectomy specimens. Due to the prognostic value of these markers, they might be of help in the identification of indolent disease. However, their value on biopsy specimens, or more specifically on biopsies with low-risk prostate cancer, is largely undetermined for most markers. In **Chapter 10**, the value of four promising immunomarkers was assessed in men eligible for active surveillance: EZH2, BMI-1, MIB-1 and p27^{kip1}. For the cohort selection, the entrance criteria of the PRIAS study were applied,⁶⁸ thus the study cohort consisted of patients who would actually be offered active surveillance.

Can immunohistochemistry be of help in improving the pre-treatment risk assessment? The results of our study described in Chapter 10 are promising: the use of immunohistochemistry for EZH2 and p27^{kip1} added prognostic value to the routinely used parameters like PSA and clinical tumour stage: the area under the curve for predicting significant disease at radical prostatectomy (according to the Epstein criteria²⁵) using clinicopathological data only was 0.74 and increased to 0.86 after addition of EZH2 and p27^{kip1} ($p < 0.05$). Apparently, in spite of the use of narrow criteria used in the selection of this cohort, some markers demonstrated a sufficient wide range of expression to separate high and low expression levels. On the other hand, the use of an expression cut off like the cut offs defined in Chapter 10, will misclassify relevant numbers of patients. The most ideal immunomarker would show expression in all cases of significant cancer and no expression in all patients with indolent cancer, but up to now such immunomarkers have not been discovered. The known immunomarkers may however be of help in identifying indolent prostate cancer using several markers simultaneously or in combination with other clinicopathological data, for example in nomograms. The use of multiple immunomarkers combined with clinicopathological data has been described in a Swedish watchful waiting cohort: the combination of immunomarkers and clinicopathological data (area under the curve 0.78) improved the prognostic value for prostate cancer mortality

of clinicopathological data alone (area under the curve 0.71, $p=0.04$) and seemed promising in distinguishing indolent from aggressive prostate cancer.⁶⁹

The use of immunomarkers may be hampered by differences in laboratory techniques and marker assays. It is important to consider this limitation when results from different centres are compared. However, standardizing techniques will solve this problem and general use of immunomarkers is possible, as has been shown for example for the basal cell markers 34BE12 and p63. These are widely accepted and used in case of doubt in the diagnosis of prostate cancer, especially in biopsy specimens.⁷⁰

Another, more general, limitation of the use of biopsy tissue is worth mentioning. Only a small sample of the malignant tissue is available for examination, while prostate cancer may show a very heterogenic aspect. The tumour may harbour aggressive features, which the prostate biopsy does not reveal. This is illustrated by the undergrading of prostate cancer at biopsy. In about 30% the actual Gleason score observed at radical prostatectomy, in which the whole tumour is examined, is higher than the biopsy tumour grade.⁷¹ Due to this sampling error, all characteristics assessed on biopsy specimens including immunomarkers should be interpreted with this possible sampling artefact in mind.

Future management of indolent and clinically insignificant prostate cancer

The current pre-treatment selection of men with indolent prostate cancer is rather inaccurate. This is demonstrated by considerable rates of possible cancer progression observed in patients enrolled in active surveillance studies, who were enrolled in the study because of a high probability of indolent disease. For example in the PRIAS study, a considerable percentage of 27% are excluded from the study after 2 years due to signs of progression, despite the selection criteria at entrance of the study.⁷² Clearly, this rate is also influenced by the definition of the trigger mechanisms, i.e. signs of progression which should prompt radical treatment and which still are under investigation. The inaccuracy of selection of indolent disease based on pre-treatment criteria was also demonstrated by the results of Chapter 10: 50% of patients suitable for active surveillance showed features of clinically significant disease according to the Epstein criteria.²⁵ However, it must be considered that all included patients were treated by radical prostatectomy and this selection might not be totally comparable to the overall population complying with the PRIAS criteria due to selection bias.

Data on the natural history of low risk prostate cancer are not available, firstly because it has long been standard practice to treat these tumours and secondly because indolent prostate cancer is a relatively new entity.⁷³ However, more accurate selection criteria for patients with indolent prostate cancer are needed and should not predict radical prostatectomy features, but the risk of morbidity and mortality. Currently, we are able to identify low risk prostate cancer, based on PSA and pathological characteristics. Although these cases are indeed more likely to show an indolent clinical course, still some of these low-risk cases are lethal even after

immediate treatment. The best alternative approach for assessing these criteria for indolent disease would probably be in active surveillance studies, in which low risk prostate cancers cases are closely monitored.^{68,74,75} Additional prognostic markers within this low-risk group, amongst which EZH2 and p27^{kip1} described in Chapter 10 seem promising,⁷⁶ are needed to further differentiate individual cancer behaviour.⁷³ In addition to the currently used criteria that aim at identifying indolent disease, it is essential to study other factors like age and co-morbidity to identify clinically insignificant disease as a whole. Although these patient related factors do not reflect tumour aggressiveness, they do influence the risk of prostate cancer mortality and thus are associated with clinical significance, and the need for radical treatment. For example, in an 80-year old patient a small degree of Gleason pattern 4 might be accepted and still be defined as clinically insignificant disease, as was suggested by Stamey.⁷⁷ Therefore, to avoid overtreatment, not only patients with indolent disease should be considered as suitable for active surveillance, but all patients harbouring clinically insignificant prostate cancer.

Epilogue

Prevention of morbidity and mortality caused by prostate cancer is the final goal in prostate cancer care. Screening was shown to contribute to reaching this goal, but with considerable adverse effects. A significant mortality reduction was observed due to screening for prostate cancer, but still 214 of 72,890 (0.3%) men in the screening arm died of prostate cancer compared to 326 of 89,353 (0.4%) in the control arm.¹⁴ Thus, of 5 prostate cancer patients who would have died of prostate cancer if they were not screened, still 4 died of their disease despite the offer of screening. Simultaneously, large numbers of screened men experience the downsides of screening, for example undergoing prostate biopsies or even unnecessary invasive treatment for indolent tumours. The balance between arguments for and against screening is not, yet, convincingly in favour of screening. The major challenge of optimizing the screening strategy will hopefully increase the benefits and especially decrease the harmful adverse effects. Only then, screening will become an established modality of care in the management of prostate cancer.

In this thesis it was shown that the reported mortality reduction in the screening arm indeed is caused by screening and not, or only marginally, by a treatment bias. This opens the possibility to focus on improvement of screening strategies.

Several studies aiming to improve screening strategies are described in this thesis. In short, the following conclusions are drawn:

- PSA velocity is not a useful biopsy indicator for prostate cancer in general, nor for possibly aggressive prostate cancer.

- The presence or absence of non-malignant biopsy features, like inflammation or biopsy core length, could not detect patients at increased risk for a prostate cancer diagnosis and were not useful in the individualization of screening intervals.
- The overall false-negative biopsy rate was estimated to be as low as 2.4%; 1.1% for prostate cancer and 1.3% for ASAP. This indicates accurate examination of the biopsies and high sensitivity for detecting prostate cancer during histopathological examination by the pathologist.

Screening will thus not be improved by the use of PSA velocity, neither by non-malignant biopsy features, nor by better histopathological examination. Further efforts in the search for optimal screening strategies should focus on other aspects to guide biopsy indication and individual screening intervals to improve screening for prostate cancer.

In addition to improving screening strategies, this thesis focuses on better definition and identification of indolent prostate cancer, which led to the following statements:

- Tumour volume at radical prostatectomy is not an independent prognostic factor in prostate cancer outcome. It is not useful to routinely report tumour volume in the pathology report.
- Indolent tumours may be larger than 0.7 ml.
- A tumour volume threshold in the definition of indolent prostate cancer is questionable.
- Determining the expression level of EZH2 and especially p27^{kip1} helps to identify patients that harbour aggressive tumour features among men with low-risk prostate cancer.

Although raising the tumour volume threshold could help avoid the unnecessary treatment of indolent prostate cancer, the use of any tumour volume threshold in the identification of indolent prostate cancer remains debatable. Pathological tumour stage and grade are still the most important criteria for indolent disease. More research on pre-treatment criteria for indolent disease is needed and should not focus on associations with prostatectomy characteristics, but on the risk of morbidity and mortality if tumours are left untreated. Immunohistochemistry may be of help in this pre-treatment stage, particular EZH2 and p27^{kip1} reported in this thesis, but prospective studies of these markers are needed to confirm their prognostic value. Finally, to avoid overtreatment, we should aim not only to identify indolent disease but all clinically insignificant prostate cancers. Age-specific criteria could improve this selection of clinically insignificant prostate cancer cases.

