

**PROSTATE CANCER
SCREENING
AND
ACTIVE
SURVEILLANCE**

can we increase survival
without doing too much harm?

LEONARD BOKHORST

**Prostate cancer screening and active surveillance:
can we increase survival without doing too much harm?**

Leonard Pieter Bokhorst

ISBN: 978-94-6169-893-3

Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

Prostate cancer screening and active surveillance: can we increase survival without doing too much harm?

Screening en een actief afwachtend beleid voor prostaatkanker:
kunnen we de overleving verbeteren zonder al te veel bijwerkingen?

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam op gezag van de
rector magnificus Prof.dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
22 juni 2016 om 11:30 uur

door

Leonard Pieter Bokhorst

geboren te Rheden

PROMOTIECOMMISSIE

Promotoren: Prof.dr. C.H. Bangma
Prof.dr. E.W. Steyerberg

Overige leden: Prof.dr. M.G.M. Hunink
Prof.dr. R.C. Pelger
Dr. D. Rizopoulos

Copromotor: Dr. M.J. Roobol

CONTENTS

	Introduction	7
	Objective	10
	Outline of research questions addressed in this thesis	10
PART 1	SCREENING	15
Chapter 1	Positive predictive value of prostate biopsy indicated by PSA-based prostate cancer screening: trends over time in a European randomized trial.	17
Chapter 2	PSA-based prostate cancer screening: reduction of prostate cancer mortality after correction for nonattendance and contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC)	31
Chapter 3	Differences in treatment and outcome after treatment with curative intent in the screening and control arm of the ERSPC Rotterdam.	47
Chapter 4	Do treatment differences between arms affect the main outcome of the European Randomized study of Screening for Prostate Cancer (ERSPC) Rotterdam?	63
PART 2	ACTIVE SURVEILLANCE	79
Chapter 5	Decision support for low-risk prostate cancer	81
Chapter 6	Compliance rates with the Prostate cancer Research International: Active Surveillance (PRIAS) protocol and disease reclassification in non-compliers.	97
Chapter 7	Complications after prostate biopsies in men on active surveillance and its effect on receiving further biopsies in the Prostate cancer Research International: Active Surveillance (PRIAS) study.	111
Chapter 8	Effect of pathologic revision and Ki67 and ERG immunohistochemistry on predicting radical prostatectomy outcome in men initially on active surveillance	123
Chapter 9	A decade of active surveillance in the PRIAS study: an update and evaluation of the criteria used to recommend a switch to active treatment	137
	Summary	155
	General discussion	161
	Samenvatting	183
	About the author	189
	List of publications	191
	Dankwoord	195
	PhD Portfolio	197

Introduction

Medicine is the practice of treating disease, and, consequently, the preservation and improvement of health (adapted from [1]). This definition, published over 100 years ago, seems to largely hold today. In modern medicine disease is still being treated (supplemented by, among others, prevention) aimed at improving health. Health in itself is a somewhat subjective definition, but could be thought of as a combination of mental and physical wellbeing (e.g. expressed as quality of life) for an as long as possible amount of time (life expectancy). What is of interesting however in the old definition is the incorporation of the word *consequently*. Apparently it is assumed that treatment will always result in a preservation and improvement of health. Although this seems completely logical it might not per definition be the case, for instance if the (side) effects of treatment on overall health are worse than the expected effects of the disease. This counterintuitive phenomenon is not unthinkable in the case of prostate cancer, often summarized in the terms 'overdiagnosis' and 'overtreatment'. Both terms are elaborately discussed in this thesis. In this case the adverse effects of the disease that would be experienced are small while the harmful side effects of the treatment are considered substantial. But less straightforward examples can be thought of resulting from the subjectiveness of the definition of health since different attributes compete for dominance. An example can be given in the case of life prolonging treatment in a person expected to suffer from his/her disease. One person might argue that this intervention will extent life expectancy, albeit in suboptimal condition, and thus overall health will be improved. Another could state that extending the time spent while suffering from the disease will result in an overall decreased wellbeing for a longer time and therefore a decrease in overall health. These examples illustrate that the decision of detecting and treating a disease is not as straightforward as stated in the first mentioned definition. In modern medicine considerations should be made constantly on the balance between harms and benefits of medical interventions. As this is subject to personal perception, medical practice is no longer a one way stream in which decisions are made by the physician who 'knows' what is best for a person's health. Instead, practice is based on a shared decision making process in which the patient, with help from his doctor, is supposed to carefully weigh personal advantages and disadvantages of an intervention. It is key to have access to all relevant information for a well-informed decision. It is here that medical research plays an increasingly important role, after all how can one make a well-informed decision if there is not enough scientifically valid information to base it on? In this thesis an attempt is made to make a contribution in this respect to the information on prostate cancer screening and its challenge of minimizing harms while preserving benefits.

Starting with the latter, the most substantial evidence on the benefits of screening comes from the European Randomized study of Screening for Prostate Cancer (ERSPC). This study, initiated in the early nineties, is the largest randomized study conducted to date with inclusion of close to 200,000 men in its core age group [2]. The rationale for this study was that detection of prostate cancer at an earlier stage (facilitated by the PSA test) would allow curative rather than palliative treatment. At that time the majority of men were diagnosed with prostate cancer already present outside of the prostate and one third of men diagnosed with prostate cancer died of their

disease [3]. Taken over the entire population prostate cancer is among the top three most lethal cancers in men from most western countries [4, 5]. To put this in perspective, in the Netherlands approximately 1 in 25 to 30 men will eventually die of prostate cancer [6]. Data from the ERSPC show that systematic PSA screening can result in a reduction of prostate cancer specific mortality of 21% after 13 years of follow-up, and a reduction of metastasis of 30% [7, 8], resulting in a net increase of quality adjusted life years [9].

These benefits are unfortunately not without its negative side effects. First, many men need to be tested, most of them will experience no benefit. It is estimated that on a lifetime basis 100 men need to be invited for screening to save the life of 1 men [9]. On top of this comes overdiagnosis of tumors that would not have given rise to any symptoms during a man's life. An estimated 5 men are unnecessarily diagnosed by screening per 100 men invited (and thus per 1 men saved) [9]. Overdiagnosis, besides from the effects on health of a cancer finding, would be less problematic if not subsequently most of these cancers are being treated (overtreatment), resulting in side effects with subsequent reduction of quality of life [9, 10]. A strategy developed to reduce overtreatment is the use of 'active surveillance'. In active surveillance men likely to have an overdiagnosed cancer are not directly treated but instead monitored only to switch to active treatment in the case of tumor reclassification (signs of higher risk disease). But in order for active surveillance to be effective in reducing the harms of screening it must be able to select men likely to have overdiagnosed cancer at entrance, selectively filter out those with signs of more aggressive disease during follow-up, and do so before the tumor becomes beyond the window of curability (in which case the early diagnosis would be in vain). All this should be achieved without itself being too demanding on a patient's health.

OBJECTIVE

This thesis will focus on two main issues. The first objective is to better understand how screening works and results in a reduction of mortality. This is important in order to develop methods of maximizing the benefits and minimizing its harms of screening. The second objective focusses specifically on the reduction of overtreatment with active surveillance, and emphasizes on how best to employ active surveillance to safely maximize the benefits in terms of reducing overtreatment.

OUTLINE OF RESEARCH QUESTIONS ADDRESSED IN THIS THESIS

The first part will focus on screening and is divided into four chapters. These will center on the following research questions:

-
- What are the mechanisms that lead to the observed reduction in prostate cancer mortality?
 - What type of prostate cancers are detected and at what time during the screening process? (Chapter 1)
 - What is the effect of correction for contamination and noncompliance? (Chapter 2)
 - What is the effect of treatment on screening outcome? (Chapter 3 and 4)
 - Where does the benefit originate from? (Chapter 1 to 4)

The second part will focus on active surveillance and is divided into five chapters addressing the following research questions:

- Are we able to selectively identify men with aggressive disease?
 - By risk based selection at inclusion? (Chapter 5)
 - By pathology and biomarkers (Chapter 8)
 - Is biopsy the best method to detect aggressive disease, i.e. what are its drawbacks? (Chapter 6 and 7)
 - Should we change the current follow-up protocol? (Chapter 9)

REFERENCES

1. Young J. The Aims of Medicine. *Br Med J.* 1871;1:553-5.
2. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med.* 2009;360:1320-8.
3. Hsing AW, Tsao L, Devesa SS. International trends and patterns of prostate cancer incidence and mortality. *Int J Cancer.* 2000;85:60-7.
4. Siegel R, Ma J, Zou Z, Jemal A. Cancer Statistics, 2014. *CA Cancer J Clin.* 2014.
5. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer.* 2013;49:1374-403.
6. www.cbs.nl. Accessed 01-01-2016.
7. Schroder FH, Hugosson J, Carlsson S, Tammela T, Maattanen L, Auvinen A, et al. Screening for Prostate Cancer Decreases the Risk of Developing Metastatic Disease: Findings from the European Randomized Study of Screening for Prostate Cancer (ERSPC). *Eur Urol.* 2012;62:745-52.
8. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet.* 2014.
9. Heijnsdijk EA, Wever EM, Auvinen A, Hugosson J, Ciatto S, Nelen V, et al. Quality-of-life effects of prostate-specific antigen screening. *N Engl J Med.* 2012;367:595-605.
10. Sanda MG, Dunn RL, Michalski J, Sandler HM, Northouse L, Hembroff L, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. *N Engl J Med.* 2008;358:1250-61.

Part 1

Screening

Chapter 1

Positive predictive value of prostate biopsy indicated by PSA-based prostate cancer screening: trends over time in a European randomized trial.

Leonard P. Bokhorst^{a,b}, Xiaoye Zhu^{a,b}, Meelan Bul^a, Chris H. Bangma^a,
Fritz H. Schröder^a, Monique J. Roobol^a.

^a Erasmus University Medical Center, Department of Urology, Rotterdam, The Netherlands

^b Both authors contributed equally to this article

ABSTRACT

Objective:

- To assess the Positive Predictive Value (PPV) of prostate biopsy, indicated by a prostate-specific antigen (PSA) cut-off of ≥ 3.0 ng/ml, over time, in the Rotterdam section of the European Randomized study of Screening for Prostate Cancer (ERSPC).

Patient and methods:

- In the Rotterdam section of the ERSPC, a total of 42376 participants identified from population registries (age 55-74 yr) were randomly assigned to a screening or control arm.
- In the ERSPC men are screened with PSA at a four year interval. A total of three screening rounds were evaluated. Therefore, only men aged 55-69 yr at the first screen were eligible for this study.

Results:

- PPVs for men without previous biopsy remained equal throughout the three subsequent screens (25.5%, 22.3% and 24.8% respectively).
- Conversely, PPVs for men with previous negative biopsy dropped significantly (12.0% and 15.2% at the second and third screen respectively).
- Additionally, in men with and without previous biopsy the percentage aggressive prostate cancers (PCa) (clinical stage $>T2b$, Gleason score ≥ 7) decreased after the first round of screening from 44.4% to 23.8% in the second ($p < 0.001$) and 18.6% in the third round ($p < 0.001$).
- Repeat biopsies accounted for 24.6% of all biopsies, but yielded only 8.6% of all aggressive cancers.

Conclusions:

- In consecutive screening rounds the PPV of PSA-based screening remains equal in previous unbiopsied men.
- In men with a previous negative biopsy the PPV drops considerably, however 20% of cancers detected still show aggressive characteristics.
- Individualized screening algorithms should incorporate previous biopsy status in the decision to perform a repeat biopsy with the goal to further reduce unnecessary biopsies.

Trial registration: ISRCTN49127736

INTRODUCTION

Prostate-specific antigen (PSA) can be used as a biomarker for the early detection of prostate cancer (PCa) [1]. In the European Randomized Study of Screening for Prostate Cancer (ERSPC) men are screened for PCa with PSA. Results of the ERSPC have shown that PSA-based screening can reduce PCa mortality by up to 29% at eleven years of follow-up after adjustment for noncompliance [2].

Although screening with PSA can reduce the PCa mortality, its use has limitations as a result of the lack of specificity, especially in low PSA ranges [3]. Consequently, if a large group of men is biopsied based on a PSA cut-off, only a modest proportion of men will have PCa. In the ERSPC the PPV of a lateralized sextant prostate biopsy indicated by PSA is approximately 25% at initial screening [4, 5]. Already, multivariable risk calculators have been developed to improve the risk stratification and select men at high risk of PCa for conducting biopsies [6-8]. Data on cancer detection and PPV per screening round could further improve risk stratification.

In the ERSPC men are re-screened at a four year interval. In this paper we aim to assess the PPV of lateralized sextant prostate biopsy, indicated by an identical PSA cut-off value in subsequent screening rounds in the Rotterdam section of the ERSPC, stratified by age group and status of previous biopsy. We also evaluate the tumour characteristics of the diagnosed cancers. This knowledge may have implications for future screening strategies.

PATIENTS AND METHODS

The study population and protocol have been described in detail previously [9]. In summary, men aged 55-74 yr, identified from population registries of Rotterdam, were invited for screening. Men previously diagnosed with PCa were excluded [9]. In total, 42376 men who responded by returning the intake questionnaire and who provided informed consent were randomized to a screening (n=21210) or control arm (21166) from November 1993 until December 1999.

In the present study, three consecutive screening rounds were evaluated. Men aged 55-69 yr at the first screening round were eligible (16600 men). Age selection was made to provide a cohort of men eligible for at least two consecutive screening visits. Men were rescreened every four year until they reached the age of 75. A prostate biopsy was indicated for those with a PSA ≥ 4.0 ng/ml and/or abnormal digital rectal examination (DRE) and/or transrectal ultrasound (TRUS). From May 1997, a PSA threshold of ≥ 3.0 ng/ml was used as the sole screening test. In screen-positive men, sextant biopsies were indicated; they were lateralized from June 1996, as described by Eskew et al [10]. An additional biopsy was taken from any suspicious area on TRUS.

Statistical analysis

Data was stratified for age groups 55-59, 60-64 and 65-69 yr at baseline and status of previous biopsy (yes or no). Aggressive PCa was defined as clinical stage >T2b and/or Gleason score ≥ 7 as described by Roobol et al. [7].

The PPV (percentage PCa detected among all men biopsied) was calculated for each screening round and subgroup. The PPV and categorical clinical variables between groups were compared using chi-square test; for continuous variables the Mann-Whitney U test was used. All statistical tests were two sided. A p-value < 0.05 was considered significant. SPSS v.17.0 was used for statistical analysis (SPSS Inc, Chicago, IL, USA).

RESULTS

In total 16600 men, aged 55-69 yr at baseline, were screened in the first screening round, 12120 in the second round and 7740 in the third round. The median age for the whole study population at first screen was 61.1 yr. An overview of the screening rounds is shown in the flow diagram (fig. 1).

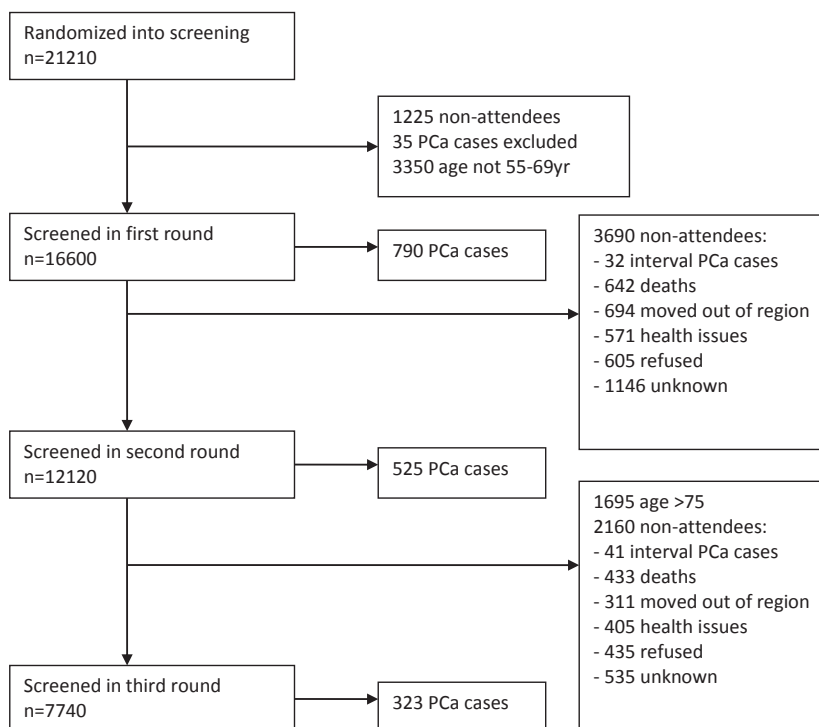


Figure 1. Consort trial flow diagram, screening rounds with a four year interval. PCa= prostate cancer

Positive predictive value

In total 7553 biopsies were performed: 3104 men were biopsied in the first round, 2789 in the second round and 1660 in the third round (96.3%, 92.1% and 93.5% of men with a biopsy indication). In addition, 288, 266 and 195 men refused a biopsy despite recommendation respectively. The numbers of cancers detected per round were 790, 525 and 323 respectively (table 1). Subsequently, the PPV of prostate biopsy in the first round was 25.5%. In the second round the PPVs for men with and without a biopsy in the first round were 22.3% and 12.0% respectively ($p < 0.001$). In the third round the PPVs for men with and without previous biopsy were 24.8% and 15.2% respectively ($p < 0.001$).

Table 1. Positive predictive values of prostate biopsies per screening round of the ERSPC Rotterdam

	First round	Second round		Third round			
	Total	Total	No Biopsy round 1	Biopsy round 1	Total	No Biopsy round 1 or 2	Biopsy round 1 and/or 2
Men screened	16600	12120	10552	1568	7740	6085	1655
Men biopsied (%screened)	3104 (18.7)	2789 (23)	1850 (17.5)	939 (59.9)	1660 (21.4)	743 (12.2)	917 (55.4)
PCa, No	790	525	412	113	323	184	139
% aggressive ^a	44.4%	23.8%	24.5% ^b	21.2% ^c	18.6%	20.7% ^b	15.8% ^c
PPV	25.5%	18.8%	22.3%	12.0% ^d	19.5%	24.8%	15.2% ^d

PCa = prostate cancer; PPV = positive predictive value; ERSPC = European Randomized Study of Screening for Prostate Cancer ^a Defined as clinical stage >T2b and/or Gleason score ≥ 7 ; ^b $p < 0.001$ (as to first round); ^c > 0.05 (as to no biopsy); ^d $p < 0.001$ (as to no biopsy)

In the first round the PPV of prostate biopsy was higher in the oldest age group (28.8% in 65-69 yr) compared to the younger age groups (23.1% in 55-59 yr, $p < 0.01$; 23.2% in 60-64 yr, $p < 0.01$). In the second and third round the differences between age groups did not reach statistical significance (table 2).

Tumour characteristics

Tumour characteristics per screening round are shown in table 3. The median PSA level and prostate volume, measured by TRUS, were significantly different in the second and third round of screening for men with or without previous biopsy (all $p < 0.001$). The percentage of aggressive PCa (defined as clinical stage >T2b and/or Gleason ≥ 7) was significantly higher in the first round compared to the second and third round (44.4%, 23.8% and 18.6% respectively; both $p < 0.001$). No significant difference in percentage aggressive PCa was seen between men with or without previous biopsy in both the second and third screening round (21.2% vs. 24.5% in second round respectively, $p = 0.549$; 15.8% vs. 20.7% in the third round respectively, $p = 0.337$). In total 536 aggressive cancers were found, of which 65.5% were found in the first screen, 25.9% in subsequent screens in men without previous biopsy and 8.6% in men with a previous biopsy. In the first

Table 2. Positive predictive value per age group at baseline and round of screening

	First round	Second round		Third round			
	Total	Total	No Biopsy round 1	Biopsy round 1	Total	No Biopsy round 1 or 2	Biopsy round 1 and/or 2
55 – 59 yr							
Men screened	6498	5061	4630	431	4004	3280	724
Men biopsied	792	937	703	234	747	360	387
PCa, No	183	168	142	26	154	91	63
% aggressive ^a	37.2%	23.2%	23.9%	19.2%	17.5%	13.2%	23.8%
PPV	23.1%	17.9%	20.2%	11.1%	20.6%	25.3%	16.3%
60 – 64 yr							
Men screened	5336	3946	3373	573	2873	2184	689
Men biopsied	1032	958	612	346	667	289	378
PCa, No	239	186	145	41	135	76	59
% aggressive ^a	39.3%	24.7%	24.1%	26.8%	20.0%	27.6%	10.2%
PPV	23.2%	19.4%	23.7%	11.8%	20.2%	26.3%	15.6%
65 – 69 yr							
Men screened	4766	3113	2549	564	863	621	242
Men biopsied	1280	894	535	359	246	94	152
PCa, No	368	171	125	46	34	17	17
% aggressive ^a	51.4% ^b	23.4%	25.6% ^c	17.4%	17.6%	29.4% ^c	5.9%
PPV	28.8% ^b	19.1%	23.4% ^c	12.8%	13.8%	18.1% ^c	11.2%

PCa = prostate cancer; PPV = positive predictive value; ^a Defined as clinical stage >T2b and/or Gleason score >=7; ^b p<0,01 (as to 55-59 yr); ^c p>0,05 (as to 55-59 yr)

round the percentage aggressive PCa was significantly higher in the oldest age group (65-69 yr) compared to the youngest age group (55-59 yr; 51.4% vs. 37.2% respectively, p=0.002). In all age groups the percentage aggressive PCa decreased after the first screening round as shown in table 2. In the first round 79.6% of PCa were clinically organ-confined (<=cT2). In the second and third round this number increased to 96.2% and 98.4% respectively. No statistically significant difference was seen between men with or without previous biopsy.

Table 4 outlines the characteristics at the time of the preceding round of men without previous biopsy, who were diagnosed in later screens. In the second round 46.6% of these men had a PSA of 2.0-2.9 ng/ml in the first round. In the third round a similar amount (48.4%) had a PSA of 2.0-2.9 ng/ml in the second round.

Table 3. Tumour characteristics per round of screening

	First round	Second round			Third round		
	Total	Total	No Biopsy round 1	Biopsy round 1	Total	No Biopsy round 1 or 2	Biopsy round 1 and/or 2
PCa, No	790	525	412	113	323	184	139
Age, median	64.6	66.7	66.4	68.2	68.6	68.4	69.3
PSA (ng/ml), median	5.6	3.9	3.6	5.7 ^d	4.2	3.7	5.3 ^d
Prostate volume (cc), median	35.9	38	36	50.5 ^d	42.5	37.2	50.9 ^d
Aggressive ^a (%)	351 (44.4)	125 (23.8)	101 (24.5) ^b	24 (21.2) ^c	60 (18.6)	38 (20.7) ^b	22 (15.8) ^c
Clinical stage							
T1 (%)	325 (41.1)	356 (67.8)	290 (70.4)	66 (58.4)	229 (70.9)	127 (69)	102 (73.4)
T2 (%)	304 (38.5)	149 (28.4)	107 (26)	42 (37.2)	88 (27.2)	54 (29.3)	34 (24.5)
T3 (%)	155 (19.6)	20 (3.8)	15 (3.6)	5 (4.4)	6 (1.9)	3 (1.6)	3 (2.2)
T4 (%)	6 (0.8)	-	-	-	-	-	-
Gleason							
<=6 (%)	531 (67.2)	417 (79.4)	324 (78.6)	93 (82.3)	268 (83)	148 (80.4)	120 (86.3)
7 (%)	202 (25.6)	93 (17.7)	78 (18.9)	15 (13.3)	39 (12.1)	26 (14.1)	13 (9.4)
>=8 (%)	50 (6.3)	15 (2.9)	10 (2.4)	5 (4.4)	14 (4.3)	9 (4.9)	5 (3.6)

PSA = prostate-specific antigen; PCa = prostate cancer; ^a Defined as clinical stage >T2b and/or Gleason score >=7; ^b p<0.001 (as to first round); ^c p>0.05 (as to no biopsy); ^d p>0.001 (as to no biopsy)

DISCUSSION

Although results of the ERSPC have shown to reduce PCa mortality [2], the US preventive Services Task Force recently released an updated recommendation against PSA screening, as the authors concluded that the harms outweigh the benefits [11, 12]. Moreover, a meta-analysis by Djulbegovic et al. [13] concluded that the existing evidence does not support the routine use of screening for prostate cancer. In addition to overdiagnosis, unnecessary biopsies triggered by false-positive screening results could be considered as one of the most important harms, leading to infections and hospital admissions [14].

In this study we assessed the PPV of a PSA indicated prostate biopsy throughout subsequent screening rounds of the ERSPC, section Rotterdam. This gives insight in the screening efficacy of the current algorithm and may be valuable in the development of future screening strategies. Our results demonstrate that during screening rounds the PPV of men without a previous biopsy remained equal (25.5%, 22.3% and 24.8% in first, second and third round respectively). The PPV of prostate biopsy indicated by a PSA cut-off dropped considerably to 12.0%-15.2% in men with a previous biopsy; 20% of cancers detected however still show aggressive characteristics.

Because the PPV depends on the underlying prevalence and the first screening round was performed in a relatively unscreened population, one would expect a decline in PPV after the first

round considering the slow natural course of PCa [15, 16]. However, data from the PCPT trial has shown that 23.9% of men with a PSA 2.1-3.0 ng/ml harbor PCa [17]. In the current analysis, almost half of the cancers detected in men without previous biopsy originated from the 2.0-2.9 ng/ml PSA group as shown in table 4. The PSA in these men increased and subsequently surpassed the biopsy threshold during the four year screening interval, resulting in equal PPVs of approximately

Table 4. Characteristics of men at the time of the preceding round and of prostate cancers which were eventually diagnosed

	Biopsy	PCa, No (%)	% PCa / Biopsies	Aggressive ^a , No (%)	% Aggressive ^a /PCa
Second round, no biopsy first round					
Total	1850	412 (100)	22.3	101 (100)	24.5
Age at baseline					
55-59 yr	703	142 (34.5)	20.2	34 (33.7)	23.9
60-64 yr	612	145 (35.2)	23.7	35 (34.7)	24.1
65-69 yr	535	125 (30.3)	23.4	32 (31.7)	25.6
PSA first round (ng/ml)					
<1.0 ng/ml.	212	29 (7)	13.7	11 (10.9)	37.9
1.0-1.9 ng/ml.	660	133 (32.3)	20.2	27 (26.7)	20.3
2.0-2.9 ng/ml.	755	192 (46.6)	25.4	42 (41.6)	21.9
>=3.0 ng/ml	223	58 (14)	26.0	21 (20.8)	36.2
Reason no biopsy first round					
Medication	9	2 (0.5)	-	-	-
Refused biopsy	11	5 (1.2)	-	3 (3)	-
DRE and TRUS normal	203	51 (12.4)	25.1	18 (17.8)	35.3
Third round, no previous biopsy					
Total	743	184 (100)	24.8	38 (100)	20.7
Age at baseline					
55-59 yr	360	91 (49.5)	25.3	12 (31.6)	13.2
60-64 yr	289	76 (41.3)	26.3	21 (55.3)	27.6
65-69 yr	94	17 (9.2)	18.1	5 (13.2)	29.4
PSA second round (ng/ml)					
<1.0 ng/ml.	153	22 (12)	14.4	4 (10.5)	18.2
1.0-1.9 ng/ml.	269	67 (36.4)	24.9	10 (26.3)	14.9
2.0-2.9 ng/ml.	300	89 (48.4)	29.7	24 (63.2)	27.0
>=3.0ng/ml	21	6 (3.3)	28.6	-	-
Reason no previous biopsy					
Medication	4	1 (0.5)	-	-	-
Refused biopsy	17	5 (2.7)	-	-	-

PSA = prostate-specific antigen; PCa = prostate cancer; DRE = digital rectal examination; TRUS = transrectal ultrasound; ^a Defined as clinical stage >T2b and/or Gleason score >=7;

25%. If we would assume that these cancers were already detectable at the previous screening round, these men may have been diagnosed when the biopsy threshold was set at a PSA of 2.0 ng/ml. However, a lower cut-off would also increase the number of overdiagnosed cancers and unnecessary biopsies [17, 18]. Lowering the biopsy threshold to a PSA of 2.0 ng/ml would have increased the number of biopsies with 64%-72% in the current study (data not shown). Applying a shorter screening interval in men with a PSA of 2.0-2.9 ng/ml may be another option. Future research should further address this problem and study the origin of cancers detected in previously screened but unbiopsied men, with the goal to reduce unnecessary biopsies, overdiagnosis and mortality.

Even more important than the actual number of PCa detected, are the characteristics of the cancers. In the first round of screening, we found almost half of the cancers to be aggressive (Gleason score ≥ 7 and/or clinical stage $>T2b$). Even though the PPV in the second and third round remained equal in men without previous biopsy, the proportion of aggressive PCa decreased to 20.7%-24.5%. Almost all cancers in the second and third round were clinically organ-confined (96.4%-98.4%). If these cancers were detectable in the first screening round, they did not progress to a stage where they became incurable. The low number of cancers detected in the interval period, as described previously [19, 20], supports this assumption.

Still, overdiagnosis is one of the major drawbacks of PCa screening. A simple solution to reduce the number of low risk PCa, which could be considered overdiagnosed, is to raise the PSA cut-off for a biopsy indication [21]. Indeed, if only men with a PSA ≥ 4.0 ng/ml were biopsied, the described PPVs in men without a previous biopsy in the first, second and third round would increase to 26.5%, 28.6% and 34.1% respectively (data not shown). Additionally 32.3%, 66.9% and 64.4% of the non-aggressive PCa would not have been detected, possibly sparing these men the burden of PCa and its treatments. However, and this is undesirable, with this strategy 19.1%, 38.6% and 47.4% of all aggressive cancers would also have been missed in the first, second and third round respectively. The drawbacks of a single PSA cut-off emphasize the need for better risk stratification tools. Already different multivariable risk calculators have been developed to improve risk stratification [22]. An external evaluation of the ERSPC risk calculator step 3 (www.prostatecancer-riskcalculator.com) showed both an improvement of PPV to 64% and an improved selection of aggressive PCa (personal communication with H.A. van Vugt, Erasmus University Medical Center, manuscript in preparation). Risk calculators will play an important role until better biomarkers and imaging techniques are validated. Several studies have already demonstrated the additional value of MRI in the diagnosis of PCa [23, 24].

In men with a previous biopsy a drop in PPV was seen at repeat screening. However, there are still cancers detected. Two explanations can be given. First, it is known that a sextant prostate biopsy does not detect all cancers. In a literature review by Schröder et al. [25], the average proportion of cancers missed with a lateralized prostate biopsy was 19%. Possibly a group of PCa was missed in the first screening round and emerged at repeat biopsy. Although some might suggest a more extended biopsy scheme, only a limited reduction in disease-specific mortality

can be expected [25]. Second, some of the cancers that were detected at repeat screening may have developed during the screening interval. This would lower the number of cancers that potentially could have been detected earlier on.

Furthermore, we found that the prostate volume of men with a previous biopsy was significantly higher than men without a previous biopsy. Because a larger prostate is associated with a higher PSA value, these men were more likely to be biopsied. Previous studies have shown a negative association between prostate volume and the risk of PCa [26, 27]. This could attribute to the relatively lower PPV in men with a previous biopsy. On the other hand, there are still cancers detected and although the PPV is lower, the percentage aggressive PCa is comparable to men without a previous biopsy. The number of aggressive PCa detected in men with a previous biopsy only accounted for 8.6% of the total number of aggressive PCa found, whereas the number of biopsies in previously biopsied men accounted for 24.6% of the total biopsies. This emphasizes the need for a more individualized screening approach, in which a previous negative biopsy should be taken into account. Already, previous biopsy status is incorporated in step 4 of the ERSPC risk calculator (www.prostatecancer-riskcalculator.com). External validation of this risk calculator in a Canadian and European cohort showed previous biopsy status to be a significant predictor of PCa in multivariable analysis [28, 29].

In the first round of screening, tumours detected in the oldest age group were of a higher grade than in the younger age groups. This poorer differentiation in older men was reported before [30, 31]. After the first round this difference is less obvious. Because in the first round the tumours in older men had a longer time to develop, this could be expected. This concurs with a previous study by Boevee et al. [32] and could be seen as an effect of screening.

Some limitations should be mentioned. First, the PPVs provided in this study are calculated for only those who actually underwent biopsy; men with a positive screening test who did not have a biopsy were not included in the analysis. However, in our cohort the compliance to a biopsy indication was > 90%, and there is no reason to assume the PPV in men who had an indication but did not undergo biopsy would be significantly different from those who actually had a biopsy. Second, sextant prostate biopsy, either classical or lateralized, will miss 23% or 19% of biopsy-detectable PCa [25]. Therefore, the PPV in this study may be underestimated. However, because the number of biopsies remained equal throughout screening rounds a comparison between screens was possible and was not affected by a change in protocol. Last, the biopsy indication has been modified over time: in the first screening round men were initially biopsied based on the results of a PSA test, DRE and TRUS; half way the first round the use of DRE and TRUS as a biopsy indication was omitted, because of limited additional value [33, 34]; in the second and third round some men were screened in side studies with different biopsy indications. Even so, a subanalysis in men who were biopsied with PSA ≥ 3 ng/ml as the sole biopsy indication showed only a negligible change in PPVs. Therefore all side studies were included in the current analysis.

In conclusion, the results of this study show that the PPV of PSA-based PCa screening remains equal in previous unbiopsied men. In men with a previous biopsy the PPV drops considerably,

however 20% of cancers detected still show aggressive characteristics. In both groups a decline in aggressive PCa is seen after the first screening round. This study indicates that previous biopsy status should definitely be considered in the decision to perform a repeat biopsy. Also, in men without an initial biopsy and a PSA value of 2.0-2.9 ng/ml earlier repeat screening could be considered. Furthermore, future research should study the origin of PCa in men without a previous biopsy. Knowing the origin of these cancers could change the way men are screened, further reducing the PCa mortality and overdiagnosis.

REFERENCES

1. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, et al. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med*. 1991;324:1156-61.
2. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med*. 2012;366:981-90.
3. Thompson IM, Ankerst DP, Chi C, Lucia MS, Goodman PJ, Crowley JJ, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. *JAMA*. 2005;294:66-70.
4. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360:1320-8.
5. Schroder FH, Denis L, Roobol MJ. Epilogue: different approaches for prostate cancer screening in the EU? *Eur J Cancer*. 2010;46:3120-5.
6. Schroder F, Kattan MW. The comparability of models for predicting the risk of a positive prostate biopsy with prostate-specific antigen alone: a systematic review. *Eur Urol*. 2008;54:274-90.
7. Roobol MJ, van Vugt HA, Loeb S, Zhu X, Bul M, Bangma CH, et al. Prediction of Prostate Cancer Risk: The Role of Prostate Volume and Digital Rectal Examination in the ERSPC Risk Calculators. *Eur Urol*. 2011;61:577-83.
8. Kransé R, Roobol M, Schroder FH. A graphical device to represent the outcomes of a logistic regression analysis. *Prostate*. 2008;68:1674-80.
9. Roobol MJ, Kirkels WJ, Schroder FH. Features and preliminary results of the Dutch centre of the ERSPC (Rotterdam, the Netherlands). *BJU international*. 2003;92 Suppl 2:48-54.
10. Eskew LA, Bare RL, McCullough DL. Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate. *The Journal of urology*. 1997;157:199-202; discussion -3.
11. Chou R, Crosswell JM, Dana T, Bougatsos C, Blazina I, Fu R, et al. Screening for prostate cancer: a review of the evidence for the u.s. Preventive services task force. *Ann Intern Med*. 2011;155:762-71.
12. Moyer VA, on behalf of the USPSTF. Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2012.
13. Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, et al. Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. *BMJ*. 2010;341:c4543.
14. Loeb S, van den Heuvel S, Zhu X, Bangma CH, Schroder FH, Roobol MJ. Infectious Complications and Hospital Admissions After Prostate Biopsy in a European Randomized Trial. *Eur Urol*. 2012;61:1110-4.
15. Gann PH, Hennekens CH, Stampfer MJ. A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. *JAMA*. 1995;273:289-94.
16. Fang J, Metter EJ, Landis P, Chan DW, Morrell CH, Carter HB. Low levels of prostate-specific antigen predict long-term risk of prostate cancer: results from the Baltimore Longitudinal Study of Aging. *Urology*. 2001;58:411-6.
17. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, 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.
18. Raaijmakers R, Blijenberg BG, Finlay JA, Rittenhouse HG, Wildhagen MF, Roobol MJ, et al. Prostate cancer detection in the prostate specific antigen range of 2.0 to 3.9 ng/ml: value of percent free prostate specific antigen on tumor detection and tumor aggressiveness. *The Journal of urology*. 2004;171:2245-9.
19. van der Crujisen-Koeter IW, van der Kwast TH, Schroder FH. Interval carcinomas in the European Randomized Study of Screening for Prostate Cancer (ERSPC)-Rotterdam. *J Natl Cancer Inst*. 2003;95:1462-6.

20. Zhu X, van Leeuwen PJ, Bul M, Otto SJ, de Koning HJ, Bangma CH, et al. Disease-specific survival of men with prostate cancer detected during the screening interval: results of the European randomized study of screening for prostate cancer-Rotterdam after 11 years of follow-up. *Eur Urol.* 2011;60:330-6.
21. Otto SJ, Moss SM, Maattanen L, Roobol M, Zappa M, Nelen V, et al. PSA levels and cancer detection rate by centre in the European Randomized Study of Screening for Prostate Cancer. *Eur J Cancer.* 2010;46:3053-60.
22. Zhu X, Albertsen PC, Andriole GL, Roobol MJ, Schroder FH, Vickers AJ. Risk-Based Prostate Cancer Screening. *Eur Urol.* 2011;61:652-61.
23. Hoeks CM, Barentsz JO, Hambrock T, Yakar D, Somford DM, Heijmink SW, et al. Prostate cancer: multi-parametric MR imaging for detection, localization, and staging. *Radiology.* 2011;261:46-66.
24. Hoeks CM, Schouten MG, Bomers JG, Hoogendoorn SP, Hulsbergen-van de Kaa CA, Hambrock T, et al. Three-Tesla Magnetic Resonance-Guided Prostate Biopsy in Men With Increased Prostate-Specific Antigen and Repeated, Negative, Random, Systematic, Transrectal Ultrasound Biopsies: Detection of Clinically Significant Prostate Cancers. *Eur Urol.* 2012.
25. Schröder FH, van den Bergh RC, Wolters T, van Leeuwen PJ, Bangma CH, van der Kwast TH, et al. Eleven-Year Outcome of Patients with Prostate Cancers Diagnosed During Screening After Initial Negative Sextant Biopsies. *Eur Urol.* 2010;57:258-66.
26. Garzotto M, Hudson RG, Peters L, Hsieh YC, Barrera E, Mori M, et al. Predictive modeling for the presence of prostate carcinoma using clinical, laboratory, and ultrasound parameters in patients with prostate specific antigen levels ≤ 10 ng/mL. *Cancer.* 2003;98:1417-22.
27. Makarov DV, Loeb S, Getzenberg RH, Partin AW. Biomarkers for prostate cancer. *Annu Rev Med.* 2009;60:139-51.
28. Cavadas V, Osorio L, Sabell F, Teves F, Branco F, Silva-Ramos M. Prostate cancer prevention trial and European randomized study of screening for prostate cancer risk calculators: a performance comparison in a contemporary screened cohort. *Eur Urol.* 2010;58:551-8.
29. Trottier G, Roobol MJ, Lawrentschuk N, Bostrom PJ, Fernandes KA, Finelli A, et al. Comparison of risk calculators from the Prostate Cancer Prevention Trial and the European Randomized Study of Screening for Prostate Cancer in a contemporary Canadian cohort. *BJU international.* 2011;108:E237-44.
30. Alibhai SM, Krahn MD, Fleshner NE, Cohen MM, Tomlinson GA, Naglie G. The association between patient age and prostate cancer stage and grade at diagnosis. *BJU international.* 2004;94:303-6.
31. Cremers RG, Karim-Kos HE, Houterman S, Verhoeven RH, Schroder FH, van der Kwast TH, et al. Prostate cancer: trends in incidence, survival and mortality in the Netherlands, 1989-2006. *Eur J Cancer.* 2010;46:2077-87.
32. Boevee SJ, Venderbos LD, Tammela TL, Nelen V, Ciatto S, Kwiatkowski M, et al. Change of tumour characteristics and treatment over time in both arms of the European Randomized study of Screening for Prostate Cancer. *Eur J Cancer.* 2010;46:3082-9.
33. Gosselaar C, Roobol MJ, Roemeling S, de Vries SH, Crujjsen-Koeter I, van der Kwast TH, et al. Screening for prostate cancer without digital rectal examination and transrectal ultrasound: results after four years in the European Randomized Study of Screening for Prostate Cancer (ERSPC), Rotterdam. *Prostate.* 2006;66:625-31.
34. Gosselaar C, Roobol MJ, van den Bergh RC, Wolters T, Schroder 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;55:139-46.

Chapter 2

PSA-based prostate cancer screening: reduction of prostate cancer mortality after correction for nonattendance and contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC)

Leonard P. Bokhorst^a, Chris H. Bangma^a, Geert J.L.H. van Leenders^b, Jan J. Lous^c,
Sue M. Moss^d, Fritz H. Schröder^a, Monique J. Roobol^a

^a Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands

^b Department of Pathology, Erasmus University Medical Center, Rotterdam, The Netherlands

^c STAR- Medical diagnostic Center, Rotterdam, The Netherlands

^d Centre for Cancer Prevention, Wolfson Institute for Preventive Medicine, Queen Mary
University of London, London, United Kingdom

ABSTRACT

Background: Large randomized screening trials provide an estimation of the effect of screening on a population based level. The effect of screening for individuals is however diluted by nonattendance and contamination in the trial arms.

Objective: To determine the prostate cancer (PCa) mortality reduction from screening after adjustment for nonattendance and contamination.

Design, setting, and participants: A total of 34833 men in the core age group of 55-69 years were randomized to a screening or control arm in the Rotterdam section of the ERSPC. PSA testing was offered to all men in the screening arm at a four year interval. A prostate biopsy was offered to men with an elevated PSA. The primary end-point was PCa specific mortality.

Outcome measurements and statistical analysis: Nonattendance was defined as non-participation in the screening arm. Contamination in the control arm was defined as receiving asymptomatic PSA testing or a prostate biopsy in the absence of symptoms. Relative risks (RR) were calculated with an intention to screen (ITS) analysis and after correction for nonattendance and contamination using a method which preserves the benefits obtained by randomization.

Results and limitation: The ITS analysis resulted in a RR of 0.68 (95% confidence interval (CI) 0.53-0.89) in favour of screening at a median follow-up of 13 years. Correction for both nonattendance and contamination resulted in a RR of 0.49 (95% CI 0.27-0.87) in favour of screening.

Conclusion: PCa screening as conducted in the Rotterdam section of the ERSPC can reduce the risk of dying from PCa with up to 51% for an individual man choosing to be screened repeatedly as compared to a man that was not screened. These benefits of screening should be balanced against the harms of overdiagnosis and subsequent overtreatment.

INTRODUCTION

Recently, the 13 year follow-up results of the Dutch centre of the European Randomized study of Screening for Prostate Cancer (ERSPC) were published, showing a prostate cancer (PCa) specific mortality reduction of 32% in favour of screening with prostate-specific antigen (PSA) [1]. Although the conventional intention to screen (ITS) analysis provides the best estimation of the PCa specific mortality reduction on a population based level, the potential effect of screening for an individual choosing to be screened needs to be corrected for nonattendance in the intervention arm and contamination (e.g. PSA testing/prostate biopsy) in the control arm. This adjustment should however not influence the benefits obtained by randomization (namely, the same baseline risk of PCa mortality in both arms). A simple comparison of men who actually receive screening (attenders), against those who do not (non-attenders), could be biased since the baseline risk of having PCa for attenders and non-attenders may be different. In order to correct for nonattendance and contamination without creating a difference in baseline risk in the two compared groups a method developed by Cuzick et al.[2] was used. This method was previously applied to correct for nonattendance and contamination at the 9 year follow-up results of the whole ERSPC [3].

The aim of this paper is to determine the PCa specific mortality reduction from PSA-based PCa screening, adjusted for nonattendance and contamination in the ERSPC, section Rotterdam, with a median follow-up of 13 years and to give detailed data on PSA and biopsy use in the control arm. Results will provide a more accurate estimation of PCa specific mortality reduction for those men who choose to be screened as compared to an ITS analysis.

MATERIALS AND METHODS

The study population and protocol have been described in detail previously [4, 5]. In summary, in the Rotterdam section of the ERSPC 17,443 men were randomized to the screening arm and 17,390 to the control arm in the core age group of 55-69 year (at time of randomization). Randomization for this study started in 1993. Men in the Rotterdam section of the ERSPC were randomized after providing written informed consent. In the screening arm men were offered PSA testing with a 4 year interval until the age of 75. Initially, a prostate biopsy was offered in men with a PSA level ≥ 4.0 ng/ml and/or an abnormal digital rectal examination (DRE). From May 1997 onwards, a PSA level ≥ 3.0 ng/ml was the only indication for sextant prostate biopsy. The primary endpoint of the ERSPC is PCa specific mortality.

Data on PCa of all men diagnosed outside the screening protocol (both in the screening and the control arm) were collected through linkage with the national cancer registry and subsequent patient chart review of all men with PCa. Cause of death of all men with PCa was assessed by an independent monitoring committee according to a predefined algorithm and blinded for study arm [6].

Follow-up for the current analysis ended December 31, 2010. The study was approved by the medical ethical committee (trial registration: ISRCTN49127736).

Nonattendance in the screening arm

In the screening arm two groups were defined: non-attenders, men refusing PSA testing at the first screening round (men refusing participation were no longer invited for subsequent screening rounds), and attenders, men attending at least the first screening round.

