Prostate Cancer Screening

The effect on prostate cancer mortality and incidence

Pim J. van Leeuwen

Explanation of cover illustration: At first glance, deciding whether to get the PSA screening test for prostate cancer seems to be pretty straightforward and attractive. It's a simple blood test that can pick up the prostate cancer long before your symptoms appear. After all, your prostate cancer is earlier treated resulting in cure and better outcome. Therefore prostate cancer screening seems to be suitable for public commercials. However, many of the cancers that will be detected by screening are so slow-growing that might never cause problems during mens life. Moreover, their diagnosis by biopsy, and treatment, might be worse than the disease itself. Therefore, prostate cancer screening looks favourable; however, watch out for the prostate cancer bomb.

Prostate Cancer Screening: the effect on prostate cancer mortality and incidence

Email: p.vanleeuwen@erasmusmc.nl

ISBN 978-94-6169-227-6

Cover design: Ernst Mutgeert
Lay out: Optima Grafisch Communicatie
Printed by: Optima Grafisch Communicatie

This thesis is based on articles published in different scientific journals. The exact wording of the text of this thesis may differ from the text of the published version of the articles due to editorial changes.

PROSTATE CANCER SCREENING

The effect on prostate cancer mortality and incidence

Prostaatkanker screening

Het effect op de prostaatkanker sterfte en incidentie

Proefschrift

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

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

De openbare verdediging zal plaatsvinden op

woensdag 16 mei 2012 om 15.30 uur

door

PIM JOHANNES VAN LEEUWEN
geboren te Heemstede

ERASMUS UNIVERSITEIT ROTTERDAM

PROMOTIECOMMISSIE

Promotor:	Prof.dr. C.H. Bangma
Overige leden:	Prof.dr. L.A.L.M. Kiemeney
	Prof.dr. H.J. de Koning
	Prof.dr. E.J. Kuipers
	Prof.dr. T.M. de Reijke
Co-Promotoren:	Dr. M.J. Roobol
	Prof.dr. F.H. Schröder

The studies reported in this thesis were performed at the department of Urology of the Erasmus Medical Centre Rotterdam, the Netherlands and at the Northern Ireland Cancer Registry Belfast, Northern Ireland. The studies form an integrated part of the European Randomized study of Screening for Prostate Cancer (ERSPC).

The ERSPC Rotterdam is supported by grants from the Dutch Cancer Society (KWF 94-869, 98-1657, 2002-277, 2006-3518), The Bonnema foundation, The Netherlands Organization for Health Research and Development (002822820, 22000106, 50-50110-98-311), 6th Framework Program of the EU: P-Mark: LSHC-CT-2004-503011, Beckman Coulter Hybritech Inc and of Europe against Cancer (SOC 95 35109, SOC 96 201869 05F02, SOC 97 201329, SOC 98 32241). The ERSPC received Erasmus MC and Ministry of Health institutional review board approval.

The Northern Ireland Cancer Registry Belfast, Northern Ireland is funded by the Public Health Agency, Northern Ireland.

Financial support for the reproduction of this thesis was generously funded by the following institutions/companies (alphabetically): Abbot, Amgen, Astellas, Bayer Nederland, Beckman Coulter, Coloplast, Chipsoft, Erasmus University Rotterdam, Eurocept Groep, Ferring, GlaxoSmithKline, IPSEN, J.E. Jurriaanse Stichting, Novartis, Sanofi Aventis, Stichting Urologisch Wetenschappelijk Onderzoek, Stichting Wetenschappelijk Onderzoek Prostaatkanker, Zambon.

The whole of science is nothing more than a refinement of everyday thinking.
Albert Einstein 1879-1955

Contents

LIST OF ABBREVIATIONS

AS	Active Surveillance
AUC	Area Under the Curve
BPH	Benign Prostate Hyperplasia
DRE	Digital Rectal Examination
EAU	European Association of Urology
e.g.	Exampli Gratia
ERSPC	European Randomized study of Screening for Prostate Cancer
i.e.	Id Est
LUTS	Lower Urinary Tract Symptoms
NI	Northern Ireland
NNI	Number Needed to Investigate
NNS	Number Needed to Screen
NNT	Number Needed to Treat
NPV	Negative Predictive Value
OR	Odds Ratio
PC	Prostate Cancer
PCa	Prostate Cancer
PCA3	Prostate Specific Antigen 3
PLCO	Prostate, Lung, Colorectal and Ovary cancer
PPV	Positive Predictive Value
PRIAS	Prostate cancer Research International Active Surveilance
PSA	Prostate Specific Antigen
PSADT	Prostate Specific Antigen Doubling Time
PSAV	Prostate Specific Antigen Velocity
QOL	Quality of Life
QALY	Quality Adjusted Life Year
RCT	Randomized Controlled Trial
REDUCE	REduction by DUtasteride of prostate Cancer Events
ROC	Receiver Operator Characteristic
RP	Radical Prostatectomy
RR	Relative Risk
SPCG	Scandinavian Prostate Cancer Group
TRUS	Transrectal Ultrasound
TURP	Trans-Urethral Resection of the Prostate
WHO	World Health Organization
5ARI	5-Alfa Reductase Inhibitors


CRUISIN' FOR SCREENING
PROSTATE CANCER

Part I

General Introduction


CRUISIN' FOR SCREENING
PROSTATE CANCER

1

The Prostate

Pim J. van Leeuwen

1.1 INTRODUCTION

The prostate is a gland in men which is located immediately below the bladder and just in front of the bowels (Figure 1)[1]. Its main function is to produce a fluid that usually constitutes 25-30% of the volume of the semen. Additionally, the prostate contains some smooth muscles that help expel semen during ejaculation. In younger men the prostate is about the size of a walnut, however, the prostate usually enlarges with age. Growth of the prostate gland can eventually result in the most common prostate problem, i.e., benign prostatic hyperplasia (BPH). Other health problems of the prostate gland are prostatitis and prostate cancer (PC).

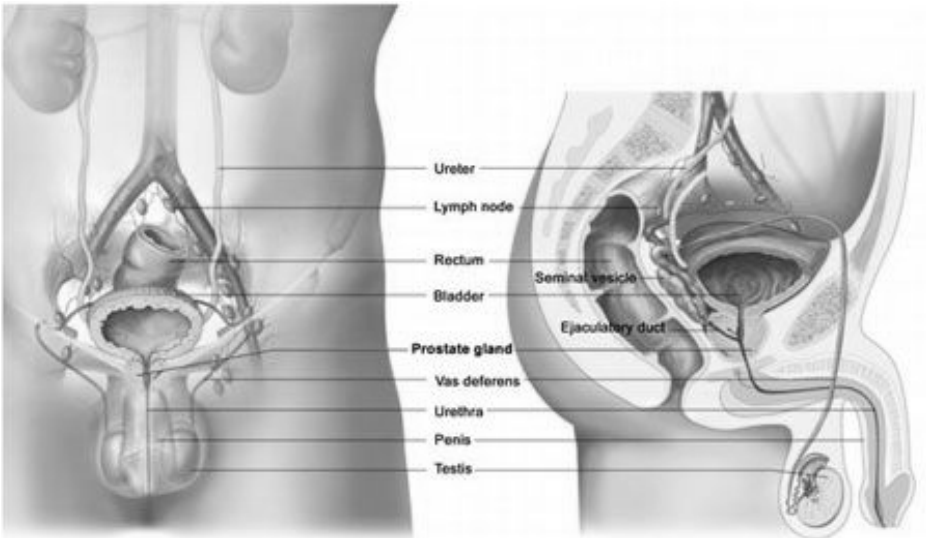


Figure 1. Anatomy of the male reproductive and urinary systems, showing the prostate, urethra, seminal vesicle, ejaculatory duct, testicles, bladder, and other organs.

1.2. BENIGN PROSTATIC HYPERPLASIA (BPH)

Benign Prostatic Hyperplasia (BPH) is a common problem in older aged men. BPH is present in more than 50% of men aged over 60 years[2]. Between 15% and 30% of these men have lower urinary tract symptoms (LUTS), however, not all such symptoms are caused by the hyperplasia, and many are attributable to various types of dysfunction of smooth muscle (detrusor) of the bladder[3]. LUTS due to BPH occur if the enlargement is sufficient to squeeze the urethra which passes through the prostate. The main complaints consist of awaking frequently at night to urinate, sudden or urgent need to urinate, difficulty in starting to urinate and slow flow of urine and difficulty in stopping[4]. BPH is a benign

condition, however, which can be treated effectively. Treatment of BPH may require specific drugs, or, in more developed cases, an operation to widen the urethral passage. The most common procedure is a transurethral resection of the prostate (TURP). Under a general anaesthetic, an instrument is passed up the urethra through the penis and some of the prostate is removed to improve urine flow.

1.3. PROSTATITIS

Prostatitis is a benign condition which is caused by usually bacterial inflammation of the prostate. It can cause discomfort deep inside the pelvis continuously, when passing urine or during ejaculation. It can be painful and can spread to other areas of the pelvis. Other complaints could consist of a sudden or urgent need to urinate, discomfort during urinating, painful ejaculation, presence of blood in the urine or semen. If the prostatitis is caused by an infection it will be treated with antibiotics.

1.4. PROSTATE CANCER

Prostate cancer (PC) is the most commonly diagnosed malignancy (excluding non-melanoma skin cancer) and the third leading cause of death from cancer in men in Western countries (after lung and colorectal cancer)[5-6]. The lifetime risk of a PC diagnosis is 15.8% for an individual man in the United States and approximately 9% for a man in Western Europe[7-9]. The lifetime risk of dying from PC is low relative to the lifetime risk of a PC diagnosis, i.e. 2.8% in the United States and 3.1% in Western Europe[7-9]. Overall, these incidence and mortality rates give PC an important public health relevance[10].

The introduction and widespread use of prostate specific antigen (PSA) testing for the early detection of PC has led to major changes in PC incidence, the tumour grade and stage at diagnosis, treatment, and the mortality from PC over the past two decades. The effects of PSA testing and PC screening are described in chapter 2 and 3. An increase in PSA testing has increased the detection of PC during the last two decades, and have lead to the diagnosis of cancers that rather should not have been diagnosed, as their detection and subsequent treatment is unlikely to benefit patients, or even might harm them. Related to this, the term 'overdiagnosis' is used. The clinical definition of overdiagnosis is the diagnosis of a tumour that would otherwise remain clinically unrecognized until the individual died from other causes. Overdiagnosis occurs when screening detects small tumours that would otherwise remain clinically unrecognized until the individual dies from other causes. Overdiagnosis appears to be especially harmful when it results

in invasive treatment of the tumours that would unlikely to be harmful. This is called overtreatment.

1.4.1 Symptoms and signs

As the majority of PC originates in the peripheral zone of the prostate gland (Figure 2), they tend not to cause local symptoms until they are large tumours[11]. Invasion of a PC into the urethra or bladder neck may cause lower urinary tract symptoms, which are usually voiding symptoms such as poor flow, straining or hesitancy although storage symptoms such as frequency, nocturia or urgency can also occur[12]. In the pre-PSA era, more men presented with evidence of locally advanced or metastatic disease such as back or pelvic pain due to bony metastasis, lower limb oedema due to lymph node involvement or compression of the iliac veins, or with renal failure due to bilateral ureteric obstruction[13]. Rare presentations include spinal cord compression or disseminated intravascular coagulation[14-15].