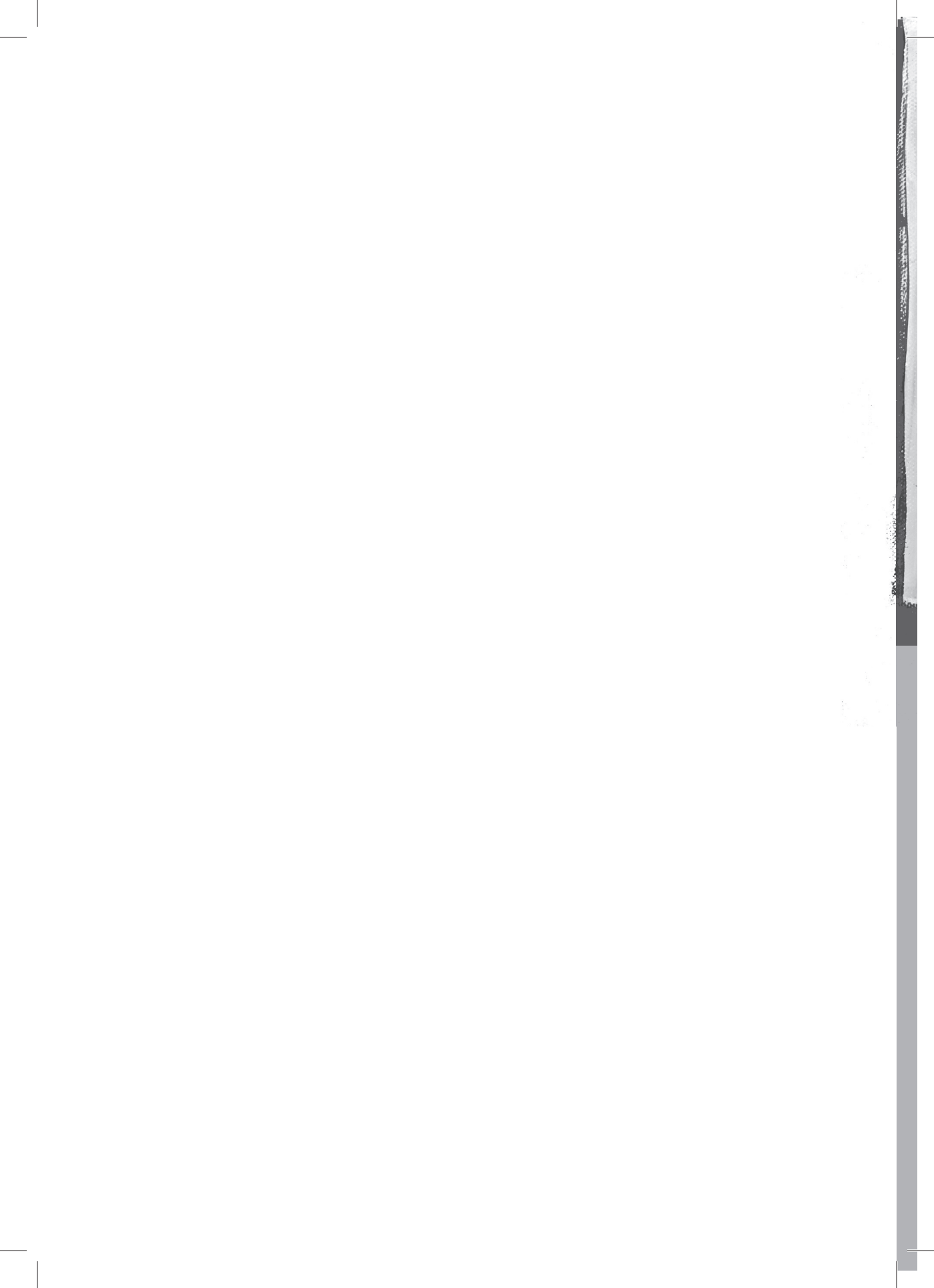
References

1. Sakr WA, Grignon JD, Heilbrun LK, Cassin BJ, Pontes JJ, Haas GP. High grade prostatic intraepithelial neoplasia (HG-PIN) and prostatic adenocarcinoma between the ages of 20-69: an autopsy study of 249 cases. *In Vivo* 1994;8:439-43
2. Sánchez-Chapado M, Olmedilla G, Cabeza M, Donat E, Ruiz A. Prevalence of prostate cancer and prostatic intraepithelial neoplasia in Caucasian Mediterranean males: an autopsy study. *Prostate* 2003;54:238-47
3. Soos G, Tsakiris I, Szanto J, Turzo C, Haas PG, Dezso B. The prevalence of prostate carcinoma and its precursor in Hungary: an autopsy study. *Eur Urol* 2005;48:739-44
4. Haas GP, Delongchamps NB, Jones RF, Chandan V, Serio AM, Vickers AJ, Jumbelic M, Threatte G, Korets R, Lilja H, de la Roza G. Needle biopsies on autopsy prostates: sensitivity of cancer detection based on true prevalence. *J Natl Cancer Inst* 2007;99:1484-9
5. Delongchamps NB, Mao K, Theng H, Zerbib M, Debré B, Peyromaure M. Outcome of patients with fortuitous prostate cancer after radical cystoprostatectomy for bladder cancer. *Eur Urol* 2005;48:946-50
6. Ruffion A, Manel A, Massoud W, Decaussin M, Berger N, Paparel P, Morel-Journel N, Lopez JG, Champetier D, Devonec M, Perrin P. Preservation of prostate during radical cystectomy: evaluation of prevalence of prostate cancer associated with bladder cancer. *J Urol* 2005;65:703-7
7. Rocco B, de Cobelli O, Leon ME, Ferruti M, Mastropasqua MG, Matei DV, Gazzano G, Verweij F, Scardino E, Musi G, Djavan B, Rocco F. Sensitivity and detection rate of a 12-core trans-perineal prostate biopsy: preliminary report. *Eur Urol* 2006;49:827-33
8. Montironi R, Mazzucchelli R, Santinelli A, Scarpelli M, López-Beltran A, Bostwick DG. Incidentally detected prostate cancer in cystoprostatectomies : pathological and morphometric comparison with clinically detected cancer in totally embedded specimens. *Human Pathology* 2005;36:646-54
9. Pettus JA, Al-Ahmadie H, Barocas DA, Koppie TM, Herr H, Donat SM, Dalbagni G, Reuter VE, Olgac S, Bochner B. Risk assessment of prostatic pathology in patients undergoing radical cystoprostatectomy. *Eur Urol* 2008;53:370-5
10. Revelo MP, Cookson MS, Chang SS, Shook MF, Smith JA, Shappell SB. Incidence and location of prostate and urothelial carcinoma in prostates from cystoprostatectomies: implications for possible apical sparing surgery. *J Urol* 2004;171:646-51
11. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Copen DE, Yuan JJ, Petros JA, Andriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991;324:1156-61
12. Catalona WJ, Smith DS, Ornstein DK. Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/mL and benign prostate examination. Enhancement of specificity with free PSA measurements. *JAMA* 1997;277:1452-5
13. Hodge KK, McNeal JE, Terris MK, Stamey TA. Random systematic versus directed ultrasound guided transrectal core biopsies of the prostate. *J Urol.* 1989;142:71-4
14. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Määttänen L, Bangma CH, Aus G, Villers A, Rebillard X, Van der Kwast ThH, Blijenberg BG, Moss SM, De Koning HJ, Auvinen A. Screening and prostate-cancer mortality in a randomized European Study. *N Engl J Med* 2009;360:1320-8
15. Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, Fouad MN, Gelmann EP, Kvale PA, Reding DJ, Weissfeld JL, Yokochi LA, O'Brien B, Clapp JD, Rathmell JM, Riley TL, Hayes RB, Kramer BS, Izmirlian G, Miller AB, Pinsky PF, Prorok PC, Gohagan JK, Berg CD; PLCO Project Team. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009;360:1310-9
16. Postma R, van Leenders AG, Roobol MJ, Schröder FH, van der Kwast TH. Tumour features in the control and screening arm of a randomized trial of prostate cancer. *Eur Urol* 2006;50:70-5
17. Bolla M, Collette L, Blank L, Warde P, Dubois JB, Mirimanoff RO, Storme G, Bernier J, Kuten A, Sternberg C, Mattelaer J, Lopez Torecilla J, Pfeffer JR, Lino Cutajar C, Zurlo A, Pierart M. Long-term results with immediate androgen suppression and external irradiation in patients with locally advanced prostate cancer (an EORTC study): a phase III randomised trial. *Lancet* 2002;360:103-6.