Contamination in the control arm

Two definitions were used for contamination in the control arm. First, contamination in the control arm was defined as having at least one PSA test in the absence of symptoms (opportunistic screening). Through linkage of the ERSPC Rotterdam database to the central laboratory of the Rotterdam region, the Netherlands, PSA testing of men in the control arm could be retrieved. The central laboratory covered 77.7% of all Dutch participants [7, 8]. Data was therefore extrapolated to the entire cohort. An analysis based on self-reported PSA testing of men in the screening arm, showed the 23.3% of GPs not covered by the laboratory were not biased for demanding PSA tests (data not shown). To determine which men received PSA testing for clinical reasons (symptomatic testing) and which men received PSA testing for screening purposes (true contamination), a survey was conducted among general practitioners (GPs) of a random sample of men without PCa. Furthermore, the reason to be referred to the urologist for all men with PSA testing and PCa was known through medical records. These data could then be used to determine the true contamination rate for all Dutch participants in the control arm.

As a screening test can only be seen as such if an abnormal test leads to an additional test to confirm the diagnosis (in the case of PCa screening a prostate biopsy) the second definition of contamination was defined as: having a prostate biopsy at least once in the absence of symptoms (and thus only because of an elevated PSA test). Through linkage with the nationwide network and registry of histo- and cytopathology in the Netherlands (PALGA), data on prostate biopsy of men in the control arm could be retrieved [8, 9]. The PALGA database covers all pathology reports in the Netherlands since 1991 and correct linkage is achieved in up to 98% of cases [10]. True contamination was then defined in the same way as with PSA testing, using reason for referral to the urologist for all men with PCa and the reason to be tested by the GP for all men without PCa.

Both data on PSA testing and prostate biopsy were available until the end of follow-up.

Contamination was however defined as a PSA test or biopsy more than two years before the end of follow-up (before the end of 2008).

Statistical analysis

The effect of screening on the PCa specific mortality for the ITS analysis and adjusted analysis was calculated as relative risk (RR). For the adjustment of nonattendance and contamination the method of Cuzick et al. [2] was applied (figure 1). Three methods for adjustment have been

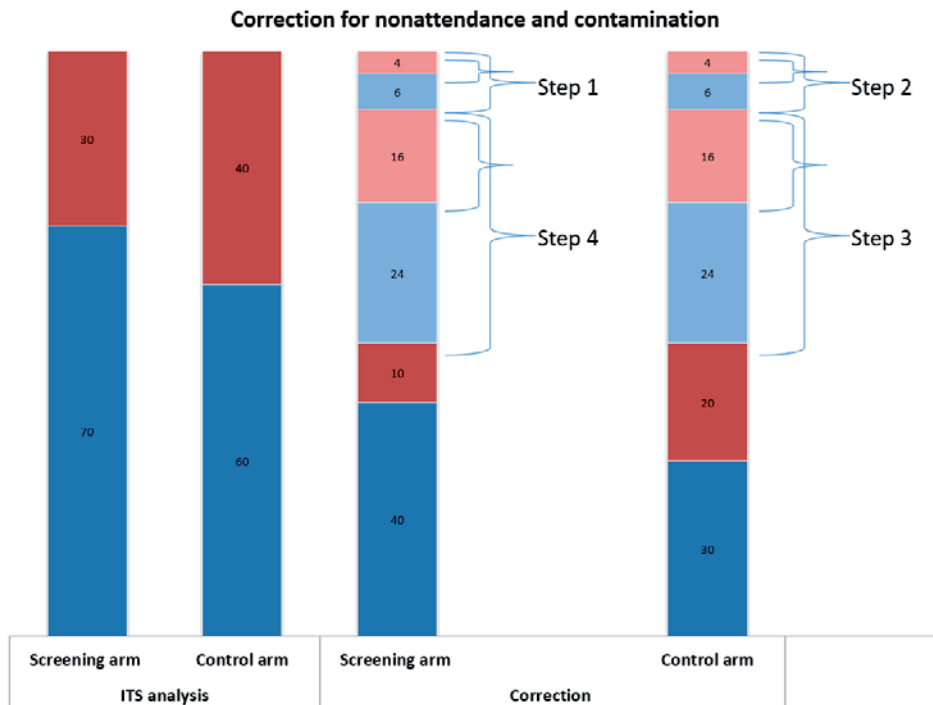


Figure 1. Example of the Cuzick method for the correction of nonattendance and contamination (numbers are fictitious): On the left side the intention to screen (ITS) analysis is shown, with in red the proportion of the endpoint, i.e. prostate cancer mortality. In this example there would be a 25% lower risk of the endpoint in the screening arm versus the control arm (30% versus 40%).

Correction: Step 1: The proportion of non-attenders (10% of participants) with the corresponding endpoint (40% of non-attenders), i.e. prostate cancer mortality, is determined and subtracted from the screening arm. Step 2: Due to the randomization process we can assume that equal numbers of individual per arm are prone to non-attendance and that these individuals as a group have the same baseline risk of dying from prostate cancer. Therefore a similar proportion of participants (10%) in the control arm, with the same rate in endpoints reached (40%), are assumed to would have been noncompliant if not randomized to the control arm. These so called 'potential' non-attenders are therefore subtracted from the control arm. Step 3: The proportion of contaminators in the control arm is determined (40% of which 40% reached the endpoint). Step 4: Due to the randomization it can again be assumed that equal numbers of individual are prone to contamination. Therefore a similar group of participants (40%) with the same rate of endpoints reached (40%) should be subtracted from the control arm (these are the 'potential' contaminators, that would have opted for screening if not randomized to the screening arm). The two remaining groups are the participants adhering to the allocated protocol. Furthermore, due to the randomization and the correction with the method described above, they still have the same baseline risk of dying from prostate cancer. Comparing these groups gives an estimate of the difference in endpoints between the screening and the control arm if nonattendance and contamination would not have occurred. After correction there would be a 50% lower risk of the endpoint in the screening arm versus the control arm (20% versus 40%) in this example. This correction necessitates the assumption that the 'treatment' given has the same effect in all groups. This assumptions might not hold as the screening offered in the screening arm is likely to be more regular than screening as done in the contaminators. Furthermore, men in the screening arm not attending all screening visits (partial compliers) are still classified as 'full' attenders. The correction for contamination and non-attendance could therefore be both under- and overestimated.

described previously: a binary analysis; a Poisson analysis, taking into account time to PCa death, nonattendance and contamination; and a semiparametric Cox proportional hazard analysis, assuming nonattendance and contamination occurred at randomization. Here, the binary analysis was used because all models gave very similar results, as described by Kerkhof et al [11].

Although the ERSPC section Rotterdam was not designed as a stand-alone trial, a separate power calculation was done as described previously [1].

RESULTS

The total number of men in the core age group was 34,833. At a median follow-up of 13.0 year 2,226 men were diagnosed with PCa in the screening arm (cumulative incidence, 12.8%) and 96 men died of their disease. In the control arm 1,152 men were diagnosed with PCa (cumulative incidence, 6.6%) and 140 died of their disease at a median follow-up of 13.0 year. A detailed description of the PCAs found is given in [1] and in table 1 and 2. Using a binary ITS analysis (no correction for attendance and contamination) the PCa specific mortality reduction in the screening arm compared to the control arm was 0.68 (95% confidence interval (CI), 0.53-0.89) at the end of follow-up.

Table 1. Clinical characteristics of prostate cancer cases in the screening arm for non-attenders and attenders

	Non-attenders (%)	Attenders (%)	Total (%)
Age at baseline, median	60.9	62.4	62.4
PSA, median	11.9	4.9	4.9
T-stage			
1	18 (38.3)	1257 (57.7)	1275 (57.3)
2	12 (25.5)	659 (30.2)	671 (30.1)
3	12 (25.5)	222 (10.2)	234 (10.5)
4	2 (4.3)	16 (0.7)	18 (0.8)
missing	3 (6.4)	25 (1.1)	28 (1.3)
Gleason score			
<=6	22 (46.8)	1586 (72.8)	1608 (72.2)
7	15 (31.9)	445 (20.4)	460 (20.7)
>=8	6 (12.8)	141 (6.5)	147 (6.6)
missing	4 (8.5)	7 (0.3)	11 (0.5)
M+			
0/X	41 (87.2)	2143 (98.3)	2184 (98.1)
1	6 (12.8)	36 (1.7)	42 (1.9)
Total	47 (100)	2179 (100)	2226 (100)

Table 2. Clinical characteristics of prostate cancer cases in the control arm for true biopsy contaminators and non- true biopsy contaminators

	True biopsy contaminators	Non- true biopsy contaminators	Total (%)
Age, median	62.8	63.2	63.1
PSA, median	10.7	11.0	11.0
T-stage			
1	220 (52.4)	352 (48.1)	572 (49.7)
2	104 (24.8)	180 (24.6)	284 (24.7)
3	72 (17.1)	145 (19.8)	217 (18.8)
4	15 (3.6)	38 (5.2)	53 (4.6)
missing	9 (2.1)	17 (2.3)	26 (2.3)
Gleason score			
<=6	225 (53.6)	354 (48.4)	579 (50.3)
7	133 (31.7)	203 (27.7)	336 (29.2)
>=8	59 (14.0)	163 (22.3)	222 (19.3)
missing	4 (1.0)	11 (1.5)	15 (1.3)
M+			
0/X	387 (92.1)	629 (85.9)	1016 (88.2)
1	33 (7.9)	103 (14.1)	136 (11.8)
Total	420 (100)	732 (100)	1152 (100)

Nonattendance in the screening arm

In total 16,502 men (94.6%) attended at least one screening round. The numbers of men per group and the numbers of (PCa) deaths are given in Table 3. Overall mortality was higher in non-attenders (37.6%) versus attenders (24.4%).

PSA contamination in the control arm

Of the 17,443 men in the Dutch control arm of the ERSPC, 6,880 (39.5%) had at least one PSA test before December 31, 2008 (extrapolated from a 77.7% coverage, according to Roemeling et al. [7]). Of these men, 660 were diagnosed with PCa (9.6%) and 81 men (12.3%) died of their disease at the end of follow-up (December 31, 2010) (table 4). Subsequently, a survey was conducted among GPs of a random sample of 671 men without PCa (10% of men with PSA testing). A total of 585 questionnaires (87.2%) were returned. Reason for PSA testing was unknown in 117 cases, of the remaining 468 cases reason for PSA testing was diagnostic (voiding problems, other relevant problems or suspicious digital rectal examination) in 50.2% and for screening purposes (patient requested testing or test was part of general blood exam) in 49.8%. So, of men without PCa but with PSA testing, 49.8% were defined as contaminators. After careful investigation of reason for referral of all men with PCa and PSA testing, 275 of 660 men (41.7%) were classified as contaminators (no symptoms or suspicious DRE). A total of 27 out of 81 PCa deaths (33.9%) were attributed to these true contaminators. The number of total true contaminators hence was 3,372

Table 3. Nonattendance in the screening arm of the European Randomized study of Screening for Prostate Cancer (ERSPC) section Rotterdam

Mortality follow-up until the end of 2010			
	Non-attenders	Attenders	Total
n (% of Total)	941 (5.4)	16502 (94.6)	17443 (100)
PCa diagnosis (% of n)	47 (5.0)	2179 (13.2)	2226 (12.8)
Death overall (% of n)	354 (37.6)	4021 (24.4)	4375 (25.1)
PCa death (% of n)	7 (0.74)	89 (0.54)	96 (0.55)

PCa = prostate cancer

Table 4. Contamination in the control arm of the European Randomized study of Screening for Prostate Cancer (ERSPC) section Rotterdam using two definitions: a PSA test before the end of 2008 and a prostate biopsy before the end of 2008

Mortality follow-up until the end of 2010			
	PSA test	Biopsy	Total
n (% of Total)	6880 (39.6)	2422 (13.9)	17390 (100)
PCa diagnosis (% of n)	660 (9.6)	923 (38.1)	1152 (6.6)
Death overall (% of n)	1440 (20.9)	603 (24.9)	4355 (25.0)
PCa death (% of n)	81 (1.18)	126 (5.20)	140 (0.81)
True contaminators* (% of total)	3372 (19.4)	1071 (6.2)	17390 (100)
PCa diagnosis (% of true)	275 (8.2)	420 (39.2)	
PCa death (% of true)	27 (0.81)	46 (4.33)	

PCa = prostate cancer; * PSA testing/biopsy in the absence of symptoms (opportunistic screening)

((49.8% of men without PCa = 3,097) + (41.7% of men with PCa=275)). The PSA contamination rate in the control arm was therefore 19.4% (3,372/17,390) (table 4).

Biopsy contamination in the control arm

A total of 2,422 men in the control arm received at least one prostate biopsy before the end of 2008 (13.9% of all men in the control arm). A total of 923 (38.1%) of these men were diagnosed with prostate cancer. In total 46 PCa deaths and 420 PCa cases in the control arm were classified as true contaminators (table 4).

Of the 896 men diagnosed with prostate cancer in the control arm until the end of 2008, 74 were not detected through linkage with the PALGA database (896-822=74). Chart review showed 15 were diagnosed by prostate biopsy (and were thus missed with the PALGA linkage) and 59 were diagnosed in another way (TURP, cystoprostatectomy, other) (These were not detected because linkage was only done for biopsies). The correct detection of the linkage with the PALGA database for men with PCa was 98.2% (822/(896-59)).

Disease specific mortality adjusted for nonattendance and contamination

Adjustment for nonattendance and biopsy contamination using the method of Cuzick et al. [2] is shown in figure 2. The correction for nonattendance and biopsy contamination resulted in a reduction of the PCa specific mortality of 51% in favour of screening (RR 0.49, 95% CI 0.27-0.87). Correction for non-attendance alone had a small effect (RR 0.68 versus RR 0.67)(table 5).

In figure 3 the absolute risk of dying from PCa per arm is given for different years of ending follow-up, with or without correction for non-attendance and PSA contamination. At the end of follow-up (2010), 5.5 PCa deaths/ 1000 men occurred in the screening arm versus 8.1/ 1000 in the control arm (ITS analysis). After correction for nonattendance and PSA contamination there were 4.7 PCa deaths/1000 men in the screening arm versus 8.1/1000 in the control arm, a difference of 3.4/1000.

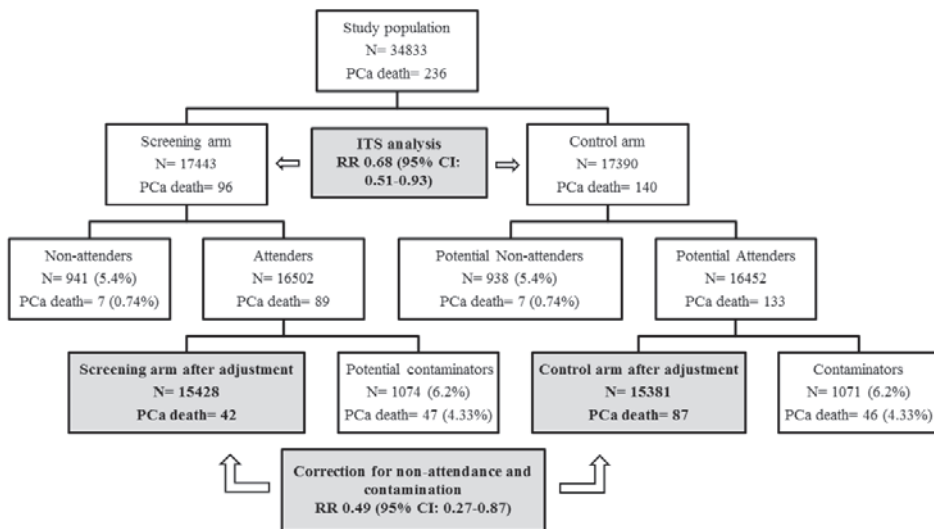


Figure 2. Correction for nonattendance and biopsy contamination in the Rotterdam section of the ERSPC using the method of Cuzick et al. [2]. Mortality follow-up until the end of 2010

DISCUSSION

PSA based PCa screening as conducted in the Dutch centre of the ERSPC (4 year interval, PSA cut-off ≥ 3.0 ng/ml and lateralised sextant biopsy) resulted in a reduction of PCa mortality of 32% using an ITS analysis after a median of 13 years of follow-up. This can be regarded as the effect of screening on a population based level. However, for an individual man who attended screening as mentioned above the current study shows that the risk of dying from PCa can be reduced with up to 51% as compared to a man not screened at all. This information might serve

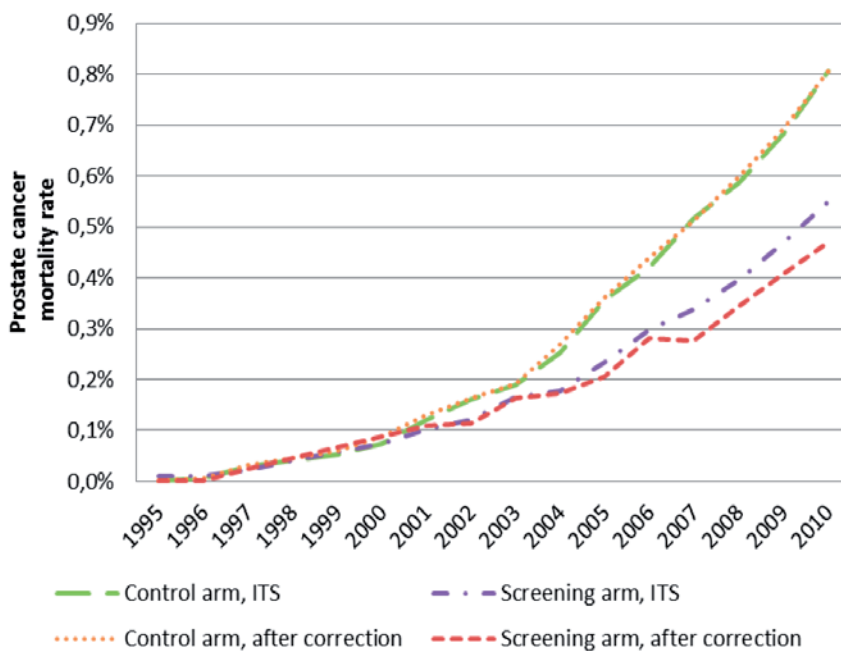


Figure 3. Absolute risk of dying from prostate cancer per arm for the intention to screen (ITS) analysis and after correction for non-attendance and PSA contamination. Results are given for different years of ending follow-up.

men who face the dilemma of PCa screening to make a more balanced judgement between the harms and benefits.

The amount of men that attended at least one screening round was high (95%) in the current study. Therefore the effect of correction for only non-attenders was minimal. The PCa mortality

Table 5. Reduction of prostate cancer (PCa) specific mortality from screening (relative risk (RR)) for the intention to screen (ITS) analysis, correction for nonattendance and correction for contamination.

Mortality follow-up until the end of 2010			
	RR	95% CI	p-value
ITS	0.68	(0.53-0.89)	0.004
Correction for nonattendance			
Adjustment for non-attenders	0.67	(0.51-0.88)	0.004
Correction for contamination			
PSA contamination	0.61	(0.42-0.88)	0.008
Biopsy contamination	0.53	(0.32-0.88)	0.014
Correction for nonattendance and contamination			
Non-attenders + PSA	0.58	(0.39-0.86)	0.007
Non-attenders + Biopsy	0.49	(0.27-0.87)	0.015

rate of the non-attenders was higher than the PCa mortality rate of the entire screening arm. It was however lower than the PCa mortality rate of the control arm. This could be explained by the higher overall mortality rate in the non-attenders, which was seen in other ERSPC centres as well [12]. It seems these men had a worse overall health status at the beginning of the trial and this could have been the reason for not complying with the screening protocol. It also shows that there is a baseline difference between the different groups. A simple subtraction of these men without using the method of Cuzick et al.[2] could therefore have misinterpreted the risk of PCa death for the men that did comply with the screening protocol [13].

During follow-up 19.4% of men in the control arm had asymptomatic PSA testing at least once until the end of 2008. Correction for this so-called PSA contamination had a relatively small effect on the mortality reduction as compared to the biopsy contamination. The PCa mortality rate in men having PSA testing in the control arm was lower than the total PCa mortality rate of the control arm. It is however still higher than in the screening arm. Although groups could be dissimilar in baseline risk, results suggest unorganized screening is not as effective as organized screening (as conducted in the ERSPC) in reducing the prostate cancer mortality. This is in contrast with the results published from the PLCO trial [14]. Reason for this difference in effect of screening could be that men in the screening arm of the ERSPC were advised to undergo prostate biopsy if PSA was ≥ 3.0 ng/ml. More than 90% of men complied with this biopsy recommendation [15]. Outside the study protocol it is not common practice to perform a biopsy if PSA is ≥ 3.0 ng/ml. In a study by Otto et al. [8] it was shown that only 7.7% of men with a PSA ≥ 3.0 ng/ml in the control arm got a subsequent prostate biopsy within 6 months. In the PLCO trial the decision to perform a prostate biopsy was not protocol based, but was left over to the participant and his health-care provider [14].

Furthermore, only 60% of all cancers detected in the control arm until 2008 (546/896) were detected in men with a PSA test ordered by the GP. The remaining men were presumably diagnosed after direct referral to the urologist. In the current analysis PSA testing at the urologist was not assessed. The PSA contamination as assessed in the current study is therefore likely to underestimate the true PSA contamination rate in the control arm.

With the biopsy contamination 98% of all men with PCa were detected. Therefore a more complete assessment could be made which men were true contaminators (biopsied only because of an elevated PSA) and which men were biopsied for diagnostic purposes. Of all biopsies in the control arm 44% were done in the absence of symptoms and thus for screening purposes. Correction for biopsy contamination hence resulted in the largest increase in RR (RR 0.53 versus RR 0.68 for the ITS analysis). The PCa detection rate of 39.2% in the true biopsy contaminators might seem relatively high for asymptomatic men. True contaminators were defined as having an asymptomatic biopsy at least once. However, more than 30% of these men (data not shown) had more than one biopsy, which resulted in the relatively high detection rate.

The absolute difference in PCa deaths between the arms after correction for nonattendance and PSA contamination was 3.4/1000 men with a median follow-up of 13 years. This number is

however highly dependent on the percentage of men in both arms that died [16] and therefore changes over time. Until the end of follow-up (2010) only 25% of participating men had died. The real absolute difference can therefore only be given if all participants have died.

In the contamination group, especially in the PSA contaminators, the overall mortality rate was lower than in the rest of the control arm (20.9%). This could indicate that men in the control arm who chose to undergo a PSA test have a better life expectancy. As death from PCa may increase even after 15 years, especially in men with a longer life expectancy, longer follow-up could further increase the effect of the correction [17].

The diluting effect of contamination and nonattendance is not exclusively limited to the effect of screening on PCa mortality. The downside of screening, overdiagnosis (with subsequent overtreatment), is watered down as well. It should therefore be recognized that although the effect of screening as conducted in this study has a major effect on PCa mortality of an individual man, PSA based screening is still far from ideal. If screening is still asked for, already available tools, such as risk calculators, should be used to reduce the overtesting and overdiagnosis [18-22]. Furthermore, a man demanding a PSA test for screening purposes should be informed about the large risk of detecting low-risk prostate cancer with screening beforehand and treatment options like active surveillance to reduce the side effects of radical treatment without effecting oncological outcome [23-25].

The current analysis is limited by the extrapolation of the PSA contamination data and by the assessment of the reason for a PSA test using a questionnaire of a random part of men without PCa. Furthermore, the correction for biopsy contamination is most dependent on the determination of the amount of PCa deaths that are classified as true contaminators. For follow-up until the end of 2010, 36.8% of all men in the control arm with a biopsy and who died of prostate cancer were classified as contaminators (this translates into 46 men or 4.33% of contaminators). If this 36.8% was overclassified with 5% points the chance of dying from prostate cancer in the screening arm would be 48% lower after correction instead of 51%.

The current analysis is only done in a single centre of the ERSPC, in a single country, limiting the possibility to extrapolate the results. In the Rotterdam section of the ERSPC, data on reason for referral of men with PCa was however complete making the determination of the true contamination rate in these men very precise. Last, the current analysis was limited to the predefined core age group of the ERSPC [26]. A sub-analysis in men outside the core age group (70-74 years) showed no significant benefit of screening in this subgroup (RR 1.23, 95% CI 0.70-2.16)(data not shown).

CONCLUSION

The effect of screening for an individual man choosing to be screened repeatedly with PSA tests is higher than the effect provided by large randomised trials, which provide an estimation of the

effect on a population based level. Screening as conducted in the Dutch centre of the ERSPC can reduce the risk of PCa death with up to 51% in men who undergo organized screening. This information can be helpful in informed decision making on PSA based screening for prostate cancer.

ACKNOWLEDGEMENT

The authors acknowledge M. Kerfhof and S. Roemeling for performing parts of the conducted questionnaire.

REFERENCES

1. Roobol MJ, Kranse R, Bangma CH, van Leenders GJLH, Blijenberg BG, van Schaik RHN, et al. Screening for Prostate Cancer: Results of the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer (ERSPC). *Eur urol.* 2013;In press.
2. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med.* 1997;16:1017-29.
3. Roobol MJ, Kerkhof M, Schroder FH, Cuzick J, Sasieni P, Hakama M, et al. 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;56:584-91.
4. Roobol MJ, Schröder FH. European Randomized study of Screening for Prostate Cancer (ERSPC): rationale, structure and preliminary results. *BJU international.* 2003;92: Suppl 2:1-122.
5. Roobol MJ, Kirkels WJ, Schroder FH. Features and preliminary results of the Dutch centre of the ERSPC (Rotterdam, the Netherlands). *BJU international.* 2003;92 Suppl 2:48-54.
6. De Koning HJ, Blom J, Merkelbach JW, Raaijmakers R, Verhaegen H, Van Vliet P, et al. Determining the cause of death in randomized screening trial(s) for prostate cancer. *BJU international.* 2003;92 Suppl 2:71-8.
7. Roemeling S, Roobol MJ, Otto SJ, Habbema DF, Gosselaar C, Lous JJ, et al. Feasibility study of adjustment for contamination and non-compliance in a prostate cancer screening trial. *Prostate.* 2007;67:1053-60.
8. Otto SJ, van der Crujisen IW, Liem MK, Korfage IJ, Lous JJ, Schroder FH, et al. Effective PSA contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer. *Int J Cancer.* 2003;105:394-9.
9. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* 2007;29:19-24.
10. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. *Int J Epidemiol.* 1990;19:553-8.
11. Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schroder FH, et al. Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomized screening trial (ERSPC section Rotterdam). *Int J Cancer.* 2010;127:2639-44.
12. Bergdahl AG, Aus G, Lilja H, Hugosson J. Risk of dying from prostate cancer in men randomized to screening: differences between attendees and nonattendees. *Cancer.* 2009;115:5672-9.
13. Kranse R, van Leeuwen PJ, Hakulinen T, Hugosson J, Tammela TL, Ciatto S, et al. Excess all-cause mortality in the evaluation of a screening trial to account for selective participation. *J Med Screen.* 2013; Epub ahead of print.
14. Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst.* 2012;104:125-32.
15. Bokhorst LP, Zhu X, Bul M, Bangma CH, Schroder FH, Roobol MJ. Positive predictive value of prostate biopsy indicated by prostate-specific-antigen-based prostate cancer screening: trends over time in a European randomized trial*. *BJU international.* 2012;110:1654-60.

16. Roobol MJ, Kranse R, Bangma CH, Otto SJ, van der Kwast TH, Bokhorst LP, et al. Reply from Authors re: Michael Baum. Screening for Prostate Cancer: Can We Learn from the Mistakes of the Breast Screening Experience? *Eur Urol*. In press.: Screening for Prostate Cancer: We Have Learned and Are Still Learning. *Eur Urol*. 2013;In press.
17. Popiolek M, Rider JR, Andren O, Andersson SO, Holmberg L, Adami HO, et al. Natural history of early, localized prostate cancer: a final report from three decades of follow-up. *Eur Urol*. 2013;63:428-35.
18. Roobol MJ, Carlsson SV. Risk stratification in prostate cancer screening. *Nat Rev Urol*. 2013;10:38-48.
19. Roobol MJ, Zhu X, Schroder FH, van Leenders GJ, van Schaik RH, Bangma CH, et al. A Calculator for Prostate Cancer Risk 4 Years After an Initially Negative Screen: Findings from ERSPC Rotterdam. *Eur Urol*. 2013;63:627-33.
20. Roobol MJ, van Vugt HA, Loeb S, Zhu X, Bul M, Bangma CH, et al. Prediction of Prostate Cancer Risk: The Role of Prostate Volume and Digital Rectal Examination in the ERSPC Risk Calculators. *Eur Urol*. 2011;61:577-83.
21. Roobol MJ, Schröder FH, Crawford ED, Freedland SJ, Sartor AO, Fleshner N, et al. A framework for the identification of men at increased risk for prostate cancer. *The Journal of urology*. 2009;182:2112-20.
22. Zhu X, Albertsen PC, Andriole GL, Roobol MJ, Schroder FH, Vickers AJ. Risk-Based Prostate Cancer Screening. *Eur Urol*. 2011;61:652-61.
23. Venderbos LD, Bokhorst LP, Bangma CH, Roobol MJ. Active surveillance: oncologic outcome. *Curr Opin Urol*. 2013;23:268-72.
24. Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, et al. Active Surveillance for Low-Risk Prostate Cancer Worldwide: The PRIAS Study. *Eur Urol*. 2012;63:597-603.
25. Dall'Era MA, Albertsen PC, Bangma C, Carroll PR, Carter HB, Cooperberg MR, et al. Active surveillance for prostate cancer: a systematic review of the literature. *Eur Urol*. 2012;62:976-83.
26. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med*. 2012;366:981-90.

Chapter 3

Differences in treatment and outcome after treatment with curative intent in the screening and control arm of the ERSPC Rotterdam.

Leonard P. Bokhorst^a, Ries Kranse^b, Lionne D.F. Venderbos^a, Jolanda W. Salman^a, Geert J.L.H. van Leenders^c, Fritz H. Schröder^a, Chris H. Bangma^a, Monique J. Roobol^a, the ERSPC Rotterdam study group.

^a Department of urology, Erasmus University Medical Center, Rotterdam, the Netherlands

^b Comprehensive Cancer Center, Rotterdam, the Netherlands

^c Department of pathology, Erasmus University Medical Center, Rotterdam, the Netherlands

ABSTRACT

Screening for prostate cancer (PCa) results in a favorable stage shift. However even if screening did not result in a clinically apparent lower stage or grade, it might still result in less disease recurrence after treatment with curative intent (radical prostatectomy (RP) and radiation therapy (RT)), because the tumor had less time to develop outside the prostate. In addition, the outcome after treatment could differ because of differences in treatment quality (e.g. radiation dosage/adjuvant hormonal therapy). To test these hypothesis we compare differences in treatment quality of the screening and control arm of the European Randomized study of Screening for Prostate Cancer (ERSPC) Rotterdam and disease free survival after curative treatment in PCa patients with similar stage and grade. In total 2595 men were initially treated with RP or RT. In the control arm, RT was more often combined with hormonal therapy and treatment dosage was more often ≥ 69 Gy. This resulted most likely from changes over time in treatment which coincided with the later detection in the control arm. Disease free survival was higher in the screening arm in all risk-groups. After correction for lead-time these differences were however minimal. We concluded that treatment quality differed between the screening and control arm of the ERSPC, Rotterdam. Especially radiotherapy quality was superior in the control arm with higher dosages and more often radiotherapy in combination with hormonal therapy. Despite these differences, favoring the control arm, disease free survival differences were minimal.

Trial registration: ISRCTN49127736.

Patient summary: In this report we looked at differences in prostate cancer treatment and outcome after prostate cancer treatment in men diagnosed after screening and men diagnosed after normal clinical practice. Treatment differed with superior treatment given in men diagnosed in normal clinical practice. In this paper we have proposed a likely explanation for this, at first sight, counter intuitive finding (progressive insight combined with an, on average, later detection of tumors in unscreened men). Despite the fact that unscreened men received better treatment this advantage seemed to be outweighed by the advantage associated with the, on average, earlier detection of the tumor in screened men.

Screening as done in the Rotterdam section of the European Randomized study of Screening for Prostate Cancer (ERSPC) has shown to reduce prostate cancer (PCa) specific mortality with 32% after 13 years of follow-up, and with up to 51% after correction for nonattendance and contamination [1, 2].

PCa screening may achieve this positive effect in multiple ways. Most likely as a result of the favorable stage shift in the screening arm, resulting e.g. in less men with advanced and metastatic disease and thus more curative rather than palliative treatment [1, 3]. But, even if screening did not result in a clinically noticeable lower stage or grade for an individual patient, earlier detection with its associated earlier treatment could still have prevented micro-metastatic tumor development outside the prostate, resulting in higher curation rates after treatment in men with screen detected PCa compared to men with clinically detected cancer with a similar clinical stage and grade. This principle is visualized in figure 1. In addition, treatment between arms could differ, this includes differences in treatment modality (e.g. radical prostatectomy (RP) versus radiation therapy (RT)), but also differences in treatment quality of similar treatment modalities (e.g. radiation dosage/adjuvant hormonal therapy in men receiving RT). We therefore aimed to compare differences in treatment quality between the screening and control arm of the ERSPC Rotterdam and to compare the outcome of curative treatment in PCa patients within similar prognostic groups, based on stage and grade at time of diagnosis, to test the hypothesis that screen detected and clinically detected men with similar characteristics might still have a different prognosis.

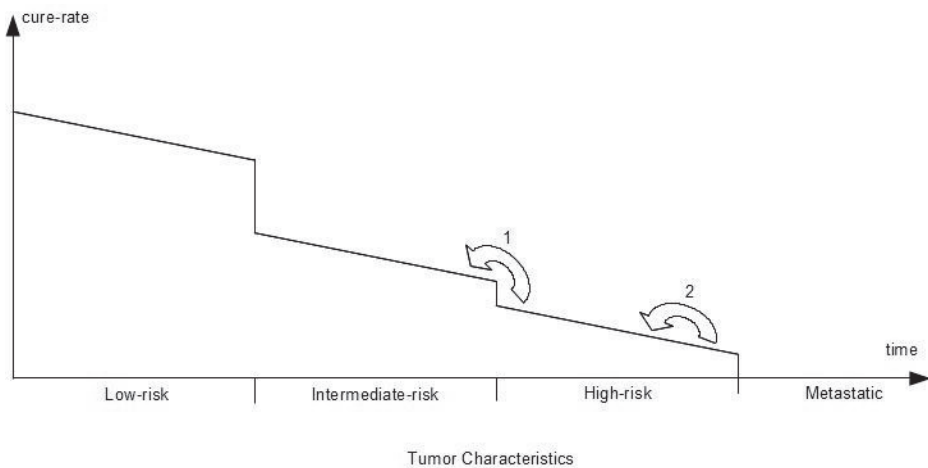


Figure 1. Model on how screening could result in better outcome

- 1: Screening could result in a stage shift, resulting in better prognosis.
- 2: If screening did not result in a clinical apparent stage shift, earlier diagnosis and treatment in time could still have resulted in a better prognosis, for instance because of less time for the tumor to develop outside the prostate.

The screening protocol and study population of the ERSPC Rotterdam have previously been described in detail [4, 5]. For this analysis all men receiving RP or RT as initial treatment were compared. Characteristics of treatment quality were studied. For comparison, men were divided by clinical characteristics into four risk groups: low-, intermediate- and high-risk PCa, based on the criteria of D'Amico et al.[6], and a separate group for men with metastatic PCa (M1 and/or PSA \geq 100 ng/ml). The last group was not further assessed in the current analysis as treatment was not with curative intent. Furthermore, disease free survival (DFS) defined as no biochemical recurrence (BCR), i.e. a PSA value 2 times $>$ 0.2ng/ml after RP, or a PSA value \geq 2.0ng/ml above the PSA nadir after RT, no local progression, no distant metastasis, no PCa death and no additional treatment during follow-up was compared between the arms using the Kaplan-Meier method. The between arm survival curve comparisons were done for equal clinical parameters and most important treatment quality characteristics, as these could affect results. The comparison was certainly affected by lead-time in the screening arm (notably for low and intermediate risk disease). For PCa, this lead-time was estimated to range from 12.2 to 2.9 years depending on tumor characteristics [7]. Therefore, a method described by Duffy et al.[8] was used to correct survival times in the screening arm for tumor characteristic specific lead-times.

Of all PCa cases in the ERSPC Rotterdam, 2595 (62.6%) were initially treated with RP or RT (supplemental table 1). Within the pre-defined groups, tumor characteristics at RP, as extra capsular extension, were less favorable in the control arm (supplemental table 2a). Furthermore, surgical margins were more often positive in the control arm in the low- and intermediate-risk group. These differences indicate that even within similar risk groups tumors were more advanced in the control arm. Most likely due to the later detection in time.

TREATMENT COMPARISON

If looked at treatment itself, especially in men receiving RT, treatment was superior in the control arm. In the high-risk group of the control arm more men received hormonal therapy (HT) in addition to RT (50.4% versus 12.0%)(supplemental table 2b). The addition of HT to RT for men with higher risk tumors was proven to increase overall survival in a randomized trial first published in 2002 [9]. The majority of men in the screening arm were diagnosed and treated before the first publication of this trial, resulting in very low rates of men receiving HT in addition to RT in the screening arm. In the control arm diagnosis was more often after 2002, meaning these men could benefit from the improved treatment (RT + HT). In addition to the combination of HT and RT, radiation dosages given in the control arm were significantly higher than in the screening arm (supplemental table 2b). This again is most likely a result of the later diagnosis in time of men in the control arm as treatment dosages gradually increased during the course of the trial. The superior RT in the control arm could have improved PCa-specific survival in the control arm.

In the RP group, surgical technique (e.g. open or laparoscopic) and individual surgeons case load/hospital volume could have differed between the screening and control arm, but both were not available for analysis. It could be expected that surgical technique might have changed over time in favor of the control arm, as with RT. Individual surgeon case load/hospital volume could have effected surgical margin status. In the Netherlands, differences between hospitals are however smaller than in some other countries, with the largest hospital performing 240 RP in 2010 [10], it might therefore be expected that its effect is small.

DISEASE FREE SURVIVAL COMPARISON

Without correction for lead-time DFS rates were higher in the screening arm in both men receiving RP and RT. After correcting the screening arm for lead-time, differences were however less apparent (figure 2 and supplemental figure 3). Correction (based on an assumed exponential model [8]) seems not perfect as it resulted in lower DFS rates in the screening arm directly after diagnosis. At the end of the survival curves DFS in the screening arm (corrected for lead-time) was higher than in the control arm. Point estimate comparison (Z-test) at the end of the survival curves only resulted in significant differences in the Gleason score 7 group in men receiving RP with positive surgical margins (DFS 38% in the screening arm versus 13% in the control arm at 9.8 years ($p=0.046$)) and in men receiving RT with a dosage <69 Gy (DFS 47% in the screening arm versus 9% in the control arm at 7.9 years ($p<0.001$)). This corroborates the hypothesis that early detection and treatment reduces PCa development outside the prostate and therefore increases DFS. More detailed data on outcome after treatment and additional treatment given can be found in supplemental table 3.

The current analysis was further limited by overdiagnosis [11]. Overdiagnosed cancers could have resulted in more favorable outcomes in the screening arm, especially in the group of low-risk PCa. This makes interpretation of the difference between the screening and control arm in especially the low-risk PCa group difficult.

CONCLUSION

The fact that the ERSPC is a randomized study does not imply that treatments in both study arms are comparable. Therefore this study aimed to compare differences in treatment quality in men receiving treatment with curative intent in the screening and control arm of the ERSPC Rotterdam, as well as comparing the outcome of men with clinically similar tumor characteristics. We found difference in the quality of similar treatments between the screening and control arm of the ERSPC, Rotterdam. Especially RT quality was superior in the control arm with higher dosages and more often RT in combination with HT. We provided a reasonable explanation (progressive

insight combined with later detection in control arm). Despite these quality differences in treatment that favored the control arm, disease free survival in men with similar treatment and tumor characteristics was marginally better in the screening arm.

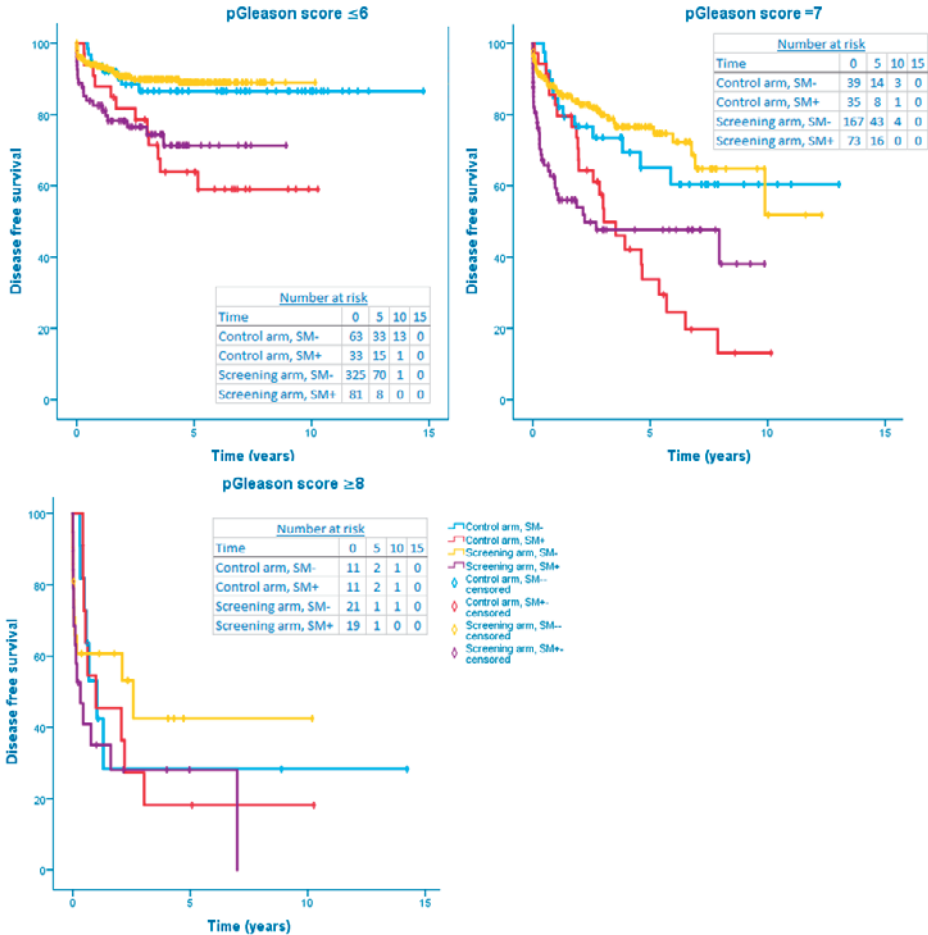


Figure 2. Disease free survival after radical prostatectomy for men with T1-T2 disease, stratified by pGleason score (≤ 6 , 7 , ≥ 8), surgical margin (SM)(positive (+)/negative(-)), and arm, after correction for lead-time in the screening arm. Time in the screening arm was corrected for lead time using a method previously described [8].

REFERENCES

1. Roobol MJ, Kranse R, Bangma CH, van Leenders AG, Blijenberg BG, van Schaik RH, et al. Screening for Prostate Cancer: Results of the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol.* 2013;64:530-9.
2. Bokhorst LP, Bangma CH, van Leenders GJ, Lous JJ, Moss SM, Schroder FH, et al. Prostate-specific Antigen-Based Prostate Cancer Screening: Reduction of Prostate Cancer Mortality After Correction for Nonattendance and Contamination in the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol.* 2014;65:329-36.
3. Schroder FH, Hugosson J, Carlsson S, Tammela T, Maattanen L, Auvinen A, et al. Screening for Prostate Cancer Decreases the Risk of Developing Metastatic Disease: Findings from the European Randomized Study of Screening for Prostate Cancer (ERSPC). *Eur Urol.* 2012;62:745-52.
4. Roobol MJ, Kirkels WJ, Schroder FH. Features and preliminary results of the Dutch centre of the ERSPC (Rotterdam, the Netherlands). *BJU international.* 2003;92 Suppl 2:48-54.
5. Roobol MJ, Schröder FH. European Randomized study of Screening for Prostate Cancer (ERSPC): rationale, structure and preliminary results. *BJU international.* 2003;92: Suppl 2:1-122.
6. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA.* 1998;280:969-74.
7. Wever EM, Heijnsdijk EA, Draisma G, Bangma CH, Roobol MJ, Schroder FH, et al. Treatment of local-regional prostate cancer detected by PSA screening: benefits and harms according to prognostic factors. *British journal of cancer.* 2013;108:1971-7.
8. Duffy SW, Nagtegaal ID, Wallis M, Cafferty FH, Houssami N, Warwick J, et al. Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival. *Am J Epidemiol.* 2008; 168:98-104.
9. Bolla M, Collette L, Blank L, Warde P, Dubois JB, Mirimanoff RO, et al. 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.
10. Rapportage prostaatanker SCK rapport 2014. Published online 27-01-2014 at: www.wikin.nl.
11. Draisma G, Boer R, Otto SJ, van der Crujisen IW, Damhuis RA, Schroder FH, et al. Lead times and overdiagnosis 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.

APPENDIX.

Table 1. Baseline characteristics.