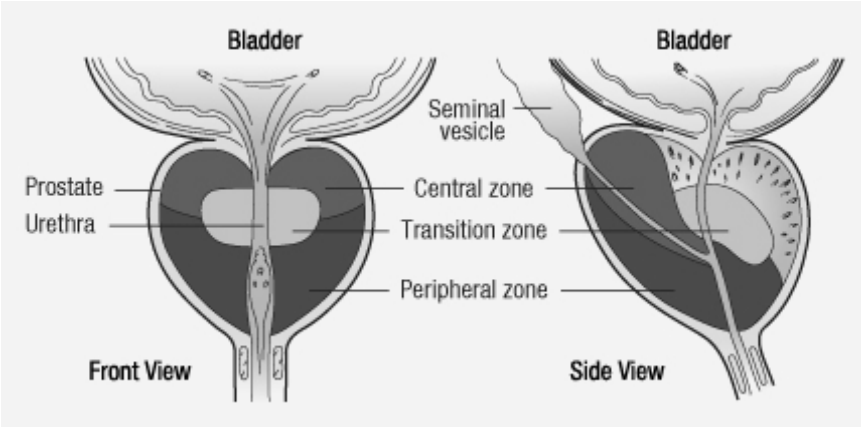


Figure 2. Anatomy prostate; PC originate 70-80% of cases in peripheral zone, 10-20% in transition zone, 2.5% in central zone, obtained from[18]

Apart from the signs and symptoms mentioned above, there are few other physical signs associated with localised PC. As many tumours arise in the peripheral zone, they may be palpated by digital rectal examination (DRE). In advanced disease, physical signs may be related to local invasion such as haematuria, impotence or lower urinary tract obstruction. Symptoms can also be related to metastatic spread such as pathological fracture or lower limb oedema. With the widespread use of PSA testing, the majority of men are asymptomatic at diagnosis after having been investigated due to an abnormal PSA result[7, 16-17].

1.4.2 Diagnosis of prostate cancer

DRE, PSA and transrectal ultrasound (TRUS) are the three main modalities for the early detection of PC.

1.4.2.1 *DRE*

DRE is possible due to the prostate anatomic position in the pelvis, with easy access for palpation using a finger placed per rectum (Figure 3). Prior to the introduction of PSA testing, DRE was the primary modality for diagnosing prostate cancers. Jewett found that approximately 50% of palpable prostate nodules were diagnosed as prostate cancers on prostate biopsy[18]. However, DRE findings are only moderately reproducible, even amongst experienced urologists[19-20]. Further, DRE tended to diagnose PC when they are pathologically advanced and therefore less likely to be curable by radical prostatectomy[21-22]. The value of DRE as an early detection modality will be discussed in more detail in chapter 2.

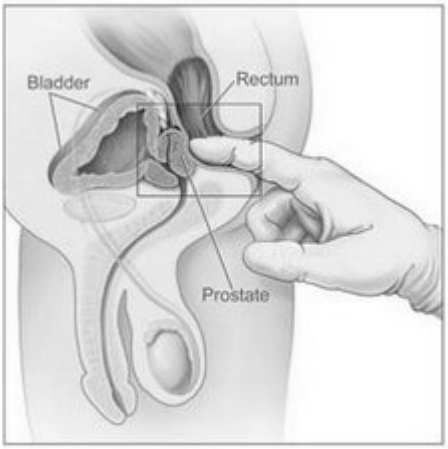


Figure 3. Digital Rectal Examination of the Prostate, obtained from[24]

1.4.2.2 *Prostate Specific Antigen*

PSA is a protease, almost exclusively produced by the prostatic ductal and acinar epithelium, which may leak into the blood stream[23]. PSA is in widespread use in urological oncology as a marker for the risk assessment before diagnosis and for the oncologic evaluation after treatment. An increased serum PSA level indicates an increased prostate cancer risk. However, an increase in the serum PSA may also have other causes: 1) an increase in normal prostate glands like BPH or 2) an increased leakage of PSA into the bloodstream due to infectious processes or obstruction[24-26]. The value of PSA for the early detection of PC will be discussed in more detail in chapter 2.

1.4.2.3 *TRUS and prostate needle biopsy*

TRUS, with associated prostate needle biopsy, has shown increasing popularity since the mid-1980s and is considered a gold standard for investigation of PC. With the development of improved high-frequency transducers, the prostate can be visualised with greater resolution to aid in the identification of prostate cancer. TRUS has the advantage of facilitating more accurate measurements of prostate size, which may help interpretation of PSA results[27-28]. Prior to TRUS, biopsies were digitally directed or 'blind'. Hodge et al. introduced systematic sextant biopsies which proved to be superior to blind biopsies[29].

1.4.3 Tumour Staging

The extent of the disease is classified according to the Tumour/Node/Metastasis (TNM) classification (Table 1). There are two main goals for an accurate staging of PC. Firstly, to predict overall prognosis and secondly, to select appropriate treatment based on the extent of the disease. The local stage of the tumour determined by DRE, TRUS and/or Magnetic Resonance Imaging (MRI), is known as the clinical tumour stage. The definive stage, or pathological tumour stage, can only be obtained after a radical prostatectomy.

1.4.4 Histopathology of prostate cancer

Over 95% of PC are adenocarcinomas[30]. The rare subtypes that occur tend to be more aggressive than adenocarcinomas. Mucinous adenocarcinoma and ductal adenocarcinoma occur in less than 1% of the cases but both have a more aggressive behaviour[31-32]. Transitional cell carcinomas occur in 1 to 4% of all PC and tend to be locally advanced at the time of presentation[33]. Other subtypes include small cell carcinoma and squamous cell carcinoma, both of which have a very poor prognosis[34-35].

1.4.4.1 *Tumour grade*

Tumour grade describes the degree of cellular differentiation as assessed by light microscopy. A number of grading methods has been described for PC although the majority of pathologists now use the Gleason grading system. This system gives a grade of differentiation ranging from 1 to 5, where grade 1 is very well differentiated whilst grade 5 is poorly differentiated or anaplastic (Figure 3)[36]. The Gleason score is the sum of the primary Gleason grade (the most common pattern) and the secondary grade, which is the next most common pattern (but which should comprise of greater than 5% of the total tumour tissue). In cases where only one pattern is identified, the primary grade is doubled i.e. 3+3 = 6. The Gleason score therefore ranges from 2 to 10.

Since the introduction of the Gleason grading system more than 40 years ago many aspects of PC. The system was updated at a 2005 consensus conference of interna-

Table 1. Tumour, node, metastasis (TNM) classification of prostate cancer (2009 version)

T: Evaluation of the (primary) tumour (a)	
TX: Primary tumour cannot be assessed	
T0: no evidence of tumour	
T1: tumour present, but not detectable clinically or with imaging	T1a: tumour was incidentally found in less than 5% of prostate tissue resected (for other reasons)
	T1b: tumour was incidentally found in greater than 5% of prostate tissue resected
	T1c: tumour was found in a needle biopsy performed due to an elevated serum PSA
T2: the tumour can be felt (palpated) on examination, but has not spread outside the prostate	T2a: the tumour is in half or less than half of one of the prostate gland's two lobes
	T2b: the tumour is in more than half of one lobe, but not both
	T2c: the tumour is in both lobes
T3: the tumour has spread through the prostatic capsule (if it is only part-way through, it is still T2)	T3a: the tumour has spread through the capsule on one or both sides including microscopic bladder neck involvement (b).
	T3b: the tumour has invaded one or both seminal vesicles
T4: Tumour is fixed or invades adjacent structures other than seminal vesicles: external sphincter, rectum, levator muscles, and/or pelvic wall.	
It should be stressed that the designation "T2c" implies a tumour which is *palpable* in both lobes of the prostate. Tumours which are found to be bilateral on biopsy only but which are not palpable bilaterally should not be staged as T2c.	
N: Evaluation of the regional lymph nodes (c)	
NX: cannot evaluate the regional lymph nodes	
N0: there has been no spread to the regional lymph nodes	
N1: there has been spread to the regional lymph nodes	
M: Evaluation of distant metastasis (d)	
MX: cannot evaluate distant metastasis	
M0: there is no distant metastasis	
M1: there is distant metastasis	M1a: the cancer has spread to lymph nodes beyond the regional ones
	M1b: the cancer has spread to bone
	M1c: the cancer has spread to other sites (regardless of bone involvement)

(a) Tumour found in one or both lobes by needle biopsy, but not palpable or visible by imaging, is classified as T1c.
(b) Invasion into the prostatic apex, or into (but not beyond) the prostate capsule, is not classified as pT3, but as pT2.
(c) Metastasis no larger than 0.2 cm can be designated pN1 mi.
(d) When more than one site of metastasis is present, the most advanced category should be used.

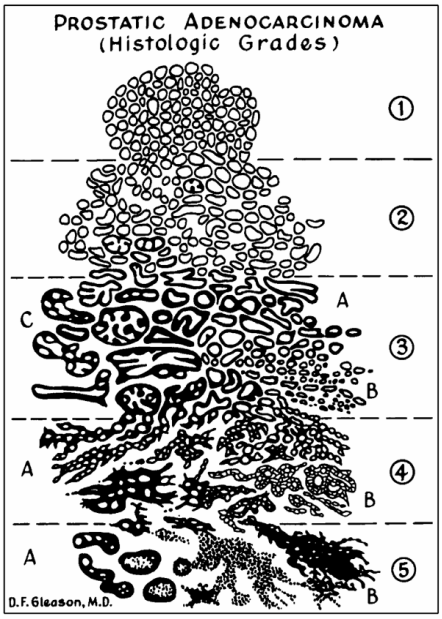


Figure 4. Gleason grading for prostate adenocarcinoma.

tional experts in urological pathology, under the auspices of the International Society of Urological Pathology. Gleason score 2–4 should rarely if ever be diagnosed on needle biopsy, certain patterns (i.e., poorly formed glands) originally considered Gleason pattern 3 are now considered Gleason pattern 4 and all cribriform cancer should be graded pattern 4[37]. These changes have resulted in disease upgrading. Comparing the original and modified Gleason systems on needle biopsy material, Gleason score 6 cancers decreased from 48.4% to 22% of the total, whereas Gleason score 7 increased from 25.5% to 67.9%[38]. For the most part Gleason's original pattern 5 has remained unchanged in modern practice. Consequently it is difficult to compare prostate cancer data sets over time because survival rates seem to improve due to changes in classification. In the older literature Gleason score 6 included admixed cases that today would be diagnosed as Gleason pattern 4, with a correspondingly worse prognosis. Today cases of Gleason score 6 are a homogenous group of tumors lacking cribriform and poorly formed glands with a better prognosis. This artificial change in prognosis has been referred to as the Will Rogers phenomenon[39].