18. Widmark A, Klepp O, Solberg A, Damber JE, Angelsen A, Fransson P, Lund JA, Tasdemir I, Hoyer M, Wiklund F, Fosså SD; Scandinavian Prostate Cancer Group Study 7; Swedish Association for Urological Oncology 3. Endocrine treatment, with or without radiotherapy, in locally advanced prostate cancer (SPCG-7/SFUO-3): an open randomised phase III trial. *Lancet* 2009;373:301-8
19. Raaijmakers R, Wildhagen MF, Ito K, Pàez A, de Vries SH, Roobol MJ, Schröder FH. Prostate-specific antigen change in the European Randomized Study of Screening for Prostate Cancer, section Rotterdam. *Urology* 2004;63:316-20
20. Stamey TA, Caldwell M, McNeal JE, Nolley R, Hemenez M, Downs J. The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? *J Urol.* 2004;172:1297-301
21. Prostate Risk Indicator (based on ERSPC). Available at: <http://www.prostate-riskindicator.com/via.html>.
22. Carter HB, Pearson JD, Metter EJ, Brant LJ, Chan DW, Andres R, Fozard JL, Walsh PC. Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. *JAMA.* 1992;267:2215-20
23. d'Amico AV, Chen MH, Roehl KA, Catalona WJ. Preoperative PSA velocity and the risk of death from prostate cancer after radical prostatectomy. *N Engl J Med.* 2004;351:125-35
24. d'Amico AV, Renshaw AA, Sussman B, Chen MH. Pretreatment PSA velocity and risk of death from prostate cancer following external beam radiation therapy. *JAMA* 2005;294:440-7
25. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumour extent of nonpalpable (stage T1c) prostate cancer. *JAMA* 1994;271:368-74
26. Vickers AJ, Wolters T, Savage CJ, Cronin AM, O'Brien MF, Pettersson K, Roobol MJ, Aus G, Scardino PT, Hugosson J, Schröder FH, Lilja H. Prostate-Specific Antigen Velocity for Early Detection of Prostate Cancer: Result from a Large, Representative, Population-based Cohort. *Eur Urol.* 2009, epub ahead of print, doi: 10.1016/j.eururo.2009.07.047
27. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA Jr. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst.* 2006;98:529-34
28. Roobol MJ, Kranse R, de Koning HJ, Schröder FH. Prostate-specific antigen velocity at low prostate-specific antigen levels as screening tool for prostate cancer: results of second screening round of ERSPC (ROTTERDAM). *Urology* 2004;63:309-13
29. Schröder FH, Carter HB, Wolters T, van den Bergh RC, Gosselaar C, Bangma CH, Roobol MJ. Early detection of prostate cancer in 2007. Part I: PSA and PSA kinetics. *Eur Urol* 2008;53:468-77
30. Jansen FH, Roobol M, Jenster G, Schröder FH, Bangma CH. Screening for prostate cancer in 2008 II: the importance of molecular subforms of prostate-specific antigen and tissue kallikreins. *Eur Urol* 2009;55:563-74
31. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644-8
32. Hessels D, Schalken J. The use of PCA3 in the diagnosis of prostate cancer. *Nat Rev Urol* 2009;6:255-61
33. Haese A, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PF, Huland H, Abbou CC, Remzi M, Tinzl M, Feyerabend S, Stillebroer AB, van Gils MP, Schalken JA. Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. *Eur Urol* 2008;54:1081-8
34. Panteleakou Z, Lembessis P, Sourla A, Pissimissis N, Polyzos A, Deliveliotis C, Koutsilieris M. Detection of circulating tumour cells in prostate cancer patients: methodological pitfalls and clinical relevance. *Mol Med.* 2009;15:101-14
35. Jansen FH, Krijgsveld J, van Rijswijk A, van den Bemd GJ, van den Berg MS, van Weerden WM, Willemsen R, Dekker LJ, Luijckx TM, Jenster G. Exosomal secretion of cytoplasmic prostate cancer xenograft-derived proteins. *Mol Cell Proteomics.* 2009;8:1192-205
36. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. *Clin Cancer Res* 2007;13:5103-8

37. Roobol MJ, Roobol DW, Schröder FH. Is additional testing necessary in men with prostate-specific antigen levels of 1.0 ng/mL or less in a population-based screening setting? (ERSPC, section Rotterdam). *Urology* 2005;65:343-6
38. Vis AN, Hoedemaeker RF, Roobol M, Schröder FH, van der Kwast TH. The predictive value for prostate cancer of lesions that raise suspicion of concomitant carcinoma: an evaluation from a randomized, population-based study of screening for prostate cancer. *Cancer* 2001;92:524-34
39. Postma R, Roobol M, Schröder FH, van der Kwast TH. Lesions predictive for prostate cancer in a screened population: first and second screening round findings. *Prostate*. 2004;61:260-6
40. Merrimen JL, Jones G, Walker D, Leung CS, Kapusta LR, Srigley JR. Multifocal high grade prostatic intraepithelial neoplasia is a significant risk factor for prostatic adenocarcinoma. *J Urol* 2009;182:485-9
41. De Nunzio C, Trucchi A, Miano R, Stoppacciaro A, Fattahi H, Cicione A, Tubaro A. The number of cores positive for high grade prostatic intraepithelial neoplasia on initial biopsy is associated with prostate cancer on second biopsy. *J Urol* 2009;181:1069-74
42. Epstein JI, Herawi M. Prostate needle biopsies containing prostatic intraepithelial neoplasia or atypical foci suspicious for carcinoma: implications for patient care. *J Urol* 2006;175:820-34
43. Wolters T, Van Leenders GJHL, Roobol MJ, Van den Bergh RCN, Van Leeuwen PJ, Bangma CH, Schröder FH. How to handle a minor focus of high-grade prostatic intraepithelial neoplasia (HG-PIN) on prostate needle biopsy in a screening setting? *Eur Urol Suppl* 2009;8:219
44. Lilja H, Ulmert D, Björk T, Becker C, Serio AM, Nilsson JA, Abrahamsson PA, Vickers AJ, Berglund G. Long-term prediction of prostate cancer up to 25 years before diagnosis of prostate cancer using prostate kallikreins measured at age 44 to 50 years. *J Clin Oncol*. 2007;25:431-6.
45. Ulmert D, Cronin AM, Björk T, O'Brien MF, Scardino PT, Eastham JA, Becker C, Berglund G, Vickers AJ, Lilja H. Prostate-specific antigen at or before age 50 as a predictor of advanced prostate cancer diagnosed up to 25 years later: a case-control study. *BMC Med*. 2008;6:6.
46. Froehner M, Koch R, Litz RJ, Oehlschlaeger S, Twelker L, Hakenberg OW, Wirth MP. Detailed analysis of Charlson comorbidity score as predictor of mortality after radical prostatectomy. *Urology*. 2008;72:1252-7
47. Chen YC, Page JH, Chen R, Giovannucci E. Family history of prostate and breast cancer and the risk of prostate cancer in the PSA era. *Prostate* 2008;68:1582-91
48. Roemeling S, Roobol MJ, de Vries SH, Gosselaar C, van der Kwast TH, Schröder FH. Prevalence, treatment modalities and prognosis of familial prostate cancer in a screened population. *J Urol* 2006;175:1332-6
49. Evans S, Metcalfe C, Ibrahim F, Persad R, Ben-Shlomo Y. Investigating Black-White differences in prostate cancer prognosis: A systematic review and meta-analysis. *Int J Cancer* 2008;123:430-5
50. Roobol MJ, Steyerberg EW, Kranse R, Wolters T, van den Bergh RC, Bangma CH, Schröder FH. A Risk-Based Strategy Improves Prostate-Specific Antigen-Driven Detection of Prostate Cancer. *Eur Urol*. 2009, epub ahead of print, doi:10.1016/j.eururo.2009.08.025
51. Cancer Risk Calculator for Prostate Cancer (Based on the PCPT). Available at: <http://www.compass.fhrc.org/edrnci/bin/calculator/main.asp>.
52. Thompson IM, Goodman PJ, Tangen CM, Lucia MS, Miller GJ, Ford LG, Lieber MM, Cespedes RD, Atkins JN, Lippman SM, Carlin SM, Ryan A, Szczepanek CM, Crowley JJ, Coltman CA Jr. The influence of finasteride on the development of prostate cancer. *N Engl J Med*. 2003;349:215-24
53. Andriole G, Bostwick D, Brawley O, Gomella L, Marberger M, Tindall D, Breed S, Somerville M, Rittmaster R; REDUCE Study Group. Chemoprevention of prostate cancer in men at high risk: rationale and design of the reduction by dutasteride of prostate cancer events (REDUCE) trial. *J Urol* 2004;172:1314-7
54. Pallwein L, Mitterberger M, Pelzer A, Bartsch G, Strasser H, Pinggera GM, Aigner F, Gradl J, Zur Nedden D, Frauscher F. Ultrasound of prostate cancer: recent advances. *Eur Radiol* 2008;18:707-15
55. Puech P, Huglo D, Petyt G, Lemaitre L, Villers A. Imaging of organ-confined prostate cancer: functional ultrasound, MRI and PET/computed tomography. *Curr Opin Urol* 2009;19:168-76
56. Villers A, Lemaitre L, Haffner J, Puech P. Current status of MRI for the diagnosis, staging and prognosis of prostate cancer: implications for focal therapy and active surveillance. *Curr Opin Urol* 2009;19:274-82