	Control n=791 (% of column total)	Screening n=1804 (% of column total)	Total n=2595 (% of column total)
Age at randomization (years), median	63.2	63.7	63.5
Age at diagnosis (years), median	70.1	67.4	68.2
PSA at diagnosis (ng/ml), median	10.3	5.6	6.8
Time randomization-diagnosis, median	6.6	0.6	4.2
Follow-up time after diagnosis, median	5.2	8.7	7.6
Charlson comorbidity score at diagnosis			
0	336 (42.5)	802 (44.5)	1138 (43.9)
1-2	360 (45.5)	666 (36.9)	1026 (39.5)
3-4	92 (11.6)	331 (18.3)	423 (16.3)
>=5	3 (0.4)	5 (0.3)	8 (0.3)
cT-stage at diagnosis			
1	379 (47.9)	878 (48.7)	1257 (48.4)
2	238 (30.1)	648 (35.9)	886 (34.1)
3	161 (20.4)	263 (14.6)	424 (16.3)
4	8 (1)	8 (0.4)	16 (0.6)
Missing	5 (0.6)	7 (0.4)	12 (0.5)
Gleason score at diagnosis*			
<=6	393 (49.7)	1169 (64.8)	1562 (60.2)
7	262 (33.1)	486 (26.9)	748 (28.8)
>=8	136 (17.2)	147 (8.1)	283 (10.9)
Missing	0 (0)	2 (0.1)	2 (0.1)
Risk-group at diagnosis			
Low	182 (23)	799 (44.3)	981 (37.8)
Intermediate	258 (32.6)	455 (25.2)	713 (27.5)
High	336 (42.5)	532 (29.5)	868 (33.4)
Meta	8 (1)	4 (0.2)	12 (0.5)
Missing	7 (0.9)	14 (0.8)	21 (0.8)
Total	791 (100)	1804 (100)	2595 (100)

*In 84 men (3% of total) Gleason scores at diagnosis were missing. Instead tumor grade was recorded as Gleason score group (Grade 1 = Gleason <=6, Grade 2 = Gleason 7, Grade 3 = Gleason >=8)

Table 2. Characteristics of initial treatment, stratified by initial treatment modality, risk group, and arm.
Table 2a. Radical prostatectomy

Risk group:	Low-risk		Intermediate-risk		High-risk		
	Control n=75 (% of column total)	Screening n=415 (% of column total)	Control n=89 (% of column total)	Screening n=194 (% of column total)	Control n=58 (% of column total)	Screening n=165 (% of column total)	p-value*
Time randomization-diagnosis, median	6.0	0.3	6.2	1.2	5.4	0.2	0.000
LN Dissection	39 (52.0)	338 (81.4)	66 (74.2)	168 (86.6)	48 (82.8)	146 (88.5)	0.265
pN1	0 (0)	0 (0)	2 (2.2)	0 (0)	2 (3.4)	4 (2.4)	0.654
Tumor size		0.070			0.146		0.112
< half lobe	9 (12.0)	108 (26)	5 (5.6)	24 (12.4)	4 (6.9)	31 (18.8)	
>= half lobe	6 (8.0)	32 (7.7)	10 (11.2)	14 (7.2)	3 (5.2)	6 (3.6)	
Bilateral	50 (66.7)	256 (61.7)	68 (76.4)	150 (77.3)	43 (74.1)	112 (67.9)	
Missing	10 (13.3)	19 (4.6)	6 (6.7)	6 (3.1)	8 (13.8)	16 (9.7)	
Capsule		0.007		0.022			0.405
capsular invasion	13 (17.3)	148 (35.7)	20 (22.5)	74 (38.1)	11 (19.0)	44 (26.7)	
Extracapsular extension	12 (16.0)	36 (8.7)	28 (31.5)	41 (21.1)	19 (32.8)	55 (33.3)	
missing	17 (22.7)	52 (12.5)	12 (13.5)	25 (12.9)	8 (13.8)	21 (12.7)	
SVI	2 (2.7)	9 (2.2)	10 (11.2)	15 (7.7)	11 (19.0)	13 (7.9)	0.014
missing	11 (14.7)	22 (5.3)	1 (1.1)	5 (2.6)	6 (10.3)	12 (7.3)	
pt4	1 (1.3)	8 (1.9)	1 (1.1)	9 (4.6)	4 (6.9)	10 (6.1)	0.683
missing	22 (29.3)	79 (19.0)	18 (20.2)	26 (13.4)	12 (20.7)	20 (12.1)	
PSM	21 (28.0)	79 (19.0)	41 (46.1)	55 (28.4)	21 (36.2)	54 (32.7)	0.531
missing	7 (9.3)	7 (1.7)	0 (0)	4 (2.1)	5 (8.6)	10 (6.1)	
Total	75 (100)	415 (100)	89 (100)	194 (100)	58 (100)	165 (100)	

LN= lymph node; SVI = seminal vesicle invasion; PSM= positive surgical margins

* Chi-square test for categorical variables and the Mann-Whitney U test for continuous variables.

Table 2b. Radiation therapy

Risk group:	Low-risk		Intermediate-risk		High-risk		p-value*
	Control n=107 (% of column total)	Screening n=384 (% of column total)	Control n=169 (% of column total)	Screening n=261 (% of column total)	Control n=278 (% of column total)	Screening n=367 (% of column total)	
Time randomization-diagnosis, median	6.0	4.2	7.3	3.2	6.3	0.2	.000
RT type							.492
EBRT	62 (57.9)	253 (65.9)	137 (81.1)	221 (84.7)	259 (93.2)	335 (91.3)	
Brachytherapy	30 (28.0)	51 (13.3)	19 (11.2)	14 (5.4)	4 (1.4)	2 (0.5)	
Brachytherapy + 45Gy	11 (10.3)	61 (15.9)	4 (2.4)	9 (3.4)	1 (0.4)	2 (0.5)	
Missing	4 (3.7)	19 (4.9)	9 (5.3)	17 (6.5)	14 (5.0)	28 (7.6)	
Combination RT+HT	7 (6.5)	5 (1.3)	23 (13.6)	14 (5.4)	140 (50.4)	44 (12)	.000
Dosage (EBRT only, % of total EBRT)							.000
<69	33 (53.2)	203 (80.2)	51 (37.2)	140 (63.3)	109 (42.1)	252 (75.2)	
>=69	29 (46.8)	50 (19.8)	85 (62.0)	81 (36.7)	150 (57.9)	83 (24.8)	
Missing	0 (0)	0 (0)	1 (0.7)	0 (0)	0 (0)	0 (0)	
LN irradiation	0 (0)	2 (0.5)	0 (0)	1 (0.4)	7 (2.5)	5 (1.4)	.283
Missing	4 (3.7)	15 (3.9)	9 (5.3)	11 (4.2)	14 (5.0)	20 (5.4)	
Total	107 (100)	384 (100)	169 (100)	261 (100)	278 (100)	367 (100)	

RT= radiation therapy; EBRT= external beam radiation therapy; HT= hormonal therapy; LN= lymph node

* Chi-square test for categorical variables and the Mann-Whitney U test for continues variables.

Table 3. Results after treatment, divided by initial treatment modality, risk group, and arm.

Risk group:	Low-risk		Intermediate-risk		High-risk		p-value*	p-value*
	Control n=75 (% of column total)	Screening n=415 (% of column total)	Control n=89 (% of column total)	Screening n=194 (% of column total)	Control n=58 (% of column total)	Screening n=165 (% of column total)		
Follow-up time after treatment	4.8	9.3	6.9	9.5	5.4	10.1	0.000	0.001
Results of treatment								
BCR (2x PSA >0.2ng/ml)	7 (9.3)	57 (13.7)	34 (38.2)	44 (22.7)	20 (34.5)	51 (30.9)	0.007	0.615
Local progression	4 (5.3)	3 (0.7)	6 (6.7)	10 (5.2)	4 (6.9)	9 (5.5)	0.591	0.687
Distant metastasis	0 (0)	4 (1.0)	7 (7.9)	8 (4.1)	9 (15.5)	14 (8.5)	0.192	0.130
Additional treatment								
No additional treatment during Follow-up	64 (85.3)	378 (91.1)	58 (65.2)	160 (82.5)	30 (51.7)	126 (76.4)	0.001	0.000
Cured*	60 (80.0)	341 (82.2)	44 (49.4)	137 (70.6)	25 (43.1)	103 (62.4)	0.001	0.010
PCa death	0 (0)	6 (1.4)	3 (3.4)	5 (2.6)	6 (10.3)	11 (6.7)	0.708	0.364
Total	75 (100)	415 (100)	89 (100)	194 (100)	58 (100)	165 (100)		

BCR= biochemical recurrence; *Defined as no tumor progression (no biochemical recurrence (BCR), no local progression, no distant metastasis, and no PCa death) and no additional treatment during follow-up

* Chi-square test for categorical variables and the Mann-Whitney U test for continues variables.

Table 3b. Radiation therapy without hormonal therapy

Risk group:	Low-risk		Intermediate-risk		High-risk		p-value*	Screening n=323 (% of column total)	p-value*	Screening n=138 (% of column total)	p-value*	Screening n=247 (% of column total)	p-value*	Screening n=323 (% of column total)
	Control n=100 (% of column total)	Screening n=379 (% of column total)	Control n=146 (% of column total)	Screening n=146 (% of column total)	Control n=138 (% of column total)	Screening n=138 (% of column total)								
Follow-up time after treatment	5.6	8.2	5.2	8.2	6.0	8.4	0.000	8.2	0.000	6.0	8.4	0.000	8.2	0.094
Results of treatment														
BCR (PSA >2ng/ml above PSA nadir)	13 (13.0)	39 (10.3)	35 (24.0)	65 (26.3)	72 (52.2)	158 (48.9)	0.438	65 (26.3)	0.606	72 (52.2)	158 (48.9)	0.606	65 (26.3)	0.522
Local progression	1 (1.0)	8 (2.1)	4 (2.7)	10 (4.0)	7 (5.1)	35 (10.8)	0.467	10 (4.0)	0.499	7 (5.1)	35 (10.8)	0.499	10 (4.0)	0.049
Distant metastasis	1 (1.0)	3 (0.8)	9 (6.2)	13 (5.3)	27 (19.6)	53 (16.4)	0.839	13 (5.3)	0.707	27 (19.6)	53 (16.4)	0.707	13 (5.3)	0.412
Additional treatment														
No additional treatment during Follow-up	90 (90)	357 (94.2)	125 (85.6)	211 (85.4)	84 (60.9)	220 (68.1)	0.135	125 (85.6)	0.958	84 (60.9)	220 (68.1)	0.958	125 (85.4)	0.133
Cured*	81 (81)	328 (86.5)	101 (69.2)	171 (69.2)	53 (38.4)	154 (47.7)	0.163	101 (69.2)	0.991	53 (38.4)	154 (47.7)	0.991	101 (69.2)	0.067
PCa death	0 (0)	5 (1.3)	4 (2.7)	8 (3.2)	27 (19.6)	51 (15.8)	0.248	4 (2.7)	0.781	27 (19.6)	51 (15.8)	0.781	4 (2.7)	0.322
Total	100 (100)	379 (100)	146 (100)	247 (100)	138 (100)	323 (100)		146 (100)		138 (100)	323 (100)		146 (100)	

BCR= biochemical recurrence; *Defined as no tumor progression (no biochemical recurrence (BCR), no local progression, no distant metastasis, and no PCa death) and no additional treatment during follow-up

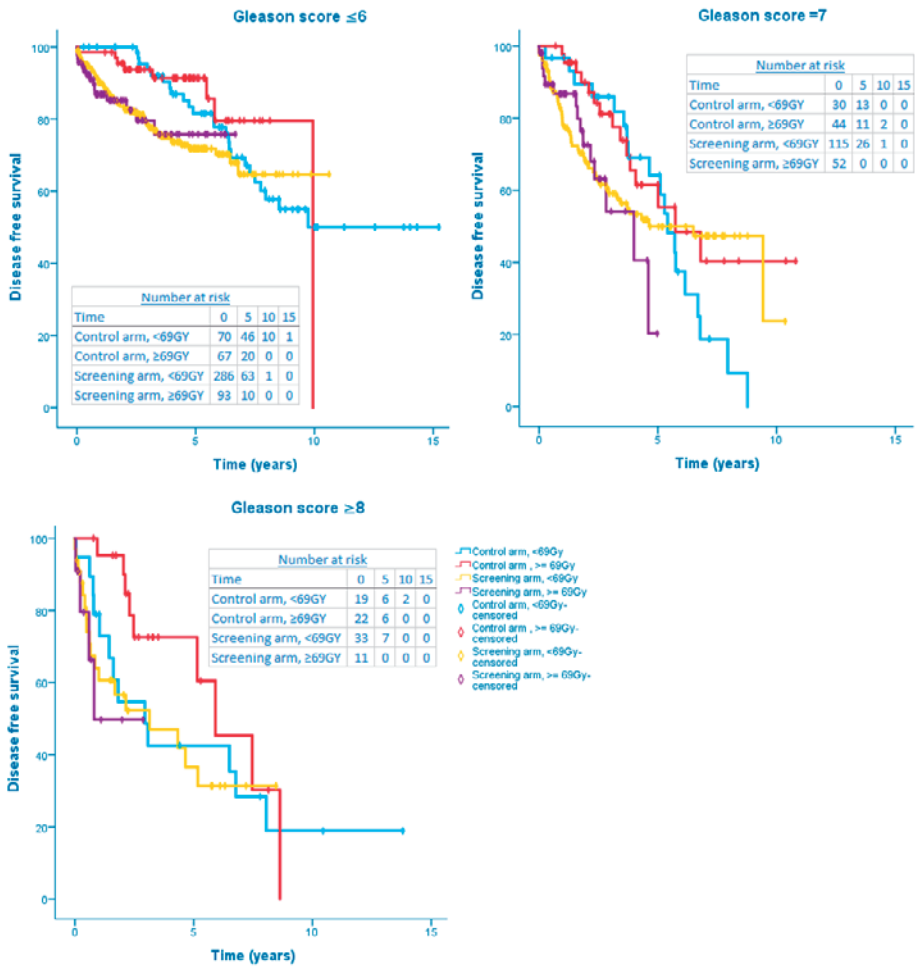
* Chi-square test for categorical variables and the Mann-Whitney U test for continuous variables.

Table 3c. Radiation therapy with hormonal therapy

Risk group:	Low-risk			Intermediate-risk			High-risk		
	Control n=7 (% of column total)	Screening n=5 (% of column total)	p-value*	Control n=23 (% of column total)	Screening n=14 (% of column total)	p-value*	Control n=140 (% of column total)	Screening n=44 (% of column total)	p-value*
Follow-up time after treatment	6.7	7.5	0.755	4.5	4.8	0.229	4.6	4.5	0.841
Results of treatment									
BCR (PSA >2ng/ml above PSA nadir)	0 (0)	0 (0)		5 (21.7)	1 (7.1)	0.243	32 (22.9)	11 (25)	0.770
Local progression	0 (0)	0 (0)		1 (4.3)	0 (0)	0.429	3 (2.1)	0 (0)	0.328
Distant metastasis	0 (0)	0 (0)		0 (0)	0 (0)		19 (13.6)	3 (6.8)	0.228
Additional treatment									
No additional treatment during Follow-up**	7 (100)	5 (100)		21 (91.3)	12 (85.7)	0.595	115 (82.1)	36 (81.8)	0.961
Cured*	7 (100)	5 (100)		15 (65.2)	10 (71.4)	0.695	84 (60.0)	26 (59.1)	0.915
PCa death	0 (0)	0 (0)		0 (0)	0 (0)		16 (11.4)	3 (6.8)	0.381
Total	7 (100)	5 (100)		23 (100)	14 (100)		140 (100)	44 (100)	

BCR= biochemical recurrence; *Defined as no tumor progression (no biochemical recurrence (BCR), no local progression, no distant metastasis, and no PCa death) and no additional treatment during follow-up (hormonal therapy within 3 year of initial treatment excluded), **hormonal therapy within 3 year of initial treatment excluded

* Chi-square test for categorical variables and the Mann-Whitney U test for continuous variables.



Supplemental figure 3. Disease free survival after radiotherapy (no hormonal treatment, no brachytherapy) for men with T1-T2 disease, stratified by Gleason score (≤ 6 , 7 , ≥ 8), dosage (< 69 Gy, ≥ 69 Gy), and arm, after correction for lead-time in the screening arm. Time in the screening arm was corrected for lead time using a method previously described[8].

Chapter 4

Do treatment differences between arms affect the main outcome of the European Randomized study of Screening for Prostate Cancer (ERSPC) Rotterdam?

Leonard P. Bokhorst^a, Lionne D.F. Venderbos^a, Fritz H. Schröder^a, Chris H. Bangma^a,
Ewout W. Steyerberg^b, Monique J. Roobol^a

^a Department of urology, Erasmus University Medical Centre, Rotterdam, the Netherlands

^b Department of public health, Erasmus University Medical Centre, Rotterdam, the
Netherlands

ABSTRACT

Purpose: To assess differences in treatment between the screening and control arm of European Randomized study of Screening for Prostate Cancer (ERSPC) Rotterdam and to study if possible treatment differences explained the positive study outcome.

Materials and Methods: In ERSPC Rotterdam, men aged 55-74 were randomized between a screening (n=21210) and a control arm (n=21166). Treatment after diagnosis was left to the care provider of the patients choice. Initial treatment was compared within four risk groups.

The relation between prostate cancer incidence and prostate cancer mortality was assessed per risk group by correlating the relative risk (RR) of prostate cancer incidence and the RR of prostate cancer mortality. A direct relation would support a stage shift as the main cause of changes in prostate cancer mortality.

Results: Initial treatment differed between the arms in the low-, intermediate-, and high-risk groups, but not in the metastatic group. RR of prostate cancer incidence and prostate cancer mortality per risk group were 1:1 related (slope of regression line 1.00, 95% confidence interval (CI); 0.30-1.74) and 94% of the changes in prostate cancer mortality could be explained by changes in prostate cancer incidence. This makes differences in treatment unlikely as the reason for the observed prostate cancer mortality reduction.

Conclusion: Differences in treatment between the screening and control arm of ERSPC Rotterdam were unlikely to explain the differences in prostate cancer mortality. Instead results are consistent with a reduction in prostate cancer mortality as a result of a favourable stage through screening.

Trial registration: ISRCTN49127736.

INTRODUCTION

A 32% prostate cancer specific mortality reduction was seen in the screening arm of the European Randomized study of Screening for Prostate Cancer (ERSPC) section Rotterdam in the age group 55-69 years, and a 20% reduction for ages 55-75 years at a follow-up time of 13 years. [1] In a secondary analysis, correction for nonattendance in the screening arm and contamination (opportunistic PSA testing resulting into biopsy) in the control arm, the benefit for a man choosing to be screened versus a man choosing not to be screened was estimated at 51%. [2] A large reduction in metastatic disease at diagnosis was also shown using a similar correction. [3] These reductions in mortality and metastatic disease resulted from a screening strategy with PSA testing every 4 years until the age of 75 and a subsequent biopsy if PSA was abnormal.

After prostate cancer was detected, treatment was left to the care provider of the patients choice in both the screening and control arm. Because treatment was not standardized in the screening protocol, differences in treatment might have arisen between the screening and control arm. Possible differences in treatment could potentially have affected the main outcome of the trial (i.e. prostate cancer specific mortality).

The aim of the current analysis is to assess differences in treatment between the screening and control arm of the ERSPC Rotterdam and to study whether possible treatment differences explained the positive study outcome.

MATERIALS AND METHODS

Between 1993 and 1999, 42376 men aged 55-74 year were randomized between a screening and control arm after providing written informed consent. In the screening arm, men were invited for PSA testing every 4 years until the age of 75. In the first half of the first screening round, men were offered a prostate biopsy if PSA was ≥ 4 ng/ml or if DRE was abnormal. Apart from this first half of the first screening round a PSA value ≥ 3 ng/ml was the only indication for prostate biopsy in all subsequent screening rounds. [4] After detection of prostate cancer, both men in the screening and the control arm received further diagnostics and treatment from their local care provider of choice. Data on all men with prostate cancer (in both screening and control arm) were collected through linkage with the national cancer registry and chart review every 6 months. Cause of death for all men with prostate cancer in the screening and control arm was determined by the cause of death committee (CODC) using medical records based on a pre-defined algorithm and blinded for study arm. [5] The protocol of the ERSPC Rotterdam has been described in detail previously. [4, 6] Analyses of the main outcome (prostate cancer specific mortality) were done on the entire cohort and on the pre-defined core age group (55-69 year). [1] Follow-up for the current analysis ended December 31, 2010.

Initial treatment was coded as radical prostatectomy (RP), radiotherapy (RT), radiotherapy combined with hormonal therapy (RT+HT), hormonal therapy (HT), and a combined group for watchful waiting and active surveillance (WW/AS) as both treatment modalities were not well distinguishable in our database.

Because treatment choice is highly dependent on tumour characteristics, comparison of treatment between the screening and control arm was done within 4 risk groups. Risk groups were defined as: low-risk (clinical stage \leq T2a, and Gleason score (GS) \leq 6, and PSA \leq 10 ng/ml), intermediate-risk (clinical stage T2b, and/or GS =7, and/or PSA $>$ 10 and \leq 20 ng/ml) and high-risk prostate cancer (clinical stage \geq T2c, or GS \geq 8, or PSA $>$ 20 ng/ml), based on the criteria of D'Amico et al.[7] (with the addition of clinical stage T3 and T4 to the high-risk group), and a separate group for men with metastatic prostate cancer (M1 or PSA \geq 100 ng/ml).

STATISTICAL ANALYSIS

Descriptive statistics and the chi-square tests were used to compare initial treatment between the screening and control arm within the risk groups. To assess if possible differences in treatment within risk groups could have an impact on total prostate cancer mortality, all mortality rates were scaled on the total prostate cancer mortality rate in the control arm (i.e. total prostate cancer mortality in the control arm is set at 100). We then adopted a method that was previously applied in breast cancer screening studies to assess the relationship between prostate cancer incidence and prostate cancer mortality within the risk groups.[8] A direct relation would support a stage shift as the main cause of changes in prostate cancer mortality, making a large effect of differences in treatment unlikely.

We regressed the natural logarithm (ln) of the relative risk (RR) of prostate cancer incidence per risk group on the ln(RR) of prostate cancer mortality per risk group. The inverse of the variance of the ln(RR) of prostate cancer mortality was used to weight each risk group. The regression line was forced through zero. By doing so it is assumed that if the RR of prostate cancer incidence is equal to the RR of prostate cancer mortality, treatment did not affect prostate cancer mortality, but changes in prostate cancer incidence (caused by screening) did. This concept is further explained in figure 1.

Because of overdiagnosis in the screening arm, the effect of prostate cancer incidence on prostate cancer mortality using this method could be misinterpreted. As overdiagnosis is most common in the low-risk tumour group, this group was not used in the initial analysis, but added in a sensitivity analysis. In addition, the effect of excluding different age groups at randomization and division of the high-risk group based on extra-prostatic extension (cT3-cT4) was evaluated, as choice of treatment might be particularly dependent on these characteristics.

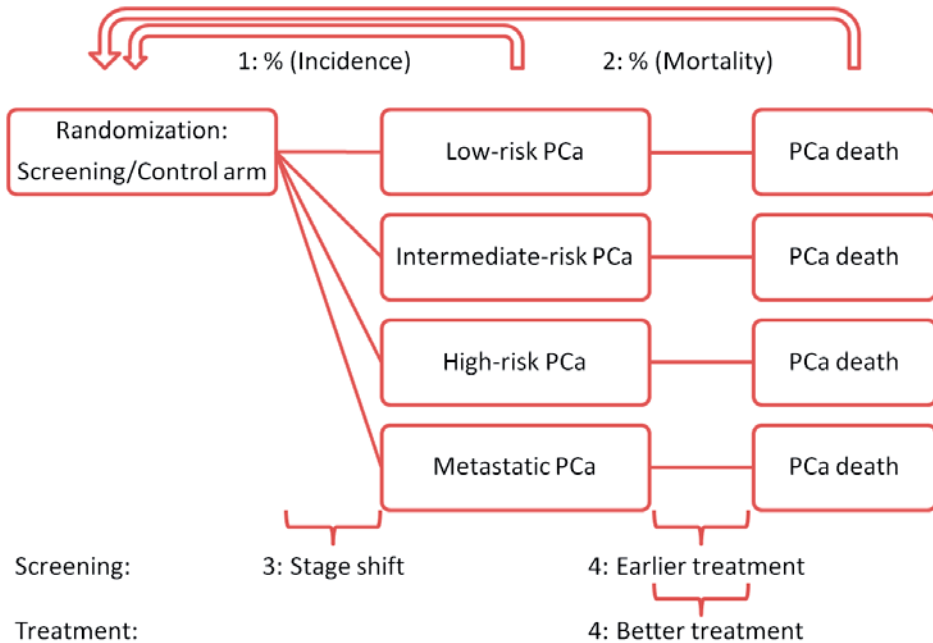


Figure 1. Relationship of prostate cancer incidence and prostate cancer mortality to determine the effect of treatment on outcome.

At randomization men in both arms have the same risk of prostate cancer death. During follow-up men can be diagnosed in one of four pre-defined risk groups (1: (incidence)). After diagnosis in one of the risk groups men could die of prostate cancer (2: (mortality)). Differences in incidence and mortality per risk group between the arm can be expressed as a relative risk (RR). During follow-up screening can affect the prostate cancer mortality either because of a stage shift affecting tumour characteristics at diagnosis (which affects incidence)(3), or by earlier detection (without effecting the risk group at diagnosis) and therefore earlier treatment (4). Differences in treatment between arms can only affect the risk of death after diagnosis (which does not affect incidence)(4). If changes in mortality within a risk group are directly related to changes in incidence in that risk group (e.g. both are reduced by 50%), it can be assumed that this is a direct result of a stage shift caused by screening, therefore excluding differences in treatment between the arms as likely cause of the mortality reduction.

RESULTS

During a median follow-up period of 12.8 years 1444 prostate cancers were diagnosed in the control arm and 2699 in the screening arm, resulting in a cumulative incidence during this period of 6.8% and 12.7% respectively. Prognostic features of prostate cancers found in the screening arm were more favourable (table 1).

In the low-risk group RP was more often given in the screening arm than in the control arm (29.9% versus 21.4%, Fig 2). In the intermediate-risk group RP was again more often used in the screening arm (35.1% versus 24.1%) instead of WW/AS (13.9% versus 24.9%). In the high risk group, RT was more often combined with HT in the control arm (27.3% versus 6.8% for the screening arm)

as compared to the screening arm in which RT alone was more common (49.9% versus 27.0% for the control arm). In the metastatic risk group treatment differences between the arms were small with most men receiving HT (90.7% and 87.7%).

Table 1. Clinical characteristics per arm.

	Control, n=1444 (% of column total)	Screening, n=2699 (% of column total)
Age at diagnosis, median [25-75]	72.1 [67.5-75.8]	68.7 [64.8-72.5]
PSA ng/ml at diagnosis, median [25-75]	11.8 [7.2-27.5]	5.2 [3.5-9.1]
cT-stage at diagnosis		
1	700 (48.5)	1504 (55.7)
2	379 (26.2)	819 (30.3)
3	272 (18.8)	319 (11.8)
4	65 (4.5)	28 (1)
Missing	28 (1.9)	29 (1.1)
Gleason score at diagnosis*		
<=6	701 (48.5)	1863 (69)
7	426 (29.5)	592 (21.9)
>=8	292 (20.2)	223 (8.3)
missing	25 (1.7)	21 (0.8)
M-stage at diagnosis		
0/X	1312 (90.9)	2643 (97.9)
1	132 (9.1)	56 (2.1)
Risk group at diagnosis		
Low	350 (24.2)	1386 (51.4)
Intermediate	371 (25.7)	553 (20.5)
High	512 (35.5)	647 (24)
Metastatic	182 (12.6)	73 (2.7)
Missing	29 (2)	40 (1.5)
Charlson comorbidity at diagnosis		
0	561 (38.9)	1146 (42.5)
1-2	650 (45)	1051 (38.9)
3-4	217 (15)	489 (18.1)
>=5	16 (1.1)	13 (0.5)
Deaths (from all causes)		
	490 (33.9)	781 (28.9)
Prostate cancer deaths	188 (13)	151 (5.6)
total	1444 (100)	2699 (100)

* In 163 men (141 screening arm and 22 control arm) with missing Gleason scores (3.9% of total men with prostate cancer) tumour grades 1, 2, and 3 were re-coded into Gleason score groups <=6, 7, and >= 8 respectively to allow more complete analysis. 46 men had missing Gleason scores and tumour grades.

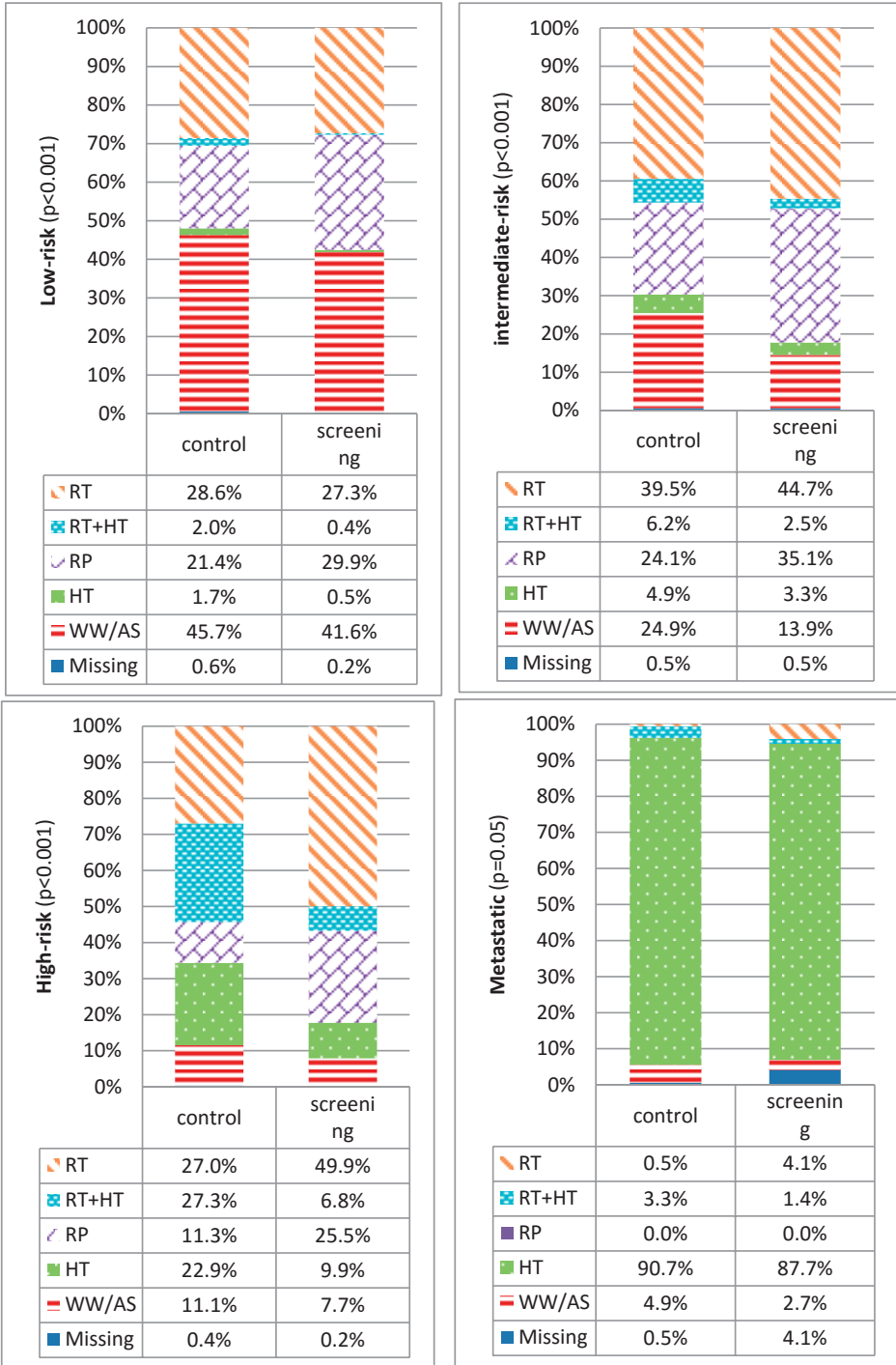
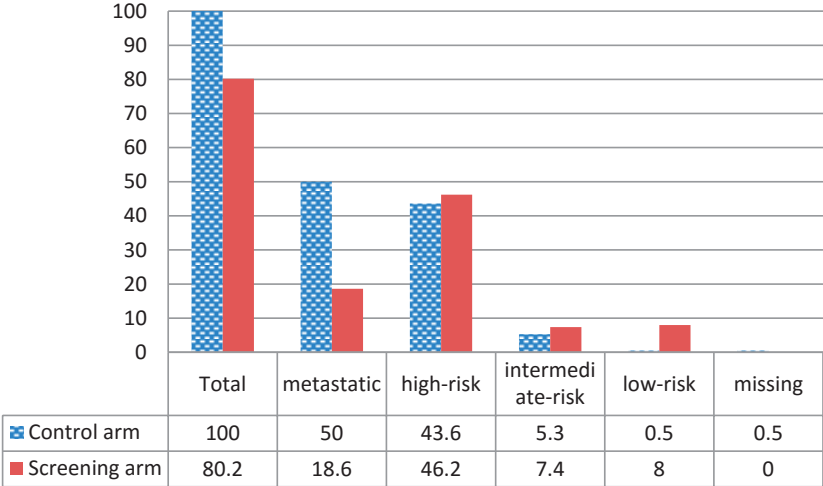


Figure 2. Initial treatment per arm, divided per risk group.

The prostate cancer mortality in the screening arm relative to the control arm was substantially reduced in the metastatic risk group (Fig 3a). In the low, intermediate and high-risk groups prostate cancer mortality was higher the screening arm relative to the control arm (Fig 3a). Prostate cancer incidence was increased in all risk groups in the screening arm except for the metastatic risk group (Fig 3b).

3a. PCa mortality per risk-group



3b. PCa incidence per risk-group

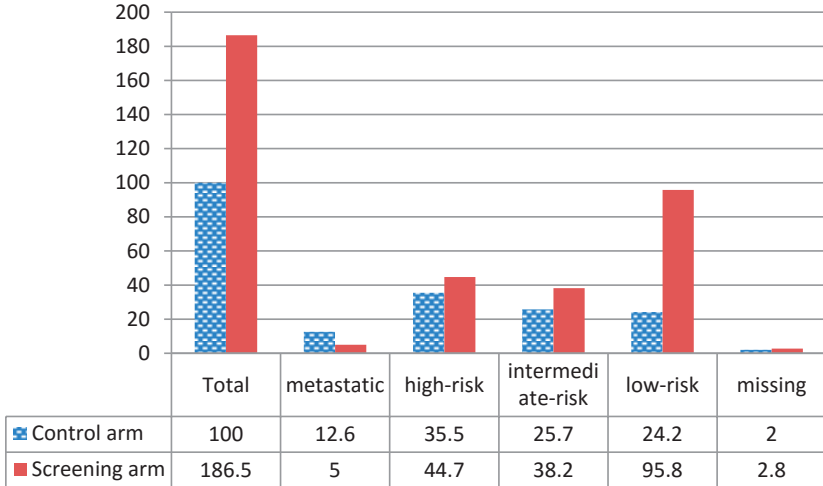


Figure 3. Prostate cancer mortality and incidence relative to the overall risk in the control arm (scaled as 100) by risk group. For comparison data were indexed on the total prostate cancer mortality and incidence in the control arm: 3A: Distribution of prostate cancer mortality in the screening and control arm per risk group. 3B: Prostate cancer incidence in the screening and control arm per risk group.

Table 2. Prostate cancer (PCa) incidence and mortality in the screening and control arm, divided per risk group. (RR=relative risk)

	Absolute PCa incidence		PCa Incidence	PCa mortality	Absolute PCa mortality		p-value
	Control (% of total men control arm)	Screening (% of total men screening arm)	RR (screening arm versus control arm)	RR (screening arm versus control arm)	Control (% of total men control arm)	Screening (% of total men screening arm)	(RR PCa incidence versus RR PCa mortality)
Missing	29 (0.1)	40 (0.2)	1.38	-	1 (0.005)	0 (0)	-
Low	350 (1.7)	1386 (6.5)	3.95	14.97	1 (0.005)	15 (0.07)	0.198
Intermediate	371 (1.8)	553 (2.6)	1.49	1.40	10 (0.05)	14 (0.07)	0.881
High	512 (2.4)	647 (3.1)	1.26	1.06	82 (0.39)	87 (0.41)	0.287
Metastatic	182 (0.9)	73 (0.3)	0.40	0.37	94 (0.44)	35 (0.17)	0.758
Total	1444 (6.8)	2699 (12.7)	1.87	0.80	188 (0.89)	151 (0.71)	<0.001

Within the risk groups, none of the RRs of prostate cancer incidence differed significantly from the RRs of prostate cancer mortality (Table 2). The slope of the regression line of $\ln(\text{RR})$ prostate cancer incidence on $\ln(\text{RR})$ prostate cancer mortality of the intermediate, high and metastatic risk group was 1.00 (95% confidence interval (CI); 0.30-1.74), with a R^2 of 0.94 (figure 4). This means that prostate cancer incidence and prostate cancer mortality were 1:1 related in these groups and 94% of the changes in prostate cancer mortality could be explained by differences in prostate cancer incidence. Addition of the low-risk group did not change the results (slope 1.06, 95% CI 0.44-1.68, R^2 0.91). Results for the core age group (55-69 years) were similar (slope 1.12, 95% CI -0.20-2.43, R^2 0.87), as were result if the high-risk group was split based on extra-capsular extension (slope 0.99, 95% CI 0.49-1.49, R^2 0.93).

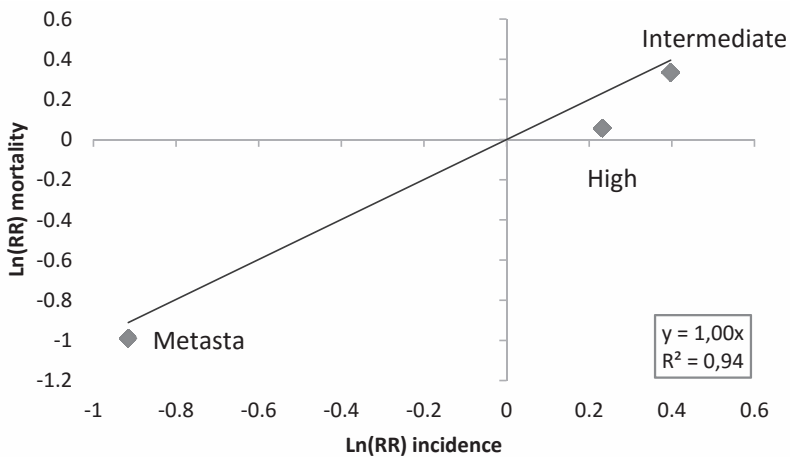


Figure 4. Relationship between the prostate cancer incidence and prostate cancer mortality in the intermediate, high, and metastatic risk groups.

DISCUSSION

In the current analysis we found differences in initial treatment between the screening and control arm of ERSPC Rotterdam. In addition, a favourable stage shift resulting in, among others, a large reduction of metastatic disease at diagnosis in the screening arm was seen. Comparison of the changes in prostate cancer incidence and mortality by risk group showed that changes in prostate cancer mortality were consistent with the changes in prostate cancer incidence. This observation supports stage shift through screening as the main reason for reduced prostate cancer mortality in the screening arm of the ERSPC Rotterdam, and thus makes treatment differences unlikely to have played a large contributing role. This observation confirms the earlier report of Wolters et al. where trial arm had only a minor role in treatment choice compared to other variables.[9]

Initial treatment differed between the screening and control arm within the low, intermediate, and high risk groups. Because 93.6% of all prostate cancer deaths in the control arm and 80.8% of all prostate cancer deaths in the screening arm occurred in the metastatic or high-risk group, differences in treatment in these groups can be seen as most relevant. Starting with the high-risk group two main differences were seen. In the control arm 53% of men receiving RT got HT in addition to their RT, compared to only 12% in the screening arm. This difference in treatment will benefit the control arm as RT in combination with HT was shown to improve survival.[10-12] In addition, a previous report showed that men with RT in the control arm received higher dosages which again improved survival.[13] Both are most likely explained by the fact that overall diagnosis of prostate cancer occurred later in time in the control arm at the moment that adjuvant endocrine therapy for high risk disease and higher radiation dosages were included in international guidelines for treatment.[13] The second difference in the high-risk group was that approximately 15% more men received RP in the screening arm, while in the control arm these men received HT. This difference remained after division of the high-risk group based on extra capsular extension (data not shown). Differences in treatment could also be based on patient characteristics such as comorbidity status and personal preference. The latter is also applicable to the physician who is in general heavily involved in treatment decisions. On multivariate logistic regression analysis men with higher PSA, age, and Charlson comorbidity score at diagnosis were more likely to receive HT. After correction for these variables, the study arm to which the men were randomized (screening arm versus control arm) was not a significant predictor of HT. This means that the difference in treatment between the arms in the high risk group was caused by differences in clinical characteristics, rather than by a difference in treatment preference in one of the two study arms (data not shown). In the metastatic risk group there were no large differences in treatment between the two arms, almost all men received HT as initial treatment.

Interestingly, we noted a relative increase in mortality in the screening arm as compared to the control arm in all risk groups except for the metastatic risk group. This indicates that the prostate cancer mortality reduction seen in the entire screening arm can largely be attributed to a reduc-

tion in prostate cancer mortality from men with metastatic disease at diagnosis. This last group was the only group where treatment did not differ between the screening and control arm. The only remaining and most plausible explanation for the reduction in prostate cancer mortality is a simple decrease in prostate cancers being diagnosed with metastasis as a result of a stage shift to more favourable tumour characteristics at diagnosis. This seems to be confirmed by the almost identical relative reduction in the screening arm of prostate cancer incidence (60%) and prostate cancer mortality (63%) in the metastatic risk group. Including all risk and age groups, changes in prostate cancer incidence (stage shift) could explain 90% of the changes in prostate cancer mortality. This support stage shift as main cause of the observed prostate cancer mortality reductions and makes a large effects of the observed differences in treatment between arms unlikely.

A stage shift is most likely also the reasons why the mortality in the screening arm was higher in especially the intermediate and low-risk groups as compared to the control arm. Screening reduced the amount of men diagnosed with metastatic prostate cancer at diagnosis by 60%. It is likely that a (substantial) part of these men will still develop metastatic disease and die. It is often assumed that men not diagnosed with metastatic disease because of screening will "stage shift" to the high-risk group. It could, however, well be that these men were not diagnosed in the high-risk group but in the intermediate or even the low-risk group. If so, the risk of prostate cancer death in these groups will increase due to the enrichment by higher risk men compared to the control arm. So, although screening led to an earlier diagnosis, cancer was still detected too late in these men. Besides screening algorithm related causes (e.g. inadequate screening tests, too long screening intervals, nonattending),[14] age at which screening was started might affect its result.[13] Simply intensifying the screening protocol or starting screening at an earlier age for all men will however also increase overdiagnosis. Individual risk adapted screening strategies, for instance based on nomogram predictions, could help to overcome this problem.[15]

Other explanations for the relative increase in mortality in the screening arm in the low and intermediate risk groups could be due to changes in the grading and staging of prostate cancer. The Gleason scoring system was changed during the course of the trial resulting in a shift to higher Gleason scores.[16] Moreover, the number of biopsy cores taken as standard, gradually increased in common practice (control arm) while in the screening arm the number of biopsy cores taken remained 6. As men in the control arm were diagnosed later in time, these changes in staging could have resulted in a shift towards higher risk groups in the control arm, which, as compared to the screening arm, would have a lower risk of mortality (Will Rogers phenomenon).

Last, more men were diagnosed with prostate cancer in the screening arm. This means that potentially more men could have died of prostate cancer. Attribution of cause of death was done by an independent committee according to a predefined algorithm blinded for study arm.[5] The committee reviewed, however, only men diagnosed with prostate cancer. This could have resulted in an underreporting of prostate cancer deaths in the control arm as less men were diagnosed with prostate cancer. Analysis using an excess mortality approach, which circumvents this problem, however showed similar results concluding this was not a large issue.[17, 18]

Some limitations should be mentioned. The regression method used to assess the relation between prostate cancer incidence and prostate cancer mortality does not take into account overdiagnosis. Overdiagnosis causes men to be diagnosed, who would not have been diagnosed if not screened. Overdiagnosis causes higher incidence rates in the screening arm, but overdiagnosed men will by definition never develop metastatic disease or die of prostate cancer. It can therefore be expected that the increase in incidence in the low, intermediate and even high-risk group is higher than the increase in mortality. Last, because the regression analysis was based on only few risk groups, its stability may be low.

CONCLUSION

A favourable stage shift, with less metastatic disease at diagnosis, was seen in the ERSPC Rotterdam. The changes in mortality observed within the ERSPC Rotterdam were consistent with the changes in prostate cancer incidence per risk group initiated through screening. This observation supports a stage shift with subsequent earlier treatment as the main reason for lower prostate cancer mortality in the screening arm of the ERSPC Rotterdam, excluding a large effect of the observed differences in treatment between arms on the primary outcome.

REFERENCES

1. Roobol MJ, Kranse R, Bangma CH, van Leenders AG, Blijenberg BG, van Schaik RH, et al. Screening for Prostate Cancer: Results of the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol.* 2013;64:530-9.
2. Bokhorst LP, Bangma CH, van Leenders GJ, Lous JJ, Moss SM, Schroder FH, et al. Prostate-specific Antigen-Based Prostate Cancer Screening: Reduction of Prostate Cancer Mortality After Correction for Nonattendance and Contamination in the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol.* 2014;65:329-36.
3. Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schroder FH, et al. Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomized screening trial (ERSPC section Rotterdam). *Int J Cancer.* 2010;127:2639-44.
4. Roobol MJ, Kirkels WJ, Schroder FH. Features and preliminary results of the Dutch centre of the ERSPC (Rotterdam, the Netherlands). *BJU international.* 2003;92 Suppl 2:48-54.
5. De Koning HJ, Blom J, Merkelbach JW, Raaijmakers R, Verhaegen H, Van Vliet P, et al. Determining the cause of death in randomized screening trial(s) for prostate cancer. *BJU international.* 2003;92 Suppl 2:71-8.
6. Roobol MJ, Schröder FH. European Randomized study of Screening for Prostate Cancer (ERSPC): rationale, structure and preliminary results. *BJU international.* 2003;92: Suppl 2:1-122.
7. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA.* 1998;280:969-74.
8. Autier P, Hery C, Haukka J, Boniol M, Byrnes G. Advanced breast cancer and breast cancer mortality in randomized controlled trials on mammography screening. *J Clin Oncol.* 2009;27:5919-23.
9. Wolters T, Roobol MJ, Steyerberg EW, van den Bergh RC, Bangma CH, Hugosson J, et al. The effect of study arm on prostate cancer treatment in the large screening trial ERSPC. *Int J Cancer.* 2010;126: 2387-93.
10. Bolla M, Collette L, Blank L, Warde P, Dubois JB, Mirimanoff RO, et al. 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.
11. Widmark A, Klepp O, Solberg A, Damber JE, Angelsen A, Fransson P, et al. 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.
12. Bolla M, Van Tienhoven G, Warde P, Dubois JB, Mirimanoff RO, Storme G, et al. External irradiation with or without long-term androgen suppression for prostate cancer with high metastatic risk: 10-year results of an EORTC randomised study. *Lancet Oncol.* 2010;11:1066-73.
13. Bokhorst LP, Kranse R, Venderbos LD, Salman JW, van Leenders GJ, Schroder FH, et al. Differences in Treatment and Outcome After Treatment with Curative Intent in the Screening and Control Arms of the ERSPC Rotterdam. *Eur Urol.* 2014.
14. Zhu X, van Leeuwen PJ, Bul M, Bangma CH, Roobol MJ, Schroder FH. Identifying and characterizing "escapes"-men who develop metastases or die from prostate cancer despite screening (ERSPC, section Rotterdam). *Int J Cancer.* 2011;129:2847-54.
15. Roobol MJ, Zhu X, Schroder FH, van Leenders GJ, van Schaik RH, Bangma CH, et al. A Calculator for Prostate Cancer Risk 4 Years After an Initially Negative Screen: Findings from ERSPC Rotterdam. *Eur Urol.* 2013;63:627-33.