1.4.5 Treatment

Treatment for PC depends on patient's tumour characteristics, age and co-morbidities. In the management of localised PC it is impossible to state that one therapy is clearly superior over another because of the lack of randomized controlled trials. Localised PC may be treated with radical surgery (open/ laparoscopic or robot-assisted laparoscopic

prostatectomy) or radiation therapy (external beam radiation therapy or brachytherapy). A third option for patients with localized PC is active surveillance. Several factors should be considered when selecting a treatment option. One would be of failure based on clinical stage, PSA level, and Gleason score. Specific initial therapies are recommended according to whether the risk category is low, intermediate or high, referring to the patient's risk of recurrence after therapy. Low is defined by D'Amico et al. as men with stage T1-T2a PC, a Gleason score 2-6, and a PSA value ≤ 10 ng/ml. Intermediate is defined as any T2b PC or any Gleason 7 score or PSA value between 10 and 20 ng/ml. High risk is defined as those with ≥T2c PC, a Gleason scor 8-10 or a PSA value > 20 ng/ml. Other factors to consider when recommending treatment options include patient's life expectancy, underlying medical conditions, and patient preference.

To date, it is impossible to state that one therapy for localized PC is clearly superior over another because of the lack of randomised controlled trials in this field. However, based on the available literature, some recommendations can be made. Expectant management, which is also called active surveillance consists of initially withholding radical treatment such as surgery or radiation therapy, but monitoring the disease instead according to a fixed pattern of frequent investigations[40]. When the first indications arise that disease progression occurs, the switch to radical treatment with curative intent and within the window of curability is advised. Since approximately 50% of men with screen detected PC are overdiagnosed, meaning, their cancer would never have caused any symptoms, active surveillance is an attractive approach to the management of early PC. This may save men the side effects of treatment, without compromising survival. There are no data from randomized trials on the safety and efficacy of active surveillance; however, encouraging results have been reported from several centers[41-42].

Radiation therapy is another option for the treatment of localized PC. Previously, the use of nonconformal radiation therapy necessitated lower doses of radiation to avoid an unacceptably high risk of side effects; this resulted in a higher likelihood of cancer recurrence[43]. Currently, radiation therapy is most commonly delivered by means of conformal, externally applied techniques. Either three-dimensional imaging is used to localize the prostate and the beams are shaped to match the contour of the prostate, or radioactive iodine-125 or palladium-103 seeds are implanted directly into the prostate. Prospective studies have shown that higher doses of radiation can be delivered safely with the use of conformal techniques, with better cancer control than is achieved with the use of nonconformal techniques[44-45]. The advantages of radiation therapy are that it is noninvasive or minimally invasive and it is less likely than radical prostatectomy to cause certain complications such as severe urinary incontinence[46]. In addition, radiation therapy can be used in the care of men with nonmetastatic prostate cancer of various degrees of severity, including men who are at higher risk for extraprostatic extension. There are no data from well-controlled, randomized trials comparing the treatment out-

comes of radiation therapy and surgery. Nonetheless, observational data suggest that the long-term disease control achieved with contemporary radiation therapy is similar to that achieved with radical prostatectomy[47-48].

Radical prostatectomy involves removal of the entire prostate and seminal vesicles along with sufficient surrounding tissue to obtain a negative surgical margin; this procedure is often accompanied by a bilateral pelvic lymph-node dissection. The perceived advantage of radical prostatectomy is that there is no better way to cure a cancer that is completely confined to the prostate than total surgical removal. Radical retropubic and perineal prostatectomies are performed through open incisions or laparoscopically, sometimes with robot-assisted methods. As compared with other approaches, laparoscopic approaches are associated with less blood loss during surgery, but this reduction in blood loss has not led to a reduction in the need for transfusion, nor has it led to a decrease in pain or the duration of hospitalization[49]. Observational data indicate that, as compared with earlier surgical approaches, the anatomical approach results in less blood loss, a 30-day mortality after surgery that is 10 times lower (0.2 to 0.4%), and, in the hands of an experienced surgeon using the nerve-sparing technique, reductions in the rates of the two most common surgical complications: clinically significant incontinence (3%) and impotence (30%)[50-51]. However, other estimates from studies in the United States have been less promising, with rates of incontinence as high as 74% and rates of impotence as high as 90%[52-53]. Thus, patients considering surgery should be referred to surgeons with considerable experience in order to optimize the likelihood of effective cancer control and to minimize the likelihood of complications. One randomized controlled trial is known comparing radical prostatectomy with expectant management of localized PC. It has shown improved cancer-specific survival rates in favour of radical surgery[54-55]. This study concluded that it is effective to treat men with early stage PC who have a life expectancy of at least 10 years or more. However, most of the patients in this study did not harbour screen-detected PC, so these data cannot be automatically transferred into daily routine practice.

Metastasized PC cannot be cured and will in time always lead to death unless comorbidity causes interfere earlier. Temporary suppression of the disease is possible using different types of hormonal therapy; chemotherapy is an option in the terminal phase of the disease[56-57].


CRUISIN' FOR SCREENING
PROSTATE CANCER

2

Prostate Cancer Screening

Based on:

Chapter 26. Early Detection and Screening for Prostate Cancer
Evidence-Based Urology. 2010 Jul; (1):243-254
Pim J. van Leeuwen, Monique J. Roobol, Fritz H. Schröder

The implementation of screening for prostate cancer
Prostate Cancer Prostatic Dis. 2010 Sep;13(3):218-27
Pim J. van Leewen, H.A. van Vugt and C.H. Bangma

2.1 BACKGROUND

Secondary prevention through screening would be an appealing option for PC. This chapter aims to present the evidence available on the efficacy of screening of asymptomatic men for PC.

2.2 EVIDENCE ACQUISITION

A MEDLINE search was performed using the term "prostate cancer" and "screening" with other relevant keywords. Randomized controlled trials of screening versus no screening for PC were included. Studies with inadequate randomization were excluded. Studies that evaluated the complications, and the sensitivity and specificity of the screening test used, were eligible for this review. The searches were limited to English-language articles.

2.3 SCREENING FOR CANCER

The objective of screening is to identify a disease at a stage in its natural history where treatment can be applied to prevent death or suffering[58]. Screening aims to avoid deaths from cancer by preventing the development of advanced disease. Therefore, effective treatment of early staged disease is essential to attain the aims of screening. Although screening may lead to an earlier diagnosis, screening tests will not always benefit the person being screened; overdetection (detected cancers that would not have been diagnosed in the absence of screening) with potential resultant of overtreatment (treatment of cancers that would not have been diagnosed in the absence of screening), increased costs, side effects and complications are potential adverse effects of screening[58-59].

The final endpoint of a cancer screening trial is cancer specific mortality. However, there are more criteria that have to be fulfilled before screening can be adopted in a public health program. A total of ten WHO-criteria for appraising the validity of a screening program were developed by Wilson and Jungner, Table 1[60]. These criteria were listed in 1968 and still upheld today as "classic" and "the gold standard" of screening assessment[61]. Nevertheless, these criteria have been discussed to be too vague or theoretical, and an exchange of views regarding screening policies has occurred over the last decades[62-63]. This has resulted in several adoptions to the classic criteria, which are emerged in ten new criteria, Table 2. The majority of the more recent criteria overlap with the classic criteria, particularly with regard to screening for health conditions at an early stage, where there exist effective interventions to improve outcomes compared

Table 1. Ten criteria by Wilson and Jungner, 1968[62]

1. Condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case finding should be a continuing process and not a "once and for all" project

Table 2. The ten updated criteria by Andermann et al.[65]

1. The screening programme should respond to a recognized need.
2. The objectives of screening should be defined at the outset.
3. There should be a defined target population.
4. There should be scientific evidence of screening programme effectiveness.
5. The programme should integrate education, testing, clinical services and programme management.
6. There should be quality assurance, with mechanisms to minimize potential risks of screening.
7. The programme should ensure informed choice, confidentially and respect for autonomy.
8. The programme should promote equity and access to screening for the entire target population.
9. Programme evaluation should be planned from the outset.
10. The overall benefits of screening should overweight the harm.

to clinical care. Finally, many Western European countries have their own updated legal regulations on mass screening[64-67].

2.4 SCREENING TESTS FOR PROSTATE CANCER

The main tool in screening for PC is the Prostate Specific Antigen (PSA) test. PSA is a human protein that is secreted by prostate epithelial cells[23-24]. The PSA test seems to be acceptable to the population as a screening procedure since the participation and adherence to screening in subsequent screening rounds is overall high[68]. PSA is a specific organ marker, but not strictly a tumour marker, since prostatitis and benign prostate hyperplasia (BPH) can also increase the serum PSA[69-70]. In part due to this, no clear PSA threshold level exists for the detection of PC[71].

Thompson et al. demonstrated this in a study in which men with a PSA ≤ 3.0 ng/ml and a normal Digital Rectal Examination (DRE) were randomized to an antiandrogen

and placebo[72]. After 7 years, in all men with a PSA < 4.0 ng/ml and a normal DRE biopsies were performed and in 15% of these men PC was detected. In 15% of these PCs a tumour with a Gleason score ≥ 7 was detected[72]. According to these study results, a physician who would like an 80% confidence in not missing a PC, should apply a PSA cut-off value of 1.1 ng/ml as indication for biopsy, which would result in 60% unnecessary (negative) biopsies[73]. In Table 3, the continuum of PC risk for different PSA ranges is presented as a result of the Prostate Cancer Prevention Trial (PCPT) and the European Randomised Study of Screening for Prostate Cancer (ERSPC)[72, 74]. As shown, sensitivity decreases with the increasing PSA level, while specificity increases with the increasing PSA level. Consequently, lowering PSA cut-off levels leads to a higher detection rate of PC, but also leads to an increase of negative (unnecessary) biopsies and of the diagnosis of cancers which might otherwise never bother their carrier (potentially overdiagnosed cancers)[75].