57. Cooperberg MR, Lubeck DP, Meng MV, Mehta SS, Carroll PR. The changing face of low-risk prostate cancer: trends in clinical presentation and primary management. *J Clin Oncol*. 2004;22:2141-9
58. Stamey TA, Freiha FS, McNeal JE, Redwine EA, Whittemore AS, Schmid HP. Localized prostate cancer. Relationship of tumour volume to clinical significance for treatment of prostate cancer. *Cancer*. 1993;71(3 Suppl):933-8
59. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, Partin AW. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. *JAMA* 2005;294:433-9
60. Han M, Partin AW, Zahurak M, Piantadosi S, Epstein JI, Walsh PC. Biochemical (prostate specific antigen) recurrence probability following radical prostatectomy for clinically localized prostate cancer. *J Urol* 2003;169:517-23
61. Stamey TA, McNeal JE, Yemoto CM, Sigal BM, Johnstone IM. Biological determinants of cancer progression in men with prostate cancer. *JAMA* 1999;281:1395-1400
62. McNeal JE, Bostwick DG, Kindrachuk RA, Redwine EA, Freiha FS, Stamey TA. Patterns of progression in prostate cancer. *Lancet* 1986;1:60-3
63. Epstein JI, Carmichael M, Partin AW, Walsh PC. Is tumour volume an independent predictor of progression following radical prostatectomy? A multivariate analysis of 185 clinical stage B adenocarcinomas of the prostate with 5 years of follow-up. *J Urol* 1993;149:1478-81
64. Epstein JI, Chan DW, Sokoll LJ, Walsh PC, Cox JL, Rittenhouse H, Wolfert R, Carter HB. Nonpalpable stage T1c prostate cancer: prediction of insignificant disease using free/total prostate specific antigen levels and needle biopsy findings. *J Urol* 1998;160:2407-11
65. Steyerberg EW, Roobol MJ, Kattan MW, van der Kwast TH, de Koning HJ, Schröder FH. Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. *J Urol* 2007;177:107-12
66. Thorson P, Humphrey PA. Minimal adenocarcinoma in prostate needle biopsy tissue. *Am J Clin Pathol* 2000;114:896-909
67. Schlegel PN. "Limited" prostate cancer on biopsy: a note of caution. *Cancer* 2008;112:966-7
68. van den Bergh RC, Roemeling S, Roobol MJ, Roobol W, Schröder FH, Bangma CH. Prospective validation of active surveillance in prostate cancer: the PRIAS study. *Eur Urol* 2007;52:1560-3
69. Mucci LA, Pawitan Y, Demichelis F, Fall K, Stark JR, Adami HO, Andersson SO, Andrén O, Eisenstein A, Holmberg L, Huang W, Kantoff PW, Kim R, Perner S, Stampfer MJ, Johansson JE, Rubin MA. Testing a multigene signature of prostate cancer death in the Swedish Watchful Waiting Cohort. *Cancer Epidemiol Biomarkers Prev* 2008;17:1682-870
70. Humphrey PA. Diagnosis of adenocarcinoma in prostate needle biopsy tissue. *J Clin Pathol* 2007;60:35-42
71. Cohen MS, Hanley RS, Kurteva T, Ruthazer R, Silverman ML, Sorcini A, Hamawy K, Roth RA, Tuerk I, Libertino JA. Comparing the Gleason prostate biopsy and Gleason prostatectomy grading system: the Lahey Clinic Medical Center experience and an international meta-analysis. *Eur Urol* 2008;54:371-81.
72. Van den Bergh RC. Short-term outcomes of the prospective multi-centre PRIAS study (Prostate Cancer Research International: Active Surveillance). *BJU Int* 2009, accepted for publication
73. O'Donnell H, Parker C. What is low-risk prostate cancer and what is its natural history? *World J Urol*. 2008;26:415-22
74. Choo R, Klotz L, Danjoux C, Morton GC, DeBoer G, Szumacher E, Fleshner N, Bunting P, Hruby G. Feasibility study: watchful waiting for localized low to intermediate grade prostate carcinoma with selective delayed intervention based on prostate specific antigen, histological and/or clinical progression. *J Urol* 2002;167:1664-9
75. Carter HB, Kettermann A, Warlick C, Metter EJ, Landis P, Walsh PC, Epstein JI. Expectant management of prostate cancer with curative intent: an update of the Johns Hopkins experience. *J Urol* 2007;178:2359-64
76. Samarasinghe H, Epstein JI. What is the molecular pathology of low-risk prostate cancer? *World J Urol*. 2008;26:431-6
77. Stamey TA. Editorial: More information on prostate specific antigen and prostate cancer. *J Urol* 2003;170:457-8





CHAPTER 12

Summary

Samenvatting (Nederlands)

Curriculum vitae

Dankwoord

List of publications

Appendix: color images

PhD portfolio

Summary

The European Randomized Study of Screening for Prostate Cancer (ERSPC) has shown that screening for prostate cancer reduces prostate cancer specific mortality. In Part I of this thesis, background information on prostate cancer (**Chapter 1**) and the ERSPC (**Chapter 2**) is provided. Some issues concerning screening for prostate cancer currently prevent the population-wide introduction of screening. The scope of this thesis, outlined in **Chapter 3**, was to improve screening strategies by addressing the following issues:

1. the role of study arm on treatment choice
2. improving the current screening protocol
3. improving the identification of indolent prostate cancer to prevent overtreatment

In **Chapter 4** (Part II) was shown that study arm played a statistically significant but minor role in treatment selection. This small difference in treatment between the study arms of the ERSPC was particularly observed in patients with high-risk PC: a control subject was more likely to receive radiotherapy, active surveillance or hormonal therapy instead of radical prostatectomy than a screening subject, but no major differences in other treatment choices were seen. This indicates that an effect of different treatment between arms on PC mortality may be possible but probably will be small. Therefore, these results show that a mortality reduction in the ERSPC based solely on unequal treatment in both arms is very unlikely.

Part III focuses on the improvement of the current screening protocols. The value of PSA velocity in a screening setting was assessed in **Chapter 5**. In our screened cohort, PSA velocity was not an independent predictor of clinically relevant prostate cancer. Using a PSA velocity cut-off level as a biopsy indicator would miss an important proportion of clinically relevant prostate cancer cases. Therefore, PSA velocity could not improve the ERSPC screening protocol.

The predictive value for prostate cancer detection in the subsequent screening round was evaluated for high-grade prostatic intraepithelial neoplasia (HG-PIN), chronic and active inflammation and biopsy core length and glandular core length (**Chapter 6**). In this study, none of these features proved to be significantly predictive of subsequent prostate cancer detection and therefore no high or low risk subgroups of PC-detection 4 years later, which could improve screening protocols, could be identified.

In **Chapter 7**, adenocarcinoma was identified in 16 (8.2%) patients and atypical lesions suspicious for malignancy in 19 patients (9.7%) in a cohort of 196 screening participants whose biopsy specimens were originally classified as benign. After correction for patient selection was estimated that in 2.4% of all biopsied screening participants a (possibly) malignant lesion was missed during histopathological examination of the biopsy specimens; 1.1% for prostate cancer and 1.3% for atypical small lesions suspicious for malignancy. Men with a false-negative biopsy seemed to be in a curable stage at the time of PC diagnosis and the low overall rate of false-negative biopsies indicates accurate examination of the biopsy specimens.

In Part IV, an effort is made to improve the differentiation between indolent prostate cancer and clinically relevant prostate cancer. First, we showed that tumour volume has no independent prognostic value in men treated with radical prostatectomy (**Chapter 8**), after correction for tumour stage, grade and surgical margin status. Additionally to the lack of independent prognostic value, tumour volume is a time-consuming measurement and is highly dependent on laboratory processes and methods of measurement. We therefore recommend that tumour volume should not routinely be reported in the pathological reports.