-
16. Epstein JI, Allsbrook WC, Jr., Amin MB, Egevad LL, Committee IG. 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. Zappa M, Puliti D, Hugosson J, Schroder FH, van Leeuwen PJ, Kranse R, et al. A Different Method of Evaluation of the ERSPC Trial Confirms That Prostate-specific Antigen Testing Has a Significant Impact on Prostate Cancer Mortality. *Eur Urol.* 2014.
 18. van Leeuwen PJ, Kranse R, Hakulinen T, Hugosson J, Tammela TL, Ciattoy S, et al. Impacts of a population-based prostate cancer screening programme on excess total mortality rates in men with prostate cancer: a randomized controlled trial. *J Med Screen.* 2013;20:33-8.

Part 2

Active surveillance

Chapter 5

Decision support for low-risk prostate cancer

Leonard P. Bokhorst^a, Ewout W. Steyerberg^b, Monique J. Roobol^a

^a Department of urology, Erasmus University Medical Center, Rotterdam, the Netherlands

^b Department of public health, Erasmus University Medical Center, Rotterdam,
the Netherlands

ABSTRACT

Since the introduction of PSA for the early detection of prostate cancer, overdiagnosis and subsequent overtreatment have become more and more apparent. Conservative treatment options, specifically active surveillance, are therefore becoming increasingly used in an attempt to reduce the morbidity of radical treatment. Several clinical prediction models were developed to assess an individual's risk of having indolent prostate cancer. Those with high probabilities of indolent cancer are best suited for conservative management. Prediction models readily provide detailed risk estimates when presented as nomograms, which is preferable over simpler presentations as rules. Incorporation of prediction models into decision aids is a good way of providing the best of care for men facing a treatment decision of low-risk prostate cancer with improved patient knowledge and more conscious treatment choices. In this chapter different currently available nomograms are discussed, with considerations to select the most appropriate nomogram for an individual patient.

INTRODUCTION

Over the past decades, increasing numbers of men are faced with the diagnosis of prostate cancer.[1, 2] Conversely, prostate cancer mortality rates have decreased.[3] In addition, men are more often diagnosed at an earlier stage of the disease. The obvious reason for these profound changes is the introduction of the prostate-specific antigen (PSA) test for the early detection of prostate cancer. Early detection of prostate cancer at an more favorable stage, allows the cancer to be treated with curative, rather than palliative, intend. Screening for prostate cancer using PSA was indeed shown to reduce the prostate cancer specific mortality rate by 21% after 13 years of follow-up in the largest randomized prostate cancer screening trial.[4] A major disadvantage is the detection of many additional prostate cancers, with an increase in prostate cancer incidence by over 50%.[4] Many of these additionally identified cancers are likely not to result in any symptoms of death during a man's lifetime. In fact, it is estimated that 50% of men diagnosed through screening have these so-called "overdiagnosed" prostate cancer.[5] The presence of many non-aggressive tumors has already been observed in autopsy studies which showed that prostate cancer was present in a large percentage (10-30% depending on age) of men who died of other causes.[6] Because of the introduction of the PSA test resulting in subsequent prostate biopsies, this large reservoir of prostate cancers is now detected. Hence, radical treatment of PSA detected tumors will for most men not result in any survival benefit.[7] It may however lead to side-effects such as impotency or incontinence, decreasing quality of life in a substantial number of men.[8] A way of avoiding the side-effects of radical treatment is to offer these patients active surveillance. This treatment option aims to avoid or delay treatment for most, while by monitoring tumor progression be able to offer radical treatment for those who benefit. Men diagnosed today with early detected prostate cancer are thus facing the difficult choice of radical treatment with its likely low benefit, but substantial risks of side-effects, versus conservative treatment, which might risk losing the benefit obtained by early detection. In this chapter we will discuss some of the tools that are available to help men to choose between active treatment of their prostate cancer or to opt for the more conservative approach of active surveillance.

DEFINING INDOLENT PROSTATE CANCER

Indolent prostate cancer is defined as a tumor that will not result in symptoms or death during a man's lifetime if left untreated. The dynamic aspect of this definition makes it difficult to operationalize. A tumor that might be moderately progressive over time could be no threat for a man with a short life expectancy, but might become problematic (e.g. metastasize or cause symptoms) for a man with a longer life. Furthermore, no single parameter yet will provide definitive information on future tumor development. Last, as tumor progression in a low risk prostate cancer group is relatively rare and takes usually at least a decade to develop, statistical evaluation

of this end-point is difficult. Therefore several other definitions were proposed to indicate tumors that are latent and have low probabilities of developing symptoms, based on parameters that do not require follow-up in time. The most common definition of indolent disease is an organ confined tumor, $\leq 0.5 \text{ cm}^3$, with no Gleason grade 4 or 5. This definition requires the removal of the prostate for pathological evaluation. In a conservative treatment strategy low-risk tumors are therefore defined as tumors with a high probability of being indolent on radical prostatectomy. Because the tumor is not removed and progression could thus occur, re-evaluation of its low-risk character through time will be necessary. A summary of the different definitions used in the context of indolent/low-risk prostate cancer is given in table 1.

Several clinical definitions were proposed to select men with high likelihood of indolent prostate cancer on radical prostatectomy. These men are best suitable for a conservative treatment approach. Most definitions use a combination of low tumor grade (Gleason ≤ 6), localized disease (cT1c-cT2c), low PSA or PSA density (PSA ≤ 10 -15, PSA-density ≤ 0.15 -0.2), and small tumor volume (≤ 2 cores positive on prostate biopsy, $\leq 50\%$ tumor involvement per core). Different definitions used for the selection of men on active surveillance studies are shown in table 2. Although commonly used in clinical practice, these rule-based definitions have some obvious disadvantages. The most important limitation of these criteria is that much of the predictive value of individual prognostic factors is lost. For example a men with a PSA of 2ng/ml and a men with a PSA of 10ng/ml would both meet the same definition if the criterion is: PSA ≤ 10 ng/ml. Using the definition would assume these men are similar in risk of having indolent disease, while clearly the men with the lower PSA will have a higher probability of indolent disease. Moreover,

Table 1. Definitions used in the context of indolent/low-risk prostate cancer.

	Indolent disease over time	Indolent disease at radical prostatectomy	Low-risk disease
Definition	A tumor that will not result in symptoms or death during a man's lifetime if left untreated	Organ confined, no Gleason grade 4 or 5, and a tumor volume $\leq 0.5 \text{ cm}^3$	A tumor with a high probability of being indolent at radical prostatectomy and/or over time, based on its clinical characteristics.
Advantage	Most optimal definition of indolent disease	Does not require follow-up	Does not require surgical excision
Disadvantage	<ul style="list-style-type: none"> • Can only be determined in retrospect • Requires long follow-up • Dependent on a man's life expectancy 	<ul style="list-style-type: none"> • Requires pathological examination • Could potentially, even with radical treatment, still progress and give rise to symptoms • Definition might be too restrictive for men with a short life expectancy • A tumor volume $\leq 0.5 \text{ cm}^3$ might be too restrictive[43] 	<ul style="list-style-type: none"> • Uncertainty of prediction, caused by underestimation of Gleason score, tumor volume or T-stage • Tumor is not removed and could progress over time, reevaluation is therefore necessary

Table 2. Inclusion criteria for different active surveillance studies.

Active surveillance study	Criteria for inclusion
Royal Marsden[44]	Gleason $\leq 3+4$ (primary Gleason grade ≤ 3); PSA ≤ 15 ng/ml; cT1c-2a; $\leq 50\%$ of cores positive
University of Miami[45]	Gleason ≤ 6 ; PSA ≤ 15 ng/ml; cT1c-2c; \leq two cores positive; $\leq 20\%$ of any core positive
Johns Hopkins[46]	Gleason $\leq 3+3$; PSA density $\leq 0,15$ ng/ml/ml; cT1c; \leq two cores positive; $\leq 50\%$ of any core positive
University of California San Francisco[47]	Gleason $\leq 3+3$; PSA ≤ 10 ng/ml; cT1c-2c; $\leq 33\%$ of cores positive; $\leq 20\%$ of any core positive
University of Toronto[48]	Gleason ≤ 6 ; PSA ≤ 10 ng/ml (until January 2000, for men age >70 years: Gleason $\leq 3+4$; PSA ≤ 15 ng/ml)
Prostate cancer Research International Active Surveillance (PRIAS)[49]	Gleason $\leq 3+3$; PSA ≤ 10 ng/ml; PSAD $\leq 0,2$ ng/ml/ml; cT1c-2c; \leq two cores positive (age >70 years: Gleason $\leq 3+4$, maximum 10% tumor per cores)

using a rule-based definition of indolent disease exerts equal value to all individual risk factors. Both a man with a Gleason score of 7 and a man with a PSA of 10.5ng/ml would not fit the rule based definition. The first might have a truly higher chance of having aggressive disease and therefore not be suitable for conservative management, while the second may only marginally differ in risk of being indolent and could therefore still consider a non-radical approach. Combining risk factors into an individual risk estimation (risk-based), instead of “eligible” or “non-eligible” (rule-based), may better inform the patient and his physician and help to make a more conscious decision on treatment choice.

INDIVIDUALIZED PREDICTIONS

Individual risk estimation using a risk-calculator, or nomogram, is frequently applied to other areas of prostate cancer care. For instance in the decision to perform a prostate biopsy. Multiple risk-calculators were developed by several study groups to predict the chance of having a positive prostate biopsy, as it was realized that performing a prostate biopsy only based on a single PSA cut-off was suboptimal.[9] These risk-calculators use different risk factors such as the age, PSA value, digital rectal examination, transrectal ultrasound findings, prostate volume, and prior biopsy status to calculate an individual risk of having a positive prostate biopsy.[10] It was estimated that using these risk-calculators to guide biopsy decision could potentially reduce the number of unnecessary biopsy with 30% without missing important prostate cancers (defined as Gleason grade >3 , PSA >20 ng/ml, T-stage 3 or 4, $>50\%$ positive cores, >20 mm cancer in all cores, or <40 mm benign tissue in all cores) as compared to a single PSA cut-off approach.[11] Risk-calculators could even be used to calculate the risk of having a positive biopsy up to eight years in the future.[12] After been validated in other cohorts,[13-17] the effect of using risk-

calculators in clinical practice has been assessed.[18] In total 83% of patients complied with the recommendation provided by the risk-calculator. If a biopsy was recommended 96% complied with the recommendation versus 64% of men with a recommendation against prostate biopsy. Of men not complying with a negative biopsy recommendation only 3% were found to have a relevant tumor (Gleason >6).[19] The main reason for not complying with a negative biopsy recommendation were a PSA \geq 3ng/ml for urologists or wanting certainty for patients.[18] In the decision to perform a prostate biopsy risk-calculators seem easy tools (most risk-calculators today can be found and used online as a web based tool, e.g. www.prostatecancer-riskcalculator.com, or even downloaded as an app for your mobile phone, e.g. Rotterdam prostate cancer risk calculator (Google play store and Apple app store)) to increase patient participation and reduce unnecessary examinations.

Several risk-calculators were developed to predict the risk of having an indolent prostate cancer as defined by radical prostatectomy characteristics (table 3). Kattan et al. developed a prediction model based on a clinical cohort of 409 men with cT1c or cT2a Gleason \leq 6 prostate cancer who received radical prostatectomy.[20] In total 20% of men had indolent disease (defined as organ confined, Gleason score \leq 6, prostate cancer with a tumor volume \leq 0.5). The model, including PSA, primary and secondary Gleason grade, clinical stage, ultrasound prostate volume, and mm cancer and mm noncancerous tissue, could reasonably predict indolent disease with a receiver operating characteristics (ROC) area under the curve (AUC) of 0.79. As already noted by the authors of this nomogram the percentage of men with indolent disease increased over time in their population. As the percentage of indolent prostate cancers in the population increases the nomogram predictions might underestimate the chance of having a indolent prostate cancer. In 2007, the nomogram was therefore updated on their website (www.nomograms.org) to better fit a more contemporary population.

Because the underlying prevalence of indolent disease in the population the model is developed on makes a substantial difference in predicted risks, a model was developed to better apply to a more intensively PSA screened population.[21] Steyerberg et al. adapted the model developed by Kattan et al. based on 278 men detected in the screening arm of the European Randomized study of Screening for Prostate Cancer (ERSPC) Rotterdam. At radical prostatectomy 49% of men had indolent disease. The new model uses the same predictors as the previous model (PSA, primary and secondary Gleason grade, clinical stage, ultrasound prostate volume, and mm cancer and mm noncancerous tissue) and is again able to predict indolent disease moderately well with an AUC of 0.76 (online available at www.prostatecancer-riskcalculator.com). [21]Based on this nomogram it was estimated that in a PSA screening setting 30% of men would have indolent disease and could be suitable for conservative management.[22] Because the model was developed based on men diagnosed with sextant prostate biopsy, length of prostate cancer and length of noncancerous tissue might not be accurate for men diagnosed with more contemporary extended biopsy core schemes. Correction factors were therefore calculated to be able to accurately predict the risk of men diagnosed with 12- or 18- core biopsies (mm cancer

should be divided by 2.03 or 2.72 and mm non-cancerous tissue by 2 or 3 for 12- or 18 core biopsies respectively when applying the nomogram).[23]

Additional nomograms to predict indolent disease at radical prostatectomy (again defined as organ confined, Gleason score ≤ 6 , prostate cancer with a tumor volume ≤ 0.5) were developed by Nakanishi et al., Chun et al., and O'Brien et al.[24-26] The first was developed on a cohort of 258 men with only 1 positive biopsy core on an extended biopsy scheme (10 to 13 cores).[24] Because of this strict selection, the percentage of men with indolent disease at radical prostatectomy in the study by Nakanishi et al. was higher than previous studies at 52%. Using age, PSA-density, and mm cancer tissue the model could moderately well predict indolent disease with an AUC of 0.73. Chun et al. developed a nomogram on a European cohort of men of all Gleason scores that were most likely not extensively PSA screened.[25] This was underlined by the very low rate of indolent disease at radical prostatectomy of only 6%. A model including PSA, biopsy Gleason sum score, length of cancer tissue, and % of positive biopsy cores was developed which had a AUC of 0.90. In the same analysis the Kattan et al. nomogram was validated which had an AUC of 0.81 in this cohort.[25] The nomogram developed by O'Brien et al. was developed in an Australian cohort of men with all Gleason scores and again a very low rate of indolent disease of 6%.[26] AUC was again high at 0.93. In addition, the nomogram slightly outperformed two rule-based inclusion criteria for active surveillance programs. Although both the Chun et al. and the O'Brien et al. nomogram may have very high predictive capabilities, results need to be interpreted with caution. Both cohorts used for model development included a substantial part of men with biopsy Gleason scores >6 (41% and 68% for Chun et al. and O'Brien et al. respectively). These men almost per definition do not have indolent disease at radical prostatectomy (in the study of O'Brien et al 0.5% of men with Gleason score >6 on prostate biopsy had indolent disease due to Gleason score downgrading at radical prostatectomy[26]). Inclusion of these men for model development tends to inflate the predictive capability of the model as indicated by the AUC. In addition, adding these men might alter the estimated prognostic effect of individual parameters for men with Gleason scores of 6.

All previously described nomograms are designed to predict the presence of indolent disease at immediate radical prostatectomy. However, as previously described this is used as a surrogate for a tumor that would not cause any symptoms or death during a man's life. The last nomogram that will be discussed takes a different approach and aims to predict prostate cancer specific survival after 10 years if conservative treatment is opted for (i.e. watchful waiting, which differs from active surveillance in that it does not attempt to offer curative treatment, but only palliative treatment if symptoms occur).[27] The model was based on 1310 men diagnosed with prostate cancer either by biopsy or transurethral resection of the prostate between 1990 and 1996. Cox regression analysis was used to predict 10 year prostate cancer mortality rates using clinical stage, method of diagnosis, % cancer tissue, PSA, age, Gleason sum score, and the use of early hormonal treatment (within 6 months).The concordance index (similar to the AUC, but for censored data) was moderate (0.73).[27] Low 10 year disease specific mortality rates obtained

from this model, combined with a short to intermediate life expectancy, could be used to select men for watchful waiting.

WHICH PREDICTION MODEL TO USE?

Clinicians and patient thus have several different nomograms at their disposal to help differentiate between indolent disease, most likely suitable for conservative management, or less indolent disease, which might require more aggressive treatment. But which of these nomograms should be used and which one is most suitable? Several aspects need to be addressed. First, a nomogram developed on a specific cohort might perform well on that cohort, but have limited predictive capabilities outside this setting.[28] External validation of a model is therefore essential.[29, 30]

Both the Kattan et al. nomogram and the Steyerberget al. nomogram were validated in an external population of 296 men with Gleason score 6, localized disease.[31] At radical prostatectomy 27% had indolent prostate cancer. Both models performed equally well in predicting indolent disease with an AUC of 0.77, which is similar to the predictive accuracy of the development cohort, indicating good generalizability. A second validation was done of all 5 nomograms described above in a contemporary cohort of 370 men with Gleason 6 disease on transrectal prostate biopsy.[32] In 38% of patients indolent disease was present on radical prostatectomy. Result indicated both the Kattan et al. and the Steyerberg et al. nomograms significantly outperformed the Nakanishi et al. nomogram, which in its turn outperformed both the Chun et al. and the O'Brien et al. nomograms. Predictive capabilities were again moderately well for the Kattan et al. and the Steyerberg et al. nomogram with an AUC of 0.77.[32] These two nomograms also showed good calibration and the highest net benefit.[32] It was noted that all models were most accurate at low predictive capabilities, indicating that these models are best at excluding indolent disease rather than accurately identifying it. One of the reasons for the lower performance at higher predictive values was the presence of anterior and apical tumors.[32] Both located at areas not frequently sampled with standard transrectal biopsy schemes. The nomogram predicting 10 year disease free survival has not yet been externally validated.[27]

As is shown the specific population at external validation may affect results. This is illustrated in an example of predicted probabilities of indolent disease (table 3). If the chance of having indolent disease for a 65-year-old men with 5 mm prostate cancer in 1 biopsy core is calculated with all nomograms, predictions range from 1-79%. Ten year disease free survival is calculated at 90-96%. Predicted probabilities seem very dependent on the percentage of men with indolent disease in the development cohort. Both the Chun et al. and the O'Brien et al. nomogram, which were developed in a group of men with often Gleason scores >6, seem not well able to identify indolent disease in men with lower Gleason scores. Other important differences include the development in a clinical cohort or a screening cohort. The latter having more men with indolent disease. Furthermore, most cohorts were developed in white European or American

Table 3. Overview of risk prediction tools for indolent prostate cancer.

	Indolent prostate cancer (organ confined, Gleason score 6, tumor volume $\leq 0.5 \text{ cm}^3$) at radical prostatectomy					10 year disease free survival
	Kattan et al.[20]	Steyerberg et al.[21]	Nakanishi et al.[24]	Chun et al.[25]	O'Brien et al.[26]	Kattan et al.[27]
Cohort origin	Clinical	Screening	Clinical	Clinical	Clinical	Clinical
Number of men used for nomogram development	409	247	254	1132	2525	1310
Percentage with indolent disease	20%	49%	52%	6%	6%	-
Example of risk prediction (Man, age 65 year, PSA 5ng/ml, prostate volume 50cm³, cT1c, Gleason 3+3, 1 of 12 cores positive, 5mm cancerous tissue, 200mm benign tissue, no early hormonal therapy)	48%	79%	37%	4-10%*	1-5%*	96-90%*

*Risk predictions were based on graphical devices and therefore presented as an interval.

men. Applicability to Asian or African men might be limited. Clinicians should be aware of these differences and try to select a nomogram best suitable for their patient. Overall, the Kattan et al. and Steyerberg et al. nomograms seem to be the most widely applicable, and outperform other nomograms.[31, 32] In addition, both nomograms can be easily applied, using an online tool (www.prostatecancer-riskcalculator.com and www.nomograms.org) which simply calculates the predicted probability after provide parameter data, greatly enhancing clinical usability.

CLINICAL APPLICABILITY

Although many nomograms are developed to help clinicians and patients in treatment decisions, very few make it into clinical practice. Physicians may be reluctant to trust nomogram predictions and rather choose to follow well established preconceptions. An example is provided by a study on the implementation of a risk-calculator to aid in the decision of prostate biopsy.[18] Although the risk-calculator (i.e. the ERSPC risk-calculator[10, 11]) was proven superior over a single PSA cut-off, 36% of men were biopsied against a negative biopsy recommendation provided by the risk-calculator. When asked for the reason of ignoring the advice, 78% of times a PSA ≥ 3.0 ng/ml was replied. Of men ignoring the negative biopsy recommendation only 3% showed aggressive prostate cancer on biopsy.[19] The remaining 97% received an unnecessary biopsies because of the prejudice that a PSA-value ≥ 3.0 ng/ml should trigger further investigation. That said, it seems vital to not only conduct studies on how to improve selection, but also to better implement successful tools in clinical practice. This can be done by simple presenting the risk provided by a nomogram, but can also be more elaborate e.g. by combine individual risk scores

with information on its meaning, prognosis, and the advantages and disadvantages of different treatment options into a (personalized) decision aid.

One study aimed to investigate the impact of using a nomogram to advise men on active treatment or active surveillance.[33] The Steyerberg et al. nomogram was used to predict the presence of indolent disease in 240 men diagnosed with prostate cancer in five Dutch hospitals. As a rule of thumb, a probability cut-off of $\geq 70\%$ was chosen to advise men on active surveillance. With this cut-off 82% of patients adhered to the recommendation to choose active surveillance. Surprisingly 29% of men with a probability $< 70\%$ of indolent disease also chose an initial active surveillance strategy. The main reason being the patients' preference to delay physical side effect of active treatment.[33] Measurements of the decisional conflict scale were low in this study, indicating that patients felt well informed by the nomogram and certain in their choice of treatment. Two other randomized trials have shown that using decision aids (not including individual risk assessments) for treatment choice in localized prostate cancer not only helped patients to make more informed decisions on treatment, but also increase satisfaction with the decision made.[34, 35] In addition, the better information provided and higher patient participation might have an effect on the treatment that is selected.[35, 36] Decision aids, including personalized risk assessments, seem good tools not only to improve the selection of men with indolent disease, but will also increase patient understanding, participation and satisfaction in the treatment chosen for the management of their prostate cancer.

FUTURE PERSPECTIVE

Although the use of nomograms to help treatment decision in the increasing number of men diagnosed with low-risk localized prostate cancer seems preferable over rule-based decision supports to reduce overtreatment, there are some limitations. As with all rule-based criteria, none of the presented nomograms is able to perfectly predict the presence of indolent disease. In fact the nomograms seem more suitable to exclude the presence of significant disease. Most likely the restrictions of currently used blind biopsy sampling, often missing anterior and apical tumors, contribute to this. Improvement of current nomograms is therefore essential. Promising and most likely to be quickly incorporated is the use of MRI. MRI seems especially useful in visualizing higher grade prostate cancers.[37, 38] MRI visualized lesions could trigger targeted biopsies which might better represent tumor grade and volume. Nomogram predictions are therefore likely to improve if data on targeted biopsies could be added. In addition to the information provided by targeted biopsies, the MR images themselves could provide new parameters on tumor characteristics. These not only include tumor volume as can be measured on MRI, but also water diffusion coefficients which seem to correlate with tumor aggressiveness.[39] MRI, using spectroscopic imaging, could also be used to obtain information on a molecular level, which again might help to predict tumor aggressiveness.[40, 41] Next to imaging, genomic and histological

information could potentially provide better information on tumor behavior and help to decide on the most appropriate treatment strategy. A recent study genotyped 242,221 single nucleotide polymorphisms (SNPs) in blood DNA of men with Gleason 6 prostate cancer.[42] Fifteen SNPs were found to be able to predict Gleason score upgrading on radical prostatectomy, however only one SNP remained predictive if other clinical information was added. The addition of the SNP to a clinical model significantly improved the predictive accuracy.[42] Future studies should validate if these findings remain significant and could improve the prediction of indolent disease.

CONCLUSION

Several nomograms may aid men in assessing their risk of having an indolent tumor, which maybe most suitable for conservative management. These decision aids, although not perfect in their prediction of indolent disease, are preferable over commonly used rule-based selection criteria for active surveillance, because they provide a more individual risk-assessment, which helps to better inform men facing treatment decision. For clinicians it is important to choose a nomogram that is most accurate, externally validated and best fits the patients characteristics. Well validated nomograms with reasonable accuracy can be found online for a clinical population (www.nomograms.org) or for a more intensive PSA screened population (www.prostatecancer-riskcalculator.com). Future developments, as MRI and new genetic markers, will likely improve current nomograms. Implementation into clinical practice is however already shown valuable. The time has therefore arrived to start using these prediction tools in clinical practice to provide the best of care for the large number of men diagnosed with low-risk prostate cancer today.

REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer Statistics, 2014. *CA Cancer J Clin*. 2014.
2. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*. 2013;49:1374-403.
3. Center MM, Jemal A, Lortet-Tieulent J, Ward E, Ferlay J, Brawley O, et al. International Variation in Prostate Cancer Incidence and Mortality Rates. *Eur Urol*. 2012.
4. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014.
5. Draisma G, Boer R, Otto SJ, van der Crujnsen IW, Damhuis RA, Schroder FH, et al. Lead times and overdiagnosis 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.
6. Rich AR. On the frequency of occurrence of occult carcinoma. *Journal of Urology*. 1935;33:215-23.
7. Wilt TJ, Brawer MK, Jones KM, Barry MJ, Aronson WJ, Fox S, et al. Radical prostatectomy versus observation for localized prostate cancer. *N Engl J Med*. 2012;367:203-13.
8. Sanda MG, Dunn RL, Michalski J, Sandler HM, Northouse L, Hembroff L, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. *N Engl J Med*. 2008;358:1250-61.
9. Zhu X, Albertsen PC, Andriole GL, Roobol MJ, Schroder FH, Vickers AJ. Risk-Based Prostate Cancer Screening. *Eur Urol*. 2011;61:652-61.
10. Krane R, Roobol M, Schroder FH. A graphical device to represent the outcomes of a logistic regression analysis. *Prostate*. 2008;68:1674-80.
11. Roobol MJ, Steyerberg EW, Krane R, Wolters T, van den Bergh RC, Bangma CH, et al. A risk-based strategy improves prostate-specific antigen-driven detection of prostate cancer. *Eur Urol*. 2010;57:79-85.
12. Roobol MJ, Zhu X, Schroder FH, van Leenders GJ, van Schaik RH, Bangma CH, et al. A Calculator for Prostate Cancer Risk 4 Years After an Initially Negative Screen: Findings from ERSPC Rotterdam. *Eur Urol*. 2013;63:627-33.
13. Oliveira M, Marques V, Carvalho AP, Santos A. Head-to-head comparison of two online nomograms for prostate biopsy outcome prediction. *BJU international*. 2011;107:1780-3.
14. Ouzaid I, Yates DR, Hupertan V, Mozer P, Chartier-Kastler E, Haertig A, et al. A direct comparison of the diagnostic accuracy of three prostate cancer nomograms designed to predict the likelihood of a positive initial transrectal biopsy. *Prostate*. 2012;72:1200-6.
15. Roobol MJ, Schroder FH, Hugosson J, Jones JS, Kattan MW, Klein EA, et al. Importance of prostate volume in the European Randomised Study of Screening for Prostate Cancer (ERSPC) risk calculators: results from the prostate biopsy collaborative group. *World J Urol*. 2012;30:149-55.
16. Trottier G, Roobol MJ, Lawrentschuk N, Bostrom PJ, Fernandes KA, Finelli A, et al. Comparison of risk calculators from the Prostate Cancer Prevention Trial and the European Randomized Study of Screening for Prostate Cancer in a contemporary Canadian cohort. *BJU international*. 2011;108:E237-44.
17. van Vugt HA, Roobol MJ, Krane R, Maattanen L, Finne P, Hugosson J, et al. Prediction of prostate cancer in unscreened men: external validation of a risk calculator. *Eur J Cancer*. 2011;47:903-9.
18. van Vugt HA, Roobol MJ, Busstra M, Kil P, Oomens EH, de Jong IJ, et al. Compliance with biopsy recommendations of a prostate cancer risk calculator. *BJU Int*. 2012;109:1480-8.

19. van Vugt HA, Kranse R, Steyerberg EW, van der Poel HG, Busstra M, Kil P, et al. Prospective validation of a risk calculator which calculates the probability of a positive prostate biopsy in a contemporary clinical cohort. *Eur J Cancer*. 2012;48:1809-15.
20. Kattan MW, Eastham JA, Wheeler TM, Maru N, Scardino PT, Erbersdobler A, et al. Counseling men with prostate cancer: a nomogram for predicting the presence of small, moderately differentiated, confined tumors. *The Journal of urology*. 2003;170:1792-7.
21. Steyerberg EW, Roobol MJ, Kattan MW, van der Kwast TH, de Koning HJ, Schroder FH. Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. *The Journal of urology*. 2007;177:107-12; discussion 12.
22. Roemeling S, Roobol MJ, Kattan MW, van der Kwast TH, Steyerberg EW, Schroder FH. Nomogram use for the prediction of indolent prostate cancer: impact on screen-detected populations. *Cancer*. 2007;110:2218-21.
23. Bul M, Delongchamps NB, Steyerberg EW, de la Roza G, van Leeuwen PJ, Zhu X, et al. Updating the prostate cancer risk indicator for contemporary biopsy schemes. *Can J Urol*. 2011;18:5625-9.
24. Nakanishi H, Wang X, Ochiai A, Trpkov K, Yilmaz A, Donnelly JB, et al. A nomogram for predicting low-volume/low-grade prostate cancer: a tool in selecting patients for active surveillance. *Cancer*. 2007;110:2441-7.
25. Chun FK, Haese A, Ahyai SA, Walz J, Suardi N, Capitanio U, et al. Critical assessment of tools to predict clinically insignificant prostate cancer at radical prostatectomy in contemporary men. *Cancer*. 2008;113:701-9.
26. O'Brien BA, Cohen RJ, Ryan A, Sengupta S, Mills J. A new preoperative nomogram to predict minimal prostate cancer: accuracy and error rates compared to other tools to select patients for active surveillance. *J Urol*. 2011;186:1811-7.
27. Kattan MW, Cuzick J, Fisher G, Berney DM, Oliver T, Foster CS, et al. Nomogram incorporating PSA level to predict cancer-specific survival for men with clinically localized prostate cancer managed without curative intent. *Cancer*. 2008;112:69-74.
28. Steyerberg EW. *Clinical Prediction Models*: Springer; 2009.
29. Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE, et al. External validation is necessary in prediction research: a clinical example. *Journal of clinical epidemiology*. 2003;56:826-32.
30. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KG. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *Journal of clinical epidemiology*. 2003;56:441-7.
31. Dong F, Kattan MW, Steyerberg EW, Jones JS, Stephenson AJ, Schroder FH, et al. Validation of pre-treatment nomograms for predicting indolent prostate cancer: efficacy in contemporary urological practice. *The Journal of urology*. 2008;180:150-4; discussion 4.
32. Iremashvili V, Soloway MS, Pelaez L, Rosenberg DL, Manoharan M. Comparative validation of nomograms predicting clinically insignificant prostate cancer. *Urology*. 2013;81:1202-8.
33. van Vugt HA, Roobol MJ, van der Poel HG, van Muilekom EH, Busstra M, Kil P, et al. Selecting men diagnosed with prostate cancer for active surveillance using a risk calculator: a prospective impact study. *BJU Int*. 2012;110:180-7.
34. Chabrera C, Zabalegui A, Bonet M, Caro M, Areal J, Gonzalez JR, et al. A Decision Aid to Support Informed Choices for Patients Recently Diagnosed With Prostate Cancer: A Randomized Controlled Trial. *Cancer Nurs*. 2014.

35. Berry DL, Halpenny B, Hong F, Wolpin S, Lober WB, Russell KJ, et al. The Personal Patient Profile-Prostate decision support for men with localized prostate cancer: a multi-center randomized trial. *Urologic oncology*. 2013;31:1012-21.
36. Auvinen A, Hakama M, Ala-Opas M, Vornanen T, Leppilahti M, Salminen P, et al. A randomized trial of choice of treatment in prostate cancer: the effect of intervention on the treatment chosen. *BJU Int*. 2004;93:52-6; discussion 6.
37. Pokorny MR, de Rooij M, Duncan E, Schroder FH, Parkinson R, Barentsz JO, et al. Prospective Study of Diagnostic Accuracy Comparing Prostate Cancer Detection by Transrectal Ultrasound-Guided Biopsy Versus Magnetic Resonance (MR) Imaging with Subsequent MR-guided Biopsy in Men Without Previous Prostate Biopsies. *Eur Urol*. 2014;66:22-9.
38. Siddiqui MM, Rais-Bahrami S, Truong H, Stamatakis L, Vourganti S, Nix J, et al. Magnetic resonance imaging/ultrasound-fusion biopsy significantly upgrades prostate cancer versus systematic 12-core transrectal ultrasound biopsy. *Eur Urol*. 2013;64:713-9.
39. Hambrock T, Hoeks C, Hulsbergen-van de Kaa C, Scheenen T, Futterer J, Bouwense S, et al. Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort. *Eur Urol*. 2012;61:177-84.
40. Thormer G, Otto J, Horn LC, Garnov N, Do M, Franz T, et al. Non-invasive estimation of prostate cancer aggressiveness using diffusion-weighted MRI and 3D proton MR spectroscopy at 3.0 T. *Acta Radiol*. 2014.
41. Kobus T, Vos PC, Hambrock T, De Rooij M, Hulsbergen-Van de Kaa CA, Barentsz JO, et al. Prostate cancer aggressiveness: in vivo assessment of MR spectroscopy and diffusion-weighted imaging at 3 T. *Radiology*. 2012;265:457-67.
42. Oh JJ, Park S, Lee SE, Hong SK, Lee S, Choe G, et al. The use of exome genotyping to predict pathological Gleason score upgrade after radical prostatectomy in low-risk prostate cancer patients. *PLoS One*. 2014;9:e104146.
43. Wolters T, Roobol MJ, van Leeuwen PJ, van den Bergh RC, Hoedemaeker RF, van Leenders GJ, et al. A critical analysis of the tumor volume threshold for clinically insignificant prostate cancer using a data set of a randomized screening trial. *The Journal of urology*. 2011;185:121-5.
44. van As NJ, Norman AR, Thomas K, Khoo VS, Thompson A, Huddart RA, et al. Predicting the probability of deferred radical treatment for localised prostate cancer managed by active surveillance. *Eur Urol*. 2008;54:1297-305.
45. Soloway MS, Soloway CT, Williams S, Ayyathurai R, Kava B, Manoharan M. Active surveillance; a reasonable management alternative for patients with prostate cancer: the Miami experience. *BJU international*. 2008;101:165-9.
46. Carter HB, Kettermann A, Warlick C, Metter EJ, Landis P, Walsh PC, et al. Expectant management of prostate cancer with curative intent: an update of the Johns Hopkins experience. *The Journal of urology*. 2007;178:2359-64; discussion 64-5.
47. Dall'Era MA, Konety BR, Cowan JE, Shinohara K, Stauf F, Cooperberg MR, et al. Active surveillance for the management of prostate cancer in a contemporary cohort. *Cancer*. 2008;112:2664-70.
48. Klotz L, Zhang L, Lam A, Nam R, Mamedov A, Loblaw A. Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. *J Clin Oncol*. 2010;28:126-31.
49. van den Bergh RC, Vasarainen H, van der Poel HG, Vis-Maters JJ, Rietbergen JB, Pickles T, et al. Short-term outcomes of the prospective multicentre 'Prostate Cancer Research International: Active Surveillance' study. *BJU international*. 2010;105:956-62.

Chapter 6

Compliance rates with the Prostate cancer Research International: Active Surveillance (PRIAS) protocol and disease reclassification in non-compliers.

Leonard P. Bokhorst^a, Arnout R. Alberts^a, Antti Rannikko^b, Riccardo Valdagni^c,
Tom Pickles^d, Yoshiyuki Kakehi^e, Chris H. Bangma^a, Monique J. Roobol^a,
for the PRIAS study group

^a Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands.

^b Department of Urology, Helsinki University Central Hospital, Helsinki, Finland

^c Prostate Cancer Program and Radiation Oncology 1, Fondazione IRCSS Istituto Nazionale dei Tumori, Milan, Italy

^d Department of Radiation Oncology, British Columbia Cancer Agency, Vancouver, Canada

^e Department of Urology, Kagawa University Faculty of Medicine, Kagawa, Japan

ABSTRACT

Background: Men with prostate cancer on active surveillance are advised to follow strict follow-up schedules and switch to definitive treatment if risk reclassification occurs. However, some men might not adhere to these strict protocols.

Objective: To determine the number of non-compliers and disease reclassification rates in men not complying with the follow-up protocol of the Prostate cancer Research International: Active Surveillance (PRIAS) study.

Design, setting, and participants: 4547 men with low-risk prostate cancer were included and prospectively followed on active surveillance. Men were regularly examined using PSA, DRE, and repeat biopsies and advised to switch to definitive treatment if disease reclassification occurred (>cT2c, Gleason score >3+3, >2 cores positive, or PSA-doubling time (PSA-DT) 0-3 year).

Outcome measurements and statistical analysis: Rates of men not complying with the follow-up visits or recommendation to discontinue active surveillance were reported. Biopsy outcome (Gleason score ≥ 7 or > 2 cores positive) between compliers and non-compliers was compared using cox proportional hazard analysis.

Results and limitations: The compliance rate with PSA visits was 91%. In contrast compliance rates with standard repeat biopsies decreased over time (81%, 60%, 53%, and 33% for 1, 4, 7, and 10 years after diagnosis respectively). Yearly repeat biopsies in men with faster rising PSA (PSA-DT 3-10 year) was low at less than 30%, although these men had higher upgrading rates at repeat biopsy (25-30% versus 16%). A PSA-DT of 0-3 year was the most common recommendation to discontinue, nevertheless 71% continued active surveillance. Men with PSADT 0-3 year were at higher risk of upgrading on repeat biopsy (HR 2.02; 95% CI 1.36-3.00) as compared to men without fast rising PSA.

Conclusion: Some men and their physician do not comply with an active surveillance follow-up protocol. Especially yearly repeat biopsies in men with fast rising PSA, are often ignored, as is the recommendation to discontinue active surveillance due to a very fast rising PSA. Although these men are at increased risk of having higher Gleason scores on repeat biopsy, the majority still presents favorable tumor characteristics. A fast rising PSA should therefore not be a recommendation to advice active treatment, but should rather serve as a criterion for stricter follow-up. In addition, we should aim to find ways of safely reducing the amount of biopsies to increase adherence to active surveillance protocols.

Patient summary: In this report we looked at the compliance with a large active surveillance protocol for low risk prostate cancer. We observed reluctance with yearly biopsies due to a fast rising PSA, despite a higher risk of disease progression. Further research should aim to safely reduce the amount of repeat biopsies in men on active surveillance, to increase protocol adherence.

INTRODUCTION

Active surveillance for prostate cancer is a treatment option aimed at reducing the negative side effects of radical treatment, while at the same time preserving the option for curative treatment. It does so by strictly following men and only offering curative treatment to those that show signs of disease progression / reclassification. However, optimal criteria for follow-up, inclusion and exclusion are currently still being investigated. Most common protocols include criteria based on a combination of PSA tests, digital rectal examinations (DRE), and repeated prostate biopsies to both include patients and define disease reclassification [1-6].

Some men and their physicians might however choose to deviate from these strict protocols, ignoring either the follow-up schedule or the advice to switch to curative treatment.

The aim of the current analysis is to determine the number of men who do not comply with the protocol of the Prostate cancer Research International: Active Surveillance (PRIAS) study. The PRIAS study is currently the largest prospective study on active surveillance, including over 100 centers in 17 countries aimed to represent a real world situation [1]. Furthermore, follow-up of men not complying with the approved protocol allows us to evaluate the protocol by investigating their intermediate term outcomes (i.e. Gleason score upgrading at repeat biopsy).

METHODS

In the PRIAS study men with low risk prostate cancer are prospectively followed on active surveillance [7]. All centers enter data on inclusion and follow-up through an online tool (www.prias-project.org), which automatically provides all recommendations for follow-up based on the protocol [7]. Criteria for inclusion are: Gleason score $\leq 3+3$, $\leq cT2c$, PSA ≤ 10 ng/ml, \leq two cores positive for prostate cancer, PSA density ≤ 0.2 ng/ml/ml, and fitness for curative treatment. A minimum number of biopsy cores taken is advised based on prostate volume (prostate volume < 40 cm³: 8 cores, 40-60cm³: 10 cores, > 60 cm³: 12 cores), but is not a strict inclusion criterion. As of 2012 men with minimal Gleason score 3+4 disease ($\leq 10\%$ core involvement) can be included if aged ≥ 70 year ($n=24$)(for follow-up all regular criteria apply except for Gleason score, which can be 3+4 on repeat biopsy).

Men are followed using PSA testing every 3 months the first 2 years and every 6 months thereafter. Digital rectal examination is advised every 6 months the first 2 years and every year thereafter. Repeat biopsies are done 1,4,7,10, and subsequent every 5 years after diagnosis. Yearly repeat biopsies are only advised if PSA doubling time (PSA-DT) is between 3 and 10 years. PSA-DT is calculated using all available PSA values since diagnosis by plotting the base 2 logarithm of the PSA values against the time since diagnosis. The PSA-DT is then calculated as the reciprocal value of the slope of the regression line through these points. PSA-DT is only used if at least 4 PSA values are available. A bone scan is recommended if PSA ≥ 20 ng/ml. Criteria used to

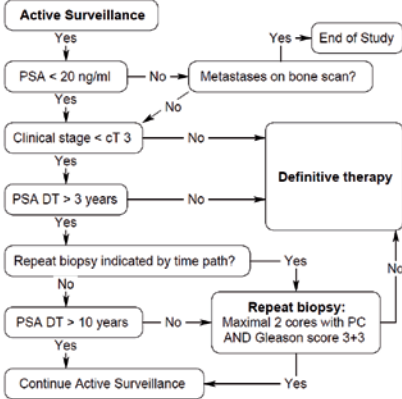
recommend a switch to definitive treatment are: Gleason score >3+3, >2 biopsy cores positive for prostate cancer, >cT2c, and a PSA-DT of 0-3 year (if at least 4 PSA values are available) on any of the follow-up visits (figure 1). Follow-up for the current analysis ended 31 December 2014.

Compliance with the follow-up schedule was studied per year of being on active surveillance. Men were defined as being compliant with the PSA visits if in the first 2 years at least 3 PSA tests were done per year and at least 1 PSA test in the years thereafter. A biopsy 6 months before or after the designated time for the biopsy was classified as being compliant with that biopsy. Standard biopsies should have been done in year 1, 4, 7, and 10 in men with >1.5, >4.5, >7.5, and >10.5 years of follow-up respectively to classify as being compliant. Men with a PSA-DT of 3-10 years within the years with no scheduled standard repeat biopsy (years 2, 3, 5, 6, 8, 9 after inclusion, see figure 1) should have had a biopsy in that year. Two definitions were used for non-compliance with a protocol based reason to discontinue active surveillance: at least 1 PSA visit or at least 1 biopsy after the protocol recommendation to discontinue.

Year	1				2				3				4				5				6				7			
Month	0**	3	6	9	12	15	18	21	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108	114	120	126	132	
PSA-test	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DRE	✓				✓				✓				✓				✓				✓				✓			
Biopsy*	✓												✓															
Evaluation	✓				✓				✓				✓				✓				✓				✓			

* repeat biopsy. Standard after 1, 4, 7 and 10 year and subsequently every 5 years.
 If PSA-DT is 0-10 years repeat biopsy every year is advised. No more than 1 biopsy per year should be performed
 ** Time of diagnosis

Active surveillance: criteria for follow-up (old protocol until



Active surveillance: criteria for follow-up (new protocol from

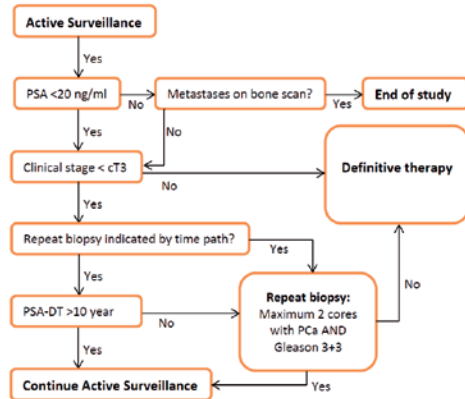


Figure 1. Follow-up schedule and criteria for follow-up (left flowchart) of the PRIAS study. As of 2015 the follow-up criteria were changed (flow chart to the right).