Currently, the suggested PSA cut-off to biopsy a man for screening differs between 2.6 and 4.0 ng/ml[68, 76-77]. Future data that include the comparison of the different studies

Table 3: The continuum of prostate cancer risk for different PSA ranges

Authors	Methods	Results							Notes
Thompson et al., 2005, PCPT[73]	Among 5587 men a PSA determination and a sextant prostate biopsy was performed to assess the sensitivity and specificity of PC detection for all PSA ranges in relation to Gleason grade.	PSA, ng/ml	PC, any grade			PC, Gleason grade ≥ 8			N= 1225 (21.9%) were diagnosed with prostate cancer
			Sen (%)	Spec (%)	LR	Sen (%)	Spec (%)	LR	
		1.1	83.4	38.9	1.4	94.7	35.9	1.5	
		2.1	52.6	72.5	1.9	86.0	65.9	2.5	
		2.6	40.5	81.1	2.1	78.9	75.1	3.2	
		3.1	32.2	86.7	2.4	68.4	81.0	3.6	
		4.1	20.5	93.8	3.3	50.9	89.1	4.7	
		6.1	4.6	98.5	3.1	26.3	97.5	10.5	
		10.1	0.9	99.7	3.0	5.3	99.5	10.6	
Schröder et al., 2008, ERSPC[76]	Among 9779 men the cancer detection rate for different PSA ranges in the ERSPC, section Rotterdam was assessed. Distribution of PSA and prostate cancers in men aged 55-74 yr biopsied (2267 men) for PSA ≥ 4.0, DRE, and TRUS are demonstrated.	PSA, ng/ml	Total biopsies (%)	Cancer (*n*)	Total cancer (%)	PPV	Biopsy (*n*) per cancer		
		0.0-0.9	36.4	4	0.8	2.2	45.8		
		1.0-1.9	31.2	45	9.5	8.8	11.4		
		2.0-2.9	12.3	30	6.3	13.6	7.4		
		3.0-3.9	7.2	44	9.3	25.3	3.9		
		4.0-9.9	10.9	241	51.0	24.5	4.1		
		≥ 10.0	2.1	109	23.0	56.5	1.8		

PCPT: Prostate Cancer Prevention Trial, ERSPC: European Randomised Study of Screening for Prostate Cancer, PC: prostate cancer, PSA: prostate specific antigen, DRE: digital rectal examination, TRUS: transrectal ultrasound, Sen: sensitivity, Spec: specificity, LR: likelihood ratio, PPV: positive predictive value

Table 4: Positive predictive of DRE for prostate cancer detection in low PSA ranges

Authors	Methods	Results	
Crawford., 1996[81]	Methods of prostate cancer early detection, to assess the positive predictive of DRE for different PSA values. N = 31 953	The positive predictive value of DRE in the lower PSA areas:	
		PSA	PPV
		0.0-4.0	15%
		4.1-9.9	34%
Schröder et al.,1998, ERSPC[82]	To assess the usefulness of DRE as a stand-alone screening test in low PSA ranges, ERSPC-Rotterdam. N= 10 523	The positive predictive value of DRE in the lower PSA areas:	
		PSA	PPV
		0.0-0.9	4%
		1.0-1.9	10%
		2.0-2.9	11%
		3.0-3.9	33%
		4.0-9.9	45%
Yamamoto et al., 2001[83]	Investigate the usefulness of DRE for prostate cancer diagnosis in subjects with PSA levels of 4.0 ng/ml or less. N = 90	The positive predictive value of DRE in the lower PSA areas:	
		PSA	PPV
		0.0-0.9	4%
		1.0-1.9	0%
		2.0-2.9	19%
		3.0-4.0	44%
Bozeman et al., 2005[84]	Men with abnormal DRE findings and a PSA level less than 4.0 ng/ml who underwent prostate biopsy to assess the positive predictive value of DRE for men PSA <4.0 ng/ml. N= 986	The positive predictive value of DRE for PSA < 4.0 ng/ml:	
		PSA	PPV
		0.0-0.9	2%
		1.0-1.9	6%
		2.0-2.9	13%
		3.0-3.9	21%
Andriole et al., 2005, PLCO[85]	Diagnostic evaluation of DRE as initial screening test in lower PSA ranges. N = 34 115	The positive predictive value of DRE in the lower PSA areas:	
		PSA	PPV
		0.0-4.0	17%
		4.1-7.0	47%

ERSPC: European Randomised Study of Screening for Prostate Cancer, PLCO: Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial, PC: prostate cancer, PSA: prostate specific antigen, DRE: digital rectal examination, Gr: group, PPV: positive predictive value, CDR: cancer detection rate

with long follow up might show the difference in mortality and morbidity outcomes using these different PSA thresholds.

Although DRE is widely used for the diagnosis of PC, the value of DRE remains controversial in screening and early detection programs for PC[78]. The acceptability of the DRE test as a screening procedure seems to be less than PSA since the participation in

a screening program with combined DRE and PSA was twice as low as with PSA alone[68]. Table 4 provides an overview of the positive predictive value for DRE in the lower PSA ranges. According to Table 4, DRE has a low sensitivity and predictive value in men with low PSA levels[79-83]. The positive predictive value of DRE is limited to 4–19% at serum PSA levels below 3.0 ng/ml. This proportion equals to the percentage of 15% cancers that were diagnosed in the study of Thompson et al. in which they performed biopsies in all men with a PSA < 4.0 ng/ml without using DRE[72]. Therefore, the studies presented in Table 4 might have found a similar PC detection rate without the use of DRE in the PSA levels 3.0 ng/ml and lower. Accordingly, it might be concluded that men with low PSA values have a 15% PC detection rate with or without the use of DRE, and that consequently the additional value of the DRE is restricted in lower PSA ranges.

In contrast, several researchers still suggest that with the use of DRE men will be screened more selectively. It is shown that men with a positive DRE are more likely to have high grade PC than men with non-palpable tumours[84-85]. For this reason, the risk of omitting DRE, and therefore the omission of biopsies at PSA <2.6, <3.0 or <4.0 ng/ml, might be that potentially aggressive tumours at these low PSA levels remain undetected at screening. Catalona et al. have confirmed this risk by showing that a substantial proportion of PC detected by DRE at PSA levels lower than 4.0 ng/ml have features associated with clinically aggressive tumours and that the omission of DRE from screening protocols might comprise treatment outcomes as omitting DRE at PSA levels less than 3.0 ng/ml would have detected 14% fewer PC overall and 7% fewer PC with a Gleason score of 7 or higher[86]. In contrast, it is shown that screening without DRE at low PSA levels (PSA<3.0 ng/ml) did not lead to the detection of significantly more (poorly differentiated) PCs 4 years later compared to screening with the use of DRE in the ERSPC[87].

2.5 COMPLICATIONS OF SCREENING TESTS

The complications of the PSA and DRE are limited, but prostate biopsies are related to clinical complications[88-89]. These complications vary between studies since they depend on the antibiotic prophylaxis used, study population and the number of prostate biopsies performed[88-89]. Djavan et al[89] reported in their literature review, that included eleven prospective PC detection studies, haematuria and haematospermia in 12.5-58.4% and 5.1-50.5% of the procedures, respectively. In these same studies, rectal bleeding was reported in 2.8-37.1% of men, 1.4-4.2% of men reported fever after prostate biopsies, 0-0.5% of men needed hospitalization due to symptoms of sepsis or prostatitis[89]. Overall, 0-1.2% of men reported urinary retention after prostate biopsies[89].

2.6 RANDOMIZED CONTROL TRIALS FOR PROSTATE CANCER SCREENING

Two leading randomized control studies are designed to evaluate the effectives of screening (Table 5). The European Randomised Study of Screening for Prostate Cancer (ERSPC) and the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO trial) were designed to evaluate whether population-based screening reduces the mortality from PC, with an acceptable level of quality-of-life aspects and the associated costs[68, 77].

The ERSPC is conducted in eight European countries (Belgium, Finland, France, Italy, the Netherlands, Spain, Sweden, Switzerland) and enrolled 267,994 men 55-74 years of age. All men with a prior diagnosis of PC were excluded. In the ERSPC men were screened in most countries with an interval of four years, however, in Sweden men were screened with an interval of two years. The screening algorithm differed among the study centres (Table 5)[90-96].

The PLCO is a trial in the United States that enrolled 155,000 women and men, 55-74 years of age, in ten screening centres. All men with a prior diagnosis of PC were excluded.

Table 5. Characteristics of the randomized control trials for prostate cancer screening

Study:	European Randomised Study of Screening for Prostate Cancer[70]
Methods:	Initiated in 1993, randomized 267,994 men.
Participants:	Male inhabitants of Belgium, Finland, France, Italy, the Netherlands, Spain, Sweden, Switzerland, aged 55–74 years. Men with a previous diagnosis of prostate cancer were excluded.
Intervention:	Screening with an interval of 4-years (except in Sweden, i.e. 2-years) versus control (not invited for screening). Screening included a PSA test or a DRE/TRUS. TRUS biopsy was performed in screen positives. Men were screen positive in France, Sweden, Switzerland, Spain by a PSA ≥3.0 ng/ml; in Finland by PSA ≥4.0 ng/ml (men PSA 3.0-3.9 had an ancillary test, i.e. DRE until 1998, free/total PSA ratio with a cut-off ≤ 0.16 from 1999 onwards); in Italy by PSA ≥ 4.0 ng/ml (men PSA 2.5-3.9 ng/ml had an ancillary test, i.e. DRE and TRUS); in the Netherlands and Belgium by positive DRE and/or TRUS and/or PSA ≥ 4.0ng/ml, but after 1997 by PSA ≥3.0 ng/ml only.

Study:	Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial[79]
Methods:	Initiated in 1993, randomized 76,693 men.
Participants:	Male inhabitants of United States, aged 55–74 years. Men with a previous diagnosis of prostate cancer were excluded.
Intervention:	Annual screening versus control (not invited for screening). 1st screening round included a PSA test and a DRE. TRUS biopsy was recommended in cases with PSA 4.0 ng/ml and/or abnormal DRE. Follow up screening included a PSA test and a DRE during the first three screenings.

In the PLCO men in the intervention arm received annual screening by DRE and serum PSA determination (Table 5)[77].

2.6.1 RESULTS

2.6.1.1 Interim results prostate cancer screening

In the ERSPC and the PLCO trial more men were diagnosed with PC in the intervention arm than in the control arm of the study[97-98]. In the ERSPC the cumulative incidence of PC was 8.2% and 4.8% for the intervention and control group respectively after 9 years[98]. In the PLCO trial, PC was diagnosed in more subjects in the intervention group (7.3%) than in the control group (6.0%) at 7 years[97].

In the ERSPC and PLCO screening trial a stage distribution was observed among men that were screened for PC. In the Rotterdam section of the ERSPC, after comparing the intervention arm with the control arm, a statistically significant migration to more favourable stages was observed in the intervention arm (Table 6)[99]. In the screening population of the PLCO the large majority of PC were stage II at diagnosis, regardless of the mode of detection in the screening group. However, overall the numbers of subjects with advanced (stage III or IV) tumours were statistically similar in the screening and the control group, with 122 men in the screening group and 135 men in the control group (Table 6)[100].

Table 6. Tumour characteristics prostate cancer detected in the intervention and the control arm of the ERSPC and the PLCO[101-102]

ERSPC study, section Rotterdam				
Tumour stage	Intervention arm No. men, (%)	Control arm No. men, (%)	1st screening round No. men, (%)	2nd screening round No. men, (%)
I	685 (40.6)	114 (33.9)	252 (30.8)	203 (60.4)
II	553 (43.6)	84 (25.0)	391 (47.8)	120 (35.7)
III	184 (14.5)	64 (19.0)	164 (20.0)	13 (3.9)
IV	13 (1.0)	13 (4.0)	11 (1.4)	-
Unknown	4 (0.3)	61 (18.1)	-	-

PLCO study				
Tumour stage	Intervention arm No. men, (%)	Control arm No. men, (%)	1st screening round No. men, (%)	2nd screening round No. men, (%)
I	18 (0.5)	15 (0.5)	2 (0.4)	2 (0.1)
II	3297 (95.5)	2790 (93.8)	516 (94.0)	1458 (97.2)
III	49 (1.4)	56 (1.9)	12 (2.2)	22 (1.5)
IV	73 (2.2)	79 (2.7)	19 (3.4)	15 (1.0)
Unknown	15 (0.4)	34 (1.1)	0 (0.0)	3 (0.2)

The reported interval cancers (those clinically diagnosed within a screening interval) in the ERSPC and PLCO trial were infrequent and in general had favourable characteristics. The ERSPC-Rotterdam reported in the first four years after initial screening 25 interval cancers. Seven of the 25 cancers were diagnosed in men who had refused a recommended biopsy at their initial screen. Of the remaining 18 cancers, all were classified as stage T1A-C or T2A, none were poorly differentiated or in a metastatic stage[101]. In the PLCO, 204 interval cancers were diagnosed. Of these cancers 96.1% were classified as stage T1A-C or T2A and 2.0% were classified as stage IV disease[100].