The lack of independent prognostic value of tumour volume resulted in a following study of tumour volume. In **Chapter 9**, the commonly used tumour volume threshold for indolent prostate cancer of 0.5 ml was reassessed in a screened population and the additional predictive value of a tumour volume cut-off for the identification of indolent prostate cancer was assessed in addition to stage and grade criteria. A threshold of 0.7 ml was observed, but tumour volume was not a useful criterion for the identification of biochemical progression in patients with organ-confined PC without Gleason pattern 4/5: patients with larger tumours were not at increased risk of biochemical progression. This suggests that a tumour volume cut off may not be effective in the identification of indolent prostate cancer in addition to stage and grade criteria.

Finally, we aimed to improve the pretreatment identification of indolent prostate cancer, using immunohistochemistry on prostate biopsy specimens containing low-risk prostate cancer (**Chapter 10**). Four immunohistochemical stainings, which all have predictive value for patient outcome when analyzed on radical prostatectomy specimens, were evaluated on biopsy specimens: EZH2, p27^{kip1}, BMI-1 and MIB-1. Although MIB-1 and BMI-1 did not show predictive value for the presence of relevant disease on radical prostatectomy specimens, p27^{kip1} and EZH2 were independent predictors of relevant prostate cancer. Therefore, these markers may improve the pretreatment identification of patients with indolent prostate cancer.

The fifth part of this thesis summarizes and discusses the studies described in the previous chapters and puts them into perspective.

Samenvatting

De European Randomized Study of Screening for Prostate Cancer (ERSPC) heeft aangetoond dat vroege opsporing, of screening, naar prostaatkanker de sterfte aan deze ziekte vermindert. In Deel I van dit proefschrift wordt achtergrondinformatie beschreven over prostaatkanker (**Hoofdstuk 1**) en de ERSPC (**Hoofdstuk 2**). Onduidelijkheden en negatieve neveneffecten van screening naar prostaatkanker verhinderen op dit moment de invoering van landelijke screeningsprogramma's. Het doel van dit proefschrift, welke wordt beschreven in **Hoofdstuk 3**, was de verbetering van de screeningsstrategieën door enkele van deze problemen omtrent screening te belichten:

1. De rol van studiearm in de keuze van prostaatkankerbehandeling
2. De verbetering van het huidige screeningsprotocol
3. De verbetering van identificatie van indolente prostaattumoren

In **Hoofdstuk 4** (Deel II) is aangetoond dat studiearm een statistisch significante, maar zeer kleine rol speelt in de therapiekeuze voor prostaatkanker binnen de ERSPC. Dit verschil werd met name gezien bij patiënten met hoogrisico prostaatkanker: een deelnemer in de controlearm had een grotere kans om behandeld te worden met radiotherapie, active surveillance of hormonale therapie dan met radicale prostatectomie vergeleken met een deelnemer in de interventiearm. Er werden geen verschillen in de overige therapiekeuzes gezien. Dit betekent dat een effect van verschil in behandeling tussen beide studiearmen op de prostaatkankermortaliteit mogelijk is, maar hoogstwaarschijnlijk zeer klein zal zijn. Een mortaliteitsreductie in de ERSPC op basis van ongelijke behandeling alleen is onwaarschijnlijk.

Deel III richt zich op de verbetering van de huidige screeningsprotocollen. De waarde van PSA velocity in een screeningssituatie wordt beschreven in **Hoofdstuk 5**. PSA velocity was geen onafhankelijke voorspeller van klinisch relevante prostaatkanker in een gescreend cohort. Het gebruik van een PSA velocity grenswaarde als indicatie voor het nemen van prostaatbiopten zou een aanzienlijk deel van de klinisch relevante prostaatkankergevallen missen. Het gebruik van PSA velocity leidde niet tot verbetering van het huidige ERSPC-screeningsprotocol.

De voorspellende waarde voor prostaatkankerdetectie in de volgende screeningsronde werd onderzocht voor de aanwezigheid van de volgende parameters in prostaatbiopten: high-grade prostatic intraepithelial neoplasia (HG-PIN), chronische en actieve ontsteking, totale bioptlengte en de bioptlengte van het deel dat daadwerkelijk glandulair weefsel bevat (**Hoofdstuk 6**). Geen van de onderzochte parameters was voorspellend voor prostaatkankerdetectie in tijdens het volgende screeningsonderzoek en er konden geen hoog- of laagrisicogroepen worden geïdentificeerd. Verbetering van het screeningsprotocol mogelijk lijkt niet haalbaar door gebruik van deze parameters.

Hoofdstuk 7 beschrijft de detectie van adenocarcinoom in 16 (8.2%) patiënten en atypische laesies verdacht voor maligniteit (ASAP) in 19 patiënten (9.7%) in een cohort van 196 deelnemers wiens bipten initieel werden geclassificeerd als benigne. Na correctie voor de patiëntselectie werd geschat dat in 2.4% van alle gebiopteerde deelnemers een (mogelijk) maligne laesie was gemist tijdens het histopathologisch onderzoek: in 1.1% een maligne laesie en in 1.3% een atypische laesie verdacht voor maligniteit. De patiënten met een gemiste laesie in een eerder prostaatbipt leken allen in een curatief ziektestadium te verkeren op het moment van daadwerkelijk prostaatankerdiagnose. De lage frequentie van vals-negatieve bipten duidt op een nauwkeurig histopathologisch onderzoek van de prostaatbipten.

Deel IV van dit proefschrift richt zich op de verbetering van de differentiatie tussen indolente en klinisch relevante prostaatanker. Allereerst werd aangetoond dat tumourvolume geen onafhankelijke prognostische factor is in mannen die behandeld zijn met radicale prostatectomie (**Hoofdstuk 8**), na correctie voor tumourstadium, graad en snijvlakstatus. Naast het gebrek aan onafhankelijke voorspellende waarde is tumourvolume een zeer tijdrovende meting en bestaat er grote variatie in gemeten volumes door verschil in laboratoriumtechnieken en meetmethodes. Ons advies is daarom om tumourvolume niet standaard op te nemen in het pathologieverslag.

Het gebrek aan onafhankelijke voorspellende waarde van tumourvolume is aanleiding geweest voor een tweede studie naar tumourvolume. In **Hoofdstuk 9** werd de traditionele tumourvolume-grenswaarde voor indolente kanker herbepaald in een gescreend cohort en de toegevoegde waarde van een tumourvolume-criterium in de definitie van indolente prostaatanker geëvalueerd. Een grenswaarde van 0.7 ml werd geobserveerd, echter, tumourvolume was geen voorspeller voor biochemische progressie in een cohort patiënten die voldeden aan de andere twee criteria voor indolente prostaatanker: geen extracapsulaire groei en geen Gleasonpatroon 4 of 5. Patiënten met een klein tumourvolume hadden evenveel risico op progressie als patiënten met een grotere tumour. Deze resultaten trekken de effectiviteit van een tumourvolumecriterium in de definitie van indolente tumouren in twijfel.

Tenslotte wordt in **Hoofdstuk 10** een studie beschreven waarin wordt getracht de identificatie van indolente prostaattumouren te verbeteren op het moment van diagnose, vóór behandeling heeft plaatsgevonden. Hierbij is gebruikt gemaakt van immunohistochemische markers op prostaatbipten met laagrisico prostaatanker. Vier kleuringen, welke allemaal voorspellende waarde hebben voor prognose indien zij gebruikt worden op prostatectomiemateriaal, werden getest op biptiemateriaal: EZH2, p27^{kip1}, BMI-1 and MIB-1. Hoewel MIB-1 en BMI-1 geen voorspellende waarde toonden, bleken p27^{kip1} en EZH2 onafhankelijke voorspellers te zijn voor de aanwezigheid van klinische relevante tumouren in het radicale prostatectomiepreparaat. Deze markers zouden de identificatie van indolente tumouren vóór behandeling kunnen verbeteren.

In het vijfde deel van dit proefschrift worden de resultaten van de voorgaande hoofdstukken samengevat en bediscussieerd.