Active surveillance: criteria for follow-up (new protocol from 2015)

Active surveillance: criteria for follow-up (old protocol until 2015)

STATISTICAL ANALYSIS

Upgrading in men not complying with a recommendation to biopsy

The number of men with upgrading (Gleason score >6 or >2 cores positive) on the second standard repeat biopsy (year 4) was compared for men without a PSA-DT between 3 and 10 in the second and third year and men with a PSA-DT between 3 and 10 during that period, but who did not receive an early repeat biopsy, using the Chi-square test. In addition, a comparison was made with the number of men with upgrading on repeat biopsy in year 2 or 3 triggered by a PSA-DT between 3 and 10. For equal comparison all men had the first scheduled repeat biopsy in year 1.

Upgrading in men not complying with a recommendation to discontinue

As the number of previous biopsies during follow-up could influence upgrading rates we only reported upgrading rates on the second repeat biopsy during follow-up for men who ignored a recommendation to discontinue active surveillance on the first repeat biopsy during follow-up (either Gleason score >6 or > 2 cores positive) or ignored a recommendation to discontinue in-between the first and second repeat biopsy (due to a PSA-DT of 0-3 year). As comparison, upgrading rates on the second repeat biopsy during follow-up for men without a previous recommendation to discontinue were reported. As time between the first two biopsies could differ between these groups we conducted a cox proportional hazard analysis with correction for age, PSA and number of positive cores at the first repeat biopsy and PSA density at diagnosis to predict upgrading on the second repeat biopsy. For this analysis PSA-DT 0-3 was assumed to be at the time of the first repeat biopsy, as most PSA-DT 0-3 occurred within 1 year of the first repeat biopsy [8]. For all analysis SPSS for windows (Version 21.0. Armonk, NY: IBM Corp.) was used.

RESULTS

Until the end of follow-up 4547 men were included and followed on active surveillance in the PRIAS study. As inclusion and follow-up is still ongoing, the median time on active surveillance for all men was only 1.5 years, but 750 men were followed for more than 4 years and 94 men for more than 7 years.

Compliance with PSA and biopsy visits

During follow-up 91% of patients complied with all PSA visits. After year 7 a slight decrease in compliance with the scheduled PSA visits was seen (figure 2). The rate of compliance with all advised biopsy visits was lower at 70%. Compliance rates with the standard biopsies (year 1, 4, 7, and 10) decreased over the years with 1867/2306 men (81%) complying with the 1 year repeat biopsy, 333/559 (60%) with the 4 year repeat biopsy, 27/51 (53%) with the 7 year repeat biopsy, and 1/3 men (33%) with the 10 year repeat biopsy. Overall compliance rates with the yearly

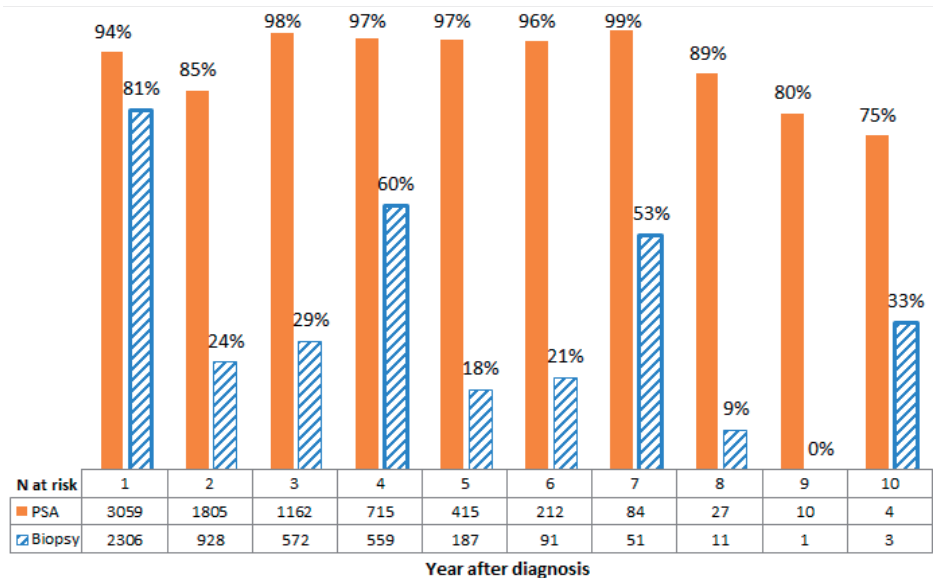


Figure 2. Percentage of men complying with PSA testing and prostate biopsies in men on active surveillance per year. (standard repeat biopsies are highlighted)

repeat biopsies due to a PSA-DT of 3-10 years was low ranging from 226/702 men (24%) in year 2 to 1/11 men (9%) in year 8 (figure 2). Of 750 men with more than 4 years of follow-up, 222 (30%) complied with all advised biopsies.

Men with a biopsy advise in year 2 or 3 (due to a PSA-DT of 3-10 years) who did not comply, more often had upgrading (Gleason >6 and/or >2 cores positive) on repeat biopsy at year 4, as compared to men without a PSA-DT of 3-10 years in year 2 or 3 (25% versus 16% respectively, $p=0.028$). Men with a PSA-DT of 3-10 in year 2 and 3 who did have a biopsy in year 2 or 3, were upgraded in 27% and 30% of cases respectively (table 1). In year 7 50 men had a repeat biopsy. Of the 22 men that fully complied with the biopsy protocol 1 (5%) had a Gleason score ≥ 7 . Of the 28 that did not fully comply, 5 (18%) had a Gleason score ≥ 7 ($p=0.15$).

Compliance with recommendation to discontinue active surveillance

During follow-up 10 men had clinical stage $\geq T3$ of which 2 continued AS (20%), 535 men had a Gleason score >6 at any repeat biopsy of which 96 continued AS (18%), 734 men had >2 cores positive for prostate cancer at any repeat biopsy of which 175 continued AS (24%), and 915 men had a PSA-DT of 0-3 year of which 651 continued AS (71%). The percentage of men continuing active surveillance were lower if a stricter definition was used (figure 3). Of all men who continued active surveillance despite a recommendation to discontinue 329 out of 839 (varying from 245/651 for PSA-DT of 0-3 year to 1/2 for clinical stage $\geq T3$) eventually switched of active

Table 1. Outcome in men complying and not complying with advised prostate biopsies.

Biopsies during follow-up	group1*		group2*		group3*		group4*	
	PSA-DT 3-10	Bx	PSA-DT 3-10	Bx	PSA-DT 3-10	Bx	PSA-DT 3-10	Bx
Year 1		Yes		Yes		Yes		Yes
Year 2	Yes	No	No	-	Yes	Yes	No	-
Year 3	Yes	No	No	-			Yes	Yes
Year 4		Yes		Yes				
Age at diagnosis, median (IQR)	64.9 (61.3-69.9)		65.9 (60.2-70.7)		64.4 (60.3-69)		64.5 (59.4-70.3)	
PSA at diagnosis, median (IQR)	5.1 (3.6-6.4)		5.7 (4.4-7.5)		5.4 (4.4-6.6)		5.5 (4.4-6.9)	
Outcome prostate biopsy in:	Year 4		Year 4		Year 2		Year 3	
no prostate cancer	58 (43%)		101 (47%)		61 (33%)		44 (37%)	
Gleason <=6	56 (42%)		97 (45%)		100 (54%)		57 (48%)	
Gleason 3+4	15 (11%)		10 (5%)		18 (10%)		12 (10%)	
Gleason 4+3	3 (2%)		5 (2%)		3 (2%)		2 (2%)	
Gleason >=8	2 (1%)		2 (1%)		4 (2%)		5 (4%)	
>2 cores positive	23 (17%)		25 (12%)		42 (23%)		28 (23%)	
Gleason >6 or >2 cores positive	34 (25%) ^a		34 (16%)		51 (27%)		36 (30%)	
Total	134 (100%)		215 (100%)		186 (100%)		120 (100%)	

*For comparison all men had a biopsy in year 1 and:

Group 1: Non-compliers: no biopsy in year 2 or 3 despite PSA-DT of 3-10 year in year 2 or 3

Group 2: Compliers: no recommendation for biopsy in year 2 or 3

Group 3: Compliers: PSA-DT of 3-10 year in year 2 and a biopsy in year 2

Group 4: Compliers: PSA-DT of 3-10 year in year 3 and a biopsy in year 3

^a: p-value as compared to group 2: 0.028.

IQR: interquartile range

surveillance after a median follow-up of 1.0 year after their recommendation to discontinue, and 510 out of 839 are still on active surveillance for a median of 1.7 year.

Men who continued active surveillance, and subsequently had a second repeat biopsy, despite >2 cores positive for prostate cancer or a PSA-DT of 0-3 year more often had a Gleason score >6 (15% and 16% respectively), as compared to men without a recommendation to discontinue (11%, table 2). After correction for other variables and time between biopsies, both a PSA-DT between 0-3 and >2 cores positive on first repeat biopsy were significant predictors of upgrading on the second repeat biopsy (table 3).

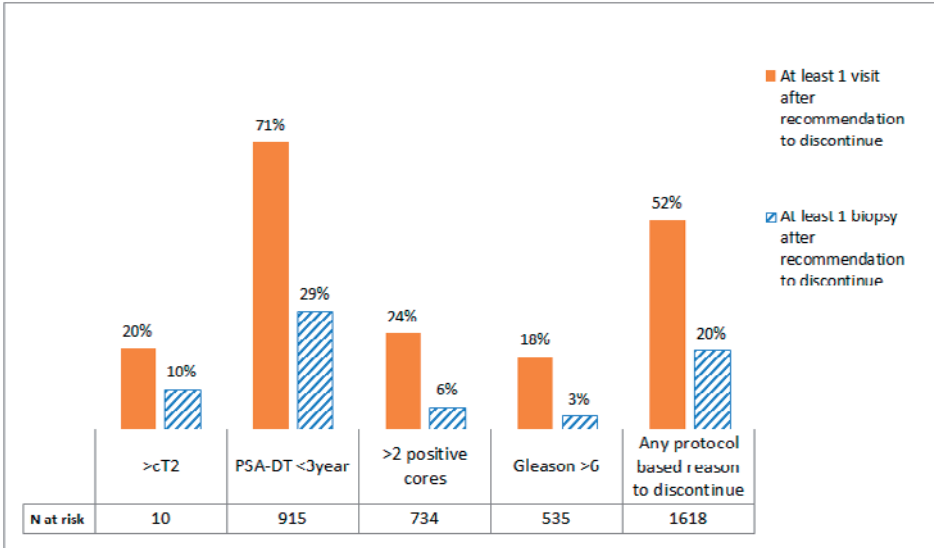


Figure 3. Percentage of men not complying with protocol based reasons to discontinue active surveillance.

DISCUSSION

The PRIAS study is currently the largest active surveillance study worldwide. It was aimed to provide a real world representation of active surveillance outside the more strictly controlled academic centers, by including both academic, non-academic and private practices across 17 countries and 4 continents. We observed a substantial proportion of men who did not comply with the repeat biopsies schedule. Especially the yearly biopsies due to a faster rising PSA were often ignored. PSA kinetics were in addition regularly put aside as recommendation to discontinue active surveillance. Both men ignoring the follow-up schedule and criteria for discontinuation of active surveillance were at increased risk of disease upgrading. Although a substantial number still presented with favorable disease characteristics on repeat biopsy.

We observed a clear decrease over time in the percentage of men receiving the standard repeat biopsies, from 81% in the first year, 60% and 53% in the fourth and seventh year, to 33% in the tenth year of follow-up. Although not recorded as standard, several common reasons for not complying with these standard repeat biopsies were recorded. Some examples included “patient does not want biopsy”, “PSA stable”, “no signs of disease progression on previous biopsy” or “complications on last biopsy”. This seems to indicate that the repeat biopsies might put a substantial strain on men. As compared to PSA testing, which was most often strictly complied with, biopsies are considered uncomfortable. In addition, several complications are recorded such as pain, hematuria, or even sepsis [9]. These complications will result in some men declining repeat biopsies [10]. Furthermore, increasing age (median age at diagnosis, 4, 7, and 10 year was 65.8, 69.5, 72.2, and 76.0 respectively) or previous negative biopsies combined with unchang-

Table 2. Outcome of second repeat biopsy in men continuing active surveillance despite protocol advice to switch to active treatment either at the first repeat biopsy (Gleason >6, >2 cores positive) or between first and second repeat biopsy (PSA-DT <=3 year), compared with men adhering to the protocol (no protocol based reason to discontinue).

Biopsies during follow-up	Continuation despite protocol advice							
	No protocol advice to discontinue		PSA-DT 0-3 year		>2 cores positive with PCa		Gleason score >6	
	Bx	PSA-DT 0-3	Bx	>2 cores	Bx	Gleason >6	Bx	
Year 1	1 st Bx	Yes	1 st Bx	Yes	1 st Bx	Yes	1 st Bx	
Year 2, 3, or 4	2 nd Bx		2 nd Bx		2 nd Bx		2 nd Bx	
Age at diagnosis, median (IQR)	64.7 (59.8-69.5)	64.5 (60.6-69.9)	65.9 (59.9-70.3)	68.4 (62.1-71.5)				
PSA at diagnosis, median (IQR)	5.6 (4.3-7)	5.1 (3.6-6.4)	4.9 (3.5-6.6)	5.6 (5-7.3)				
Time from first to second repeat biopsy, median (IQR)	2.2 (1.1-3.0)	1.5 (1.0-2.9)	1.3 (1.0-2.9)	0.5 (0.5-1.9)				
Time between PSA-DT 0-3 and second repeat biopsy, median (IQR)	-	1.4 (1.0-2.6)	-	-				
Outcome second repeat biopsy								
no PCa	267 (43%)	81 (37%)	4 (12%)	1 (8%)				
Gleason <=6	283 (46%)	103 (47%)	24 (73%)	4 (33%)				
Gleason 3+4	45 (7%)	21 (10%)	2 (6%)	6 (50%)				
Gleason 4+3	10 (2%)	6 (3%)	1 (3%)	0 (0%)				
Gleason >=8	11 (2%)	8 (4%)	2 (6%)	1 (8%)				
>2 cores positive	83 (13%)	12 (5%)	15 (45%)	5 (42%)				
Gleason >6 or >2 cores positive	120 (19%)	64 (29%)	18 (55%)	7 (58%)				
Total	616 (100%)	219 (100%)	33 (100%)	12 (100%)				

PSA-DT: PSA doubling time
 IQR: interquartile range
 PCa: prostate cancer

ing PSA values, might reassure both physicians and patients of stable disease which might not become clinically significant. This assumptions seems to be confirmed by biopsy results in men with a negative PSA-DT or a PSA-DT >10 years. In these men a biopsy 4 years after diagnosis showed Gleason score >6 in only 8% (group 2, table 1). This questions whether yearly biopsies for everyone, as used in some active surveillance studies [2, 11, 12], are justified. As active surveillance is primarily aimed at reducing the side effects of aggressive treatment to improve quality of

Table 3. Multivariable cox proportional hazard model predicting upgrading (Gleason >6 and/or >2 cores positive) on the second repeat biopsy during follow-up.

	HR (95% CI)	p-value
Age first repeat biopsy	1.00 (0.98 - 1.02)	0.9
PSA first repeat biopsy	1.00 (0.93 - 1.07)	>0.9
PSA second repeat biopsy	1.05 (0.99 - 1.11)	0.078
PSA density at diagnosis (0.1 increase)	1.26 (0.88 - 1.82)	0.2
Gleason >6 first repeat biopsy (but continued active surveillance)	3.59 (1.62 - 7.98)	0.002
Number of positive cores first repeat biopsy		<0.001
0	ref	
1	2.12 (1.48 - 3.03)	<0.001
2	3.19 (2.22 - 4.59)	<0.001
>=3 (but continued active surveillance)	4.32 (2.43 - 7.66)	<0.001
PSA-DT between first and second repeat biopsy		0.002
Always >10 years or negative	ref	
At least once from 10-3 years	1.45 (1.02 - 2.08)	0.039
At least once from 3-0 years (but continued active surveillance)	2.02 (1.36 - 3.00)	<0.001

PSA-DT: PSA doubling time

CI: confidence interval

HR: hazard ratio

life, one might argue that the small portion of men that might benefit from yearly biopsies does not outweigh the additional burden and its possible reduction of quality of life. Especially in men with a slow rising PSA (PSA-DT negative of >10 years) the risk of upgrading was low, which could trigger some patients and their physicians to switch to a watchful waiting strategy, avoiding further biopsies.

Men with a PSA-DT of 3-10 year do seem to have a higher risk of having higher grade and extent of disease. In year 2 and 3 approximately 30% of men with a biopsy due to a PSA-DT of 3-10 year were upgraded (Gleason >6 or >2 cores positive). Men who ignored the biopsy advise in year 2 or 3 seemed to have a similar rate of upgrading if biopsied in year 4. This indicates that for 10-15% of men ignoring the recommendation to have a repeat biopsy based on PSA kinetics, upgrading is delayed by 1-2 year. Despite the increased risk, which was published before [1, 13], many men do not have yearly biopsies. It seems important that during theoretical design of active surveillance follow-up schedules, practical adoption and compliance should not be disregarded. Instead we need to develop follow-up schedules that are acceptable to those who follow it. Less harmful ways of monitoring tumor progression, such as MRI [14], might be incorporated in the protocol design to improve compliance. In the PRIAS study we initiated a side study to investigate if replacing yearly biopsies in men with fast rising PSA by MRI with targeted biopsies in case of visible tumor progression could substantially reduce the amount of biopsies (protocol

available on www.prias-project.org). Even if such an approach will delay active treatment for some, the reduced strains of follow-up might outweigh the harms.

Of the 4 protocol based recommendations to discontinue active surveillance, a PSA-DT for 0-3 year occurred most frequently. At the same time this was the recommendation most often ignored. More than 70% of men (or perhaps more likely their physicians) did not comply with the recommendation to discontinue active surveillance. More men presented with a Gleason score >6 in this group as compare to men not having a PSA-DT of 0-3 years. The higher risk of upgrading remained after correction for other variables. This correlation between PSA-DT and biopsy outcome was reported before in the PRIAS study [1, 13], but not in another study [15]. Differences could be due to variances in study population or to the relatively small numbers in the later study [15]. If looked at the outcome on radical prostatectomy of men who did discontinue active surveillance, 29% of men who discontinued due to a low PSA-DT only had unfavorable outcomes (defined as Gleason $\geq 4+3$ or cT3-4 disease)[16]. Low PSA-DT was also found to be a strong predictor of biochemical recurrence after radical treatment [6]. However, despite these higher risks, a substantial part of men still had favorable disease characteristics. As many men might thus be excluded from active surveillance without having true unfavorable disease, and many men and their physicians did not follow the advice to discontinue, the recommendation to discontinue active surveillance if PSA-DT is 0-3 years was removed from the PRIAS protocol as of 2015 (see figure 1 for new follow-up schedule). Instead more frequent (yearly) repeated biopsies, preferably sampling the anterior transition zone, are advised as with a PSA-DT of 3-10 years. If available and MRI with targeted biopsies could be done to rule out large anterior tumors in men with fast rising PSA.

Another recommendation to discontinue active surveillance that occurred frequently and was sometimes ignored was >2 cores positive for prostate cancer. These men had higher rates of Gleason score >6 at repeat biopsy than men with only 1 or 2 cores positive. However, a previous analysis indicated that a substantial part of men had Gleason 6 prostate cancer if subsequently treated with radical prostatectomy [16]. As metastasis in men with true Gleason 6 disease, irrespective of tumor volume, seems very rare [17, 18], the number of cores positive for prostate cancer might currently only function as a surrogate for higher grade disease. If targeted biopsies could (partially) eliminate this undergrading problem, determining the extent of the tumor might become obsolete in the future. In the PRIAS MRI side study the number of cores is therefore omitted as a criterion to recommend active treatment (protocol available on www.prias-project.org).

Outcome in the current analysis was defined as the outcome on repeat biopsy. Although this allows for a comparison of compliers and non-compliers the effect on more definitive outcomes (e.g. prostate specific death) could currently not be assessed, as no such events were reported yet. Longer follow-up is warranted to assess the effect of non-compliance on definitive outcomes.

CONCLUSION

Some men with low risk prostate cancer on active surveillance do not comply with the schedule for follow-up and recommendations to switch to active treatment. Repeat biopsies, especially yearly biopsies in men with fast rising PSA, are often ignored, as is the recommendation to discontinue active surveillance due to a very fast rising PSA. Although these men are at increased risk of having higher Gleason scores on repeat biopsy, the majority still presents favorable tumor characteristics. A fast rising PSA should therefore not be a recommendation to advice active treatment, but should rather serve as a criterion for stricter follow-up. In addition, reducing the amount of yearly biopsies might increase the amount of men complying with the active surveillance protocol.

REFERENCES

1. Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, et al. Active Surveillance for Low-Risk Prostate Cancer Worldwide: The PRIAS Study. *Eur Urol.* 2012;63:597-603.
2. Tosoian JJ, Trock BJ, Landis P, Feng Z, Epstein JI, Partin AW, et al. Active surveillance program for prostate cancer: an update of the Johns Hopkins experience. *J Clin Oncol.* 2011;29:2185-90.
3. Soloway MS, Soloway CT, Eldefrawy A, Acosta K, Kava B, Manoharan M. Careful selection and close monitoring of low-risk prostate cancer patients on active surveillance minimizes the need for treatment. *Eur Urol.* 2010;58:831-5.
4. Selvadurai ED, Singhera M, Thomas K, Mohammed K, Woode-Amisshah R, Horwich A, et al. Medium-term Outcomes of Active Surveillance for Localised Prostate Cancer. *Eur Urol.* 2013.
5. Cooperberg MR, Cowan JE, Hilton JF, Reese AC, Zaid HB, Porten SP, et al. Outcomes of active surveillance for men with intermediate-risk prostate cancer. *J Clin Oncol.* 2011;29:228-34.
6. Klotz L, Vesprini D, Sethukavalan P, Jethava V, Zhang L, Jain S, et al. Long-Term Follow-Up of a Large Active Surveillance Cohort of Patients With Prostate Cancer. *J Clin Oncol.* 2014.
7. van den Bergh RC, Roemeling S, Roobol MJ, Roobol W, Schroder FH, Bangma CH. Prospective validation of active surveillance in prostate cancer: the PRIAS study. *Eur Urol.* 2007;52:1560-3.
8. Bokhorst L, Alberts A, Kakehi Y, Rannikko A, Pickles T, Valdagni R, et al. Frequency of PSA testing in men on active surveillance for prostate cancer. *Journal of Urology.* 2015;193:e755.
9. Loeb S, Vellekoop A, Ahmed HU, Catto J, Emberton M, Nam R, et al. Systematic review of complications of prostate biopsy. *Eur Urol.* 2013;64:876-92.
10. Mkinen T, Auvinen A, Hakama M, Stenman UH, Tammela TL. Acceptability and complications of prostate biopsy in population-based PSA screening versus routine clinical practice: a prospective, controlled study. *Urology.* 2002;60:846-50.
11. Welty CJ, Cowan JE, Nguyen H, Shinohara K, Perez N, Greene KL, et al. Extended Follow-Up and Risk Factors for Disease Reclassification from a Large Active Surveillance Cohort for Localized Prostate Cancer. *The Journal of urology.* 2014.
12. Kates M, Tosoian JJ, Trock BJ, Feng Z, Carter HB, Partin AW. Indications for Intervention During Active Surveillance of Prostate Cancer: A Comparison of the Johns Hopkins and PRIAS Protocols. *BJU international.* 2014.
13. Bul M, van den Bergh RC, Rannikko A, Valdagni R, Pickles T, Bangma CH, et al. Predictors of unfavourable repeat biopsy results in men participating in a prospective active surveillance program. *Eur Urol.* 2012;61:370-7.
14. Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz L, et al. Magnetic Resonance Imaging in Active Surveillance of Prostate Cancer: A Systematic Review. *Eur Urol.* 2014;67:627-36.
15. Ross AE, Loeb S, Landis P, Partin AW, Epstein JI, Kettermann A, et al. Prostate-specific antigen kinetics during follow-up are an unreliable trigger for intervention in a prostate cancer surveillance program. *J Clin Oncol.* 2010;28:2810-6.
16. Bul M, Zhu X, Rannikko A, Staerman F, Valdagni R, Pickles T, et al. Radical Prostatectomy for Low-Risk Prostate Cancer Following Initial Active Surveillance: Results From a Prospective Observational Study. *Eur Urol.* 2012;62:195-200.
17. Kweldam CF, Wildhagen MF, Bangma CH, van Leenders GJ. Disease-specific death and metastasis do not occur in patients with Gleason score ≤ 6 on radical prostatectomy. *BJU international.* 2014.
18. Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) ≤ 6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol.* 2012;36:1346-52.

Chapter 7

Complications after prostate biopsies in men on active surveillance and its effect on receiving further biopsies in the Prostate cancer Research International: Active Surveillance (PRIAS) study.

Leonard P. Bokhorst^a, Inari Lepistö^b, Yoshiyuki Kakehi^c, Chris H. Bangma^a, Tom Pickles^d, Riccardo Valdagni^e, Arnout R. Alberts^a, Axel Semjonow^f, Petra Strölin^g, Manuel F. Montesino^h, Viktor Bergeⁱ, Monique J. Roobol^a, Antti Rannikko^b.

^a Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands.

^b Department of Urology, Helsinki University Central Hospital, Helsinki, Finland

^c Department of Urology, Kagawa University Faculty of Medicine, Kagawa, Japan

^d Department of Radiation Oncology, British Columbia Cancer Agency, Vancouver, Canada

^e Prostate Cancer Program and Radiation Oncology 1, Fondazione IRCSS Istituto Nazionale dei Tumori, Milan, Italy

^f Department of Urology, Prostate Center, University Hospital Muenster, Muenster, Germany

^g Department of Urology, Martini Klinik, Hamburg, Germany

^h Department of Urology, Hospital Virgen del Camino, Pamplona, Spain

ⁱ Department of Urology, Oslo university hospital, Oslo, Norway

ABSTRACT

Objective: To study the risk of serial prostate biopsies on complications in men on active surveillance and determine the effect of complications on receiving further biopsies.

Materials and methods: In the global Prostate cancer Research International: Active Surveillance (PRIAS) study men are prospectively followed on active surveillance and repeat prostate biopsies are scheduled 1, 4, and 7 years after the diagnostic biopsy, or once yearly if PSA-doubling time (PSA-DT) is <10 years. Data on complications after biopsy, including infection, hematuria, hematospermia, and pain, were retrospectively collected for all biopsies done during follow-up in men from several large participating centers. Generalized estimating equations were used to test predictors of infection after biopsy. Competing risk analysis was used to compare the rates of men receiving further biopsies between men with and without previous complications.

Results: In total 2184 biopsies were done in 1164 men. Infection was reported after 55 biopsies (2.5%), and one in five men reported any form of complication. At multivariable analysis the number of previous biopsies was not a significant predictor of infection (OR 1.04, 95% CI 0.76-1.43). The only significant predictor for infection was the type of prophylaxis used. Of all men with a complication at the diagnostic or first repeat biopsy 21% did not have a repeat biopsy at the time a repeat biopsy was scheduled according to protocol, versus 12% for men without a previous biopsy complication.

Conclusion: In our cohort of men on active surveillance we found no evidence that repeated prostate biopsies in itself pose a risk of infection. Complications after biopsy were however not uncommon and after a complication men were less likely to have further biopsies. We should aim to safely reduce the amount of repeat biopsies in men on active surveillance.

INTRODUCTION

Active surveillance aims to reduce overtreatment, and its side effects, by initially monitoring men with low-risk prostate cancer instead of immediate invasive treatment (e.g. radical prostatectomy). If signs of more aggressive disease appear, men on active surveillance are offered treatment with curative intent. Monitoring usually consists of regular PSA tests, digital rectal examinations, and repeated prostate biopsies.

We recently showed that some men on active surveillance are reluctant to undergo repeat biopsies [1]. Some of the reasons stated were previous complications of the prostate biopsies. With every prostate biopsy men are at risk of harboring complications such as an infection, hematuria, or pain. It was suggested that the risk of some complications, such as infection, increases with the number of previous biopsies taken in men on active surveillance [2, 3].

We therefore set out to study the risk of complications after serial prostate biopsies in men on active surveillance in the Prostate cancer Research International: Active Surveillance (PRIAS) study and the effect of a previous complication on receiving subsequent biopsies.

MATERIALS AND METHODS

The study protocol of the PRIAS study has been described in detail previously [4, 5]. In brief, men with low risk prostate cancer (Gleason score $\leq 3+3$, $\leq cT2c$, PSA ≤ 10 ng/ml, \leq two cores positive for prostate cancer, PSA density ≤ 0.2 ng/ml/ml, and fitness for curative treatment) can be included and prospectively followed on active surveillance. All centers collect data through an online tool (www.prias-project.org). During follow-up regular PSA tests and digital rectal examinations are advised. Repeat biopsies are planned 1, 4, 7, 10, and subsequent every 5 years, after diagnosis. Yearly repeat biopsies are only indicated if PSA doubling time (PSA-DT) is between 0 and 10 years. A minimum number of biopsy cores is advised according to prostate volume (prostate volume < 40 cm³, 8 cores; 40-60cm³, 10 cores; > 60 cm³, 12 cores). Until the end of March 2015, 4749 men were included by more than 100 centers in 17 different countries.

For this analysis several large participating centers from different countries (including approximately 25% of all men included in PRIAS) retrospectively collected data through chart review on complications after biopsy for all participants in their center. This included data on infection (defined as temperature $> 38^{\circ}\text{C}/100.4^{\circ}\text{F}$ within 3 days after biopsy), hematuria (> 3 days), hematospermia, and pain. To minimize the risk of reporting bias due to selective reporting of men with infections, only men were included for whom data on infectious complications on all biopsies during follow-up were entered. Complication rates were reported per subsequent biopsy during follow-up. Generalized estimating equations (GEE) were used to study the effect of number of previous biopsies on infection rates. In GEE all biopsies are used for prediction, while at the same time accounting for possible correlations among multiple biopsies in the same man. Results were

presented before and after correction for age at biopsy, year of biopsy, biopsy route (perineal versus rectal), type of prophylaxis, and an infection on last biopsy. Competing risk analysis was used to compare the percentage of men without a subsequent biopsy between men with and without any previous biopsy complication at the time a repeat biopsy is indicated by the study protocol (competing risk being men who discontinued before their next biopsy). For the analysis SPSS for windows (Version 21.0. Armonk, NY: IBM Corp.) and the survival package of R (R Foundation for Statistical Computing, Vienna, Austria.) were used.

RESULTS

For 1164 men complication data were entered for all biopsies during follow-up and were used for analysis. In total, 2184 biopsies were done in these men during follow-up, 465 men only had 1 biopsy, 464 men had two biopsies, 149 had 3 biopsies, and 86 men had 4 or more biopsies.

Table 1. Characteristics for men with and without an infection after biopsy.

	No infection	Infection	Total
Age at biopsy			
<50	21 (100%)	0 (0%)	21 (100%)
50-59	327 (97%)	10 (3%)	337 (100%)
60-69	1092 (97.3%)	30 (2.7%)	1122 (100%)
70-79	666 (98.1%)	13 (1.9%)	679 (100%)
>=80	23 (92%)	2 (8%)	25 (100%)
Prophylaxis used			
Fluorquinolones	1830 (97.3%)	51 (2.7%)	1881 (100%)
TMP SMX	15 (100%)	0 (0%)	15 (100%)
Other	234 (99.6%)	1 (0.4%)	235 (100%)
Not used	16 (84.2%)	3 (15.8%)	19 (100%)
Missing	34 (100%)	0 (0%)	34 (100%)
Biopsy route			
Rectal	1789 (97.3%)	49 (2.7%)	1838 (100%)
Perineal	228 (97.9%)	5 (2.1%)	233 (100%)
Missing	112 (99.1%)	1 (0.9%)	113 (100%)
Country			
European	1327 (97.4%)	36 (2.6%)	1363 (100%)
Japan	592 (97.9%)	13 (2.1%)	605 (100%)
Australia/New Zealand	39 (95.1%)	2 (4.9%)	41 (100%)
Canada	171 (97.7%)	4 (2.3%)	175 (100%)
Total	2129 (97.5%)	55 (2.5%)	2184 (100%)

In total there were 55 infections after biopsy (2.5%) in 54 men. The majority of biopsies were rectal biopsies (84%) and Fluorquinolones were the most common prophylaxis used (86%). Infection rates were 2.7% for rectal biopsies and 2.1% for perineal biopsies. Nineteen men did not receive prophylaxis, 3 of whom had an infection (16%). There were no large differences in infection rates between countries or the age at which the biopsy was taken (table 1).

Infection rates slightly increased after the diagnostic biopsy (2.3%), to 2.6% and 3.8% at the first and second repeat biopsy respectively, after which it decreased to 1.2% for ≥ 3 repeat biopsies (table 2). Infection rates increase over the course of time from 0% in and before 2006 to 4.8% in 2015 (figure 1). Data on hematuria, hematospermia, and pain after biopsy were missing in many cases (not often reported in the medical records). In men for which the data was available only pain after biopsy seemed to increase with subsequent biopsies from 10% at diagnostic biopsy, to 27% at ≥ 3 repeat biopsies (table 2).

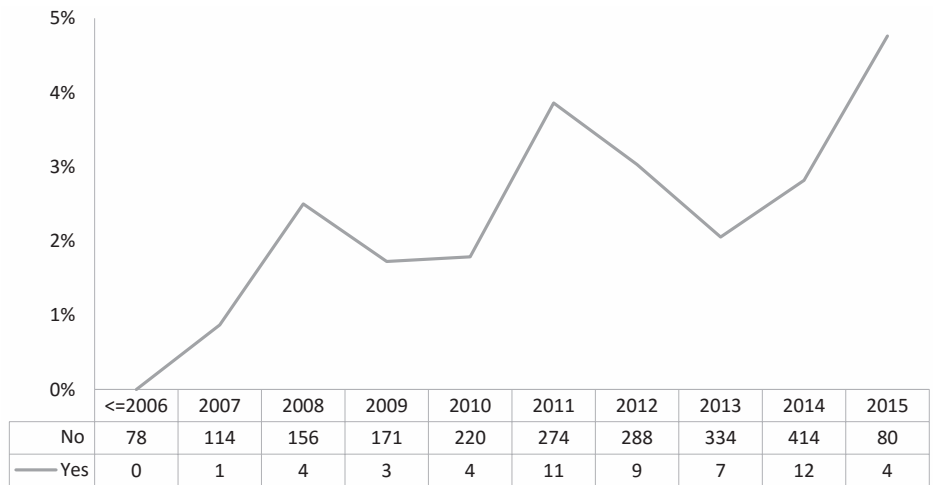


Figure 1. Infection rates per year the biopsy was taken.

In total 273 out of 1164 men (23%) had at least one complication (infection, pain, hematuria, or hematospermia) on the diagnostic biopsy. After 1.5 years of follow-up (standard repeat biopsy is planned 1 year after diagnosis) 17% of men with a complication at the diagnostic biopsy did not have a repeat biopsy yet versus only 10% for men without a biopsy complication. After the first repeat biopsy 21% of men with any complication on the diagnostic or first repeat biopsy did not go on to have a biopsy after 3.5 years of follow-up (second standard repeat biopsy is planned 3 years after first repeat biopsy) versus 12% of men without a biopsy complication. A sub analysis comparing only men with an previous infection (instead of any biopsy complication) did not show a higher percentage of men without a subsequent biopsy (9% did not receive a biopsy after 1.5 years of follow-up versus 11% of men without an infection).

Table 2. Complications per subsequent biopsy on active surveillance.

	Diagnostic biopsy	Repeat biopsy 1	Repeat biopsy 2	Repeat biopsy >=3	Total
Infection					
Yes	27 (2.3%)	18 (2.6%)	9 (3.8%)	1 (1.2%)	55 (2.5%)
% of men with filled data (yes/no)	2.3%	2.6%	3.8%	1.2%	2.5%
No	1137 (97.7%)	681 (97.4%)	226 (96.2%)	85 (98.8%)	2129 (97.5%)
Missing	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Hematuria					
Yes	148 (12.7%)	66 (9.4%)	14 (6%)	8 (9.3%)	236 (10.8%)
% of men with filled data (yes/no)	21.0%	20.3%	22.6%	42.1%	21.3%
No	556 (47.8%)	259 (37.1%)	48 (20.4%)	11 (12.8%)	874 (40%)
Missing	460 (39.5%)	374 (53.5%)	173 (73.6%)	67 (77.9%)	1074 (49.2%)
Hematospermia					
Yes	152 (13.1%)	50 (7.2%)	13 (5.5%)	3 (3.5%)	218 (10%)
% of men with filled data (yes/no)	22.6%	15.9%	22.4%	18.8%	20.5%
No	521 (44.8%)	264 (37.8%)	45 (19.1%)	13 (15.1%)	843 (38.6%)
Missing	491 (42.2%)	385 (55.1%)	177 (75.3%)	70 (81.4%)	1123 (51.4%)
Pain					
Yes	70 (6%)	39 (5.6%)	11 (4.7%)	3 (3.5%)	123 (5.6%)
% of men with filled data (yes/no)	10.1%	13.0%	20.0%	27.3%	11.6%
No	620 (53.3%)	262 (37.5%)	44 (18.7%)	8 (9.3%)	934 (42.8%)
Missing	474 (40.7%)	398 (56.9%)	180 (76.6%)	75 (87.2%)	1127 (51.6%)
Total	1164 (100%)	699 (100%)	235 (100%)	86 (100%)	2184 (100%)

On univariable analysis a small, non-significant, increase in the odds of infection was seen for every additional previous biopsy (odds ratio (OR) 1.08, 95% confidence interval (CI) 0.79-1.48). On multivariable analysis the effect of the number of previous biopsies on infection diminished (OR 1.04, 95% CI 0.76-1.43). The only significant predictor for infection after biopsy was the type of prophylaxis used (table 3).

DISCUSSION

In the current analysis we assessed complication rates after repeated prostate biopsies in men on active surveillance. We found no evidence that repeated prostate biopsies in itself pose an increased risk of infection. However, approximately one in five men report some sort of complica-

Table 3. Predictors of infection after prostate biopsy.

	Univariable		Multivariable	
	OR (95% CI)	p-value	OR (95% CI)	p-value
Age at biopsy (1 year increase)	0.99 (0.95-1.03)	.713	0.99 (0.95-1.03)	.663
Year of biopsy (1 year increase)	1.11 (1.00-1.24)	.053	1.12 (1.00-1.26)	.051
Number of previous biopsies	1.08 (0.81-1.46)	.589	1.04 (0.77-1.4)	.819
Perineal biopsy route (vs rectal)	0.83 (0.28-2.44)	.740	1.02 (0.36-2.91)	.975
Infection on last biopsy	2.17 (0.32-14.86)	.429	2.06 (0.24-17.65)	.511
Antibiotic prophylaxis used				
Fluorquinolones	1		1	
TMP SMX/Other	0.13 (0.02-0.88)	.037	0.12 (0.02-0.70)	.018
No prophylaxis used	6.73 (1.86-24.35)	.004	5.59 (1.40-22.23)	.015

tion after biopsy, with only pain seeming to increase with subsequent biopsies. Men with a previous complication were less likely to undergo a repeated biopsies while under active surveillance for prostate cancer.

In total infection was reported after 2.5% of biopsies in this cohort of men on active surveillance for prostate cancer. This reported percentage is somewhat lower than previous reports on infectious complications after prostate biopsy [6-9]. As with previous reports there was however an increase, although not statistically significant, in the rate of infection with increasing year of biopsy [6, 9, 10]. Increasing bacterial resistance to Fluorquinolones is seen as the most important reason for this increase in infection rates. In our cohort Fluorquinolones were the most common prophylaxis (86%) used. The rate of infection was higher in these men as compared to men receiving other antibiotic prophylaxis. As fluorquinolones are one of the standard recommended prophylaxis in for instance the EAU guideline [11], it could be that the other prophylaxis used were targeted based on a rectal swab. This could explain the lower infection rates, in addition to differences in resistance patterns. In men on active surveillance, who receive repeated biopsies and are therefore at increased risk of an infection, rectal swabs might be a method to reduce the risk of infection [8, 12].

Another method sometimes proposed to reduce the amount of infections after biopsy is perineal biopsies instead of rectal biopsies [13]. Although initially the infection rate for perineal biopsies were lower in our study, after correction for the type of prophylaxis used the protective effect disappeared. A recent review and meta-analysis was also not able to find a difference in complication rates between the rectal and perineal approach [14]. The lower infection rates in men receiving perineal biopsy thus seem to be a result of different antibiotic prophylaxis used instead of a different biopsy route. This suggests that there might be another pathway for prostatic infection after (perineal) biopsy, other than an infection because of pathogens passing directly through the rectal wall as result of the biopsy needle passing. Instead prostatic infections after biopsy might be caused by direct inoculation from within the urinary tract. It was

recently discovered that, opposite of what was commonly thought, urine is not sterile [15-17]. The prostate is in direct contact with urine, and consists of multiple ducts that might provide a sheltered environment for bacteria. Prostatic biopsies may then facilitate tissue infection from bacteria that are already present in the prostate at the time of the biopsy. An indirect observation supporting this hypothesis is the very high infection rate (27%) after prostate biopsy seen in men with asymptomatic bacteriuria [18]. Further research is needed to investigate the pathways causing infections after rectal and perineal prostate biopsies.

As men on active surveillance receive repeated prostate biopsies they are at increased absolute risk of an infection as a result of a biopsy. However, repeated biopsies in itself were not a risk factor for infection in our cohort. This was also shown in a large database study [19]. Furthermore, a recent study showed no increase in fluorquinolone resistance in rectal swabs before repeated biopsy in men on active surveillance [12]. This is in contrast with a previous study which did show an increased risk of infection with every subsequent biopsy in men on active surveillance [2]. This difference might be due to the low number of events in that study or because no corrections were made for confounding factors. Also, lack of data on the number of prediagnostic biopsy sessions (with benign histology) may confound the analysis. Additionally, it was shown that although previous fluorquinolones use was a risk factor for infection after biopsy it was only so if used within the previous 3 months [20]. As in the PRIAS study repeat biopsies are done 1, 4, 7 and 10 year after diagnosis the effect of previous prophylaxis might be passed.

Despite the lack of increase in infectious complications with subsequent biopsies, there was an increase in pain reported after repeated biopsies. Furthermore, men with any previous complication after biopsy (either infection, hematuria, hematospermia, or pain) were less likely to have a repeat biopsy at the time it is normally scheduled. As there was no difference in the percentage of men discontinuing active surveillance (both for protocol based reasons and other reasons) between these groups (data not shown), men with a previous complication were simply more likely to continue active surveillance without having a repeat biopsy at the designated time. We observed that the percentage of men not receiving further biopsies was even higher in a subgroup of men who apart from a previous complication on prostate biopsy presented with a stable PSA during follow-up (defined as a PSA doubling time >10 years or negative)(50% of men with a previous complication did not have a biopsy at the allocated time in this subgroup versus 14% for men without a previous complication, data not shown). This is in line with a previous report in the PRIAS study that showed that men were becoming less likely to have a repeated biopsies over time during active surveillance [1]. However, as also shown in this analysis [1], upgrading rates on biopsies later during follow-up, especially in men with stable PSA, were relatively low. It could be questioned if the risk of missing these biopsies might balance against the discomfort and risk of complications. We must therefore find ways to safely reduce the amount of biopsies (e.g. by better risk-stratifying men at increased risk of disease progression/ reclassification) or even replace the biopsies for instance with MRI [21]. Reducing the amount of biopsies at times it is relatively safe might increase the compliance at times when there is a truly higher risk of upgrading.

This analysis has some limitations. First, as this was a retrospective analysis there might be an underreporting of complications. It was noted that especially data on hematuria, hematospermia and pain were often missing from the medical records. In addition, very few data on hospitalization and resistance patterns in men with an infection were available and this was therefore not reported. Despite the possibility of underreporting there was no reason to believe that there was selective underreporting for one of the biopsy session, making an effect on the conclusions drawn highly unlikely. Second, data on comorbidity (Charlson score) was only available at the time of the diagnostic biopsy and could therefore not be added to the main analysis. When only looking at the diagnostic biopsy there was a small, statistically non-significant ($p=0.48$), lower infection rate in men with any comorbidity (1.8% had an infection) versus men without any comorbidity (2.5% had an infection)(data not shown). Third, although the overall number of events was low for the multivariable analysis, most statistically non-significant OR's were close to one indicating little predictive effect. Last, no correction could be made for the number of biopsies before the start of active surveillance.

In conclusion, we did not find evidence in this analysis that repeated biopsies in itself increase the risk of infection in men on active surveillance for prostate cancer. However, in one in five men some sort of complication was reported after biopsy. Men with a previous complication were less likely to undergo a subsequent repeat prostate biopsy as indicated by the active surveillance protocol. We should therefore aim to safely reduce the amount of biopsies in men on active surveillance to reduce the absolute risk of complications and increase compliance with active surveillance protocols.