2.6.1.2 Final results prostate cancer screening

Mortality data have been presented by the ERSPC and PLCO trial (Table 7). The ERSPC trial reported that PSA screening without digital rectal examination was associated with a 20% relative reduction in the death rate from PC at a median follow-up of 9 years[102]. The absolute reduction in the screening population was 7 PC deaths per 10,000 men that were screened. The results were associated with a number of 1410 men that needed to be screened and 48 men that needed treatment to save one death from PC death. The treatment distributions were slightly different between the two groups, however unlikely to play a major role in interpretation of the final results[103]. Data analysis of the ERSPC with adjustment for the diluting effect of nonattendance and contamination showed that the mortality effect among men was increased to 30%[104-105]. In the ERSPC, 82.2% of the men in the screening group were screened at least once and the average

Table 7. Final outcomes prostate cancer screening trials and limitations

ERSPC study[100]			
Reduction PC mortality	NNS	NNT	Limitations
Rate Ratio: 0.80 (95%CI:0.60-0.99)	1410	48	Different study protocols: The different centres of participants used (marginally) different study protocols. Contamination: In 20-30% of men in the control group a PSA test was performed.
PLCO study[99]			
Reduction PC mortality	NNS	NNT	Limitations
Rate Ratio: 1.13 (95%CI:0.75-1.70)	-	-	Contamination: In 40-52% of men in the control group a PSA test was performed and in 41-46% of men a DRE test. 9.3% of men reported having had two or more PSA tests before start of the study. 44% of the men in each study group had undergone one or more PSA tests before randomization. Low compliance prostate biopsies: in 40% of men with a positive screening test a biopsy was actually performed in the intervention arm.

ERSPC: The European Randomised Study of Screening for Prostate Cancer; PLCO: The Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; NNS: number needed to screen to save one death from prostate cancer. NNT: number needed to treat to save one death from prostate cancer

rate of compliance with biopsy recommendations was 85.8% (range, 65.4 to 90.3)[98]. The level of contamination by PSA testing in the control group was estimated in the order of 20-31%[104, 106-107].

The Gothenburg screening trial, which is part of the ERSPC, has published their own mortality outcomes. The Gothenburg trial was initiated as an independent study in 1994 as an effectiveness trial (without upfront informed) but joined the ERSPC trial shortly thereafter. Data up to 2008, after a median follow up of 14 years, showed a RR for PC death of 0.56 (95%CI 0.39-0.82, p = 0.002). This resulted in a NNS of 234 and NNT of 15. The main differences with the ERSPC as a whole are the type of randomization, younger age, a shorter screen interval, and, most importantly, a longer follow-up due to the simultaneous randomization of all participants in 1994.

The PLCO trial found no mortality benefit from combined screening with PSA testing and DRE during a median follow-up of 7-10 years comparing those screened to those that were not[97]. The incidence of PC death per 10,000 person-years was 2.0 (50 deaths) in the screening group and 1.7 (44 deaths) in the control group (rate ratio, 1.13; 95% CI, 0.75 to 1.70) after a median of 7 years follow-up. The data at 10 years were 67% complete and consistent with these overall findings. The treatment distributions were similar in the two groups within each tumour stage[97]. In the PLCO trial the compliance with the screening protocol overall was 85% for PSA testing and 86% for DRE[97]. The average rate of compliance with biopsy recommendations was 40%. The level of contamination is well established, i.e. the rate of PSA testing was 40-52% and the rate of screening by DRE ranged from 41 to 46% in the control group[97]. Approximately 44% of the men in each study group had undergone one or more PSA tests before randomization, which would have eliminated some cancers detectable on screening from the randomized population, especially in health-conscious men (who tend to be screened more often, a form of selection bias). No results are available for the effect of screening after the adjustment for the contamination, however the PC specific mortality was 25% lower among the men who were screened prior to randomization in the PLCO[97].

Whereas the ERSPC found a statistically significant reduction in PC mortality with screening, the PLCO trial did not. In the PLCO trial the contamination in the control group and compliance with the screening protocol in the intervention group is of major influence. This is highlighted in the stage distribution among the men in the control arm of the PLCO study. In comparison to the 96% of men diagnosed with a stage ≤ II tumour in the intervention arm, were 94.3% of men with a stage ≤ II tumour diagnosed in the control arm of the PLCO. Consequently, the PLCO trial is more a trial comparing two screening strategies of a different intensity and is inadequate in establishing if PC screening has the potential to reduce the PC specific mortality. Therefore, we can conclude that

systematic PC screening is not effective in terms of reducing the PC specific mortality in comparison to widespread opportunistic screening and early detection.

The ERSPC is a randomized controlled trial with an adequate methodological design, a high level of compliance in the intervention group and a relative low level of contamination in the control group of the study. Furthermore, an intention-to-screen analysis, and an analyses with adjustment for non-attendance and contamination are available. In essential, the ERSPC provides clear evidence that screening for PC has the potential to reduce the PC specific mortality.

2.7 POTENTIAL HARMS OF PROSTATE CANCER SCREENING

It is proven that screening increases the PC incidence[97-98]. The excess incidence and overtreatment are associated with a distinct pattern of change in quality of life[108-109]. Quality of life (QoL) parameters that are affected, are a change pattern in the urinary, bowel, and erectile functions, as well as the emotional distress and anxiety[108-109]. Currently, no QoL analysis is presented by the ERSPC study group, nore by the PLCO study group.

2.8 COST EFFECTIVENESS OF PROSTATE CANCER SCREENING

No results from randomized controlled trials are reported on cost effectiveness, cost utility or cost benefit of screening for PC.

2.10 OTHER CANCER SCREENING PROGRAMS

Randomized controlled trials demonstrated contradictory results that mammographic screening reduces mortality from breast cancer or not. Two trials with adequate randomisation did not show a significant reduction in breast cancer mortality: rate ratio (RR) 0.93 (95CI: 0.80 - 1.09) at 13 years[110]. In contrast, four trials with adequate randomisation showed a significant reduction in breast cancer mortality, RR: 0.75 (95%CI: 0.67 - 0.83)[110]. The RR for all six trials combined was 0.80 (95%CI: 0.73 - 0.88)[110]. Data from prospective studies have found that the NNS to save one breast cancer specific death to be 1500 and the NNT to save one cancer specific death 10[110-112]. It has been established that breast screening is cost effective (Table 8)[113].

Screening for colorectal cancer is implemented in many European countries[114]. For colorectal cancer screening data is available from four eligible randomised control trials showing that participants allocated to screening with faecal occult blood test had a 16%

Table 8. Screening for breast, colorectal, lung and cervix cancer

Breast Cancer[60-63]				
Mortality	NNS	NNT	Cost effectiveness	Comments
RR: 0.80 (95%CI: 0.73-0.88)	1500	10	Yes	- Results from RCT - Contradictory results 6 trials
Colorectal Cancer[65-68]				
Mortality	NNS	NNT	Cost effectiveness	Comments
RR: 0.84 (95%CI: 0.78-0.90)	600	2	Yes	- Reduction mortality increased to 25% after adjustment for nonattendance.
Lung Cancer[69,70]				
Mortality	NNS	NNT	Cost effectiveness	Comments
No reduction	-	-	No	- RCT with chest radiography and sputum cytology showed no reduction in mortality. - RCT with computed tomography are ongoing
Cervix Cancer[71-73]				
Mortality	NNS	NNT	Cost effectiveness	Comments
Reduction between 5-80%	-	-	Yes	- No results RCT - Mortality declined between 5 and 80% in 5 countries after start screening.

NNS: number needed to screen to save one death from cancer. NNT: number needed to treat to save one death from cancer. RCT: Randomized Controlled Trial

reduction in the relative risk of colorectal cancer mortality, RR: 0.84 (95%CI: 0.78-0.90)[115]. When adjusted for screening attendance in the individual studies, there was a 25% relative risk reduction (RR 0.75, 95%CI: 0.66 - 0.84) for those attending at least one round of screening using the faecal occult blood test[115-116]. The NNS to save one cancer specific death would be in the order of 600 cases and the NNT in the order of 2 cases[115-117]. Several studies indicated that colorectal screening is cost effective (Table 8)[118-119].

Lung cancer screening using regular chest radiography and sputum examination programs were not effective in reducing mortality from lung cancer. Studies showed that early detection of lung cancer was possible with such programs, but mortality was not improved[120]. Currently, randomized control studies using computed tomography (CT) scan are ongoing for lung cancer in high risk patients, however results are not expected before 2010 (Table 8)[121-122].

For cervical cancer, Cytological Papanicolaou (Pap) smear screening remains the best method readily available in reducing the incidence and mortality. No randomized controlled trials are performed indicating that screening for cervical cancer reduces the mortality. However, various studies (mainly time series and case-control) revealed that widespread use of cervical screening in developed countries has been associated with the substantial reduction in rate of mortality from cervical cancer. A study conducted in five northern European countries investigated the time trends in mortality from cervical cancer in relation to the extent and intensity of organised screening programmes. In all five countries the cumulative mortality rates (0-74 years) fell with 80%-5% between 1965 and 1982[123]. The NNS was in the order of 100-300 women[124]. Although there is no universal criterion that defines a threshold cost-effectiveness ratio, above which an intervention would not be considered cost-effective, the cost-effectiveness studies concluded that cervical screening falls within the acceptable limits of cost-effectiveness, i.e., less than $500 per year of life saved (Table 8)[125].

PC screening has, relative to screening for breast and colorectal cancer, a comparable impact on the relative risk reduction in cancer specific mortality. Nevertheless, the absolute benefit of PC screening is modest compared to screening for colorectal cancer. For this reason, compared to other colorectal cancer screening, a relatively high number of men would need PC screening to prevent one man from cancer specific death. Finally, the additional number of men that are diagnosed with PC is large in comparison to breast and colorectal cancer screening. On this regard, the main limitation of PC screening is the difference in natural history of the disease and/or the lack of a screening test with a high sensitivity and specificity for aggressive cancer. Finally, in comparison to screening for breast, cervical and colorectal cancer, no evidence is available that PC screening is cost effective.