Curriculum vitae

Tineke Wolters is geboren op 27 oktober 1981 in Gouda. Nadat zij haar VWO had afgerond aan de Waezenburg in Leek, begon zij in 1999 met de studie geneeskunde aan de Rijksuniversiteit te Groningen. De co-schappen liep zij in het Martiniziekenhuis in Groningen en gedurende 3 maanden een keuzecoschap urologie in het Universitair Medisch centrum Groningen, onder begeleiding van dr. I.J. de Jong. In 2005 behaalde zij cum laude haar artsenbul. Na enkele maanden te hebben rondgereisd door Midden-Amerika, werkte Tineke een half jaar als arts-assistent op de spoedeisende hulp van het Delfzichtziekenhuis in Delfzijl. In juli 2006 begon zij met haar promotietraject op het screeningsbureau

Urologie van het Erasmus MC in Rotterdam, onder leiding van professor Schröder en Monique Roobol, Arno van Leenders en professor Bangma. Op 1 januari 2010 is Tineke begonnen met haar heelkundige vooropleiding in het Reinier de Graaf Gasthuis in Delft en zij zal in 2012 starten met haar opleiding Urologie aan het Erasmus MC te Rotterdam.

List of publications

1. Roemeling S, Roobol MJ, de Vries SH, **Wolters T**, Gosselaar C, van Leenders GJ, Schröder FH. Active surveillance for prostate cancers detected in three subsequent rounds of a screening trial: characteristics, PSA doubling times, and outcome. *Eur Urol.* 2007 May;51(5):1244-50
2. Schröder FH, Carter HB, **Wolters T**, van den Bergh RC, Gosselaar C, Bangma CH, Roobol MJ. Early detection of prostate cancer in 2007. Part 1: PSA and PSA kinetics. *Eur Urol.* 2008;53:468-77. Review.
3. **Wolters T**, Roobol MJ, Schröder FH, Van der Kwast TH, Roemeling S, Van der Crujssen-Koëter IW, Bangma CH, Van Leenders GJLH; ERSPC Rotterdam section. Can non-malignant biopsy features identify men at increased risk of biopsy-detectable prostate cancer at re-screening after 4 years? *BJU Int* 2008;101:283-9
4. Van den Bergh RC, Roobol MJ, **Wolters T**, van Leeuwen PJ, Schröder FH. The Prostate Cancer Prevention Trial and European Randomized Study of Screening for Prostate Cancer risk calculators indicating a positive prostate biopsy: a comparison. *BJU Int.* 2008;102:1068-73
5. Van den Bergh RC, Roemeling S, Roobol MJ, **Wolters T**, Schröder FH, Bangma CH. Prostate-specific antigen kinetics in clinical decision-making during active surveillance for early prostate cancer--a review. *Eur Urol.* 2008;54:505-16 Review.
6. Gosselaar C, Roobol MJ, Roemeling S, **Wolters T**, van Leenders GJ, Schröder FH. The value of an additional hypoechoic lesion-directed biopsy core for detecting prostate cancer. *BJU Int.* 2008;101:685-90
7. Gosselaar C, Roobol MJ, van den Bergh RC, **Wolters T**, Schröder FH. Digital Rectal Examination and the Diagnosis of Prostate Cancer-a Study Based on 8 Years and Three Screenings within the European Randomized Study of Screening for Prostate Cancer (ERSPC), Rotterdam. *Eur Urol.* 2009;55:139-47
8. **Wolters T**, Roobol MJ, Bangma CH, Schröder FH. Is prostate-specific antigen velocity selective for clinically significant prostate cancer in screening? *Eur Urol* 2009;55:385-92
9. **Wolters T**, Roobol MJ, Steyerberg EW, Van den Bergh RCN, Bangma CH, Hugosson J, Ciatto S, Kwiatkowski M, Villers A, Luján M, Nelen V, Tammela TLJ and Schröder FH. The effect of study arm on prostate cancer treatment in a large screening trial (ERSPC). *Int J Cancer.* 2009, *epub ahead of print*, doi:10.1002/ijc.24870
10. **Wolters T**, Roobol MJ, Van Leeuwen PJ, Van den Bergh RCN, Hoedemaeker RF, Van Leenders GJLH, Schröder FH, Van der Kwast ThH. Should prostate tumour volume routinely be reported by the pathologist?: The prognostic value of tumour volume in prostate cancer. *Eur Urol* 2009, *epub ahead of print*, doi:10.1016/j.eururo.2009.07.027

11. **Wolters T**, Vissers CJ, Bangma CH, Schröder FH, Van Leenders GJLH. EZH2, p27^{kip1} and MIB-1 on Gleason 6 biopsy specimens are valuable in selecting men with indolent prostate cancer at prostatectomy. *BJU Int* 2009, *epub ahead of print*, doi:10.1111/j.1464-410x.2009.08998
12. **Wolters T**, Van der Kwast ThH, Vissers CJ, Bangma CH, Roobol MJ, Schröder FH, Van Leenders GJLH. False-negative prostate needle-biopsies: frequency, histopathologic features and follow-up. *Am J Surg Pathol*. 2009 ,*epub ahead of print*, doi: 10.1097/PAS.0b013e3181c3ece9
13. **Wolters T**, Roobol MJ, Van Leeuwen PJ, Van den Bergh RCN, Hoedemaeker RF, Schröder FH, Van der Kwast ThH. Reassessment of the tumour volume threshold defining clinically insignificant prostate cancer; results from a randomized trial. *Am J Surg Pathol* 2010;34:35-43
14. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Määtänen L, Bangma CH, Aus G, Villers A, Rebillard X, van der Kwast T, Blijenberg BG, Moss SM, de Koning HJ, Auvinen A; **ERSPC Investigators**. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360:1320-8
15. Van den Bergh RC, Essink-Bot ML, Roobol MJ, **Wolters T**, Schröder FH, Bangma CH, Steyerberg EW. Anxiety and distress during active surveillance for early prostate cancer. *Cancer*. 2009;115:3868-78
16. Van den Bergh RC, **Wolters T**, Spaander MC, Schröder FH, van Leenders GJ. Non-prostatic pathology on prostate needle-biopsy - colorectal carcinoid: a case report. *Cases J*. 2009 21;2:75
17. Van Leeuwen P, van den Bergh R, **Wolters T**, Schröder F, Roobol M. Screening: should more biopsies be taken in larger prostates? *BJU Int*. 2009, *epub ahead of print*, doi: 10.1111/j.1464-410X.2009.08627
18. Vickers AJ, **Wolters T**, Savage CJ, Cronin AM, O'Brien MF, Pettersson K, Roobol MJ, Aus G, Scardino PT, Hugosson J, Schröder FH, Lilja H. Prostate-Specific Antigen Velocity for Early Detection of Prostate Cancer: Result from a Large, Representative, Population-based Cohort. *Eur Urol*. 2009, *epub ahead of print*, doi: 10.1016/j.eururo.2009.07.047
19. Roobol MJ, Steyerberg EW, Kranse R, **Wolters T**, van den Bergh RC, Bangma CH, Schröder FH. A Risk-Based Strategy Improves Prostate-Specific Antigen-Driven Detection of Prostate Cancer. *Eur Urol*. 2009, *epub ahead of print*, doi:10.1016/j.eururo.2009.08.025
20. Van den Bergh RCN, Steyerberg EW, Khatami A, Aus G, Phil CG, **Wolters T**, Van Leeuwen PJ, Roobol MJ, Schröder FH, Hugosson J. Is delayed radical prostatectomy in men with early screen-detected prostate cancer associated with a higher risk of unfavourable outcomes? *Cancer* 2010, *epub ahead of print*, doi: 10.1002/cncr.24882

21. Van den Bergh RCN, Steyerberg EW, Van Leeuwen PJ, **Wolters T**, Aus G, Roobol MJ, Bangma CH, Hugosson J, Schröder FH. Rule-based versus probabilistic selection for active surveillance for prostate cancer and outcomes after radical prostatectomy. *Submitted*

Dankwoord

Allereerst gaat mijn dank uit naar alle mannen die vrijwillig hebben meegewerkt aan de ERSPC. Zonder hen zou onze kennis over prostaatkanker en vroegopsporing naar deze ziekte een stuk beperkter zijn. Bovendien zou dit proefschrift hier niet liggen.

Dank aan mijn beide promotoren:

Professor Schröder, dank u voor het enthousiasme dat u zo goed weet over te brengen. Het werkt inspirerend en aanstekelijk en is een goede drijfveer geweest tijdens mijn promotie-onderzoek. Ik ben trots zo nauw met u te hebben samengewerkt.

Professor Bangma, hartelijk dank voor het vertrouwen dat u in mij heeft gehad. Ik vind het fantastisch straks mijn opleiding Urologie bij u in Rotterdam te mogen volgen.

Ook mijn co-promotoren wil ik hartelijk danken voor hun enthousiaste inzet.

Monique, je hebt me de ‘ins en outs’ van de ERSPC leren kennen. Zoals bij vele van mijn collega-promovendi was jouw begeleiding essentieel voor mijn promotietraject (en voor de gehele ERSPC overigens). Dank je wel.