REFERENCES

1. Bokhorst LP, Alberts AR, Rannikko A, Valdagni R, Pickles T, Kakehi Y, et al. Compliance Rates with the Prostate Cancer Research International Active Surveillance (PRIAS) Protocol and Disease Reclassification in Noncompliers. *Eur Urol.* 2015;68:814-21.
2. Ehdai B, Vertosick E, Spaliviero M, Giallo-Uvino A, Taur Y, O'Sullivan M, et al. The Impact of Repeat Biopsies on Infectious Complications in Men with Prostate Cancer on Active Surveillance. *The Journal of urology.* 2013;191:660-4.
3. Glass AS, Hilton JF, Cowan JE, Washington SL, Carroll PR. Serial prostate biopsy and risk of lower urinary tract symptoms: results from a large, single-institution active surveillance cohort. *Urology.* 2014;83:33-8.
4. van den Bergh RC, Roemeling S, Roobol MJ, Roobol W, Schroder FH, Bangma CH. Prospective validation of active surveillance in prostate cancer: the PRIAS study. *Eur Urol.* 2007;52:1560-3.
5. Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, et al. Active Surveillance for Low-Risk Prostate Cancer Worldwide: The PRIAS Study. *Eur Urol.* 2012;63:597-603.
6. Carignan A, Roussy JF, Lapointe V, Valiquette L, Sabbagh R, Pepin J. Increasing risk of infectious complications after transrectal ultrasound-guided prostate biopsies: time to reassess antimicrobial prophylaxis? *Eur Urol.* 2012;62:453-9.
7. Wagenlehner FM, van Oostrum E, Tenke P, Tandogdu Z, Cek M, Grabe M, et al. Infective complications after prostate biopsy: outcome of the Global Prevalence Study of Infections in Urology (GPIU) 2010 and 2011, a prospective multinational multicentre prostate biopsy study. *Eur Urol.* 2013;63:521-7.
8. Loeb S, Vellekoop A, Ahmed HU, Catto J, Emberton M, Nam R, et al. Systematic review of complications of prostate biopsy. *Eur Urol.* 2013;64:876-92.
9. Loeb S, van den Heuvel S, Zhu X, Bangma CH, Schroder FH, Roobol MJ. Infectious Complications and Hospital Admissions After Prostate Biopsy in a European Randomized Trial. *Eur Urol.* 2012;61:1110-4.
10. Siddiqui MM, Rais-Bahrami S, Turkbey B, George AK, Rothwax J, Shakir N, et al. Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. *JAMA.* 2015;313:390-7.
11. Grabe M, Bartoletti R, Bjerklund-Johansen TE, Cai T, Cek M, Koves B, et al. EAU Guidelines on Urological Infections. www.uroweb.org. 2015.
12. Loeb S, Carter HB, Berndt SI, Ricker W, Schaeffer EM. Is repeat prostate biopsy associated with a greater risk of hospitalization? Data from SEER-Medicare. *The Journal of urology.* 2013;189:867-70.
13. Grummet JP, Weerakoon M, Huang S, Lawrentschuk N, Frydenberg M, Moon DA, et al. Sepsis and 'superbugs': should we favour the transperineal over the transrectal approach for prostate biopsy? *BJU international.* 2014;114:384-8.
14. Shen PF, Zhu YC, Wei WR, Li YZ, Yang J, Li YT, et al. The results of transperineal versus transrectal prostate biopsy: a systematic review and meta-analysis. *Asian J Androl.* 2012;14:310-5.
15. Wolfe AJ, Brubaker L. "Sterile Urine" and the Presence of Bacteria. *Eur Urol.* 2015;68:173-4.
16. Kogan MI, Naboka YL, Ibishev KS, Gudima IA, Naber KG. Human Urine Is Not Sterile - Shift of Paradigm. *Urol Int.* 2015.
17. Hilt EE, McKinley K, Pearce MM, Rosenfeld AB, Zilliox MJ, Mueller ER, et al. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. *J Clin Microbiol.* 2014;52:871-6.
18. Lindstedt S, Lindstrom U, Ljunggren E, Wullt B, Grabe M. Single-dose antibiotic prophylaxis in core prostate biopsy: Impact of timing and identification of risk factors. *Eur Urol.* 2006;50:832-7.

-
19. Loeb S, Carter HB, Berndt SI, Ricker W, Schaeffer EM. Complications after prostate biopsy: data from SEER-Medicare. *The Journal of urology*. 2011;186:1830-4.
 20. Taylor S, Margolick J, Abughosh Z, Goldenberg SL, Lange D, Bowie WR, et al. Ciprofloxacin resistance in the faecal carriage of patients undergoing transrectal ultrasound guided prostate biopsy. *BJU international*. 2013;111:946-53.
 21. Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz L, et al. Magnetic Resonance Imaging in Active Surveillance of Prostate Cancer: A Systematic Review. *Eur Urol*. 2014;67:627-36.

Chapter 8

Effect of pathologic revision and Ki67 and ERG immunohistochemistry on predicting radical prostatectomy outcome in men initially on active surveillance

Leonard P. Bokhorst^a, Monique J. Roobol^a, Chris H. Bangma^a,
Geert J.L.H. van Leenders^b

^a Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands.

^b Department of Pathology, Erasmus University Medical Center, Rotterdam, The Netherlands

(Submitted for publication)

ABSTRACT

Background: Discordance between biopsy and radical prostatectomy tumor grade are not uncommon, but are unwanted in men on active surveillance for prostate cancer.

Objective: To investigate if pathologic biopsy reevaluation and implementation of immunohistochemical biomarkers could improve prediction of radical prostatectomy outcome in men initially on active surveillance.

Methods: Biopsy specimens from diagnosis until switching to radical prostatectomy in men initially on active surveillance in the Dutch part of the Prostate cancer Research International Active Surveillance (PRIAS) study were collected and revised by a single pathologist. Original and revised biopsy Gleason score were compared and correlated with radical prostatectomy Gleason score. Biopsy specimens were immunohistochemically stained for Ki67 and ERG. Predictive ability of clinical characteristics and biomarkers on Gleason ≥ 7 or $\geq pT3$ on radical prostatectomy was tested using logistic regression and ROC curve analysis.

Results: A total of 150 biopsies in 95 men were revised. In 13% of diagnostic or second-to-last biopsies and 20% of the last biopsies on active surveillance revision of Gleason score resulted in change of recommendation (i.e. active treatment or active surveillance). Concordance with Gleason on radical prostatectomy was however similar for both the revised and original Gleason on biopsy. Ki67 and ERG were not statistically significant predictors of Gleason ≥ 7 or $\geq pT3$ on radical prostatectomy.

Conclusions: Although interobserver differences in pathology reporting on biopsy could result in a change of management strategy in approximately 13-20% of men on active surveillance, both pathological revision and tested biomarkers (Ki67 and ERG) did not improve prediction of outcome on radical prostatectomy. Undersampling of most aggressive tumor remains the main focus in order to increase accurate grading at time of treatment decision making.

INTRODUCTION

Discrepancies in prostate cancer grade at biopsy and radical prostatectomy specimen are not uncommon (1, 2). This is especially an issue in the context of active surveillance, as underestimation of true tumor characteristics might lead to an unwanted delay, and overestimation to an unnecessary switch to active treatment.

While sampling error is considered the most important cause for grading discrepancy, inter-observer variability and guideline differences for grading at biopsy and prostatectomy might also play a role (3). Pathologic biopsy reevaluation might change outcome and increase grading concordance. Secondly, immunohistochemical staining of biopsy specimens with markers known to correlate with more aggressive characteristics could help better to define the true tumor state (4).

The objective of this study is to investigate the effect of pathologic biopsy reevaluation and implementation of immunohistochemical biomarkers on predicting radical prostatectomy outcome in men initially on active surveillance for prostate cancer.

METHODS

For this study we selected all Dutch men initially on active surveillance in the Prostate cancer Research International Active Surveillance (PRIAS) study who switched to radical prostatectomy before July 2014. The PRIAS study was initiated in 2006 and includes men with low risk prostate cancer (Gleason score $\leq 3+3$, $\leq cT2c$, PSA ≤ 10 ng/ml, ≤ 2 cores positive for prostate cancer, PSA density ≤ 0.2 ng/ml/cm³, and fitness for curative treatment) on active surveillance (5). Criteria used to recommend a switch to active treatment were Gleason score >6 , >2 positive biopsy cores, a PSA-doubling time of 0-3 years, and $>cT2$.

We retrieved all available biopsy specimens from diagnosis until end of active surveillance as well as pathology reports from all biopsy and radical prostatectomy specimens from the participating hospitals. All biopsy specimens were re-evaluated for Gleason score (ISUP 2014 (6)), number of positive biopsies cores, and mm and percentage tumor per core, by a single pathologist (GvL) blinded for the initial biopsy report.

Immunohistochemical staining

Per biopsy session (both diagnostic biopsy and all repeat biopsies) one core, with either the highest Gleason score or the largest tumor size in mm, was selected for immunostaining with ERG and Ki67.

Sections were cut (4 μ m) and deparaffinized in xylene and dehydrated in ethanol. Endogenous peroxidase was blocked by 0.3% hydrogen peroxide/phosphate buffered saline (PBS) for 20 minutes. Slides were placed in a microwave in pH9 Tris (hydroxymethyl)aminomethane-EDTA buffer for 15 minutes and incubated with primary antibody overnight at 4 degrees Celsius. Secondary

antibody (antimouse, Envision kit) was applied, followed by chromogenic visualization (Envision, DAKO kit), and counterstaining with hematoxylin. ERG was scored as either positive or negative, representative for ERG-gen fusion (7). Ki67 was scored for the percentage of positive tumor cells.

Statistical analysis

Concordance between initial and revised number of positive biopsy cores and the Gleason score were analyzed, and compared to the Gleason score on radical prostatectomy. Univariable logistic regression analysis was used to assess the predictive value of immunohistochemical biomarker and other tumor characteristics (PSA at the time of the biopsy, biopsy Gleason (≤ 6 or ≥ 7), number of positive biopsy cores (≤ 2 or > 2), and $> 50\%$ single core involvement (yes or no)) on unfavorable outcome (Gleason score ≥ 7 or $\geq pT3$) on radical prostatectomy. Multivariable analysis was not conducted due to the low number of events. ROC curves were used to test clinical performance. All analysis were done separately for the diagnostic or second-to-last biopsy, and for the last biopsy before switching to radical prostatectomy. Analysis were done using SPSS for windows (Version 23.0. Armonk, NY: IBM Corp.).

RESULTS

Patients characteristics

A total of 150 biopsy sessions of 95 participants who switched to radical prostatectomy after initial active surveillance were received and available for analysis. Of these, 71 biopsies were diagnostic biopsies, 11 biopsies were repeat biopsies not being the last biopsy, and 89 biopsies were last biopsy before switching to radical prostatectomy (21 men only had 1 biopsy during active surveillance, in these men the diagnostic and last biopsy are the same). The reason for discontinuation of active surveillance was protocol based progression of disease in 63 (66%) men and non-protocol based in 32 (34%) men (i.e. anxiety, patient request, or other). At diagnosis median age was 64 years (IQR 60-67, median PSA 6.1 ng/ml (IQR 5.0-7.4), and all men had Gleason 6 on biopsy.

Pathologic review

Six out of the 82 diagnostic or second-to-last biopsies, and 4 out of the 89 last biopsies could not be revised as not all representative slides were received, leaving a total of 76 and 85 biopsy sessions for review, respectively.

On diagnostic or second-to-last biopsy the original Gleason score was ≤ 6 in all men while the revised score was ≥ 7 in 10 men (13%) (table 1a), and in 8 men (11%) originally ≤ 2 cores were scored positive while revision reported > 2 cores to be positive (table 2a). Based on pathologic revision of Gleason score and number of positive biopsy cores, a total of 17 men (22%) would either not fulfill the criteria for inclusion or should have switched off active surveillance at an earlier time-point.

Table 1a. Original versus revised Gleason score of diagnostic and second-to-last biopsy on active surveillance before switching to radical prostatectomy.

		Revised Gleason score					Total
		No PCa	3+3	3+4	4+3	4+4	
Original Gleason score	No PCa	4	1	-	-	-	5
	2+2	-	1	-	-	-	1
	2+3	1	3	-	-	-	4
	3+2	-	1	-	-	-	1
	3+3	2	51	8	1	1	63
	Not reported	-	1	1	-	-	2
	Total	7	58	9	1	1	76

* Of the 82 diagnostic of second-to-last biopsies: 6 could not be assessed due to too little remaining biopsy tissue for analysis, 21 were also the last biopsy.

Table 1b. Original versus revised Gleason score of last biopsy on active surveillance before switching to radical prostatectomy.

		Revised Gleason score							Total
		No PCa	3+3	3+4	4+3	4+4	5+3	5+5	
Original Gleason score	No PCa	3	-	-	-	-	-	-	3
	2+3	-	1	-	-	-	-	-	1
	3+3	2	46	9	1	-	-	-	58
	3+4	-	8	9	-	-	-	-	17
	4+3	-	-	2	-	-	-	1	3
	4+4	-	-	-	1	1	-	-	2
	5+4	-	-	-	-	-	1	-	1
	Total	5	55	20	2	1	1	1	85

* Of the 89 last biopsies: 4 could not be assessed due to too little remaining biopsy tissue for analysis, 21 were also the diagnostic biopsy.

On the last biopsy before switching to radical prostatectomy, 10 men (12%) had an original Gleason score ≤ 6 while the revised score was ≥ 7 , and 8 men (9%) had an original Gleason score =7 with a revised Gleason score ≤ 6 (table 1b). The number of positive cores on last biopsy was initially reported ≤ 2 in 6 men (7%) while revision reported >2 cores positive and in 6 men was initially reported >2 cores positive while revision reported ≤ 2 cores to be positive (table 2b). Based on pathologic revision of Gleason score and number of positive biopsy cores, 6 (7%) men would still be eligible for active surveillance at last biopsy. The number of positive biopsy cores was initially not reported or could not accurately be determined due to tissue fragmentation, on diagnostic or second-to-last biopsy, or last biopsy in 6 and 7 men, respectively (table 2a and 2b). Overall in only 13% of cases all biopsy cores were put in separate containers ensuring accurate determination of total number of positive biopsies.

Table 2a. Original versus revised number of cores positive on diagnostic and second-to-last biopsy on active surveillance before switching to radical prostatectomy.

		Revised number of cores positive						Total
		0	1	2	3	4	Cannot be determined*	
Original number of cores positive	0	4	1	-	-	-	-	5
	1	2	33	4	2	1	-	42
	2	-	-	16	5	-	1	22
	3	-	-	-	1	-	1	2
	4	-	-	-	-	1	-	1
	Not in report	-	1	-	1	-	2	4
	Total	6	35	20	9	2	4	76

* Cannot be determined accurately due to multiple biopsies on one slide.

** Of the 82 diagnostic of second-to-last biopsies: 6 could not be assessed due to too little remaining biopsy tissue for analysis, 21 were also the last biopsy.

Table 2b. Original versus revised number of cores positive on last biopsy on active surveillance before switching to radical prostatectomy.

		Revised number of cores positive						Total	
		0	1	2	3	4	>4		Cannot be determined*
Original number of cores positive	0	3	-	-	-	-	-	3	
	1	-	19	-	-	-	-	19	
	2	1	-	12	4	2	-	1	20
	3	-	-	6	7	5	1	-	19
	4	-	-	-	3	7	1	-	11
	>4	-	-	-	-	-	7	1	8
	Not in report	-	3	1	1	-	-	-	5
Total	4	22	19	15	14	9	2	85	

* Cannot be determined accurately due to multiple biopsies on one slide.

** Of the 89 last biopsies: 4 could not be assessed due to too little remaining biopsy tissue for analysis, 21 were also the diagnostic biopsy.

The clinico-pathologic characteristics of radical prostatectomy specimens is given in table 3. On the diagnostic or second-to-last biopsy on active surveillance (original Gleason ≤ 6 for all men), 11 men (14%) were revised as Gleason ≥ 7 , 5 of whom had a Gleason ≥ 7 on radical prostatectomy (table 4a). On last biopsy upgrading from a Gleason ≤ 6 to a Gleason ≥ 7 on radical prostatectomy occurred in 40% of men based on the original score and in 45% of men based on the revised score. Downgrading from a Gleason ≥ 7 to ≤ 6 occurred in 13% and 28% for the original and revised score respectively (table 4b).

Table 3. Characteristics on radical prostatectomy

	n (%)
T-stage	
pT2	71 (75%)
pT3a	14 (15%)
pT3b	0 (0%)
pT4	2 (2%)
Missing	8 (8%)
Gleason score*	
<=6	36 (38%)
3+4	36 (38%)
4+3	5 (5%)
>=8	10 (11%)
Missing	8 (8%)
Surgical margin	
Negative	68 (72%)
Positive	16 (17%)
Missing	11 (12%)
Total	95 (100%)

* Gleason score on radical prostatectomy was not revised. Original Gleason score is reported here.

Table 4a. Comparison of the original and revised Gleason score of the diagnostic and second-to-last biopsy on active surveillance with the Gleason score on radical prostatectomy.

		Gleason score radical prostatectomy			
Gleason score on biopsy	Original (O) versus Revised (R)	<=6	>=7	Unknown*	Total
	O and R equal: <=6	24	36	4	64
O <=6, R >=7	5	5	-	10	
O not reported, R <=6	-	1	-	1	
O not reported, R >=7	1	-	-	1	
Total	30	42	4	76	

* Radical prostatectomy data were not available.

Table 4b. Comparison of the original and revised Gleason score of the last biopsy on active surveillance with the Gleason score on radical prostatectomy.

		Gleason score radical prostatectomy			
Gleason score on biopsy	Original (O) versus Revised (R)	<=6	>=7	Unknown*	Total
	O and R equal: <=6	25	20	7	52
O and R equal: >=7	2	13	-	15	
O <=6, R >=7	5	5	-	10	
O >=7, R <=6	1	7	-	8	
Total	33	45	7	85	

* Radical prostatectomy data were not available.

Immunohistochemical analysis

For 48 out of the 87 diagnostic or second-to-last biopsies and 58 out of the 85 last biopsies enough tissue was remaining for immunohistochemical analysis. For all biopsies the median percentage of Ki67 positive tumor cells was 2% (IQR 2-4%), while ERG was scored positive in 67% of men. On univariable analysis percentage of tumor cells positive for Ki67 and ERG positivity were not statistically significant predictors of Gleason ≥ 7 or \geq pT3 on radical prostatectomy on both diagnostic or second-to-last biopsy, and on last biopsy (table 5a and 5b). Gleason score on last biopsy and $>50\%$ tumor in a single core were the only significant predictors of Gleason ≥ 7 or \geq pT3 on radical prostatectomy (table 5b). The area under the ROC curve for Gleason ≥ 7 or \geq pT3 on radical prostatectomy was 0.66 for Gleason score on last biopsy, 0.75 with the inclusion of $>50\%$ tumor involvement in a single core, and 0.79 if Ki67 and ERG were added to Gleason on last biopsy and $>50\%$ tumor in a single core. On diagnostic or second-to-last biopsy only ERG expression showed some discriminative ability with an area under the ROC curve of 0.58.

Table 5a. Correlation between biopsy characteristics on diagnostic and second-to-last biopsy and unfavorable outcome (pT3 or Gleason score ≥ 7) on radical prostatectomy.

	Univariable	
	OR (95% CI)	p-value
Original biopsy Gleason*	-	-
$>50\%$ involvement in a single core	1.4 (0.4-4.5)	.621
>2 cores positive*	-	-
PSA, ng/ml	1.0 (0.8-1.3)	.896
Ki67	1.0 (0.8-1.3)	.670
ERG	0.3 (0.1-1.5)	.149

* All Gleason ≤ 6 and ≤ 2 cores positive on biopsy

Table 5b. Correlation between biopsy characteristics on last biopsy and unfavorable outcome (pT3 or Gleason score ≥ 7) on radical prostatectomy.

	Univariable	
	OR (95% CI)	p-value
Original biopsy Gleason	6.5 (1.8-24.3)	.005
$>50\%$ involvement in a single core	4.3 (1.5-12.5)	.007
>2 cores positive	1.8 (0.7-4.7)	.226
PSA, ng/ml	1.1 (0.8-1.4)	.597
Ki67	1 (0.9-1.2)	.960
ERG	0.7 (0.2-2.3)	.541

DISCUSSION

Discordance between tumor characteristics at biopsy and subsequent radical prostatectomy is a well-known caveat in prostate cancer management. In men with Gleason score 6 on biopsy upgrading on radical prostatectomy occurs in approximately 40% of men (8). In this analysis we studied the effect of pathologic biopsy reevaluation and implementation of immunohistochemical biomarkers on predicting radical prostatectomy outcome in men initially on active surveillance for prostate cancer. Although we found differences in Gleason score or number of positive cores leading to a change in management in approximately 20% of cases, there was no improvement in correlation with final pathology. Suggesting that only by reducing undersampling on biopsy better predictions can be made on radical prostatectomy outcome.

Although revision by a dedicated uro-pathologist did thus not improve the correlation with pathology on radical prostatectomy, clinical implications can still be large. In the current analysis revision resulted in a change of management (either active treatment or active surveillance) in 13% and 20% of men at diagnosis or on last available biopsy. This was even higher than a previous analysis which reported changes in treatment strategy in 10% of men on active surveillance (9). Inter-observer variability in Gleason grading is most prominent in distinguishing Gleason score 6 from 7 (6). Gleason grade 4 encompasses a heterogeneous group of tumor growth pattern, of which ill-formed and fused glands are particularly sensitive for interpretation difficulties. Current grading guide-lines state that any amount of higher Gleason grade should be incorporated in the biopsy Gleason score, implying that even one single atypical glandular structure interpreted as Gleason grade 4 would lead to a Gleason score of 7 at biopsy.

As revision did not improve correlation with final pathology, undersampling of most aggressive tumor parts remains as main explanation for the discordance between biopsy and final pathology. For correct determination of tumor characteristics the focus should thus be on improving accurate tumor sampling. Simply increasing the number of biopsy cores taken or the frequency with which they are taken might not be the best method, as even with saturation biopsies there is substantial undergrading and it does not selectively targets the tumor (10-12). Alternatively, methods of targeting biopsies to areas suspicious for higher grade cancer could be a way to decrease under sampling. Currently, MRI seems the best option to direct biopsies in men on active surveillance (13). At inclusion, men with negative MRI have a very low probability of having higher grade tumor (13). Once the absence of higher grade tumor is established, progression rates to a higher Gleason while being on active surveillance are estimated to be low (14). The frequency at which men on active surveillance are tested could thus be reduced if initial sampling is more accurate. This might help to increase the acceptance and follow-up of men on active surveillance as biopsies are considered the most important drawbacks of active surveillance (15).

Secondly we studied if biomarker staining on biopsy tumor tissue could help in improving predictions of final pathology. The tested markers were selected as they previously have shown to correlate with Gleason score in radical prostatectomy specimens (16, 17). The lack of correla-

tion of these markers on biopsy with final pathology found in the current study could be related to again undersampling. Other than with Gleason grading it could however be theorized that biomarkers might still hold their predictive value in the presence of undersampling. Two main pathways in which a higher grade tumor is expected to develop are that of an independent origin (e.g. a Gleason 3+4 tumor develops independent of an already existing Gleason 3+3 tumor) or shared origin and progression (e.g. a Gleason 3+4 develops from the already existing Gleason 3+3 tumor). In the latter case genetic and biochemical changes could occur throughout the entire tumor (both Gleason 3 and 4 parts) or in Gleason grade 3 tumor cells that are in the process of progression towards a Gleason grade 4 (16). If this is the case, tissue biomarkers might help in distinguishing true low grade tumors and tumors with a higher grade. In the current analysis no such predictive marker was identified, similar to (18), (only ERG had some discriminative ability, albeit not statistically significant and primarily on initial biopsy). Although there are genetic differences between Gleason grades most studies studying these differences report an accuracy in differentiating Gleason grade 3 from Gleason grade 4 less or equal to 70%, or do not discuss the discriminative power (19-22). It seems better markers should still be discovered before they can be used to select and follow men on active surveillance.

In the current analysis apart from biopsy Gleason the only statistically (and clinically) significant variable predictive for Gleason grade ≥ 7 or $\geq pT3$ on radical prostatectomy was the presence of $>50\%$ tumor involvement in a single core. This variable is used as an inclusion criteria in some active surveillance studies (23). Although it proved to increase predictive capability in our study, the discriminative value was not very high (a lot of men with $>50\%$ tumor involvement still have favorable disease and vice versa). Hence, we feel that % tumor involvement on biopsy is not of value in selecting or excluding men for active surveillance. It could however be used to risk stratify men for stricter or less strict follow-up during AS.

The current analysis has some limitations. Not all biopsy specimens were available for revision and if they were available for revision the remaining biopsy tissue was in some cases too small for biomarker analysis. There was however no indication of selective missing of cases. Furthermore, radical prostatectomy specimens were not revised. It is likely that concordance between biopsy and radical prostatectomy Gleason is higher if scored by a single pathologist, irrespective of the fact whether this is the 'correct' Gleason score for both.

CONCLUSION

Interobserver differences in both scoring of tumor grade and extend on biopsy could result in a change of management strategy in approximately 13-20% of men on active surveillance for prostate cancer. Despite differences when re-grading biopsy specimens there was no improvement in correlation with final pathology on radical prostatectomy. This suggests that interobserver differences are not an important determinant of discordant tumor characteristics on biopsy and

radical prostatectomy specimen. Undersampling of most aggressive tumor remains the main focus in order to increase accurate grading at time of treatment decision. Tested tissue biomarkers (Ki67, and ERG) were not able to improve prediction of final pathology.

ACKNOWLEDGEMENT

We acknowledge the pathology departments of the following hospitals for providing biopsy tissue for this study: Stichting Pathan, Maasstad, Rijnstate, Antoni van Leeuwenhoek, Jeroen Bosch, Admiraal de Ruyter, Canisius, Stichting Pamm, Diakonessenhuis Utrecht, VU Medisch Centrum, Spaarne, Tweesteden, Atrium Medisch Centrum, UMC Radboud, Meander Medisch Centrum.

REFERENCES

1. Skradski V, Bektic J, Ladurner-Rennau M, Horninger W. Correlation of biopsy gleason score and gleason score of the corresponding radical prostatectomy specimen in patients who met the inclusion criteria for active surveillance. *The Journal of urology*. 2012;187(4):e539-e40.
2. Vellekoop A, Loeb S, Folkvaljon Y, Stattin P. Population based study of predictors of adverse pathology among candidates for active surveillance with Gleason 6 prostate cancer. *The Journal of urology*. 2014;191(2):350-7. Epub 2013/09/28.
3. Egevad L, Ahmad AS, Algaba F, et al. Standardization of Gleason grading among 337 European pathologists. *Histopathology*. 2013;62(2):247-56. Epub 2012/12/18.
4. Wolters T, Vissers KJ, Bangma CH, Schroder FH, van Leenders GJ. The value of EZH2, p27(kip1), BMI-1 and MIB-1 on biopsy specimens with low-risk prostate cancer in selecting men with significant prostate cancer at prostatectomy. *BJU international*. 2010;106(2):280-6. Epub 2009/11/06.
5. Bokhorst LP, Alberts AR, Rannikko A, et al. Compliance Rates with the Prostate Cancer Research International Active Surveillance (PRIAS) Protocol and Disease Reclassification in Noncompliers. *Eur Urol*. 2015;68(5):814-21. Epub 2015/07/04.
6. Epstein JI, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol*. 2015. Epub 2015/10/23.
7. van Leenders GJ, Boormans JL, Vissers CJ, et al. Antibody EPR3864 is specific for ERG genomic fusions in prostate cancer: implications for pathological practice. *Mod Pathol*. 2011;24(8):1128-38.
8. Venderbos LD, Roobol MJ, Bangma CH, et al. Rule-based versus probabilistic selection for active surveillance using three definitions of insignificant prostate cancer. *World J Urol*. 2015. Epub 2015/07/15.
9. Thomsen FB, Marcussen N, Berg KD, et al. Repeated biopsies in patients with prostate cancer on active surveillance: clinical implications of interobserver variation in histopathological assessment. *BJU international*. 2015;115(4):599-605. Epub 2014/06/07.
10. Chung PH, Darwish OM, Roehrborn CG, Kapur P, Lotan Y. Histologic upgrading in patients eligible for active surveillance on saturation biopsy. *Can J Urol*. 2015;22(1):7656-60.
11. Thompson JE, Hayen A, Landau A, et al. Medium-term oncological outcomes for extended vs saturation biopsy and transrectal vs transperineal biopsy in active surveillance for prostate cancer. *BJU international*. 2015;115(6):884-91.
12. Linder BJ, Frank I, Umbreit EC, et al. Standard and saturation transrectal prostate biopsy techniques are equally accurate among prostate cancer active surveillance candidates. *Int J Urol*. 2013;20(9):860-4.
13. Schoots IG, Petrides N, Giganti F, et al. Magnetic Resonance Imaging in Active Surveillance of Prostate Cancer: A Systematic Review. *Eur Urol*. 2014;67(4):627-36. Epub 2014/12/17.
14. Inoue LY, Trock BJ, Partin AW, Carter HB, Etzioni R. Modeling grade progression in an active surveillance study. *Stat Med*. 2014;33(6):930-9. Epub 2013/10/15.
15. Bokhorst LP, Lepisto I, Kakehi Y, et al. Complications after prostate biopsies in men on active surveillance and its effect on receiving further biopsies in the Prostate cancer Research International: Active Surveillance (PRIAS) study. *BJU international*. 2016.
16. Sowalsky AG, Ye H, Bublely GJ, Balk SP. Clonal progression of prostate cancers from Gleason grade 3 to grade 4. *Cancer Res*. 2013;73(3):1050-5. Epub 2012/12/04.
17. Zhu G, Liu Z, Epstein JI, et al. A Novel Quantitative Multiplex Tissue Immunoblotting for Biomarkers Predicts a Prostate Cancer Aggressive Phenotype. *Cancer Epidemiol Biomarkers Prev*. 2015;24(12):1864-72. Epub 2015/09/26.

-
18. Carozzi F, Tamburrino L, Bisanzi S, et al. Are biomarkers evaluated in biopsy specimens predictive of prostate cancer aggressiveness? *J Cancer Res Clin Oncol*. 2015.
 19. True L, Coleman I, Hawley S, et al. A molecular correlate to the Gleason grading system for prostate adenocarcinoma. *Proc Natl Acad Sci U S A*. 2006;103(29):10991-6. Epub 2006/07/11.
 20. Pascal LE, Vencio RZ, Page LS, et al. Gene expression relationship between prostate cancer cells of Gleason 3, 4 and normal epithelial cells as revealed by cell type-specific transcriptomes. *BMC Cancer*. 2009;9:452. Epub 2009/12/22.
 21. Rubin MA, Girelli G, Demichelis F. Genomic Correlates to the Newly Proposed Grading Prognostic Groups for Prostate Cancer. *Eur Urol*. 2015. Epub 2015/11/14.
 22. Boutros PC, Fraser M, Harding NJ, et al. Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat Genet*. 2015;47(7):736-45.
 23. Tosoian JJ, Mamawala M, Epstein JI, et al. Intermediate and Longer-Term Outcomes From a Prospective Active-Surveillance Program for Favorable-Risk Prostate Cancer. *J Clin Oncol*. 2015;33(30):3379-85. Epub 2015/09/02.

Chapter 9

A decade of active surveillance in the PRIAS study: an update and evaluation of the criteria used to recommend a switch to active treatment

Leonard P. Bokhorst^a, Riccardo Valdagni^b, Antti Rannikko^c, Yoshiyuki Kakehi^d,
Tom Pickles^e, Chris H. Bangma^a, Monique J. Roobol^a, for the PRIAS study group

^a Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands.

^b Prostate Cancer Program and Radiation Oncology 1, Fondazione IRCSS Istituto Nazionale
dei Tumori, Milan, Italy

^c Department of Urology, Helsinki University Central Hospital, Helsinki, Finland

^d Department of Urology, Kagawa University Faculty of Medicine, Kagawa, Japan

^e Department of Radiation Oncology, British Columbia Cancer Agency, Vancouver, Canada

ABSTRACT

Background: The Prostate cancer Research International Active surveillance (PRIAS) study was initiated a decade ago to study the most optimal selection and follow-up of men on active surveillance.

Objective: We report on 10 years of follow-up of men on active surveillance in the PRIAS study and evaluate if criteria used to recommend a switch to active treatment truly predict unfavorable outcome on subsequent radical prostatectomy (RP).

Design, setting, and participants: Men with low-risk prostate cancer were included and prospectively followed on active surveillance. Follow-up consisted of regular PSA tests, digital rectal examinations (DRE) and biopsies. Men with Gleason $>3+3$, >2 biopsy cores positive, or $>cT2$ were advised to switch to active treatment (until 2014 a PSA-doubling time (PSA-DT) of 0-3 years was used as well).

Outcome measurements and statistical analysis: Reclassification rates, treatment after discontinuation, and outcome on RP after discontinuing active surveillance were reported. Regression analysis on the outcome of RP was used to evaluate the predictive value of criteria currently used to recommend a switch to active treatment. Kaplan Meier and competing risk analysis were used to report discontinuation rates over time, and long term oncological endpoints.

Results and limitations: A total of 5302 men were included in PRIAS across 18 countries. Reclassification rates remained stable on all subsequent biopsies, with 22-33% of men having either Gleason $>3+3$ or >2 cores positive on any repeat biopsy. At 5 and 10 years of follow-up 52% and 73% of men had discontinued active surveillance, the majority because of protocol based reclassification. One third of men undergoing subsequent RP had favorable pathological tumor features (Gleason 3+3 and pT2). Of the criteria used to recommend a switch to active treatment >2 cores positive and a PSA-DT of 0-3 years were not predictive of unfavorable pathological outcome on RP.

Conclusions: A substantial group of men discontinued active surveillance without subsequent unfavorable tumor features on RP. We therefore propose Gleason upgrading and cT3 as only indicators of an immediate switch to active treatment. Surrogate indicators (such as >2 cores positive and a fast rising PSA) should not trigger immediate active treatment, but further investigation to confirm the suspicion of higher risk disease.

Patient summary: In the current study we confirm the safety of active surveillance as treatment option for men with low-risk prostate cancer. However some changes could be made to the follow-up protocol to safely increase the number of men that remain on active surveillance.

INTRODUCTION

The Prostate cancer Research International Active surveillance (PRIAS) study was initiated a decade ago (2006) with the aim of providing evidence-based recommendations on how to select and follow men with low risk prostate cancer on active surveillance (1). Other than most single (academic) center active surveillance studies, the PRIAS study aims to represent a more 'real world' situation with inclusions from academic, non-academic, and private practices across the world greatly increasing the generalizability of the results. Since its introduction the PRIAS study developed into the largest prospective active surveillance study worldwide, with at present more than 150 participating centers in 18 countries. Data on the first 500 study participants were reported in 2010 (2), with an update on 2500 men in 2012 (3).

In this paper we report on more than 5000 men followed on active surveillance in the PRIAS study to date, and we specifically evaluate the criteria used to recommend a switch to active treatment by assessing their ability to predict outcome on radical prostatectomy (RP) in men discontinuing active surveillance.

METHODS

All centers prospectively enter data on inclusion and follow-up through the PRIAS website (www.prias-project.org) which automatically provides recommendations on follow-up (1). The original criteria for inclusion are Gleason score $\leq 3+3$, $\leq cT2c$, PSA ≤ 10 ng/ml, ≤ 2 cores positive for prostate cancer, PSA density ≤ 0.2 ng/ml/cm³, and fitness for curative treatment. In 2012 and 2015 inclusion criteria were adapted to include minimal Gleason 3+4, and accommodate changes in the number of positive cores caused by MRI targeted biopsies or saturation biopsies (all changes made to the study protocol are summarized in table 1). No minimum number of biopsy cores is required, but based on prostate volume the following is advised: <40 cm³: 8 cores, 40-60cm³: 10 cores, >60 cm³: 12 cores.

In the first two years of follow-up a PSA test is scheduled every 3 months and Digital Rectal Examination (DRE) every six months. Thereafter PSA is measured every six months and DRE is performed once yearly. Standard repeat biopsy are scheduled 1,4,7, and 10 years after diagnosis, and subsequently every 5 years. Yearly biopsies are only recommended if PSA doubling time (PSA-DT) is between 0 to 10 years. A bone scan is recommended if PSA ≥ 20 ng/ml.

Criteria used to recommend a switch to active treatment are Gleason $>3+3$, >2 cores positive, and $>cT2$. A PSA-DT between 0 and 3 years (if at least four PSA values are available) was used to recommend immediate active treatment until the end of 2014, but was dropped afterwards due to the low number of men complying with this recommendation and the high percentage of men receiving unnecessary treatment, as described in a recent publication (4). Criteria used to

recommend a switch to active treatment were adapted for those included with Gleason 3+4 and >2 cores based on MRI or saturation biopsies (table 1).

More information on the follow-up schedule and a flowchart of the current follow-up protocol can be found online (www.prias-project.org). Follow-up for the current analysis ended November 2015.

Table 1. Changes made and proposed to the PRIAS study protocol.

Year	2012	2015	2015	2016 (Proposed in current manuscript)	2016 (Proposed in current manuscript)
Change to	Inclusion criteria	Inclusion criteria	Follow-up criteria	Follow-up criteria	Follow-up schedule
Change	Gleason 3+4 (<=10% tumor involvement per biopsy core, maximum 2 cores positive) can be included if aged >=70 years. (Upgrading during follow-up is defined as Gleason 4+3, all other criteria still apply)	>2 cores can be positive (no maximum) if an MRI, including targeted biopsies on positive lesions, is done at inclusion. AND If saturation biopsies are done (>=20 cores), 15% of cores can be positive with a maximum of 4. (Upgrading based on the positive number of cores is adapted accordingly, all other criteria still apply)	A PSA-doubling time (PSA-DT) of 0-3 years is removed as recommendation to switch to active treatment. Instead yearly biopsies are advised as with a PSA-DT of 3-10 years. If MRI is available an MRI can be performed to rule out anterior tumors.	>2 cores positive should trigger an MRI with targeted biopsy instead of an immediate switch to active treatment.	PSA testing can be reduced to once yearly after 4 years of follow-up AND A digital rectal examination can be performed at the time of a biopsy only

STATISTICAL ANALYSIS

Descriptive statistics were used to report baseline characteristics, reclassification rates on subsequent biopsies, treatment after discontinuation, and outcome on RP after discontinuing active surveillance. Pathological outcome on RP was divided in three categories: favorable pathology (Gleason 3+3 and pT2), intermediate pathology (Gleason 3+4 and pT2), and unfavorable pathology (Gleason >= 4+3 or >=pT3), based on a previous analysis and recent reports on low metastatic potential of men with Gleason 3+3 organ confined disease on RP (5-7).

Multinomial logistic regression analysis was used to evaluate the predictive value of criteria used to recommend a switch to active treatment (Gleason >3+3, >2 cores positive, PSA-DT between 0 and 3 years) on pathological RP outcome. The predictive value of cT3 could not be assessed due to the low number of men with this characteristic. Corrections were made for clinical characteristics at the time of switching to active treatment (age and PSA).

Kaplan Meier and competing risk analysis were used to report discontinuation rates over time, (prostate cancer) mortality, and a combined endpoint of biochemical recurrence (BCR, defined as a PSA ≥ 0.2 ng/ml after RP or a PSA level 2.0 ng/ml above the nadir after radiotherapy (RT)) or local recurrence after active treatment (either RP or RT), metastasis, and prostate cancer death, whichever occurred first for all men included.

Analysis were done for all men included and for men fulfilling the original inclusion criteria, except for the multinomial analysis which was only done in men fulfilling the original inclusion criteria. For analysis SPSS for Windows (Version 21.0. Armonk, NY: IBM Corp.) and the survival package of R (R Foundation for Statistical Computing, Vienna, Austria.) were used.

RESULTS

In total 5302 men were included and prospectively followed on active surveillance in the PRIAS study across 18 countries (figure 1, supplements). Out of these men 622 were followed on active surveillance for more than 5 years, and 107 for more than 7.5 years. At diagnosis median age was



Figure 1 (Supplements). Number of inclusion in the PRIAS study per country.

65.9 years, median PSA 5.7 ng/ml, most men had 1 biopsy core positive (69%) with Gleason 3+3 (99%), and a clinical stage T1c (88%) (table 2). A total of 216 men (4%) did not fully comply with the original inclusion criteria (due to the changes in protocol or because the PRIAS website currently allows 'off protocol' inclusions). Sub analyses in men fulfilling the original inclusion criteria yielded identical results and are therefore not presented.

Table 2. Characteristics at diagnosis of all men included in the PRIAS study.

	Median (IQR)/N (%)
Age, years	65.9 (61.0-70.4)
PSA, ng/ml	5.7 (4.5-7.1)
Prostate volume, cm³	45 (35-59)
PSA density, ng/ml/ cm³	0.13 (0.09-0.16)
Number of biopsy cores	12 (10-12)
Number of cores positive	
1	3643 (69%)
2	1615 (30%)
>=3	44 (1%)
Gleason	
3+3	5271 (99%)
3+4	31 (1%)
T-stage	
cT1c	4649 (88%)
cT2a	579 (11%)
cT2b	54 (1%)
cT2c	20 (<1%)
Charlson score	
0	3745 (71%)
1	264 (5%)
2	734 (14%)
>=3	559 (11%)
Total	5302 (100%)

IQR = interquartile range

During follow-up 3379 men received at least one repeat biopsy, 1077 men two biopsies, 282 men three biopsies, 68 men four biopsies, and 15 men five biopsies. Reclassification rates remain stable on all subsequent biopsies, with 13-16% of men having a Gleason >3+3, 16-27% of men having >2 cores positive for prostate cancer, and 22-33% of men having either Gleason >3+3 or >2 cores positive (figure 2).

A total of 1768 out of the 5302 men discontinued active surveillance until the end of follow-up. The majority (n=1102) because of protocol based reclassification. Treatment after discontinuation

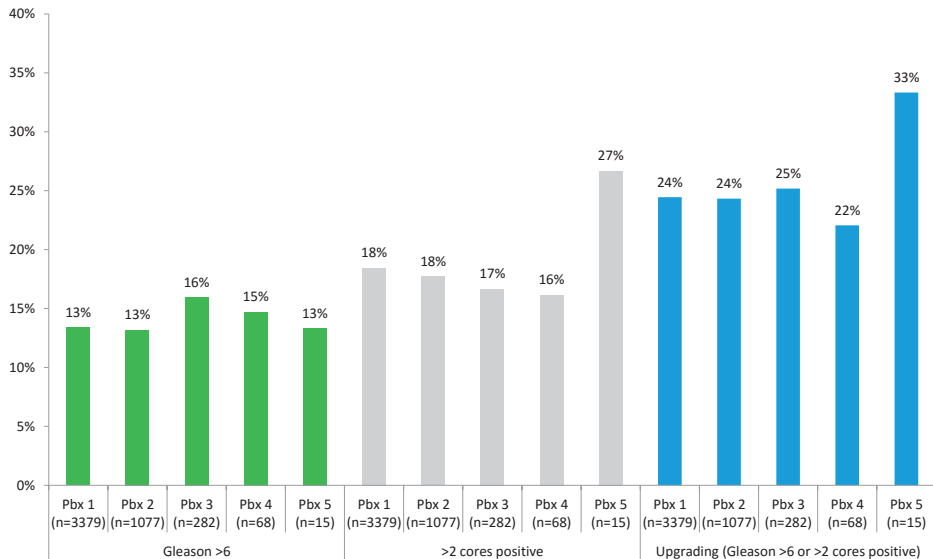


Figure 2. Reclassification rates on subsequent repeat biopsies (Pbx) during follow-up.

was RP or RT in 67% of men and only 3% received hormonal therapy (HT) as primary treatment (table 3). There were no differences in tumor characteristics (PSA, PSA-DT, Gleason, and number of positive cores on last biopsy) between men switching to RP or RT, but the latter had a 2 year higher median age at the time of discontinuation (67 vs 69 years, respectively, $p < 0.001$).

Table 3. Discontinuation and treatment after discontinuation.

	RP	RT	HT	WW	Other	Unknown	Died/Lost to FU	Total
Non-protocol based								
Anxiety/Patient request	52 (29%)	32 (18%)	2 (1%)	-	2 (1%)	89 (50%)	-	177 (100%)
Other/Unknown	108 (45%)	78 (32%)	2 (1%)	-	27 (11%)	27 (11%)	-	242 (100%)
Switch to WW	-	-	-	134 (100%)	-	-	-	134 (100%)
Died	-	-	-	-	-	-	57 (100%)	57 (100%)
Lost to FU	-	-	-	-	-	-	56 (100%)	56 (100%)
Protocol based*	496 (45%)	419 (38%)	29 (3%)	30 (3%)	30 (3%)	98 (9%)	-	1102 (100%)
Total	656 (37%)	529 (30%)	33 (2%)	164 (9%)	59 (3%)	214 (12%)	113 (6%)	1768 (100%)

* Gleason >3+3, >2 cores positive, >cT2, or PSA-DT 0-3 years. RP = radical prostatectomy, RT = radiotherapy, HT = hormonal therapy, WW = watchful waiting, FU = follow-up.