2.11 CONCLUSIONS

Based on the presented evidence so far we can conclude that in theory screening for PC is an appealing option since the disease has a number of characteristics that makes it suitable for screening. PC is an important public health problem with 20 men dying every day from PC in the Netherlands. With PSA and DRE, acceptable and moderately accurate tests are available and the complications of prostate biopsies are limited. PC has a long latent stage which provides the possibility to produce a beneficial stage shift to more favourable stages by screening. Curative treatment of localized PC reduces the PC specific mortality relative to expected management in a selected population. Screening for PC reduces the disease-specific mortality significantly. However, PC screening causes

a significant increase in the cumulative incidence of PC and no robust evidence is available regarding the economical quality of life implications of PC screening. Furthermore, there is limited understanding of the natural history of screen-detected PC, meaning that there is a lack of data indicating which specific PC will develop from indolent to significant disease and which PC will not surface clinically during lifetime if left undiagnosed and/or untreated. For these reasons, there is no agreed policy which men should be treated with aggressive curative treatment and which men might be treated by an expectant management in order to reduce the overdiagnosis.

In conclusion, currently the scientific basis for the introduction of a population-based PC screening program is insufficient since a number of essential criteria are not met[60, 63-67]. Follow up of the ongoing randomized controlled trials is crucial before recommending population based prostate cancer screening. A more convenient balance between the screening harms and benefits seems to be needed in which a reduction of the overdiagnosis is crucial. Since the decision whether to introduce a population based screening program will also be based on an appropriate economic and quality of life analysis taking into account the quality of life aspects of the screening program, introduction of a population based PC screening program depends on further progress in these areas of regards.


CRUISIN' FOR SCREENING
PROSTATE CANCER

3

Trends in Prostate Cancer Incidence, Survival and Mortality

Pim J. van Leeuwen

3.1 INTRODUCTION

Trends in prostate cancer (PC) incidence, survival and mortality have been changing worldwide during the last three decades[7, 126]. The interpretations of these trends have shown to be confusing without proper knowledge of the natural history of PC and proper statistical data on outcomes. In this chapter we explored the outcome measures (i.e., cancer incidence, survival and mortality) in general, and studied the trends in PC incidence, survival and mortality during the last decades in the Netherlands.

3.2 INCIDENCE

Cancer incidence is defined as the number of new cancers occurring in a specified population during a year, usually expressed as the number of cancers per 100,000 persons at risk[7]. Cancer incidence is an important measure since it reflects the processes that drive disease progression, the inheritance of predisposing genetic variants, and the consequences of carcinogenic exposures.

Cancer incidence trends could be misinterpreted due to potential technical diagnostic difficulties that have to be taken into account. First, incidence data have to be quantified and evaluated on incompleteness and inaccuracy[127]. Additionally, a consensus has to be made with respect to the definition of cancer. Usually cancers are defined by pathologists in terms of the extent of tumour invasion, however, studies attest that there is not a complete concordance between pathologists in the diagnosis of cancers and some types of cancer are often not histologically proven[128].

A more serious problem is that the cancer incidence is measured at a point in time at which the diagnosis takes place, somewhere in the continuum of the natural history of cancer. Histological specimens taken from individuals, in whom there were no symptoms or no clinical suspicions of cancer, can detect cancers that would not have surfaced clinically during an individual's lifetime[129]. This can occur during opportunistic or systematic screening for cancer. Screening aims to advance the diagnosis, and so when introduced it will give rise to a temporary rise in incidence[58]. Therefore, screening for cancer can be of major influence on the cancer incidence over time.

3.2.1 PROSTATE CANCER INCIDENCE

The incidence of PC has been rising during the last decades in the Western world[8]. Figure 1 presents the age-standardized rates for PC incidence between 1988 and 2006 in the Netherlands. With the introduction of PSA testing in the late 1980s, the age-adjusted

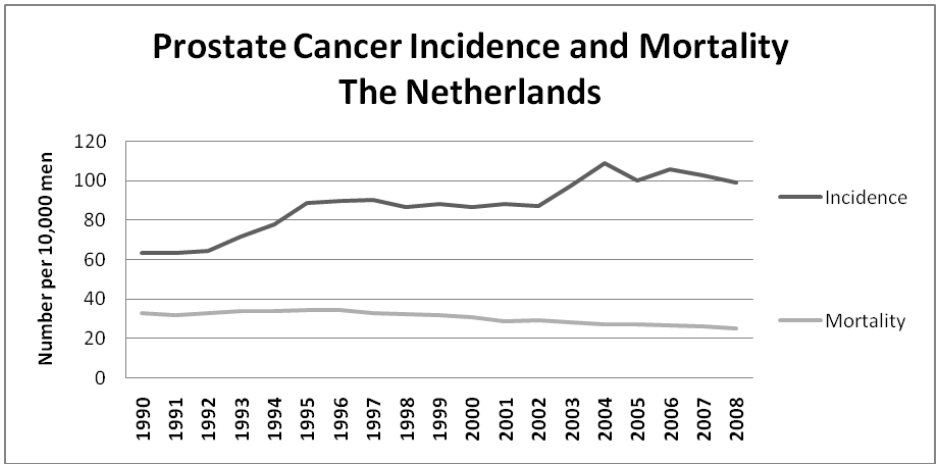


Figure 1. Prostate Cancer Incidence and Mortality 1990 – 2008, the Netherlands

incidence rates rose from 60.1 cases / 100,000 men in 1990 to more than 100 cases / 100,000 men after 2003.

It is the effect of screening which has resulted in the dramatic increase in PC incidence during the last decades[130]. Since the intensity of screening differs between different population it affects incidences in different population to a different degree. Given the very large affect of screening on incidence it is unlikely that the inheritance of predisposing genetic variants and the consequences of the change in carcinogenic exposures will be quantified by the population based PC incidence trends of the latest years.

3.2.2 EXCESS INCIDENCE AND OVERDIAGNOSIS

Screening and early detection should identify cancers that will become clinically relevant. In this case, the number of individuals diagnosed with cancer would be unaffected by screening. However, early detection and screening can be associated increases in incidence, raising the question of over-diagnosis. The clinical definition of overdiagnosis is diagnosing tumors that would otherwise remain clinically unrecognized until the individual died from other causes. Overdiagnosis occurs when screening detects small tumors that would otherwise remain clinically unrecognized until the individual dies from other causes. Thus, cancers that are found by over-diagnosis never progress to cause symptoms or death. In PC, randomized screening trials show an excess incidence in the screened group compared to control subjects. Excess incidence is considered a proxy for over-diagnosis. Excess incidence is defined as the difference between the actually observed and the expected number of cancers at a moment of follow-up.

3.3 SURVIVAL

Cancer survival is the study of the distribution of times, e.g., the time from a diagnosis to some terminal event (relapse or death)[131]. It refers to the fraction of a population which will survive past a certain time. Two methods for studying cancer survival are discussed hereunder, i.e., cause specific survival and relative survival. Cause specific survival represents the probability of surviving a given cancer at a particular point in time[132]. It allows excluding death due to unrelated causes in patients suffering from the cancer in question. As a result, the estimation of cause specific survival proportions require reliably coded information on cause of death. It is possible to reliably ascertain causes of death by the review of patients medical records with the use of predetermined clinical algorithms[133]. However, usually the causes of death are based on the primary cause of death that is recorded on the death certificate[134]. Consequently, this assessment of the cause of death may not be sufficiently reliable particularly when patients are older or have considerable co-morbidities[133, 135-137].

The second method is the relative survival analysis. Relative survival is defined as the ratio of the proportion of observed survivors in a cohort of cancer patients to the proportion of expected survivors in a comparable cohort of cancer free individuals[138]. In practise the expected survival is often calculated on the basis of a cohort of individuals from the general population matched by age, race, sex, and period[139]. There are three methods that differ regarding how long each individual is considered to be at risk for the purpose of estimating expected survival.

1. Ederer I method: the matched individuals are considered to be at risk indefinitely (even beyond the closing date of the study). The time at which a cancer patient dies or is censored has no effect on the expected survival.
2. Ederer II method: the matched individuals are considered to be at risk until the corresponding cancer patient dies or is censored.
3. Hakulinen method: if the survival time of a cancer patient is censored then so is the survival time of the matched individual. However, if a cancer patient dies the matched individual is assumed to be at risk until the closing date of the study.

For all three methods, each individual must have equal life expectancy as the patient should have in the absence of cancer. Eventually, the relative survival is calculated as the ratio of the survival probabilities for a patient and age and gender matched individuals. Advantage of relative survival analyses is that information on cause of death is not required. However, it requires accurate estimation of expected survival in a comparable population of cancer free individuals. For most types of cancer, patients diagnosed with

cancer are representative of the general population, so their expected mortality can be estimated using general population mortality rates. The most notable exceptions are smoking-related cancers where patients will have lower survival than the general population due to numerous other smoking-related conditions (e.g. cardiovascular disease)[140]. Furthermore, estimation of the expected survival may not be sufficiently reliable particularly for screening-related cancers since screening patients are often healthier and of higher social economic class (healthy screenee bias)[141].

Both survival measures are typically expressed as the proportion of patients alive at some point subsequent to the diagnosis of their cancer[131]. Thus, survival represents the difference in time between two dates and is therefore sensitive to changes in either of the dates. As a result, survival analyses could be plagued by three potential biases, i.e., lead time, length time and overdiagnosis[142]. The potential magnitude of lead time on patient's survival is graphically presented in Figure 2. In Figure 2a the diagnosis was made in a patient who presented with symptoms and the survival time is the time between this diagnosis and the date of death. However, if the cancer was detected by screening (Figure 2b) the survival time increased by an amount, called the lead time, even if the date of death remained unchanged. Now we would hope that the early diagnosis increases the potential for cure so that death is postponed and survival time further increased (Figure 2c). It is this improvement in survival resulting from postponing death that is of real interest. However, it is difficult to separate this component of

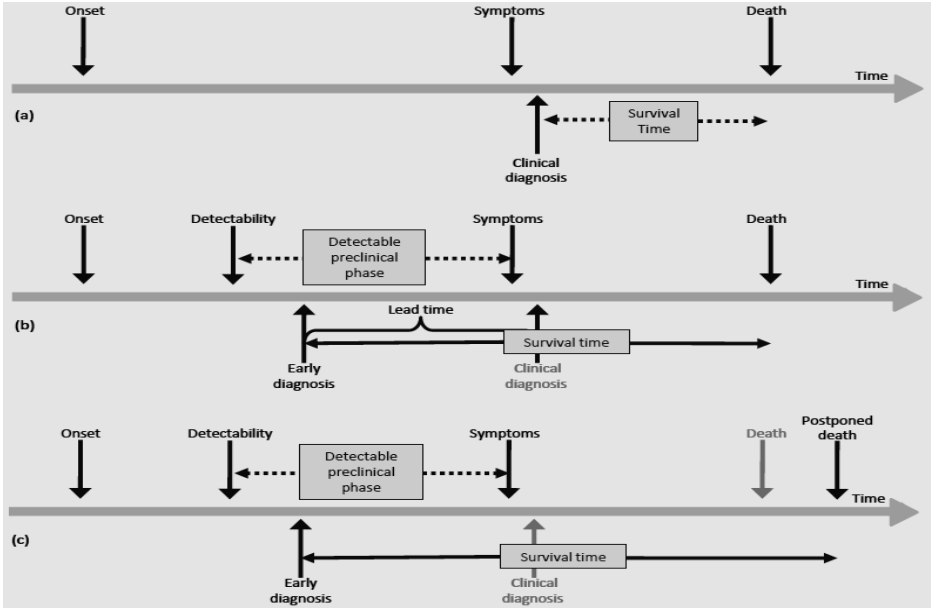


Figure 2. Potential magnitude lead time. Adapted with permission from [145].

patient survival since it cannot be easily separated from other components such as lead time.