Arno, ik had tevoren niet gedacht dat de pathologie daadwerkelijk interessant en zelfs spannend kon zijn. Jouw passie voor je vak is mooi om mee te maken. Dank je wel (ook voor de cappuccino's!).

De overige leden van mijn promotiecommissie wil ik hartelijk danken. Professor Mulders en professor Kiemeny, dank u voor het lezen en beoordelen van mijn proefschrift.

Professor Van der Kwast, bedankt voor de fantastische samenwerking. Het schrijven van een artikel wordt een feestje met uw begeleiding. Ik ben vereerd dat u uit Canada komt om zitting te nemen in mijn promotiecommissie.

Annelies, zussie, geweldig om jou naast me hebben als paranimf! Ik ben ontzettend blij met zo'n zusje.

Pim, jouw humor en eerlijkheid zijn onontbeerlijk in kamer NH-227 (en het hele screeningsbureau). Ik ben trots je dat je mijn paranimf bent!

Dank aan het hele team van het screeningsbureau.

Allereerst kamer NH-227: in het begin met Stijn en Claartje, en natuurlijk vooral de laatste jaren met Roderick en Pim. Waardevolle discussies, een goed klankbord, de soap van het screeningbureau beleven en vooral veel lol. Ook samen naar de congressen in binnen- en buitenland was een fantastische ervaring. Ik prijs mij gelukkig met zo'n promotieplek.

Maaïke, Marlies, Maevis, Lakshmi, Conja, Heidi, Lionne en Heleen, zonder jullie inzet voor de studie zou dit boekje er niet zijn gekomen. Dank je wel!

Graag wil ik de afdeling immunohistochemie bij de Urologie in het JN1 bedanken. Kees, heel erg bedankt voor je geduld en de hartelijkheid waarmee jij en je collega's mij ontvangen hebben tijdens mijn stage bij jullie op het lab. Alex, bedankt voor al je hulp bij de altijd weer fijne tumourvolumemetingen.

Ook de collega's van de afdeling Urologie van het Erasmus MC wil ik bedanken. Alle onderzoekers voor de gezellige lunches en borrels; de assistenten en urologen voor de begeleiding tijdens mijn klinische stage en natuurlijk ook de gezelligheid. Ik kijk er naar uit om na de vooropleiding bij jullie aan de slag te gaan.

Tenslotte wil ik alle vrienden en familie om mij heen bedanken. Hoewel jullie uiteraard geen inhoudelijke bijdrage aan het proefschrift hebben geleverd, is jullie aanwezigheid voor mij onmisbaar bij de tot stand koming van dit proefschrift.

Mam, je had me geen betere basis kunnen geven. Ik heb diepe bewondering voor je.

Bart, ik heb ontzettend veel zin in de rest van ons leven!

Tineke Wolters

Rotterdam, december 2009

Appendix: Color Images

Chapter 7 | Prostate needle-biopsies false-negative for cancer or suspicious lesions during histological examination

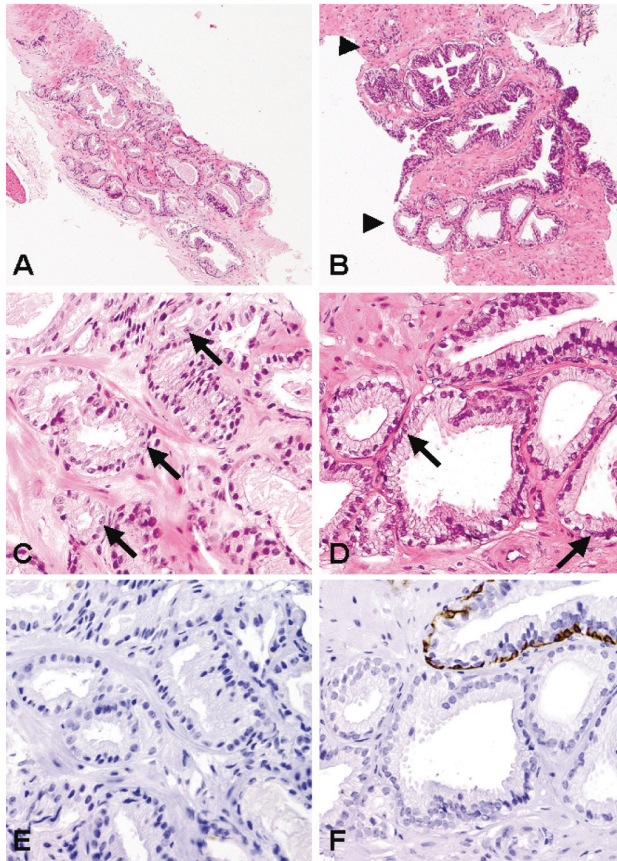


Figure 1 | False-negative prostate biopsies with adenocarcinoma. A, B: Prostate cancer glands are architecturally arranged in a nodule reminiscent of benign tissue. In B two foci (arrowheads) of atypical glands are present adjacent to normal pre-existent glands. C, D: At high magnification enlarged nuclei and conspicuous nucleoli (arrows) are visible, while cytoplasm is not conspicuous (D). E, F: Basal cells are absent (34BE12). In F the second focus of atypical glands also lacked basal cells (not shown); notice positive internal control (F). Both lesions were considered adenocarcinoma Gleason score 6 (3+3). The lesions were derived from two separate patients (A, C, E and B, D, F). Original magnifications: A, B H&E 40x; C, D H&E 200x; E 34BE12 100x; F 34BE12 200x, see page 92.

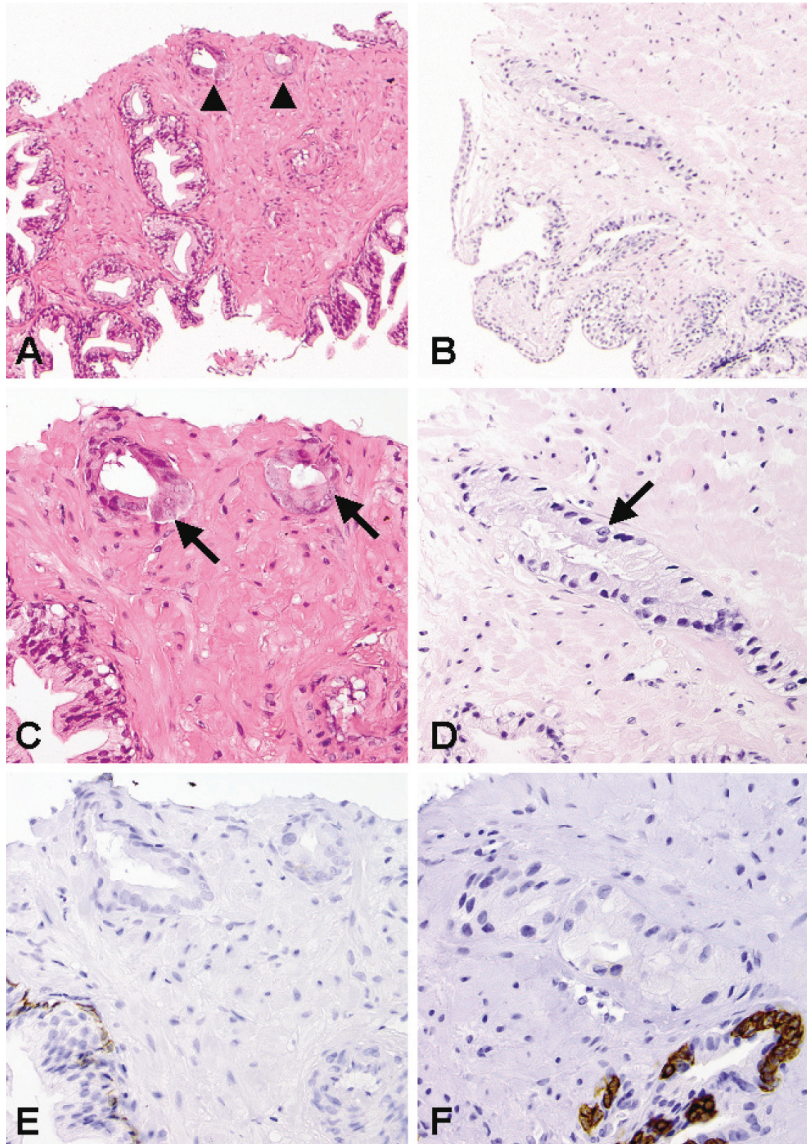


Figure 2 | False-negative prostate biopsies with atypical glands suspicious for adenocarcinoma. A, B: Two (A; arrowheads) and one (B) atypical glands were discovered at low magnification by their amphophilic cytoplasm and subtle architectural abnormality. C, D: At high magnification, the suspicious glands revealed enlarged nuclei and prominent nucleoli (arrows). E, F: The atypical glands showed lack of basal cells (34BE12). Both lesions were considered highly suspicious for malignancy. Due to a low number of atypical glands no definitive diagnosis for malignancy was given. The lesions were derived from two separate patients (A, C, E and B, D, F). Original magnifications: A H&E 40x; B H&E 100x; C, D H&E 200x; E, F 34BE12 200x, see page 93.