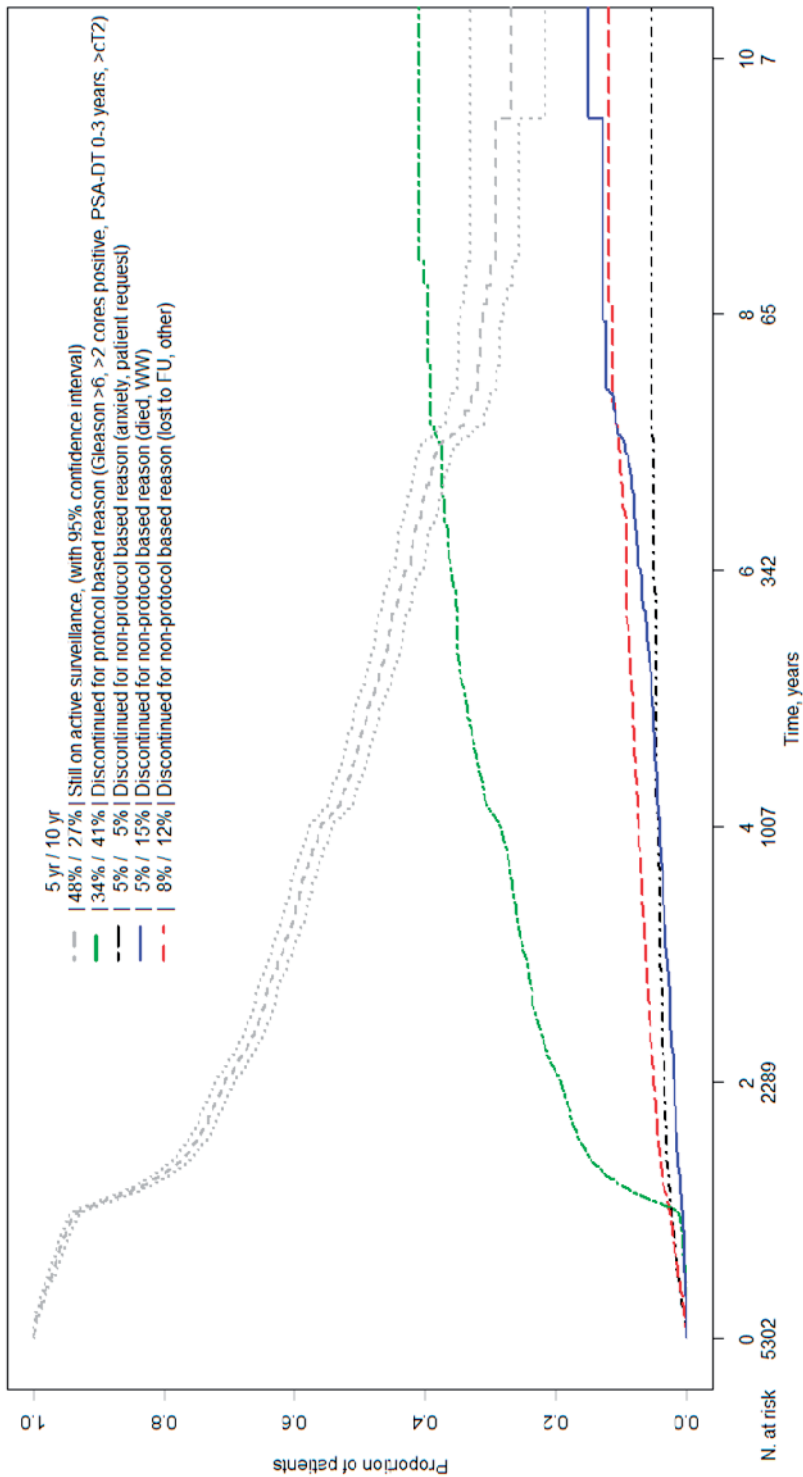


Figure 3. Discontinuation of active surveillance over time, stratified by reason for discontinuation. FU = follow-up, WW = watchful waiting, PSA-DT = PSA doubling time

After 5 and 10 years of follow-up 48% and 27% of men were still on active surveillance, 34% and 41% discontinued because of protocol based reclassification, 5% and 5% discontinued due to anxiety/patient request (without having reclassification, anxiety and patient request were equally distributed), 5% and 15% switched to watchful waiting (WW) or died of other cause (without having reclassification), and 8% and 12% discontinued due to other reasons (without having reclassification) (figure 3).

Pathology data for 360 men receiving RP after discontinuing active surveillance were available for analysis. For men who switched to RP due to anxiety, 13 (57%) had a favorable pathological outcome, 6 (26%) had an intermediate pathological outcome, and 4 (17%) had an unfavorable pathological outcome. For men who switched to RP because of protocol based reclassification, 82 (30%) had favorable pathological outcome, 85 (34%) intermediate pathological outcome, and 100 (36%) unfavorable pathological outcome. pT3a, pT3b, pT4, Gleason ≥ 8 , and N1 was found in 61,13,2,14, and 1 men (out of 119 men receiving a lymph node dissection) respectively. Large differences in distribution of outcomes were observed between different protocol based reasons to discontinue active surveillance (table 4). On regression analysis only Gleason score >6 on last biopsy was a statistically significant predictor of intermediate or unfavorable pathological outcome on RP (table 5).

Table 4. Outcome on radical prostatectomy after discontinuing active surveillance.

	Favorable (Gleason 3+3 and pT2)	Intermediate (Gleason 3+4 and pT2)	Unfavorable (Gleason $\geq 4+3$ or $\geq pT3$)	Total
Non-protocol based				
Anxiety/Patient request	13 (57%)	6 (26%)	4 (17%)	23 (100%)
Other/Unknown	28 (47%)	12 (20%)	20 (33%)	60 (100%)
Protocol based				
(1) Only Gleason $>3+3$				
Gleason 3+4	7 (27%)	15 (58%)	4 (15%)	26 (100%)
Gleason $\geq 4+3$	1 (7%)	3 (21%)	10 (71%)	14 (100%)
(2) Only >2 cores positive				
	28 (41%)	22 (32%)	18 (26%)	68 (100%)
(3) Only PSA-DT 0-3 years				
	24 (46%)	9 (17%)	19 (37%)	52 (100%)
(4) Only cT3				
	1 (50%)	-	1 (50%)	2 (100%)
Combination 1+2	1 (2%)	28 (55%)	22 (43%)	51 (100%)
Combination 1+3	2 (18%)	3 (27%)	6 (55%)	11 (100%)
Combination 2+3	16 (50%)	6 (19%)	10 (31%)	32 (100%)
Combination 1+2+3	2 (10%)	9 (43%)	10 (48%)	21 (100%)
Total	123 (34%)	113 (31%)	124 (34%)	360 (100%)

PSA-DT = PSA doubling time

Table 5. Predictors of intermediate (Gleason 3+4 and pT2) and unfavorable outcome (Gleason \geq 4+3 or \geq pT3) on radical prostatectomy (only men fulfilling the original inclusion criteria were included, n=347)

	Intermediate		Unfavorable	
	OR (95% CI)	p-value	OR (95% CI)	p-value
Age at discontinuation, years	1.04 (0.99-1.09)	.136	1.04 (0.99-1.09)	.101
Last PSA, ng/ml	0.93 (0.85-1.03)	.146	1.00 (0.91-1.09)	.925
PSA-DT 0-3 years	0.71 (0.37-1.38)	.312	1.44 (0.79-2.63)	.230
Number of positive cores >2 on last biopsy	1.50 (0.83-2.70)	.183	1.37 (0.78-2.43)	.274
Gleason >6 on last biopsy	7.44 (3.68-15.06)	<0.001	6.12 (3.04-12.32)	<0.001

OR= odds ratio; CI= confidence interval; PSA-DT = PSA doubling time
OR as compared to favorable RP outcome

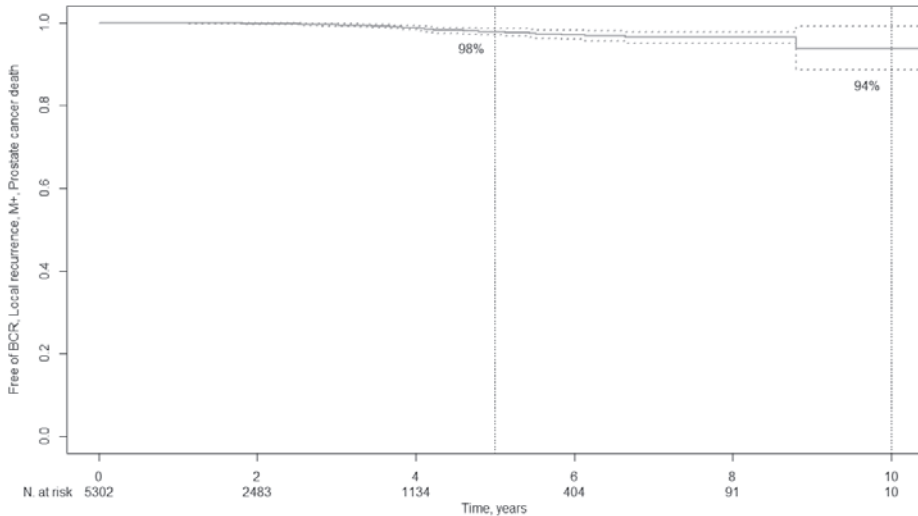


Figure 4 (Supplements). Time since diagnosis free of biochemical recurrence (BCR), local recurrence, metastasis, or prostate cancer death for all men included in PRIAS.

Until the end of follow-up out of all men included 30 men had biochemical recurrence (BCR) after active treatment (either RP or RT), 10 men had local recurrence, 8 developed metastasis, and 1 died of prostate cancer, resulting in 98% and 94% of all men included to be free of BCR, local recurrence, metastasis, and prostate cancer death 5 and 10 years after diagnosis (figure 4, supplements). Other cause mortality and disease specific mortality for all men included were 3% and <1% 5 years after diagnosis, and 11% and <1% 10 years after diagnosis (figure 5, supplements).

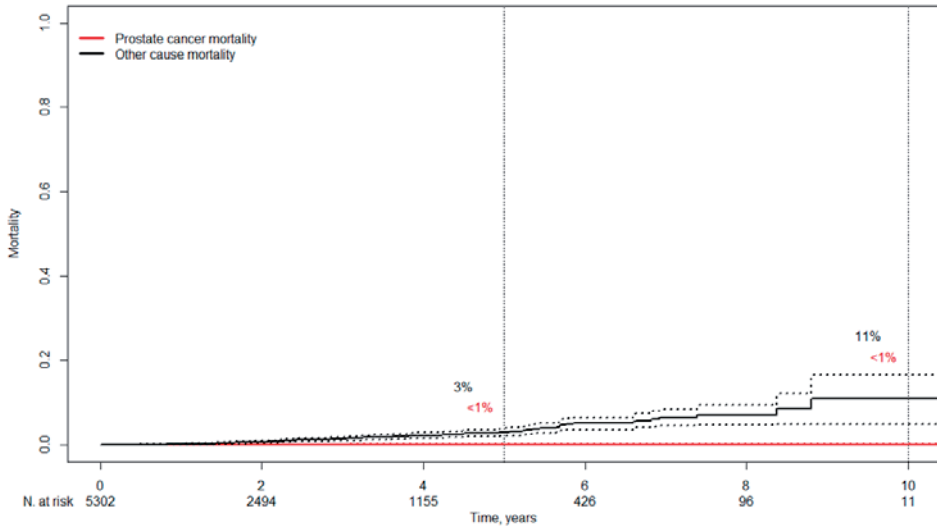


Figure 5 (Supplements). Overall and prostate cancer specific mortality since diagnosis for all men included in the PRIAS study.

DISCUSSION

Active surveillance aims to reduce overtreatment of tumors that are very unlikely to cause symptoms if left untreated. Nevertheless, the possibility remains that tumors with initial low-risk features will turn out not to be indolent and develop into lethal disease. Active surveillance is supposed to selectively filter out these men as soon as possible while avoiding invasive treatment in the majority that prove to be truly indolent. The first part of the aim of active surveillance seems partially fulfilled. In the Toronto and Johns Hopkins cohort of men on active surveillance very low prostate cancer mortality and metastasis rates were observed, comparable to those after more invasive treatment (RP and RT) (8, 9). Although only few men were followed for more than 10 years in the current analysis, results support the safety of active surveillance. Prostate cancer mortality was <1% and a combined end-point of adverse outcome (BCR, local recurrence, metastasis, or prostate cancer death) was observed in only 6% of men 10 years from diagnosis.

However, in the current analysis we found only 50% of men to still be on active surveillance after 5 years of follow-up and approximately only 25% after 10 years of follow-up, lower than reported by other active surveillance studies (8-10). Some of these differences could be explained by the setup of the PRIAS study which is not a strictly controlled single academic center study, but instead tries to represent a real world situation (e.g. resulting in more men switching to active treatment without a clear protocol based reason or more men switching to WW). However even if accounting for the 15% of men stopping active surveillance because of a switch to WW or other cause mortality, a substantial part of men (60% of the initial cohort) received a form of active treatment after 10 years of follow-up. This, together with the observation that one third of men

still had a favorable pathological outcome on RP, indicates that the criteria used to recommend a switch to active treatment are far from optimal. On regression analysis the only statistically significant predictor of intermediate or unfavorable pathological outcome on RP was a Gleason score >6 on last biopsy. In men who switched to RP due to > 2 cores positive on last biopsy and in men with a PSA-DT of 0-3 years the rates having favorable pathological outcome were close to 50%, although the rates of unfavorable pathological outcome were higher than in men who switched to active treatment without a protocol based recommendation. Both these protocol based indications thus seem to increase the risk of unfavorable pathological outcome, but are not specific enough to trigger an immediate switch to active treatment. Importantly, PSA-DT 0-3 and >2 cores positive with prostate cancer together are responsible for more than 50% of all recommendations to switch to active treatment.

To achieve a higher rate of men who continue active surveillance while selectively identifying those with unfavorable disease, we propose a change of protocol for the PRIAS study. Instead of an immediate switch to active treatment if >2 cores are positive, men should receive further investigation to confirm higher risk disease. As MRI is shown to have a negative predictive value for Gleason upgrading very close to 100% in men on active surveillance (11-13), it currently seems the best method to exclude the presence of a higher Gleason score. If the MRI is negative active surveillance can thus be continued, if a lesion is present targeted biopsies should confirm Gleason upgrading before a switch to active treatment is advised. MRI with targeted biopsies in men with increased risk is expected to detect the majority of men with truly unfavorable tumor characteristics (11, 14). This modification is in line with the recently changed recommendation in men with a PSA-DT of 0-3 years and with the changed inclusion criteria which allow inclusion of any number of positive cores if an MRI with targeted biopsies is done (4). Gleason >6 and cT3 will thus remain the only indicators for an immediate switch to active treatment. It is estimated, based on the data in this article, that because of the suggested protocol change instead of 43/100 men, 64/100 men could have stayed on active surveillance after 3 repeat biopsies.

In the (near) future MRI might even be able to replace systematic repeat biopsies altogether. Systematic biopsies currently appears one of the largest burdens for men on active surveillance (4, 15, 16), and in fact are redundant in three quarters of men as they do not show reclassification (figure 2). But, before we can definitively adopt such a change we must collect enough data on men with a negative MRI who simultaneously received systematic biopsies. We therefore plea for increased inclusion of men in e.g. the PRIAS MRI side study (www.prias-project.org) to further establish the negative predictive value of MRI in men on active surveillance. If confirmed, many systematic biopsies can be replaced by less invasive imaging.

Some changes to the number and frequency of follow-up visits remain to be discussed. PSA testing is done regularly and used to calculate PSA-DT (used to recommend more frequent biopsies) and a bone scan (if PSA >20ng/ml). We previously showed data from the PRIAS study indicating that after 4 years of follow-up both a change in PSA-DT triggering a biopsy and an absolute PSA >20 occurred very infrequently (17). Furthermore, clinical utility was low (all bone

scans were negative and biopsies were only advanced by six months). It was concluded that PSA testing can be reduced to once yearly after 4 years of follow-up.

It is sometimes suggested to use an absolute PSA value (e.g. >20ng/ml) to recommend a switch to active treatment. The PRIAS study did not include such a recommendation as it was felt that once included with a PSA \leq 10ng/ml the PSA value could only rise slowly (in which case BPH might be a more likely cause) or fast but then PSA-DT would trigger further investigation with biopsies to exclude rapid tumor development as its cause. The analysis presented in table 5 confirms that this initial assumption is now justified as the absolute PSA value does not show a positive correlation with RP outcome within the current follow-up protocol which includes regular repeat biopsies and additional biopsies in the case of fast rising PSA. We therefore do not recommend an absolute PSA cut-off to discontinue active surveillance.

As reclassification on DRE only occurred in 10 men (<1% of all reclassifications) one could argue to reduce the number of DRE's as well, e.g. only perform a DRE at the time of a biopsy, although the potential benefit of reducing this relative inexpensive and easy to perform test could be questioned. The largest benefit should come from individualizing the frequency of repeat biopsies (or in the future possibly MRI's). Currently we individualize the biopsy frequency only based on the PSA-DT. Several other predictors of reclassification were however specified (e.g. PSA density, the number of positive biopsy cores, and time since last biopsy) (3, 4, 18, 19). Risk prediction models were already developed and should be validated and updated in several cohorts to prolong the time to next biopsy in men with low risk of reclassification and increase the frequency in men with high risk (19). In the future these models can be supplemented by newly validated marker predictive of outcome (20).

Such models can also be used to assist in the timing of when to stop active surveillance and switch to WW. Simultaneous predictions on life expectancy and time until symptoms from a low-risk tumor left untreated are needed. This is one of the topics of the Movember-GAP3 project (21), which combines the majority of the worldwide available active surveillance cohorts, including the PRIAS study.

Criticisms of the current analysis are that follow-up data of men who discontinued active surveillance, including outcome of RP, was missing in several cases. Although limiting the power of some of the analyses, there was no indication of selective reporting, which could have affected the results (men with and without RP data available did not differ in terms of age, PSA, PSA-DT, Gleason, and number of positive cores on last biopsy).

CONCLUSION

After a decade of active surveillance in the PRIAS study criteria used to recommend a switch to active treatment seem not selective enough to avoid unnecessary switches to active treatment. A substantial proportion of men abandoning active surveillance based on a protocol advice do

not have unfavorable features after RP. We therefore propose Gleason score upgrading or cT3 on rectal examination as the only reasons for a direct switch to active treatment. Other factors, such as >2 biopsy cores positive and fast rising PSA, should first trigger further investigation to confirm the suspicion of higher risk disease.

REFERENCES

1. van den Bergh RC, Roemeling S, Roobol MJ, Roobol W, Schroder FH, Bangma CH. Prospective validation of active surveillance in prostate cancer: the PRIAS study. *Eur Urol.* 2007;52(6):1560-3. Epub 2007/05/29.
2. van den Bergh RC, Vasarainen H, van der Poel HG, et al. Short-term outcomes of the prospective multicentre 'Prostate Cancer Research International: Active Surveillance' study. *BJU international.* 2010;105(7):956-62. Epub 2009/10/13.
3. Bul M, Zhu X, Valdagni R, et al. Active Surveillance for Low-Risk Prostate Cancer Worldwide: The PRIAS Study. *Eur Urol.* 2012;63(4):597-603. Epub 2012/11/20.
4. Bokhorst LP, Alberts AR, Rannikko A, et al. Compliance Rates with the Prostate Cancer Research International Active Surveillance (PRIAS) Protocol and Disease Reclassification in Noncompliers. *Eur Urol.* 2015;68(5):814-21. Epub 2015/07/04.
5. Bul M, Zhu X, Rannikko A, et al. Radical Prostatectomy for Low-Risk Prostate Cancer Following Initial Active Surveillance: Results From a Prospective Observational Study. *Eur Urol.* 2012;62(2):195-200. Epub 2012/02/22.
6. Kweldam CF, Wildhagen MF, Bangma CH, van Leenders GJ. Disease-specific death and metastasis do not occur in patients with Gleason score ≤ 6 at radical prostatectomy. *BJU international.* 2015; 116(2):230-5. Epub 2014/07/26.
7. Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) ≤ 6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol.* 2012;36(9):1346-52. Epub 2012/04/26.
8. Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. *J Clin Oncol.* 2015;33(3):272-7. Epub 2014/12/17.
9. Tosoian JJ, Mamawala M, Epstein JI, et al. Intermediate and Longer-Term Outcomes From a Prospective Active-Surveillance Program for Favorable-Risk Prostate Cancer. *J Clin Oncol.* 2015;33(30):3379-85. Epub 2015/09/02.
10. Welty CJ, Cowan JE, Nguyen H, et al. Extended Follow-Up and Risk Factors for Disease Reclassification from a Large Active Surveillance Cohort for Localized Prostate Cancer. *The Journal of urology.* 2014. Epub 2014/09/28.
11. Schoots IG, Petrides N, Giganti F, et al. Magnetic Resonance Imaging in Active Surveillance of Prostate Cancer: A Systematic Review. *Eur Urol.* 2014;67(4):627-36. Epub 2014/12/17.
12. Hoeks CM, Somford DM, van Oort IM, et al. Value of 3-T multiparametric magnetic resonance imaging and magnetic resonance-guided biopsy for early risk restratification in active surveillance of low-risk prostate cancer: a prospective multicenter cohort study. *Invest Radiol.* 2014;49(3):165-72. Epub 2013/11/14.
13. Walton Diaz A, Shakir NA, George AK, et al. Use of serial multiparametric magnetic resonance imaging in the management of patients with prostate cancer on active surveillance. *Urol Oncol.* 2015;33(5): 202 e1-7. Epub 2015/03/11.
14. Schoots IG, Roobol MJ, Nieboer D, Bangma CH, Steyerberg EW, Hunink MG. Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. *Eur Urol.* 2015;68(3):438-50. Epub 2014/12/07.
15. Moore CM, Parker C. The Evolution of Active Surveillance for Prostate Cancer. *Eur Urol.* 2015;68(5): 822-3. Epub 2015/07/29.

-
16. Bokhorst LP, Lepisto I, Kakehi Y, et al. Complications after prostate biopsies in men on active surveillance and its effect on receiving further biopsies in the Prostate cancer Research International: Active Surveillance (PRIAS) study. *BJU international*. 2016.
 17. Bokhorst L, Alberts A, Kakehi Y, et al. Frequency of PSA testing in men on active surveillance for prostate cancer. *Journal of Urology*. 2015;193(4):e755.
 18. Bul M, van den Bergh RC, Rannikko A, et al. Predictors of unfavourable repeat biopsy results in men participating in a prospective active surveillance program. *Eur Urol*. 2012;61(2):370-7. Epub 2011/06/28.
 19. Ankerst DP, Xia J, Thompson IM, Jr., et al. Precision Medicine in Active Surveillance for Prostate Cancer: Development of the Canary-Early Detection Research Network Active Surveillance Biopsy Risk Calculator. *Eur Urol*. 2015;68(6):1083-8. Epub 2015/03/31.
 20. Loeb S, Bruinsma SM, Nicholson J, et al. Active Surveillance for Prostate Cancer: A Systematic Review of Clinicopathologic Variables and Biomarkers for Risk Stratification. *Eur Urol*. 2014. Epub 2014/12/03.
 21. Bruinsma SM, Bangma CH, Obbink H, Roobol MJ. Active surveillance for low risk prostate cancer: The study protocol of the Movember Global Action Plan 3 (GAP3) project. *European Urology, Supplements*. 2015;14(2):e1036-ea.

Summary

This thesis set out to provide further in-depth information on prostate cancer screening and active surveillance by addressing two main questions: “What are the mechanisms that lead to the observed reduction in prostate cancer mortality?” (part 1, screening) and “Are we able to selectively identify men with aggressive disease?” (part 2, active surveillance). To address these two main questions several sub-questions were studied.

PART 1, SCREENING

- *What type of prostate cancers are detected and at what time during the screening process? (Chapter 1)*

Cancer detection rates remained stable in men without a previous biopsy throughout the course of the screening program, but the majority of high risk cancers were detected in the first screening round. Underlying the stable cancer detection rate was the detection of cancers in men who at the previous screening had PSA values below the biopsy cut-off. Men with a previous negative biopsy only accounted for a small proportion of the cancers detected at repeat screenings, which, supported by an on average larger prostate volume, suggests that the prolonged elevation of the PSA level in these men is most likely not caused by prostate cancer.

- *What is the effect of contamination and noncompliance on the observed prostate cancer mortality reduction? (Chapter 2)*

Especially contamination (unwanted screening) in the control arm of the European screening study had a large diluting effect on the mortality reduction observed. This indicates that the effect of screening on an individual's risk of dying from prostate cancer is (substantially) larger than recorded for the entire study population. As non-compliance was low its effect on the observed mortality reduction was small.

- *What is the effect of treatment on the observed prostate cancer mortality reduction in the European screening study? (Chapter 3 and 4)*

Due to later detection in time, tumors in the control arm were able to benefit from progressing insight in how best to treat prostate cancer (chapter 3). The overall effect of treatment differences on the observed prostate cancer mortality reduction outcome was however small (chapter 4).

- *Where does the benefit of prostate cancer screening originate from? (Chapter 1 to 4)*

The benefit, in terms of prostate cancer mortality reduction, seemed to originate from all screening rounds (chapter 1), by an absolute reduction of metastatic disease at diagnosis (i.e. a stage shift: chapter 4). Treatment differences had little effect on the observed benefit (chapter 3 and 4), but to better estimate the mortality reduction for an individual a correction should be made for contamination and noncompliance, as both dilute the overall effect observed (chapter 2).

In conclusion from part 1 we can say that screening works through early detection (stage shift) and subsequent adequate treatment.

PART 2, ACTIVE SURVEILLANCE

- *Are we able to selectively identify men with aggressive disease using risk based selection at inclusion? (Chapter 5)*

Risk based selection of potential candidates for active surveillance, as opposed to rule based selection, helps them to make a more conscious and personal treatment choice by providing individualized risk estimates. However, as with rule based selection, distinction between aggressive and non-aggressive disease is not perfect yet. Advances should be made to improve the information available for risk-based selection, including forthcoming biomarkers and imaging procedures.

- *Are we able to selectively identify men with aggressive disease using information coming from pathology and biomarkers? (Chapter 8)*

The immunohistochemical biomarkers that were tested (Ki67, ERG, and P27), and reevaluation of biopsy pathology report, were not able to improve the initial evaluation of disease aggressiveness based on standard used parameters. Undersampling of most aggressive tumor fragments seems the limiting factor that should be focused on first.

- *Is prostate biopsy the best method to detect aggressive disease, i.e. what are its drawbacks? (Chapter 6 and 7)*

Prostate biopsies result in complications in a substantial part of men (chapter 7) and form a substantial burden in men on active surveillance. This might very well be the reason that many recommended biopsies during follow-up are ignored (chapter 6). Furthermore, many currently indicated biopsies do not lead to disease reclassification and might thus be avoidable.

- *Should we change the current follow-up protocol for men on active surveillance? (Chapter 9)*

After a decade of active surveillance the results of the PRIAS study underline its safety. However, a substantial number of men switch to active treatment without subsequently showing aggressive disease on final pathology. The follow-up protocol should thus be adapted to reduce the number of men that unnecessarily switch to active treatment.

In conclusion, it is possible currently to selectively identify men without aggressive disease for active surveillance. This results in very low rates of disease-specific mortality. However, current protocols are not selective enough. Especially during follow-up they result in a substantial number of men still being overtreated. Some improvements to more (and better) individualized protocols can already be made, but further research is definitively needed.

The findings of all previous chapters will be discussed in more detail in the following section.

General discussion

SCREENING (HOW DOES SCREENING RESULT IN A MORTALITY REDUCTION?)

Screening for prostate cancer results in a reduction of prostate cancer specific mortality [1]. However, the underlying mechanism is not completely clear. Furthermore, although PSA based screening for prostate cancer is effective, it is far from optimal. Mortality reduction is in the order of 20% (i.e. 80% of prostate cancer deaths are not avoided) and there are substantial side effects. Thus, optimizing screening is key, both for men who already choose to be screened and those who want screening but at the moment opt out due to the suboptimal balance between harms and benefits. So, it is crucial to understand what factors influence the way screening works and how can we use this information to optimize screening.

EXTERNAL FACTORS THAT AFFECT HOW SCREENING WORKS

External factors do not result from the working of the screening algorithm itself, but originate e.g. from issues in its practical implementation or from other factors that could influence the course of the tumor. In order to provide clear information on the results of screening if performed under optimal conditions these external factors should be corrected for.

Noncompliance

The first external factor studied in this thesis is noncompliance to the screening protocol (chapter 2). In a randomized screening study (with a intervention and control arm) there can be two types of noncompliance. The first (contamination) is if men in the control arm receive the intervention under study (in this case screening), which will dilute the observed effects (both positive and negative) of the trial. As discussed in chapter 2 correction for contamination had a substantial effect on the measured outcome of screening (from a 32% mortality reduction as calculated by the intention to treat analysis up to 47% after correction for contamination). This information is relevant for an individual deciding on screening. Although the negative effects of screening (namely overdiagnosis) are diluted as well, the relative effect of the correction on prostate cancer mortality is slightly larger than the relative effect on overdiagnosis (calculated from table 4 and figure 2, chapter 2), meaning that there is no decrease in the harms to benefits ratio. Other than providing more accurate information on the effect of screening for an individual, contamination does not need to be addressed in daily practice as there is no control arm to consider.

The second form of noncompliance is nonattendance in the screening arm. Nonattendance as defined in chapter 2 (nonattendance to all screening rounds) was relatively small, as was its effect on mortality (from a 32% mortality reduction to 33%). The low nonattendance observed in chapter 2 might be a result of the study setup as only men providing upfront consent (and thus willing to undergo screening) were randomized. Nonattendance in the general population is likely higher. However, for those willing to undergo screening the observed, low, nonattendance

has some implications. First, it implies that the screening algorithm itself is not too demanding on the participant (i.e. if nobody attended one could question if the screening algorithm itself might be the cause). Secondly as nonattendance seems limited to a specific group of men with above average comorbidity it might not be a large problem in daily practice as the effect on mortality for men with short life expectancy (generally <10 years) is considered to be limited [2]. The impact of nonattendance in an organized screening program selected by an individual is thus low and predicted effects will closely resemble the effects in real life.

These properties of an organized screening program might not be transferable to an unorganized screening program (opportunistic screening), thus making predictions on its effect far more difficult. There is evidence that suggests that in an unorganized screening setting the benefits (reduced mortality) are smaller while the harms (overdiagnosis) are larger [3]. In a sense this is logical. Evidence on how best to screen can merely come from scientific studies. Per definition the only means of performing better than the best available evidence is by pure coincidence (there might be a system behind the 'coincidence' in which case new evidence can be synthesized, but this can only be proven better after a process of empirical validation). As unorganized screening is often not strictly according to best protocol as provided by the evidence [3-6], it can at best be equally good. This pleads for an organized screening program once screening is elected by an individual, to maximize the benefits and minimize the harms.

Treatment

Treatment is essential for screening to be effective. After all if screening would not subsequently be followed by treatment it would simply advance diagnosis but not delay death. Screening is based on the principle that treatment is more effective if the tumor is diagnosed at an earlier stage. Taken to the extreme, if there is a treatment for metastasized prostate cancer that would cure in 100% of cases, the screening algorithm itself would not necessarily be ineffective, but it would be rather abundant as, compared to no screening, there would be no additional benefit. Subsequently, if improvements are made in the treatment of earlier stage prostate cancer, the observed effect of screening would also become larger.

In chapter 3 and 4 of this thesis we questioned whether there were differences in treatment between the two arms of the screening study and whether this affected the outcome of screening as observed in the screening trial. Two key observations were made. First, there was an improvement in treatment in the control arm of the study (chapter 3), most likely resulting from progressive insight in how best to treat prostate cancer over time. Second, despite these differences and differences in the treatment modality itself between the screening and control arm, the effect of treatment on the observed benefit of screening was estimated to be very low (<10% of changes in prostate cancer mortality)(chapter 4). Thus, changes in treatment which occurred during the course of the screening study (and are reflected by differences in treatment between the arms) are currently not a major determinant for the positive effect of screening in real life.

Interestingly the studies in chapter 3 and 4 also provided us with information on how the actual screening process affects disease specific mortality. This obviously warrants further discussion, but before doing so some other internal factors that underlie the mechanisms of mortality reduction through screening will be discussed.

INTERNAL FACTORS THAT DETERMINE HOW SCREENING WORKS

Internal factors control the effect of the screening algorithm itself. By determining which factors are most influential, efforts could be focused on improving these factors to maximize the return on resource and energy spent.

Screening is based on the assumption that treatment is more effective if applied at an earlier stage. This makes the effect of screening not only dependent on treatment but also on the natural development of prostate cancer (i.e. the 'stages' it runs through, and at which it can be detected). Three models of the natural history of cancer are commonly considered [7, 8]. The first describes the natural history as a stepwise process with the cancer originating from one point within the organ, always progressing through distinct steps before distant systemic spread (A). The second assumes the disease not to go through distinct steps but to be systemic directly (or at least at the time it is detectable with currently available diagnostic tools)(B). The last considers the likelihood of systemic spread (i.e. metastasis) as a direct function of tumor size, with a constant tumor growth rate (C). Although these models seem to be very distinct representations of how a tumor matures, they can all be considered variances of a continuous spectrum describing tumor development and progression, and describing the effect of intervening (detection and treatment) in this process. This is illustrated in figure 1a. On the x-axis the time from tumor development until death is plotted, on the y-axis the average time (or probability) death can be delayed by intervention (diagnosis and treatment). In the graph the last model is represented by line C (systematic spread as continuous function of time), the second to last by line B (immediate systemic spread), and the first model by line A (in this case with two distinct steps through which the tumor progresses). But essentially any line connecting the beginning and the end is a possible representation of how tumor development interacts with intervention. As described above treatment will affect these lines, so for further discussion on internal factors we assume it to be a constant (as the effect of treatment differences was small this is at the moment a reasonable representation of real life). Two internal parameters of screening determine where screening interacts in this figure of tumor development; timing of screening (starting time and frequency), and the method used for early detection. The latter limits how close to the point of tumor origin screening can interact, starting time and frequency where screening is likely to interact between this limit and prostate cancer death (figure 1b). As any line, or combination of lines, (representing a different tumor model) can be considered possible, we first need to estimate which percentage of all tumors are likely to have progressed to a state where little benefit of diagnosis and treatment

can be expected before they can be picked up with current detection methods. This group of tumors will only benefit from screening if we change the method of detection to pick them out at an even earlier stage. (In the following section this group will be named group B (figure 1c)).

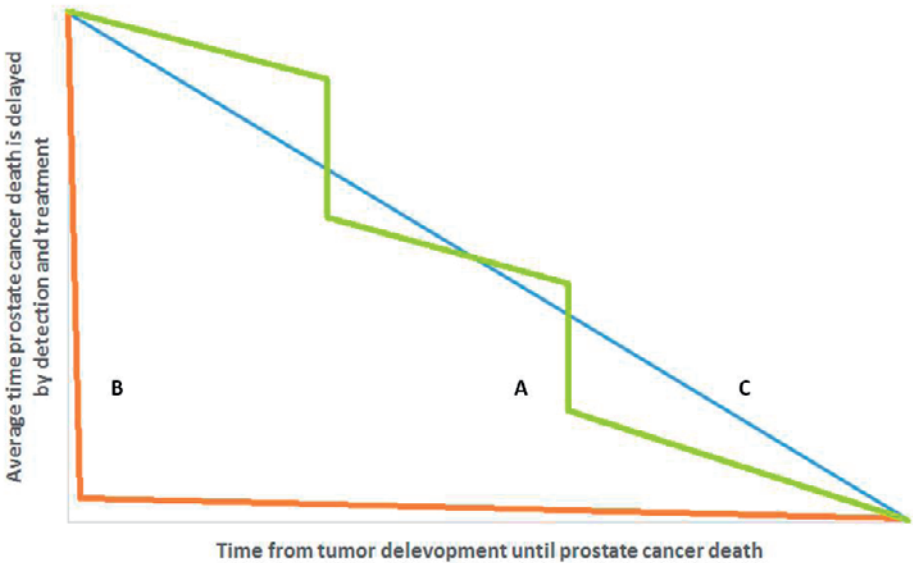


Figure 1a. The effect of detection and treatment on prostate cancer specific survival during the natural development of prostate cancer. Three examples (lines A through C) are given of possible tumor progression. *On the y-axis 'average time prostate cancer death is delayed' is interchangeable with 'probability prostate cancer death is delayed'/'percentage of men in which prostate cancer is delayed'.

Method of detection

As screening is proven to reduce mortality we can say that men in whom prostate cancer death is avoided by screening do not belong to group B. Taken the most intensive screening program tested (ERSPC Sweden [3, 9]), and maximum correction for noncompliance (chapter 2), this percentage can be estimated at 35%-51% of all prostate cancer deaths in absence of screening. Question is then which percentage of men remaining (49%-61%) do belong to group B? The current detection method of screening is based on identification by PSA (>3 ng/ml) and subsequent diagnosis by biopsy. Starting with PSA, every men who died of prostate cancer while diagnosed and treated with a PSA <3 ng/ml can be considered part of group B. In the screening arm of the ERSPC Rotterdam this were 4 men (data not shown), 2% of expected prostate cancer deaths in absence of screening at the time. (Even if men with a PSA of up to 6 ng/ml were considered (these men might still be 'incurable' if detected with half their PSA) the percentage would still be 10%.) Based on these data and having used PSA as method of detection, group B can be estimated to be relatively small. It must be noted that this is a rough estimation, based on strong assumptions. A study to directly test this estimate would however be very difficult to conduct.

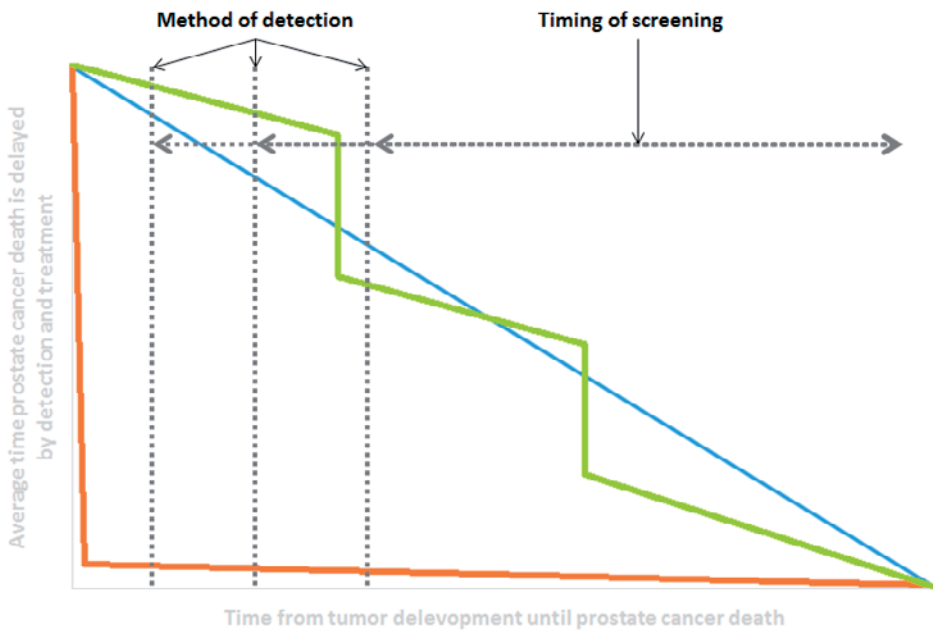


Figure 1b. Two parameters of screening (method of detection and timing of screening) that determine where screening is likely to interact during tumor development.

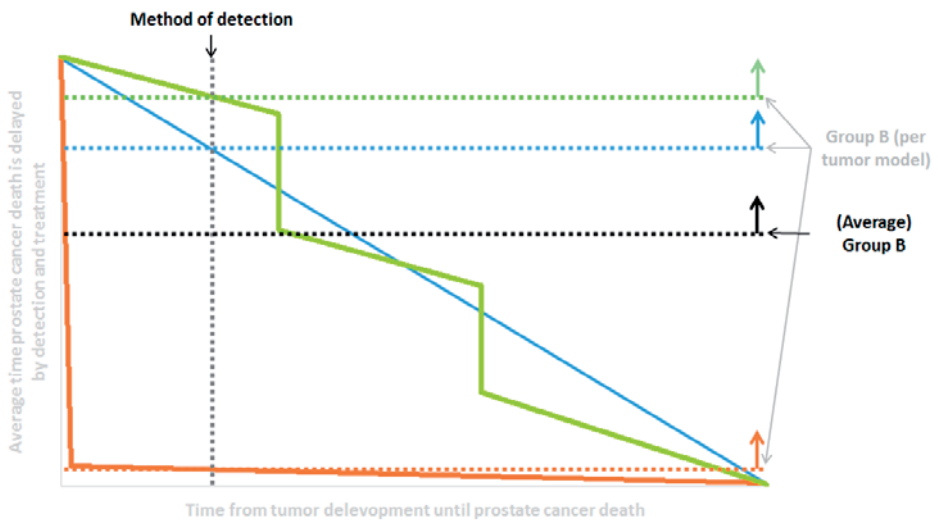


Figure 1c. Group B is defined as the percentage of men that do not benefit any more once they are detectable.

One could frequently test PSA from a very young age and remove every prostate if PSA rose above 3 ng/ml: the remaining prostate cancer deaths are group B.



The second part of early prostate cancer detection is diagnosis by systematic biopsy, which is known to miss cancer in some cases [10]. In chapter 1 we observed however that only a small proportion of cancers were detected after a previous negative biopsy (12% of high grade cancers). The majority of men were diagnosed on first biopsy. This was either in the first screening round in men with often higher PSA values, or in men with PSA rising above the biopsy threshold for the first time in subsequent screening rounds. A previous study concluded that <10% of observed prostate cancer deaths in the screening arm of the ERSPC Rotterdam at the time were in men with a previous negative biopsy and were thus potentially missed on first biopsy [11]. Taken together with the percentage of men missed by a PSA cut-off of 3 ng/ml, up to 12% of men can be considered to belong to group B. These men can potentially only benefit from screening if changes are made to the method of detection. The benefit of simply lowering the PSA threshold or increasing the biopsy intensity will however likely be disproportional to the additional harms. Focusing on group B will thus only be worthwhile if substantial improvements are made to the methods of detection.

The remaining men not belonging to group B (37%-49%) will benefit from changes in starting time and frequency of screening as this increases the likelihood of earlier intervention. So what is the most optimal timing of screening?

Timing of screening

A substantial part of men with cancer were diagnosed in the first screening round (chapter 1). If specifically looked at those men that have died of prostate cancer within the available follow-up time, >50% were diagnosed in the first screening round [12]. This suggests that for some men screening simply started too late. In the ERSPC Rotterdam the age that screening started was between 55 and 74 years [13]. To test whether the starting age should be even younger than 55 we analyze data of the 55-59 year age group at the start of the screening trial and estimated the percentage of men (out of all prostate cancer deaths) that were already 'beyond cure' at that time [14]. This percentage turned out to be relatively low (<10%-20%), implying 55-59 is a reasonable starting age in order to maximize benefit (the percentage of men 'beyond cure' for the age group 60-64, 65-69, and 70-74 are estimated using the same method at 16%-29%, 31%-64%, and 49%-70% respectively (data not shown)). Obviously the age limit at time of inclusion in ERSPC limits further analyses but one could argue that of the <10%-20% with potential benefit of starting before the age of 55 a relatively large part will belong to group B (aggressive growing tumors with quick systemic spread), the effect of starting earlier will then become increasingly smaller. Given a starting age we can focus on what the subsequent ideal frequency of testing should be.

The frequency chosen for PSA testing and biopsy in the ERSPC Rotterdam was 4 years ending if aged 75 or above. As the frequency of testing in the Swedish part of the ERSPC was 2 years, both algorithms could be directly compared [15]. Based on these data it was concluded that doubling the testing frequency indeed resulted in a stage shift with less advance (40% reduction) and more early cancers detected. Unfortunately this cannot directly be translated into the

proportion of prostate cancer deaths averted. A rough estimation based simply on the difference in mortality reduction indicates approximately up to 10% of all prostate cancer deaths benefit from a doubling in testing frequency from 4 to 2 years [13, 16]. Whether a further reduction of time between screens would result in a similar reduction of prostate cancer deaths is dependent on the average natural course. If the average course tends more towards line C (linear model) in figure 1 an equal reduction in mortality can be expected for every equal reduction in screening interval. If the average course tends more towards line A (stepwise model) in figure 1 with similar times to sharp decreases in benefit, there might be an optimal screening interval. So can we make a better estimation on the average natural course of the cancers detected by screening?

The natural course of prostate cancers detected by screening

In chapter 3, figure 1 (reprinted here as figure 2), we illustrated a model of how screening could result in mortality reduction. The figure is in fact similar to figure 1, but only depicts a stepwise tumor model with the most important steps being represented by distinct clinical characteristics (defined as low-risk, intermediate-risk, high-risk, and metastatic disease). A comparison was made in disease free survival (defined as no evidence of recurrent disease after treatment with curative intent) between these groups (arrow 1, figure 2), but also within the groups themselves by comparing the control arm and screening arm (arrow 2, figure 2). Where in the latter comparison the difference between arms represent a difference in time of detection within a specific group. Based on the data it could be concluded that there was indeed an indication for a more stepwise model (instead of a linear model), as the difference in disease free survival between groups was much larger than that within groups but with earlier detection. This is supported by the results of the study described in chapter 4. Here we specifically looked at the effect on prostate cancer mortality of changes in stage of detection. It was concluded that changes in the stage of detection almost completely (94%) account for the changes seen in prostate cancer mortality. Most prostate cancer deaths avoided by the current screening algorithm thus seem to tend towards a stepwise development (if this also accounts for the men that were not saved by screening cannot be inferred directly).

If screening works by shifting the stage at which the cancer is detected it would be interesting to know if every shift in stage is equally important in terms of the benefit on survival. If e.g. a shift from the metastatic group to the high risk group provides the largest improvement we could focus on detecting cancers in the high risk stage. If every step contributes equally then the most benefit is obtained by detecting the cancer in the earliest stage (in this case as low risk) and focusing on detection in higher stages would result in an subsequent decrease in benefit. It is unfortunately not possible to directly observe which prostate cancer deaths are prevented by screening and at which stage they are diagnosed and would have been diagnosed if not screened. But, there are indirect signs that provide some information on their origin, some of which can be found in figure 3A, chapter 4 (reprinted here as figure 3).

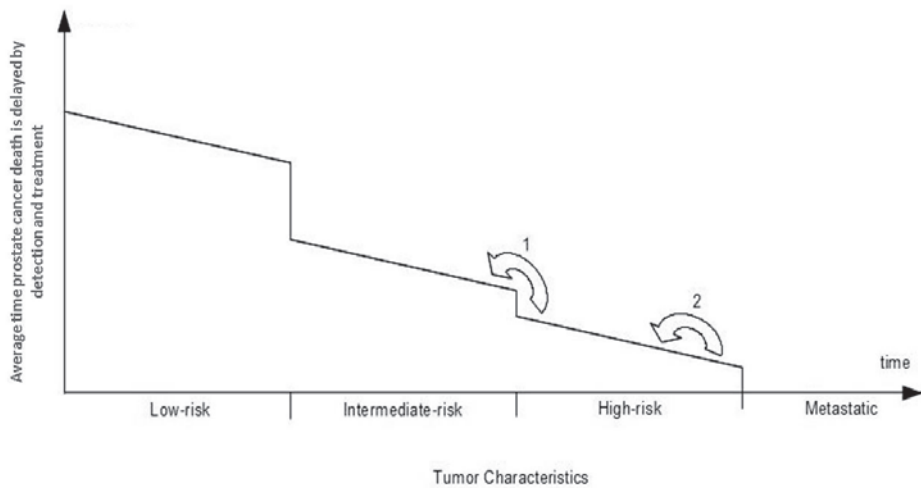


Figure 2. Model on how screening could result in better outcome (adapted from figure 1, chapter 3)

- 1: Screening could result in a stage shift, resulting in better prognosis.
- 2: If screening did not result in a clinical apparent stage shift, earlier diagnosis and treatment in time could still have resulted in a better prognosis, for instance because of less time for the tumor to develop outside the prostate.