Length time may also occur in the seting of screening for a disease. Length time reflects the effect that screening predominantly detects slow growing tumours, as they remain longer in the preclinical detectable phase than fast growing tumours[142]. Hence, slow growing tumours are likely to be over-represented in screening populations and are ironically associated with improved survival (Figure 3). As shown, slowly growing tumours remain for a relatively long period in the zone before the cancer is spread beyond the prostate while this period for fast growing tumours (patient D) stay relatively short in this zone. Consequently slowly growing tumours have a relatively increased chance in being diagnosed as screen detected cancer in a program with repeated screening and form a high percentage of all screen detected cancers. Relative overrepresentations of slowly growing and probably favourable prognostic cancers positively bias the survival rate; the survival will improve even if there is no effect of screening.

The third bias that influences the survival rate is overdiagnosis. Overdiagnosis is defined as the diagnosis of a disease that will never cause symptoms or death during a patient's lifetime[142]. Screening detects a considerable part of insignificant cancers that would never have surfaced during the lifetime of a patient. Consequently, if overdiagnosis were

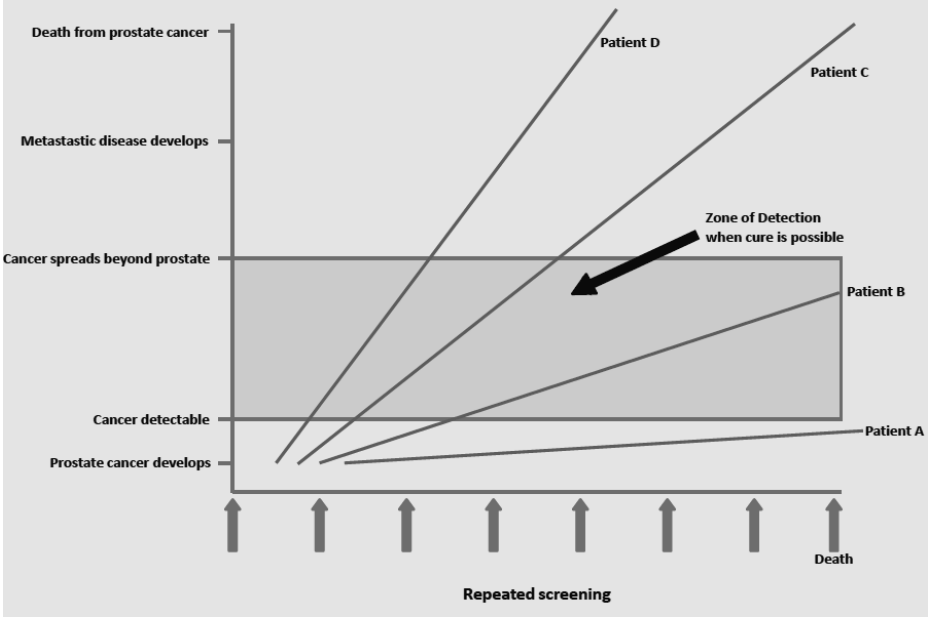


Figure 3. Potential magnitude length time. Adapted from [146]

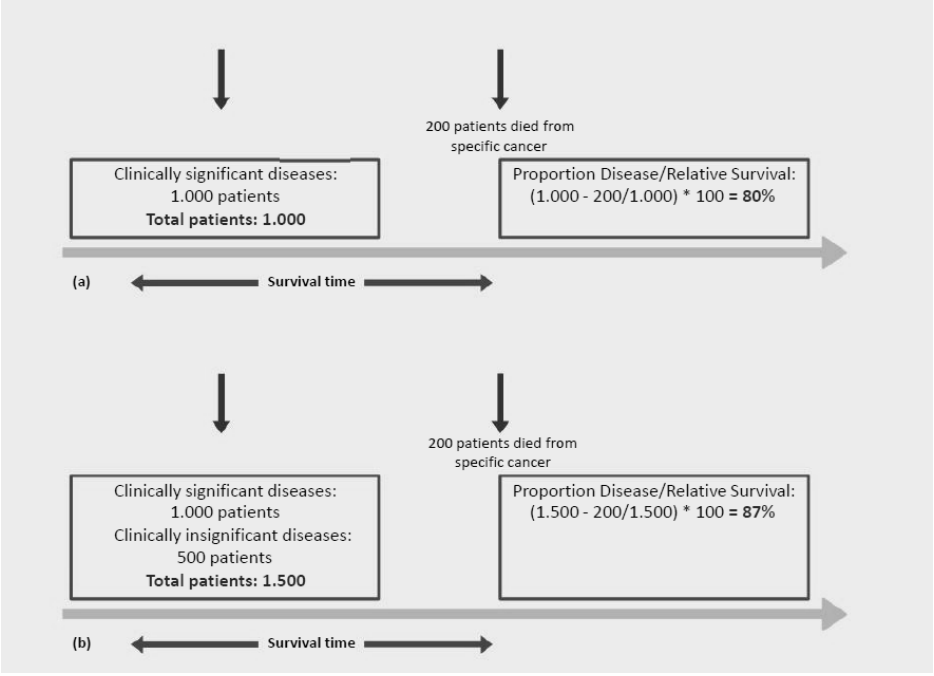


Figure 4. Potential magnitude overdiagnosis in cancer survival analysis

the only effect of screening the observed improvement in survival is caused by increasing the number of patients (most of whom would not have died of the disease in the first place). Figure 4 demonstrates two possible conditions, i.e., a situation in which only clinical significant diseases are diagnosed (Figure 4a) and in which clinically significant and insignificant diseases are diagnosed (Figure 4b). Accordingly at every point in the follow-up, in Figure 4b relative to Figure 4a, the number of men at risk to die of the disease is increased by overdiagnosis but the number of men who actually die of the disease is unaffected. Therefore the probability of survival improves while the number of men that die from the disease remains the same.

3.3.1 PROSTATE CANCER SURVIVAL

The survival rates for PC have been improving during the last 30 years in several countries. Figure 5 presents the 5-years relative survival from PC diagnosed in a part of the Netherlands between 1988 and 2005. 5-year relative survival improved 8% between 1988-1992 and 1993-1996, and more than 10% between 1993-1996 and 2001-2005. Although these improvements in survival patterns are encouraging, the survival rates are likely to have changed over time with earlier diagnoses following the advent of PSA

testing. The PSA test enables invasive PC to be indentified earlier than it might have been diagnosed clinically on the basis of symptoms (lead time). However PSA also enables the identification of latent tumours that have caused symptoms during the man's lifetime (overdiagnosis). For PC screening the mean lead times and rates of overdiagnosis depends on a man's age at screening. For a screening program with a 4-year screening interval from age 55 to 67, the estimated mean lead time is approximately 11 years (range = 10.8-12.1 years), and the overdetection rate is 48% (range = 44%-55%)[129]. Consequently, since the introduction of the PSA test, lead time and overdiagnosis have a significant influence on the changes observed in the PC survival rates. These effects are illustrated by the 16% difference in 10-years PC specific survival between men diagnosed with PC in the screening and control arm of the ERSPC section Rotterdam between 1993 and 2006 (figure 6).

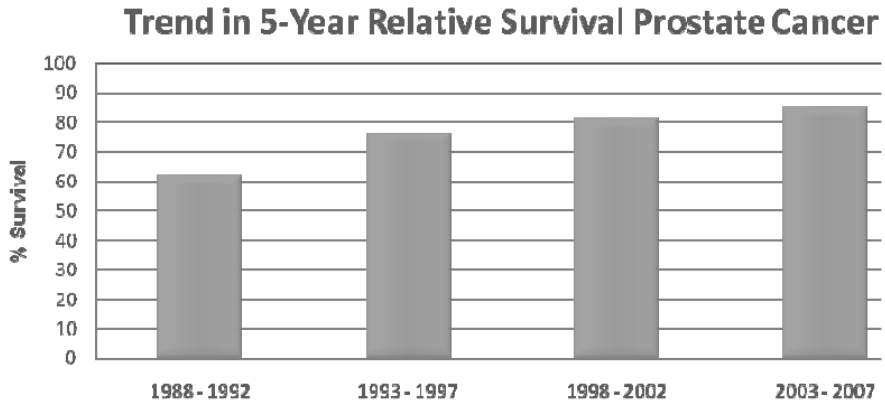


Figure 5. 5-years relative survival prostate cancer diagnosed 1988 - 2007, region Amsterdam, the Netherlands.

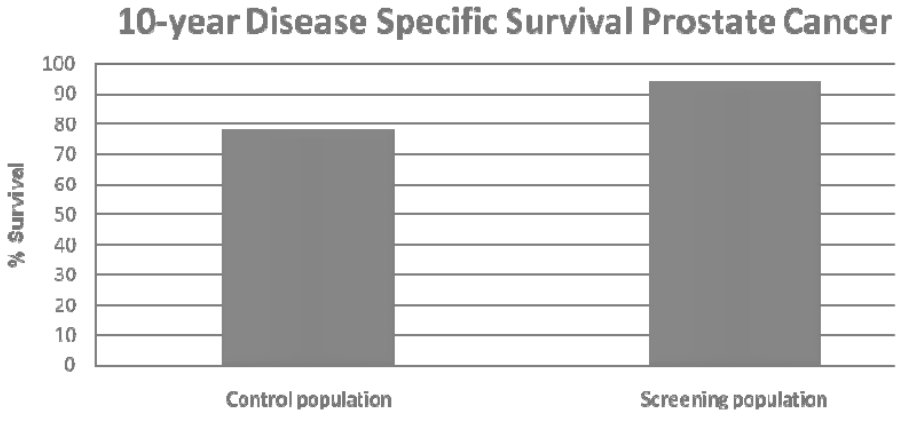


Figure 6. 10-years prostate cancer specific survival per study arm ERSPC section Rotterdam

In conclusion, patients live longer with the diagnosis of PC (more person years) due to a PC diagnosis at an earlier age and with on average a PC of a lower grade and stage during the last two decades. Since the effect in survival cannot be separated from other components such as lead time, length time and overdiagnosis which are extensively in present in these analyses, survival rates fail to demonstrate the improvement in PC survival resulting from postponing PC death. Consequently, survival in a clinical situation of increasing overdiagnosis, is not an appropriate endpoint.

3.4 MORTALITY

Three methods can be used for cancer mortality analyses, i.e., all cause mortality, a cause specific mortality and an excess mortality analysis. An all cause mortality analysis measures the total number of deaths in a population per unit time. Benefits of the all-cause mortality are that it does not require judgments about the cause of death, and that all-cause mortality can capture unexpected lethal side effects of medical care.