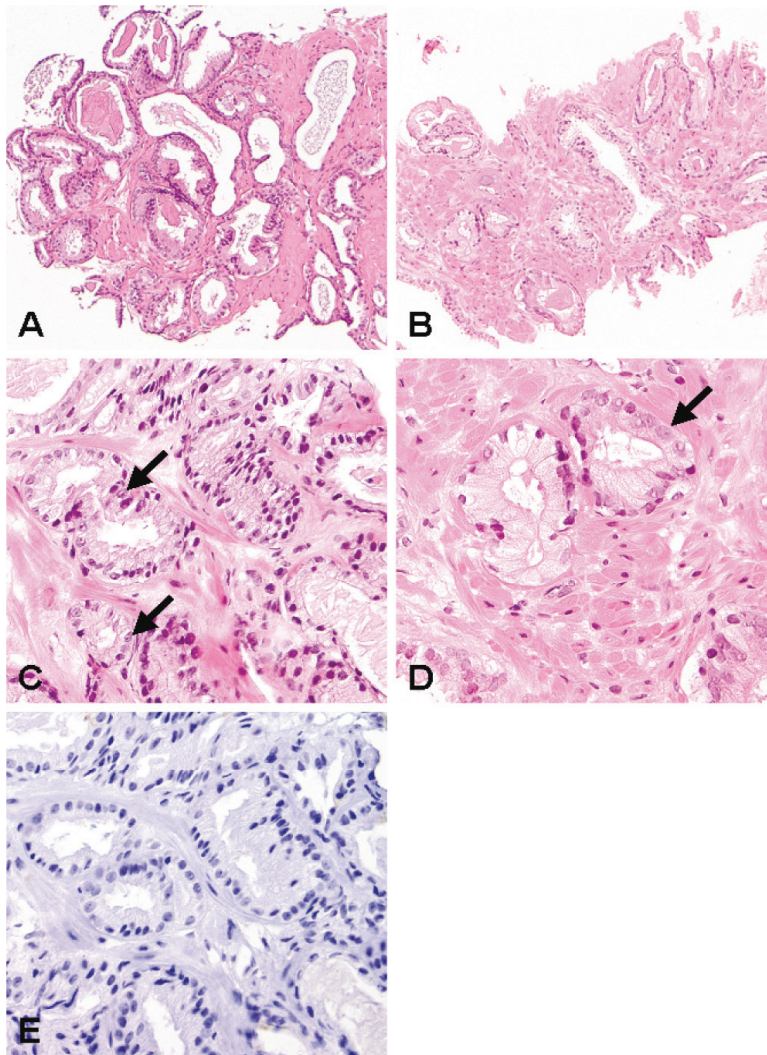


Figure 3 | False-negative prostate biopsy with pseudohyperplastic (A, C, E) and foamy gland (B, D) prostate adenocarcinoma. A, C: Dilated glands with cytoplasmic amphophilia, enlarged nuclei and conspicuous nucleoli characterize pseudohyperplastic cancer. E: Basal cells are absent in 34BE12 staining; notice positive staining in pre-existent atrophic glands. The lesion might be missed as large-sized glands can be interpreted as benign or hyperplastic glands at low magnification. B, D: Foamy gland cancer is characterized by architecturally disorganized glands with clear to foamy cytoplasm with some enlarged nuclei and sporadic nucleoli (arrow). The lesion might be missed as cytoplasmic amphophilia is not conspicuous at low magnification and only some of the nuclei are atypical with prominent nucleoli at high magnification. No immunohistochemical staining was performed on this lesion. Original magnifications: A, B H&E 40x; C, D H&E 200x; E 34BE12 200x, see page 94.

Chapter 10 | The value of EZH2, p27^{kip1}, BMI-1 and MIB-1 on biopsy specimens with low risk PC in selecting men with significant prostate cancer at prostatectomy

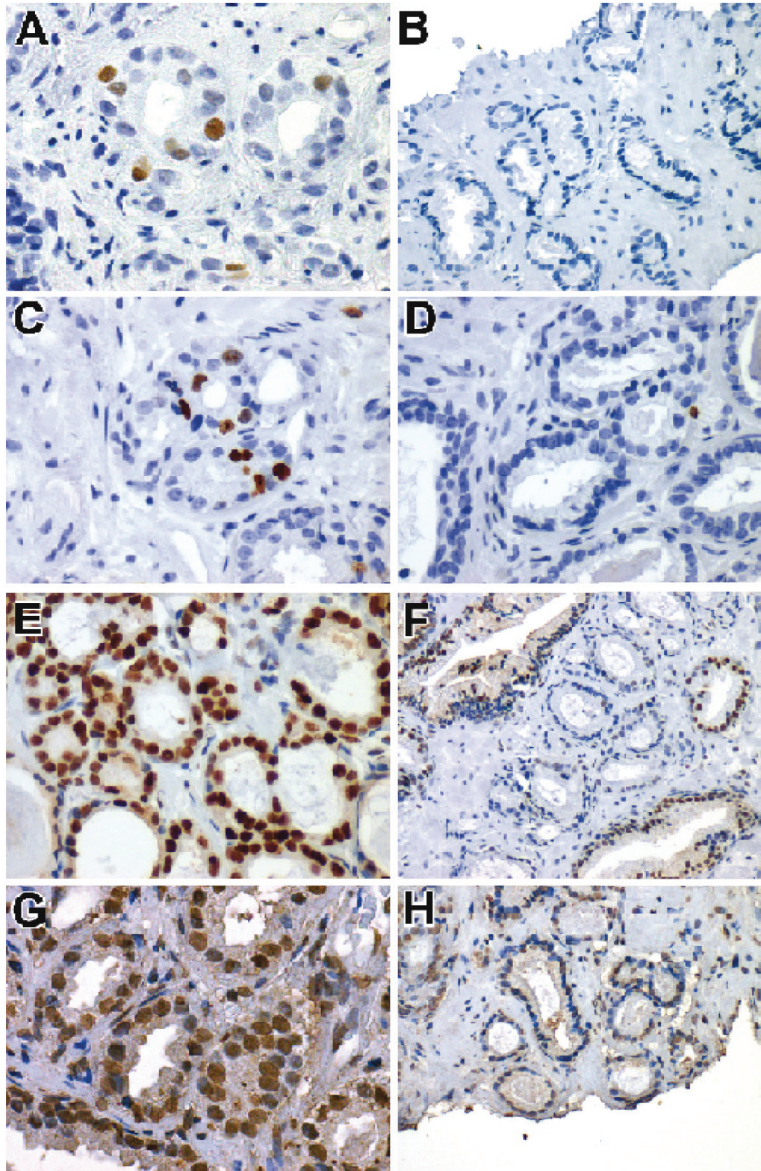


Figure 1 | Various expression levels of the immunomarkers. A: High EZH2 expression, 400x. B: Very low EZH2 expression, 200x. C: High MIB-1 expression, 400x. D: Low MIB-1 expression, 400x. E: High p27^{kip1} expression, 400x. F: Low p27^{kip1} expression, 200x. G: High BMI-1 expression, 400x. H: Low BMI-1 expression, 200x, see page 138.

PhD Portfolio

Name PhD student	Tineke Wolters
Erasmus MC	Department of Urology
Research School	-
PhD period	2006-2009
Promotors	Prof.dr. F.H. Schröder Prof.dr. C.H. Bangma
Supervisors	Dr. M.J. Roobol Dr. G.J.L.H. van Leenders

PhD training (ECTS)	Year	Workload
General courses		
Classical methods for data-analysis	2006	5.7
Biomedical English Writing and Communication	2007	3
Seminars and workshops		
Department journal club	2006-2009	2
Department “refereeravond”	2006-2009	2
Department “promovendiavond”	2006-2009	2
Presentations		
NVU, International day of the Prostate, Cytology Course for Pathologists	2006-2009	3
International conferences		
Oral an poster presentations (EAU 2007, 2008, 2009, ASCO 2008, 2009, AUA 2009, ERSPC meeting 2007, 2008, 2009)	2006-2009	9
Other		
Working at Pathology laboratory	2008	2
Working at Urology ward	2009	2
Total		30.7