To explain the origin of this evidence consider the following example (depicted in figure 4a-c): Suppose there are two distinct stages in which cancers can be diagnosed, called M- and M+, with cancers always progressing from the first to the last. In the absence of screening 200 men would be diagnosed with cancer, 100 as M- and 100 as M+, and 60 would die as a result, 10 diagnosed as M- and 50 diagnosed as M+ (figure 4a). Now we introduce an early detection program, but do not subsequently treat the men that are detected earlier (the absolute number of prostate cancer deaths thus remains similar). As a result 50 men (25 of which died of prostate cancer) previously diagnosed as M+ are now diagnosed as M- (figure 4b). Finally we also decide to immediately treat men who were diagnosed at an earlier stage and as a result 20 of the 25 men who previously died of prostate cancer (but shifted from the M+ group to the M- group) do not so any more (figure 4c). Suppose we were only in the possession of the data in figure 4a and 4c of the example, can we then still work out that exactly 20 men were 'saved' by early detection in the M- instead of M+ group? The answer is yes, if we assume that stage shift with subsequent early treatment is the only method to do so (notice that the increase in deaths in the M- group is an indirect indication of deaths being prevented by shifting to that group). This is of course a simplified example. Directly applying its principles to real life (figure 3) is not as straight forward. (Additional factors should be accounted for and further assumptions should be made to make a realistic estimation on the risk group in which men 'saved' by prostate cancer screening were diagnosed.) But looking at figure 3 it can already be seen that the largest increase in mortality in the screening arm was in the low risk group, indicating that at least some of the prostate cancer deaths prevented were diagnosed at this stage.

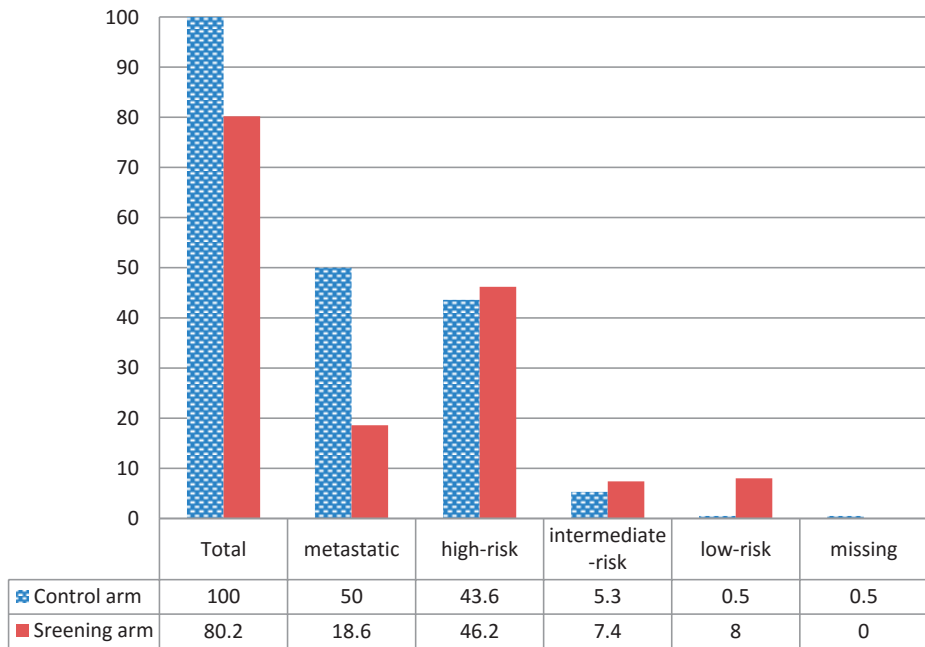


Figure 3. Distribution of prostate cancer mortality in the screening and control arm per risk group. For comparison data were indexed on the total prostate cancer mortality in the control arm. (reprinted from figure 3a, chapter 4)

HOW TO USE THIS INFORMATION TO OPTIMIZE SCREENING ALGORITHMS IN THE FUTURE?

One of the biggest challenges in prostate cancer screening today is to decrease its harms. The most important harms being overdiagnosis and subsequent overtreatment of very early stage prostate cancers. However, as set out above, due to the natural tumor development increasingly earlier detection will also expand the benefit (prostate cancer mortality reduction) of screening. More benefit thus seems to equal more harm. This implies that we either should find a way to distinguish overdiagnosed cancers from those that are not, accept some loss of benefit at the expense of reduced harm (preferable in an uneven ratio, as otherwise there is no overall gain), or deal with the effect of overdiagnosis in another matter. The latter could be achieved by e.g. reducing the side effects of detection and treatment, or reducing overtreatment altogether, although then the initial paradox might remain.

So, can we distinguish cancers that are overdiagnosed from those who are not? The answer is likely no, at least not upfront. We might be able to identify cancers that are definitively not overdiagnosed (i.e. as symptoms occur and we are thus too late), but the remaining men can still be both. The reason being that we can never definitively know how the cancer would have behaved or will behave in the future if it was not diagnosed. Furthermore, overdiagnosis is not

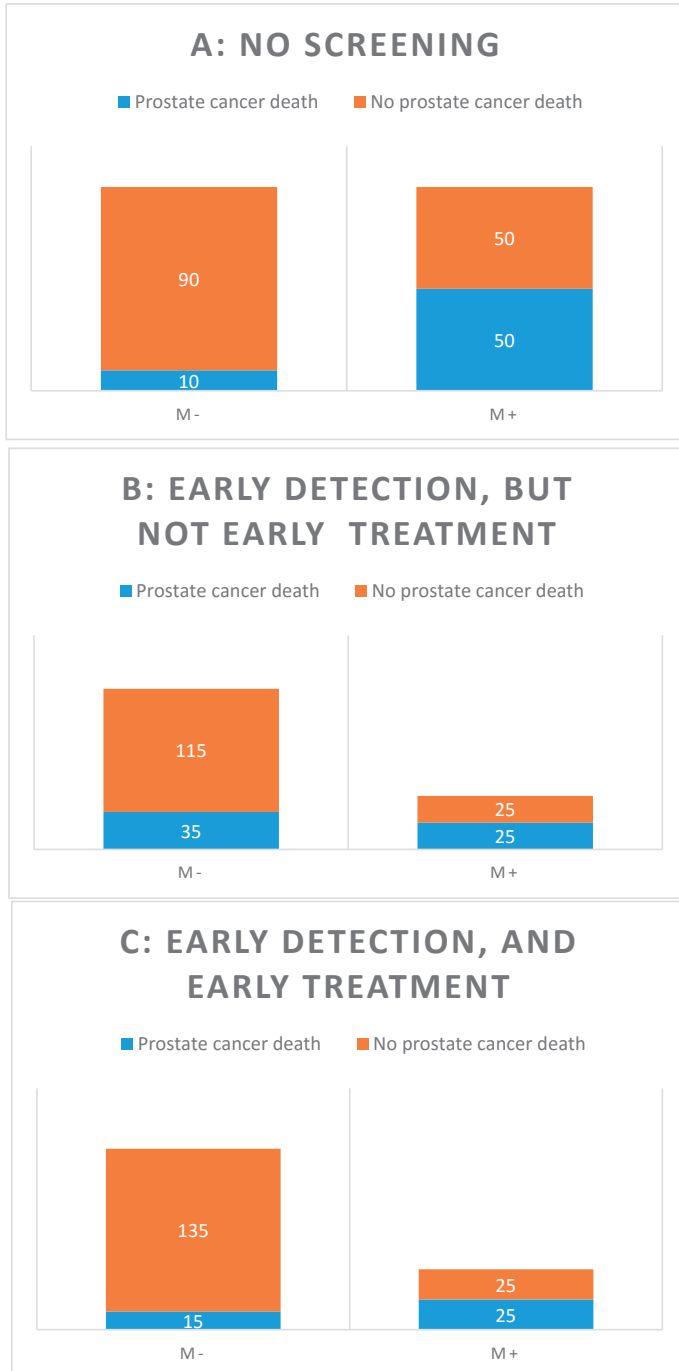


Figure 4. Example of effect of stage shift and treatment on distribution of prostate cancer and mortality

only dependent on the natural course of the tumor but also on life expectancy of its host, again a factor that cannot be ascertained upfront.

What we can do is strive to make the most accurate prediction on the likelihood of the cancer not causing symptoms if directly diagnosed and treated and weigh this against the chance of loss in benefit if not detected and treated directly. Those with high probability of being overdiagnosed might be good candidates to be monitored (either before or after diagnosis (after diagnosis is discussed below)). This will ideally result in maximization of harm reduction, at the cost of minimal loss to profit (lower mortality). One of the best currently available predictors of a cancer being overdiagnosed is the (modified) Gleason score. It has been shown that men with Gleason scores of 6, if radically treated, have a probability of close to zero to still giving rise to symptoms [17-19]. While the tumor remains at this stage diagnosis and treatment might thus not be necessary. In fact, reevaluation of biopsy Gleason score showed upgrading close to 50% in those men that died of prostate cancer, but were diagnosed in the lowest risk group in the ERSPC Rotterdam (20% in men who did not die) [20]. Predicting the presence of Gleason scores above 6 can already be done with fairly good accuracy based on currently available clinical characteristics [21-23]. Improvements should come from more accurate prostate sampling, e.g. by targeting biopsies using MRI [24]. As the negative predictive value of current risk prediction tools seem already quite good [25], the added value of MRI might mainly be in men with increased risk according to these risk prediction tools. The Gleason 6 group could potentially be extended with Gleason 3+4 that possess certain favorable characteristics (e.g. certain growth patterns). The second focus should then be on the optimal monitoring frequency in men based on their predicted risk (e.g. as shown in [26]). Development, validation, but above all implementation of tools to stratify men depending on risk for more or less frequent screening programs cannot only potentially reduce overdiagnosis, but also reduce unnecessary biopsies (close to 90% of secondary biopsies are negative (chapter 1)). Evaluation should be according to the change in the balance of harms to benefits.

New predictors, once found and validated, can then with relative ease be combined with existing predictors to improve stratification for monitoring frequency. Although one should be cautious to have too high expectations on any new predictor. An example of this is the, commonly referred to, 'golden marker'. This marker is said to selectively identify those tumors that will result in prostate cancer death, or at least systemic spread. But how valuable will such a marker be? Consider the following: Imagine different human cells all have specific properties, which are determined by DNA. The cells divide to produce new cells. Normally this goes well, but sometimes alterations in the DNA occur (e.g. by deletions, mutations, activations, rearrangements, etc.) which change the properties of the new cell. At some point there might occur an alteration that provides the new cell with the property of systemic spread (likely a combination of several other properties). From this point forward this property can be measured in the form of a marker (i.e. the 'golden marker'). But what is the value of knowing this property at this point? Wouldn't the cell with the property of systemic spread immediately spread systemically? There is after all

nothing any more that holds it back. The measurement and intervention should thus be at the moment it occurs to have any effect. Before the systemic spread property is present, we can by internal observation at best say something about the probability of it occurring based on the current properties of the cells involved (e.g. a larger tumor volume equals more division and thus a higher chance of this specific property occurring). In this view the presence of the 'golden marker' will thus not be very useful as it will only tell us that we are already too late for curative treatment, while its absence will merely tell us it might or might not be present sometime in the future. (Knowing systemic spread has occurred could possibly prevent organ specific treatment with its side effects, although the alternative treatment (systematic treatment) is often not with less harm. Furthermore there is evidence that suggests that treating the primary cancer source might still cause slowing of overall cancer growth [27].) Instead of searching for the 'golden marker' it might be better to focus on markers that can improve risk stratification at an earlier stage.

This section discussed monitoring before a diagnosis is made, but monitoring after diagnosis might also be a viable option to reduce the harms of detection. This monitoring is commonly referred to as active surveillance.

ACTIVE SURVEILLANCE (ARE WE ABLE TO SELECTIVELY FILTER OUT MEN WITH AGGRESSIVE DISEASE?)

Active surveillance might aid in addressing the paradox discussed above that by early prostate cancer detection we will reduce the chance of prostate cancer death, but simultaneously increase the risk of detecting cancers that are no threat to a man's health. In a sense this section is therefore a direct continuation of that discussed above. As we are likely not able to immediately distinguish those cancers that are overdiagnosed and those that are not, some form of monitoring should be in place to selectively filter out those that tend towards more aggressive growth. But first we should make a selection who should be monitored and for whom this is too risky and would thus be better served with direct invasive treatment.

SELECTION OF MEN FOR ACTIVE SURVEILLANCE

As discussed above Gleason score 6 tumors, if directly treated, have almost zero chance of still causing symptoms. This indicated that while the tumor remains at this grade treatment can safely be postponed. Gleason 6 thus seems a good starting point for selection of men for active surveillance, as with every increase in Gleason there might be a potentially loss of benefits from early detection. Unfortunately, the presence of solely Gleason 6 tumor can currently only be firmly established after the prostate is removed (the primary reason being under sampling on

diagnostic biopsy). We thus have to make the best possible estimation on the probability that there is indeed only Gleason 6 tumor present.

At the start of active surveillance (one to two decades ago) several predictors were selected for this aim, PSA (density), tumor extent on biopsy (either number of biopsy cores positive or a quantification of involve of a single core), and tumor stage. Often they are combined in a crude risk estimation using rule based selection criteria (chapter 5). A better solution seems however the use of risk prediction tools that provide a man's individual risk. These risk prediction tools are not necessarily better in terms of selecting men (although certainly not worse) [28], but provide more accurate information for an individual to make a decision (chapter 5). Preferably this information should be combined in a decision aid incorporating personal weights attached to the risks of harm and benefit. This is however easier said than done. Research shows that humans are poor in prediction their reaction and feelings to events they have not yet experienced (e.g. how badly a side effect will influence their daily quality of life) [29]. There thus seems a place for increasing research in how patients could be aided to make more accurate predictions on actual reactions to future events. Examples of methods that could be used are narratives of people that already have experience with such events [30], or the use of tests that determine e.g. how 'risk avoiding' patients in general are and correlate this to reactions on certain outcomes. Although there is currently little evidence to support its use, it can be theorized that new technologies such as virtual reality might be of help as they offer patients the opportunity of 'virtually' experiencing future events, increasing their ability to estimate its influence on personal wellbeing [31].

To correctly inform men we need to know how well the initially selected active surveillance entry criteria predict the presence of Gleason 6 on radical prostatectomy. The best way of testing this is to immediately operate men who fulfill the active surveillance criteria and correlate this with the radical prostatectomy outcome. If all men with Gleason 6 on biopsy would be selected for active surveillance 35%-50% will turn out to have Gleason >6 [28, 32]. If men fulfilled the PRIAS criteria this dropped to 25%-38% [28, 32]. However, this also meant that a substantial part of men were not selected by the criteria, and that the majority still had Gleason 6 on radical prostatectomy. The area under the ROC curve was therefore moderate for the PRIAS criteria at 0.6 (other active surveillance selection criteria had similar results) [28]. Of the individual parameters of the inclusion criteria only PSA, PSA density and T-stage had some predictive ability [32]. It might therefore be reasonable to expand the selection criteria and search for methods of better predicting the presence of Gleason score >6 on radical prostatectomy. In chapter 6 we proposed the use of MRI with targeted biopsies instead of ≤ 2 cores positive for prostate cancer to safely increase the number of men that can select active surveillance. The reason being that MRI with targeted biopsies has a high negative predictive value for the presence of more aggressive disease [33]. A small radical prostatectomy series of men with diagnostic MRI indeed showed a lower rate of upgrading in men with negative MRI, although still not perfect [34]. Prediction models incorporating MRI results with classic predictors might offer the best result in near future. New markers could be added if proven to increase accuracy. Currently especially blood/urine/

prostate fluid biomarkers seem the best candidates as tissue biomarkers are likely still hampered by under sampling (chapter 8), although the initially results of currently available markers offer only moderate additional value [35].

FOLLOW-UP OF MEN ON ACTIVE SURVEILLANCE

In theory follow-up should selectively filter out those that progress to a more aggressive disease state (e.g. Gleason >6). Preferably this should be done just before the tumor truly reaches such a state, and thus there must be a prediction made on the chance of it occurring in the short term to adapt the follow-up frequency accordingly. Currently follow-up seems mostly in place to find those tumors that were not adequately picked out at inclusion. Although further improvement in selection criteria might change this, there likely remains a place for confirmation of the results found at entrance. Timing of this confirmation test could be discussed, but might be primarily dependent on personal patient preference (i.e. some might want direct confirmation, others might prefer to postpone). Research indicates that postponing radical treatment up to one year in men selected for active surveillance does not change outcome [36], and hence men can safely choose any time for confirmatory testing within this period.

So, if after confirmation only men with favorable tumor features are selected, what is then the subsequent rate of progression over time? Two modeling studies addressed grade progression in prostate cancer, one before detection (in a screening setting) [37], and one during active surveillance [38]. Both concluded that there is grade progression over time, but only the latter also estimated the rate of progression (12%-24% over a 10 year time period) [38]. It can immediately be appreciated that for the majority of men (>75%) follow-up should thus not result in a change of treatment and preferably be kept to a minimum. In chapter 9 we studied a switch to active treatment in the first 10 years of active surveillance in the PRIAS study. Although one should also consider the imperfect selection (with 25%-38% under grading), the percentage of men that switched to active treatment in 10 years (60%) was somewhat higher than can be expected based on the estimated progression rates. Especially considering the fact that of these men 35% still presented with Gleason 6 on radical prostatectomy specimen (chapter 9), and more men were advised to discontinue active surveillance, but did not adhere to this recommendation (chapter 6). I.e., the recommendations used to advise a switch to active treatment are too stringent, while on the other hand potentially missing aggressive cancers (34% of men presented with unfavorable outcome on radical prostatectomy), urging the need for improvements. Two of the initial criteria used to recommend active treatment that were most frequent in occurrence, but simultaneously had the highest rate of men with Gleason 6 on radical prostatectomy, were >2 cores positive and a PSA-doubling time (PSA-DT) between 0-3 years (chapter 9 and 6). Instead of a direct switch to active treatment we therefore proposed stricter follow-up (in the form of an MRI with targeted biopsies), as there was an increased risk the presence of higher risk disease.

This change could reduce the number of men that directly switch to active treatment by up to 30% to 50% (chapter 9 and 6), lessened with the number of men that are subsequently upgraded after MRI. Some other refinements were proposed, and it can be theorized that in the future even more will change. The reason that the number of positive core was used as selection tool is that it represents tumor volume (which in itself can be used to predict tumor aggressiveness). Unfortunately, the number of positive cores seems more dependent on variations in biopsy placing between biopsies sessions than actual changes in tumor volume (the same could be said on tumor involvement in a single non-targeted biopsy core). Lesion size on MRI could potentially correlate better with actual tumor volume and might thus replace the number of positive cores as surrogate for tumor extend and growth. But MRI offers other possibilities. Not only does it produce rough anatomical maps, it can also be used to get more in-depth knowledge on functional aspects of cellular structure (e.g. diffusion weighted [39] and contrast enhanced imaging [40]), or even information on a cellular level (e.g. oxygenation status, and chemical composition [41, 42]). Much of these aspects are currently understudied in prostate cancer and active surveillance [33]. MRI might be supplemented by other imaging modalities that provide information on chemical composition and potentially behavior (e.g. nuclear imaging). If more firmly established, biopsies might then only serve as tool to confirm what was seen on the images.

More noninvasive testing in general seem to be the main focus for improving how patients experience active surveillance. Currently follow-up in men on active surveillance is for a substantial part based on repeat biopsies. They are however not without side effects, such as infection, hematuria, and discomfort (chapter 7). In fact, this resulted in men not receiving the biopsies recommended by the protocol (chapter 7), with overall compliance dropping substantially during active surveillance (chapter 6). As active surveillance is meant to reduce the side effect of prostate cancer treatment as much as possible, some trade off in the form of loss in immediate detection of higher grade cancer could be acceptable for men. Active surveillance protocols might even evolve into truly personalized management strategies in which the choice, timing, and even intensity (e.g. systematic yes or no) of the biopsy (to confirm presence of absence of higher grade disease) is up to the patient based on his weighing of risks (discussed above are some caveats that should be considered in such an approach). Research should thus aim at providing the most accurate dynamic risk predictions to base decisions on. This can already be done using available clinical characteristics, soon supplemented by MRI, and in the future with every new marker detected that provides additional predictive value.

EPILOGUE

This thesis set out to gain better insight in our understanding of early detection of prostate cancer in reducing disease specific mortality, while preventing related harms as much as possible. Information that in combination with existing knowledge, can be used to well-inform every men

deciding on what is best for this health now and in the future. As with all scientific information it is however part of an ongoing process of providing the best possible current state of knowledge. By design this means that the best state of knowledge today might not necessarily be that of tomorrow. Although this seems unfortunate, it is only by showing that our current theories are false that we are truly able to learn something new and can provoke ourselves to think of a better one.

REFERENCES

1. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014;384:2027-35.
2. Heidenreich A, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, et al. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol*. 2014;65:124-37.
3. Arnsrud Godtman R, Holmberg E, Lilja H, Stranne J, Hugosson J. Opportunistic Testing Versus Organized Prostate-specific Antigen Screening: Outcome After 18 Years in the Goteborg Randomized Population-based Prostate Cancer Screening Trial. *Eur Urol*. 2015;68:354-60.
4. Jessen K, Sondergaard J, Larsen PV, Thomsen JL. Danish General Practitioners' Use of Prostate-Specific Antigen in Opportunistic Screening for Prostate Cancer: A Survey Comprising 174 GPs. *Int J Family Med*. 2013;2013:540707.
5. D'Ambrosio GG, Campo S, Cancian M, Pecchioli S, Mazzaglia G. Opportunistic prostate-specific antigen screening in Italy: 6 years of monitoring from the Italian general practice database. *Eur J Cancer Prev*. 2010;19:413-6.
6. Bokhorst L, Roobol M. Biopsy rates after PSA testing at the general practitioner are low, even at high PSA values. *Urology*. 2014;84:S380.
7. Welch HG, Gorski DH, Albertsen PC. Trends in Metastatic Breast and Prostate Cancer--Lessons in Cancer Dynamics. *N Engl J Med*. 2015;373:1685-7.
8. Hellman S. Karnofsky Memorial Lecture. Natural history of small breast cancers. *J Clin Oncol*. 1994;12:2229-34.
9. Roobol MJ. Unorganized prostate-specific antigen-based screening for prostate cancer: more harm than benefit. When will we finally start to implement guidelines and risk assessment tools in clinical practice? *Eur Urol*. 2015;68:363-4.
10. Chun FK, Epstein JI, Ficarra V, Freedland SJ, Montironi R, Montorsi F, et al. Optimizing performance and interpretation of prostate biopsy: a critical analysis of the literature. *Eur Urol*. 2010;58:851-64.
11. Schröder FH, van den Bergh RC, Wolters T, van Leeuwen PJ, Bangma CH, van der Kwast TH, et al. Eleven-Year Outcome of Patients with Prostate Cancers Diagnosed During Screening After Initial Negative Sextant Biopsies. *Eur Urol*. 2010;57:258-66.
12. Zhu X, van Leeuwen PJ, Bul M, Bangma CH, Roobol MJ, Schroder FH. Identifying and characterizing "escapes"-men who develop metastases or die from prostate cancer despite screening (ERSPC, section Rotterdam). *Int J Cancer*. 2011;129:2847-54.
13. Roobol MJ, Kranse R, Bangma CH, van Leenders AG, Blijenberg BG, van Schaik RH, et al. Screening for Prostate Cancer: Results of the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol*. 2013;64:530-9.
14. Bokhorst LP, Alberts A, Bangma CH, Schroder FH, Roobol MJ. Should we start screening men for prostate cancer before the age of 55? (Presented at EMUC 2015, Barcelona). 2015.
15. van Leeuwen PJ, Roobol MJ, Kranse R, Zappa M, Carlsson S, Bul M, et al. Towards an optimal interval for prostate cancer screening. *Eur Urol*. 2012;61:171-6.
16. Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, et al. Mortality results from the Goteborg randomised population-based prostate-cancer screening trial. *Lancet Oncol*. 2010;11:725-32.

17. Kweldam CF, Wildhagen MF, Bangma CH, van Leenders GJ. Disease-specific death and metastasis do not occur in patients with Gleason score ≤ 6 at radical prostatectomy. *BJU international*. 2015;116:230-5.
18. Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) ≤ 6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol*. 2012;36:1346-52.
19. Haffner MC, Mosbruger T, Esopi DM, Fedor H, Heaphy CM, Walker DA, et al. Tracking the clonal origin of lethal prostate cancer. *J Clin Invest*. 2013;123:4918-22.
20. Bokhorst LP, Kweldam CF, Bangma CH, Schröder FH, Van Der Kwast TH, Leenders GJLH, et al. Gleason score reclassification of men with classical Gleason score ≤ 6 who eventually died of prostate cancer in the ERSPC Rotterdam. *European Urology, Supplements*. 2015;14:e115-ea.
21. Roobol MJ, Steyerberg EW, Kranse R, Wolters T, van den Bergh RC, Bangma CH, et al. A Risk-Based Strategy Improves Prostate-Specific Antigen-Driven Detection of Prostate Cancer. *Eur Urol*. 2010;57:79-85.
22. Roobol MJ, Schröder FH, Crawford ED, Freedland SJ, Sartor AO, Fleshner N, et al. A framework for the identification of men at increased risk for prostate cancer. *The Journal of urology*. 2009;182:2112-20.
23. Roobol MJ, Carlsson SV. Risk stratification in prostate cancer screening. *Nat Rev Urol*. 2013;10:38-48.
24. Schoots IG, Roobol MJ, Nieboer D, Bangma CH, Steyerberg EW, Hunink MG. Magnetic resonance imaging-targeted biopsy may enhance the diagnostic accuracy of significant prostate cancer detection compared to standard transrectal ultrasound-guided biopsy: a systematic review and meta-analysis. *Eur Urol*. 2015;68:438-50.
25. Alberts AR, Schoots IG, Bokhorst LP, van Leenders GJ, Bangma CH, Roobol MJ. Risk-based Patient Selection for Magnetic Resonance Imaging-targeted Prostate Biopsy after Negative Transrectal Ultrasound-guided Random Biopsy Avoids Unnecessary Magnetic Resonance Imaging Scans. *Eur Urol*. 2015.
26. Roobol MJ, Zhu X, Schroder FH, van Leenders GJ, van Schaik RH, Bangma CH, et al. A Calculator for Prostate Cancer Risk 4 Years After an Initially Negative Screen: Findings from ERSPC Rotterdam. *Eur Urol*. 2013;63:627-33.
27. Widmark A, Klepp O, Solberg A, Damber JE, Angelsen A, Fransson P, et al. 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.
28. Venderbos LD, Roobol MJ, Bangma CH, van den Bergh RC, Bokhorst LP, Nieboer D, et al. Rule-based versus probabilistic selection for active surveillance using three definitions of insignificant prostate cancer. *World J Urol*. 2015.
29. Kahneman D, Snell J. Predicting a changing taste: do people know what they will like? *Journal of Behavioral Decision making*. 1992;5:187-200.
30. Dillard AJ, Fagerlin A, Dal Cin S, Zikmund-Fisher BJ, Ubel PA. Narratives that address affective forecasting errors reduce perceived barriers to colorectal cancer screening. *Soc Sci Med*. 2010;71:45-52.
31. Chirico A, Lucidi F, De Laurentiis M, Milanese C, Napoli A, Giordano A. Virtual Reality in Health System: Beyond Entertainment. A Mini-Review on the Efficacy of VR During Cancer Treatment. *J Cell Physiol*. 2016;231:275-87.
32. Vellekoop A, Loeb S, Folkvaljon Y, Stattin P. Population based study of predictors of adverse pathology among candidates for active surveillance with Gleason 6 prostate cancer. *The Journal of urology*. 2014;191:350-7.
33. Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz L, et al. Magnetic Resonance Imaging in Active Surveillance of Prostate Cancer: A Systematic Review. *Eur Urol*. 2014;67:627-36.

34. Lee SH, Koo KC, Lee DH, Chung BH. Nonvisible tumors on multiparametric magnetic resonance imaging does not predict low-risk prostate cancer. *Prostate Int.* 2015;3:127-31.
35. Porpiglia F, Cantiello F, De Luca S, Manfredi M, Veltri A, Russo F, et al. In parallel comparative evaluation between multiparametric mri, pca3 and phi in predicting pathologically confirmed significant prostate cancer in men eligible for active surveillance. *BJU international.* 2015.
36. van den Bergh RC, Albertsen PC, Bangma CH, Freedland SJ, Graefen M, Vickers A, et al. Timing of Curative Treatment for Prostate Cancer: A Systematic Review. *Eur Urol.* 2013;64:204-15.
37. Draisma G, Postma R, Schroder FH, van der Kwast TH, de Koning HJ. Gleason score, age and screening: modeling dedifferentiation in prostate cancer. *Int J Cancer.* 2006;119:2366-71.
38. Inoue LY, Trock BJ, Partin AW, Carter HB, Etzioni R. Modeling grade progression in an active surveillance study. *Stat Med.* 2014;33:930-9.
39. Hambrock T, Somford DM, Huisman HJ, van Oort IM, Witjes JA, Hulsbergen-van de Kaa CA, et al. Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. *Radiology.* 2011;259:453-61.
40. Chen YJ, Chu WC, Pu YS, Chueh SC, Shun CT, Tseng WY. Washout gradient in dynamic contrast-enhanced MRI is associated with tumor aggressiveness of prostate cancer. *J Magn Reson Imaging.* 2012.
41. Giskeodegard GF, Bertilsson H, Selnaes KM, Wright AJ, Bathen TF, Viset T, et al. Spermine and citrate as metabolic biomarkers for assessing prostate cancer aggressiveness. *PLoS One.* 2013;8:e62375.
42. Kobus T, Hambrock T, Hulsbergen-van de Kaa CA, Wright AJ, Barentsz JO, Heerschap A, et al. In vivo assessment of prostate cancer aggressiveness using magnetic resonance spectroscopic imaging at 3 T with an endorectal coil. *Eur Urol.* 2011;60:1074-80.

Samenvatting

Dit proefschrift heeft als doel om verdere informatie te verschaffen over screening op, en een actief afwachtend beleid bij, prostaat­kanker. Twee hoofdvragen worden behandeld: “Wat zijn de mechanismen die leiden tot de waargenomen daling van prostaat­kanker sterfte bij screening?” (Deel 1, screening) en “Zijn we in staat zijn om bij een actief afwachtend beleid selectief mannen met een agressieve ziekte te identificeren?” (deel 2, active surveillance). Om deze twee hoofdvragen te beantwoorden worden een aantal sub-vragen behandeld. Hieronder worden de belangrijkste bevindingen per sub-vraag besproken:

DEEL 1, SCREENING:

- Wat voor typen prostaat­kanker worden opgespoord en op welk moment tijdens de screening? (Hoofdstuk 1)

Kanker detectie blijft stabiel gedurende elkaar opvolgende screeningsrondes bij mannen eerder biopt. Desondanks werd de meerderheid van de hoog risico kankers detecteert in de eerste screeningsronde. De stabiele kanker detectie lijkt ten komen door diagnose van kanker bij mannen die bij hun vorige onderzoek een PSA-waarden onder de biopsie afkap hadden. Slechts een klein deel van de gedetecteerde kankers is gevonden in mannen met een eerder negatief biopt. Dit ondersteund de suggestie dat de langdurige verhoging van het PSA bij deze mannen waarschijnlijk niet het gevolg is van prostaat­kanker.

- Wat is het effect van het niet naleven van het screeningsprotocol op de waargenomen reductie van prostaat­kanker sterfte in een screeningsstudie? (Hoofdstuk 2)

Vooraf ongewenste screening in de controlegroep van de Europese screenings studie had een groot verdunnend effect op de waargenomen mortaliteitsreductie. Dit indiceert dat het effect van screening op het risico van overlijden aan prostaat­kanker voor een individu (aanzienlijk) groter is dan wat gevonden wordt in de totale onderzoekspopulatie.

- Wat is het effect van verschillen in behandeling tussen de armen van de Europese screening studie op de gevonden mortaliteitsreductie? (Hoofdstuk 3 en 4)

Vanwege latere detectie in tijd konden tumoren in de controlegroep profiteren van voortschrijdend inzicht in de beste manier prostaat­kanker te behandeling (hoofdstuk 3). Het effect van verschillen in de behandeling tussen de armen op de waargenomen prostaat­kanker specifieke sterftereductie was echter klein (hoofdstuk 4).

-
- Hoe komt het dat screening op prostaatkanker resulteert in een reductie van sterfte? (Hoofdstuk 1-4)

De reductie van prostaatkankersterfte lijkt afkomstig uit alle screeningsrondes (hoofdstuk 1) en is het resultaat van verschuiving naar detectie van tumoren in een vroeger stadium (hoofdstuk 4). Verschillen in behandeling tussen de armen had weinig effect op de waargenomen sterftereductie (hoofdstuk 3 en 4).

DEEL 2, ACTIVE SURVEILLANCE:

- Zijn we in staat om selectief mannen met meer agressieve ziekte te identificeren door middel van individuele risicofratificatie bij inclusie voor een actief afwachtend beleid? (Hoofdstuk 5)

Selecteren van potentiële kandidaten voor een actief afwachtend beleid op basis van individuele risicofratificatie helpt in tegenstelling tot selectie op basis van vaste criteria om een meer bewuste en persoonlijke behandelkeuze te maken. Individuele risicofratificatie is echter nog niet accuraat genoeg om agressieve van niet-agressieve ziekte te onderscheiden. Toekomstig onderzoek moet zich richten op het vergrijpen van informatie om een betere risico inschatting te kunnen maken.

- Zijn we in staat om selectief mannen met een agressieve ziekte te identificeren op basis van informatie afkomstig van pathologie en biomarkers? (hoofdstuk 8)

De immunohistochemische biomarkers die werden getest (Ki67, ERG en P27) en herevaluatie van het biopsie pathologieverslag waren niet in staat om een betere inschatting te maken van de ziekte agressiviteit. Undersampling van de meest agressieve tumor fragmenten met systematische biopten lijkt daarvoor de belangrijkste reden.

- Is prostaatbiopsie de beste methode om agressieve ziekte op te sporen, d.w.z. wat zijn de nadelen? (Hoofdstuk 6 en 7)

Prostaatbiopsieën leiden tot complicaties in een substantieel deel van de mannen (hoofdstuk 7) en vormen een aanzienlijke last bij mannen op een actief afwachtend beleid. Dit is mogelijk een reden dat veel van de aanbevolen biopten tijdens de follow-up niet worden ondergaan (hoofdstuk 6). Daarnaast leidt het merendeel van de biopten niet tot een verandering van het beleid. Het aantal vervolgbiopten in mannen op een actief afwachtend beleid zou dus omlaag moeten.

-
- Moeten we het huidige follow-up protocol voor mannen op een actief afwachtend beleid veranderen? (Hoofdstuk 9)

Na tien jaar onderstrepen de resultaten van de PRIAS studie de veiligheid van een actief afwachtend beleid. Maar een aanzienlijk deel van de mannen dat volgens protocol overstapt naar een actieve behandeling blijken op definitieve pathologie geen aanwijzingen te hebben voor een agressieve ziekte. De follow-up protocol moet dus worden aangepast om de hoeveelheid mannen dat onterecht stopt met een actief afwachtend beleid te verminderen. Op basis van de beschikbare gegevens worden in dit proefschrift verschillende voorstellen gedaan om het protocol voor een actief afwachtend beleid nu en in de toekomst te verbeteren.

ABOUT THE AUTHOR

Leonard Pieter Bokhorst was born in Rheden on July 11, 1988. He completed high school at the 'Stedelijk Gymnasium Arnhem', after which he started medical school at the University of Utrecht. In 2013 he obtained his medical degree after which he started with PhD research at the department of Urology, Erasmus Medical Center Rotterdam. His main topics were active surveillance and screening for prostate cancer. Currently, he is working as a resident at the department of Radiation Oncology, University Medical Center Utrecht.

LIST OF PUBLICATIONS

[1] Bokhorst LP, Zappa M, Carlsson SV, et al. Correlation between stage shift and differences in mortality in the European Randomized study of Screening for Prostate Cancer (ERSPC). *BJU international*. 2016.

[2] Bruinsma SM, Bangma CH, Carroll PR, et al. Active surveillance for prostate cancer: a narrative review of clinical guidelines. *Nature reviews Urology*. 2016;13(3):151-67.

[3] Bokhorst LP, Lepisto I, Kakehi Y, et al. Complications after prostate biopsies in men on active surveillance and its effect on receiving further biopsies in the Prostate cancer Research International: Active Surveillance (PRIAS) study. *BJU international*. 2016.

[4] Alberts AR, Schoots IG, Bokhorst LP, van Leenders GJ, Bangma CH, Roobol MJ. Risk-based Patient Selection for Magnetic Resonance Imaging-targeted Prostate Biopsy after Negative Transrectal Ultrasound-guided Random Biopsy Avoids Unnecessary Magnetic Resonance Imaging Scans. *European urology*. 2015.

[5] Bruinsma SM, Bokhorst LP, Roobol MJ, Bangma CH. How Often is Biopsy Necessary in Patients with Prostate Cancer on Active Surveillance? *The Journal of urology*. 2016;195(1):11-2.

[6] Bokhorst LP, Roobol MJ. Ethnicity and prostate cancer: the way to solve the screening problem? *BMC Med*. 2015;13:179.

[7] Venderbos LD, Roobol MJ, Bangma CH, van den Bergh RC, Bokhorst LP, Nieboer D, et al. Rule-based versus probabilistic selection for active surveillance using three definitions of insignificant prostate cancer. *World J Urol*. 2015.

[8] Alberts AR, Bokhorst LP, Roobol MJ. [Guidelines on the early detection of prostate cancer]. *Ned Tijdschr Geneesk*. 2015;159:A8811.

[9] Bokhorst LP, Alberts AR, Rannikko A, Valdagni R, Pickles T, Kakehi Y, et al. Compliance Rates with the Prostate Cancer Research International Active Surveillance (PRIAS) Protocol and Disease Reclassification in Noncompliers. *Eur Urol*. 2015.

[10] Bokhorst LP, Venderbos LD, Schroder FH, Bangma CH, Steyerberg EW, Roobol MJ. Do Treatment Differences between Arms Affect the Main Outcome of ERSPC Rotterdam? *J Urol*. 2015;194:336-42.

-
- [11] Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz L, et al. Magnetic resonance imaging in active surveillance of prostate cancer: a systematic review. *Eur Urol.* 2015;67:627-36.
- [12] Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet.* 2014;384:2027-35.
- [13] Bokhorst LP, Roobol MJ, Kranse R. Reply from Authors re: Sigrid V. Carlsson, Peter C. Albertsen. Better Survival After Curative Treatment for Screen-detected Prostate Cancer Compared with Clinical Diagnosis: A Real Effect or Lead-time Bias? *Eur Urol.* 2015;68(2):183-4: Better Treatment in the Control Arm of the ERSPC Rotterdam: A Point Worth Noting? *Eur Urol.* 2014;68:184-5.
- [14] Bokhorst LP, Kranse R, Venderbos LD, Salman JW, van Leenders GJ, Schroder FH, et al. Differences in Treatment and Outcome After Treatment with Curative Intent in the Screening and Control Arms of the ERSPC Rotterdam. *Eur Urol.* 2015;68:179-82.
- [15] Bokhorst LP, Steyerberg EW, Roobol MJ. Decision support for low-risk prostate cancer. (Chapter 24; Prostate cancer: science and clinical practice, Mydlo JH, Godec CJ. 2015.)
- [16] Roobol MJ, Bokhorst LP. The ProtecT trial: what can we expect? *Lancet Oncol.* 2014;15:1046-7.
- [17] Bokhorst LP, Moss SM, Roobol MJ. Reply from Authors re: Chris Metcalfe. Can the Results of the European Randomized Study of Screening for Prostate Cancer Be Decontaminated? *Eur Urol.* 2014;65(2):337-8: Yes, by Remaining Conservative in Our Assumptions. *Eur Urol.* 2013;65:338-39.
- [18] Bokhorst LP, Bangma CH, van Leenders GJ, Lous JJ, Moss SM, Schroder FH, et al. Prostate-specific Antigen-Based Prostate Cancer Screening: Reduction of Prostate Cancer Mortality After Correction for Nonattendance and Contamination in the Rotterdam Section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol.* 2014;65:329-36.
- [19] Roobol MJ, Kranse R, Bangma CH, Otto SJ, van der Kwast TH, Bokhorst LP, et al. Reply from Authors re: Michael Baum. Screening for Prostate Cancer: Can We Learn from the Mistakes of the Breast Screening Experience? *Eur Urol.* 2013;64:540-1: Screening for Prostate Cancer: We Have Learned and Are Still Learning. *Eur Urol.* 2013;64:541-3.
- [20] Venderbos LD, Bokhorst LP, Bangma CH, Roobol MJ. Active surveillance: oncologic outcome. *Curr Opin Urol.* 2013;23:268-72.

[21] Bokhorst LP, Zhu X, Bul M, Bangma CH, Schroder FH, Roobol MJ. Positive predictive value of prostate biopsy indicated by prostate-specific-antigen-based prostate cancer screening: trends over time in a European randomized trial*. *BJU Int.* 2012;110:1654-60.

[22] Bokhorst LP, Zegers RH. [Couching then and now] Staarsteek vroeger en nu. *Ned Tijdschr Geneeskd.* 2011;155:A3283.

DANKWOORD

Het dankwoord is waarschijnlijk het meest gelezen deel van een proefschrift, ik vrees soms het enige. Maar is dat niet zonde? Zonde van alle moeite die het de onderzoeker gekost heeft om de resultaten goed en leesbaar op schrift te krijgen, zonde van het papier en zonde van de tijd die in de rest is gaan zitten? Is bovendien de weergave van het eigenlijke onderzoek en wetenschappelijke discussie in de andere hoofdstukken niet vele malen interessanter dan een opsomming van namen en verdiensten? Ik denk het wel en daarom volsta ik hier met een korte tekst, zodat ook de lezer met drukke agenda geen excuus heeft om de overige hoofdstukken terzijde te leggen (mijn dank is er niet minder om). Naast mijn promotoren, Chris en Ewout, wil ik vooral Monique bedanken (echt voor alles) en natuurlijk de overige mensen die het screeningsbureau maken tot wat het is: Conja en Marlies (voor de gezelligheid), Lionne, Arnout, Jolanda, Elaine en Frank-Jan (kamerogenoten door de jaren), Prof. Schröder en Meneer Knol ("vast meubilair" in de goede zin van het woord), Sophie en Lisette (onderzoekers op de 17e verdieping), Mathijs, Robert, Meelan, Xiaoye en Roderick (voorgangers en mede onderzoekers), Ries (voor de gedachtevorming), Arno (voor een kijkje door de microscoop), members of the ERSPC and the PRIAS steering committee, patiënten en overige personen die bijdroegen en/of bijdragen aan de PRIAS en ERSPC studies, de grote en kleine commissie, Wouter en Woody, nieuwe collega's en tot slot familie, vrienden en natuurlijk Berdine.

Gratias agimus tibi!

PhD PORTFOLIO

Name	Leonard P. Bokhorst
PhD period	January 2013 – December 2015
Erasmus MC department	Urology
Promotors	Prof. dr. C.H. Bangma, Prof. dr. E.W. Steyerberg
Supervisor	Dr. M.J. Roobol

PhD training	Year	Workload (ECTS)
General courses		
Biostatistical Methods I: Basic Principles	2013	5.7
Biostatistical Methods II: Classical Regression Models	2013	4.3
Research Integrity	2014	0.3
'Basiscurcus klinisch onderzoekers (BROK)'	2014	1
Seminars and workshops		
Workshop, targeted prostate biopsies, Paris	2013	0.5
Department of Urology journal club	2013-2015	1
Department of Urology internal course	2013-2015	1
Department of Urology PhD meeting	2013-2015	0.5
Symposium Urological tumors IKNL	2013-2015	0.5
Presentations		
Annual meeting EAU, Milan	2013	0.5
Andrology and Oncology seminar, Ehrenhausen	2013	1
ERSPC meeting, Göteborg	2013	0.5
Prostatakarzinom, Heidelberg	2014	1
Annual meeting EAU, Stockholm	2014	1
Annual meeting AUA, Orlando	2014	0.5
Annual meeting SIU, Glasgow	2014	0.5
ERSPC meeting, Antwerp	2014	0.5
Pelvic happiness for pelvic cancer patients, Lisbon	2015	1
Externe refereeravond Erasmus MC, Rotterdam	2015	0.5
Annual meeting EAU, Madrid	2015	1
Annual meeting AUA, New Orleans	2015	0.5
ERSPC meeting, Madrid	2015	0.5
EMUC, Barcelona	2015	1
Symposium 'Een leven lang screenen', Venlo	2015	0.5
ESO conference active surveillance, Milan	2016	1

Conferences

Annual meeting EAU	2013-2015	1.5
Annual meeting AUA	2012,2014,2015	1.5
Annual meeting ERSPC	2013-2015	1.5
SIU	2014	0.5
EMUC	2015	0.5
ESO conference active surveillance	2014,2016	1
Global congress on prostate cancer	2012	0.5
NVU meeting	2013-2015	0.5

Teaching

Klinisch chemici in opleiding (PSA kinetiek in active surveillance)	2013,2014	1
Prostaat echografie huisartsen	2014	0.5
Lichamelijk onderzoek co-assistenten	2014,2015	0.5

Total ECTS		35.8
-------------------	--	-------------