3.4.1 CAUSE SPECIFIC MORTALITY

A cause specific mortality analysis measures the number of deaths due to a specific cause in a population per unit time[131]. Similar to the cause specific survival, accurate information on the cause of death is required for this analysis[136]. Cause specific mortality analyses have both their advantages and disadvantages (Table 1). Cause specific mortality has the aim to measure if the mortality is directly attributable to the cancer in question. Deaths which are due to the cancer are counted as deaths (or its treatment), while all other deaths are treated as censored observations (i.e. observed until the time of death, considred alive at that moment within the context of the analysis). For this reason it is a suitable method e.g. the effect of screening on the mortality related to the cancer of interest. The concept disease specific mortality implies a dichotomy that not always reflects the clinical reality since it might be difficult to decide whether a person died either entirely due to the cancer in question or completely unrelated to the cancer

Table 1. Cause Specific mortality versus Excess Mortality

Measure	Advantage	Possible bias
Cause Specific Mortality	Measures mortality directly due to cancer	Accurate classification of cause of death
Excess Mortality	Measures mortality due to cancer, capturing both direct and indirect mortality	Accurate estimation of expected mortality in a comparable population without the disease of interest

in question. Especially for PC when there are several competing causes of death. E.g., in elderly PC patients there are often a considerable number of co-morbidities and the dichotomy could be an oversimplification of reality. Therefore, cause specific mortality has shown to be more valid in clinical studies where more effort is made to distinguish between the deaths due to cancer and deaths due to competing risks. Studies that have questioned the validity of death certificate-assigned causes of death in PC reported that between 4 and 46% of death certificates were inaccurately coded[135, 137, 143-144].

3.4.2 EXCESS MORTALITY

Excess mortality is the difference between the total (all-cause) mortality of the patients and the mortality that would be expected in the absence of cancer[145]. Excess mortality is related to relative survival, i.e., the number of observed deaths in a cohort of cancer patients to the number of expected deaths in a comparable cohort of cancer free individuals[138]. Excess mortality provides a measure of the mortality associated with a diagnosis of cancer irrespective of whether the excess mortality is directly or indirectly attributable to the cancer. The excess mortality rate is defined as the actually recorded number of cancer patient deaths in excess of the number expected on the basis of a cohort of cancer free individuals per unit of time. Excess mortality can also be expressed as a count, i.e., the actually recorded number of cancer patient deaths in excess of the number expected on the basis of a cohort of cancer free individuals during a defined time of follow-up. If the excess mortality is defined as a count it will not be affected by the above mentioned biases (lead time bias, length time and overdiagnosis). The biases mentioned above only affect the time period that an individual is identified as having the disease or the number of people that are considered to have the disease, not the probability to die of it. Note that the time is not specified if the excess mortality is expressed as a count, thus e.g., lead time bias (which occurs in the relative survival analyses) does not affect the outcome. Excess mortality analyses have both their advantages and disadvantages (Table 1). Information on cause of death is not required for an excess mortality analysis, thereby circumventing problems with the inaccuracy of the assessment of the cause of death[140]. An excess mortality analysis measures the difference between the mortality in a group of patients and the mortality in that group that would be expected in the absence of cancer. It is a suitable measure to assess the burden from a specific cancer on the total mortality since it measures all the mortality that is associated with cancer. It measures not only whether cancer leads directly to death, but it also measures whether the death is indirectly attributable to cancer. The latter cosist of amongst others deaths from "intercurrent causes", for example pneumonia, which were not taken by pneumonia if they were in the absence of cancer, or deaths due to treatment complications or suicide are

examples of deaths which may be considered indirectly attributable to the cancer of interest. For an excess mortality analysis an accurate estimation of the expected mortality is needed. For most types of cancer, patients diagnosed with cancer are representative of the general population, so their expected mortality can be estimated using general population mortality rates. The most notable exceptions are smoking-related cancers where patients have a lower survival than the general population due to numerous other smoking-related conditions (e.g. cardiovascular disease).

3.4.1 PROSTATE CANCER MORTALITY

The age-standardized rate for mortality from PC between 1988 and 2006 in the Netherlands is presented in Figure 1. Mortality from PC slowly increased through the 1980s, and levelled in the early 1990s. There has been a steady decrease in mortality between 1993 and 2002 from 39.3 cases / 100,000 men to 26.6 cases respectively (23.9% decrease). Many studies have explored the connection between interventions and PC mortality declines in the US and Western Europe. The first decrease is likely to be due to better treatment of more advanced disease and the increasing use of treatment with curative intent for localized disease before the PSA era[146-147]. Furthermore, incorrect attribution of cause of death may have made a substantial contribution to the decline. This bias in attribution may have been reduced after physicians got used to the idea that many men with PC not die of their disease when they were labelled as having the disease as a result of screening. Although conflicting opinions exist[8, 148-149], PSA screening likely accounts for a part of the most recent decline in PC mortality.

3.5 CONCLUSIONS

Trends in PC incidence are not reliable for measuring the increases in environmental risk factors since the practices in screening and early detection have changed too much during the last decades. Therefore, trends in PC incidence are likely to reflect the change in the intensity of screening and early detection activities over time.

Trends in PC survival are suspected to be unreliable because they are based on the same series of patients as incidence rates, and any inflation of incidence due to the inclusion of less malignant or nonmalignant diseases increases survival rates. As a result, PC survival rates should be interpreted with caution and with the knowledge of the effects of lead time, length time and overdiagnosis on survival outcomes.

Trends in PC mortality are the principal measures through which the success of the different interventions on population basis can be determined. Two mortality analyses

can be used, i.e., cause-specific mortality and excess mortality, of which cause specific mortality is more accepted in clinical epidemiology. Although both measures might be similar in practice, differences can exist (Table 1). In chapters 8-10 of this thesis the value of these mortality measures for PC analyses will be discussed in further detail.


CRUISIN' FOR SCREENING
PROSTATE CANCER

4

Scope and Outline of the Thesis

Pim J. van Leeuwen

4.1 SCOPE

It is widely accepted that PSA based screening has the potential to reduce the prostate cancer (PC) specific mortality. However, more evidence on the trade-off between the harms and benefits is needed before PC screening can be justified as a population based screening program. Concerns exist whether the reduction in PC specific mortality is underestimated by the widespread testing of asymptomatic men for PC who are participating in the control population of the ongoing randomized controlled trials. Given the modern tendency of testing asymptomatic men and the aggressive investigation in men with low PSA values, there will be little opportunity for new randomized control trials. Therefore populations with very low intensity of screening are useful for comparison. Northern Ireland represents a unique population for research on PC screening. It has excellent data on PC investigations and screening and early detection have not been routinely performed during the last decades. Comparing the intervention population of a large screening trial with the population in Northern Ireland provides an estimation of the benefits of prostate cancer screening after adjustment for contamination and a stratified analysis according to men's baseline serum PSA to asses whether the mortality reduction of screening is limited to men with PSA in a particular range.

Now since the outcome of a randomized PC screening trial shows an unfavorable trade-off between the harms and benefits, it is unrealistic to think that population based PC screening programs will be implemented according to the algorithm used within the trial. Approaches that aim to improve the screening algorithm in order to reduce the harms and to increase the benefits of screening and early detection are needed. Part IV of this thesis contributes to the improvement of future screening strategies, indicating that screening algorithms have to be made more specific and should be based on the individual risk stratification in men.

Finally, although disease specific mortality is a generally used endpoint of cancer screening trials, alternatives to this approach have to be investigated. It comprises the estimate of the excess mortality rate in the cancer patients in both arms of the study. No differences in the disease specific mortality and excess mortality rates are expected to be found. This is conditional on PC being the only factor affecting differences in mortality, the accurate ascertainment of the disease specific mortality and an accurate estimate of the expected mortality. Whether the disease specific mortality rate differs from the excess mortality rate in men, who are systematically screened for PC, and in men who are not systematically screened is relevant information. The potential absence of a discrepancy supports the established effect of screening on prostate cancer mortality, the presence of a significant discrepancy warrants further research.

4.2 OUTLINE

In Part II we describe the effectiveness of PC screening on the PC specific mortality comparing the intervention arm of the ERSPC section Rotterdam to the general population of Northern Ireland. Two studies are performed, both with a different aim. The first one to assess the effect of screening on the PC specific mortality after adjustment for the diluting effect of opportunistic PSA testing, in the control population of the randomized controlled trials. The second, to provide a stratified analysis according to men's serum PSA on study entry to asses whether the mortality reduction of screening was limited to men with PSA in a particular range.

In Part III the excess mortality analysis for the ERSPC is assessed. It includes four subsequent analyses. First we performed a pilot excess mortality analysis based on data of the ERSPC section Rotterdam. Secondly, we described the theory of performing an excess mortality analysis in a randomized controlled trial of screening for PC. This includes a method for validating the expected mortality and coping with the effect of non-attendance in the intervention arm of the trial. Third, an excess mortality analysis is performed based on the data of the four largest centres of the ERSPC. Finally, we have tried to quantify the results found in the pilot excess mortality analysis, which in contrast to the disease specific mortality rates showed an increased difference in the excess mortality rates between the two arms, by studying the causes of death among the excess deaths.

In Part IV, in chapter 11 the effect of using a 2 year screening interval or a 4-years screening is assessed by studying the proportional incidence of advanced screen and interval detected cancers in the ERSPC-Gothenburg, that uses a 2 year screening interval, and the ERSPC-Rotterdam, that uses a 4 year interval. Chapter 12 contains an evaluation of the use of sextant biopsies in men with larger prostate glands in men who participate in a screening program with a four year interval. Both studies aim to evaluate the effect of adjustment of the screening algorithm on the harms and benefits trade-off of PC screening. In chapter 13 of Part IV a competing risk stratification is made for men with localized PC who consider treatment with radical surgery for dying from PC or another cause of death.


CRUISIN' FOR SCREENING
PROSTATE CANCER

Part II

Prostate Cancer Screening: disease specific mortality


CRUISIN' FOR SCREENING
PROSTATE CANCER

5

Prostate cancer mortality in screen and clinically detected prostate cancer: Estimating the screening benefit

Eur J Cancer. 2010 Jan;46(2):377-83

Pim J. van Leeuwen, David Connolly, Anna Gavin, Monique J. Roobol, Amanda Black, Chris H. Bangma and Fritz H. Schröder

ABSTRACT

Background: To estimate the benefits of prostate specific antigen (PSA) screening on prostate cancer (Pca) metastasis and Pca specific mortality, we compared two populations with a well-defined difference in intensity of screening.

Methods: Between 1997 and 1999, a total of 11,970 men, aged 55-74 years, were included in the intervention arm of the European Randomized Study of Screening for Prostate Cancer (ERSPC) section Rotterdam. Control population consisted of 133,287 men, aged 55-74 years, between 1998 and 1999 in Northern Ireland (NI). Men were followed for Pca incidence, Pca metastasis and cause of death until December 31, 2006.

Results: Median age in both groups was 63 years at study entry (p=0.184). 94.2% of men in Rotterdam and 6% of men in NI underwent PSA testing. In Rotterdam, 1,153 men (9.6%) were diagnosed with Pca with median baseline PSA 5.1ng/ml. In NI, 3,962 men (3.0%, p<0.001) were diagnosed with Pca with median baseline PSA of 18.0ng/ml (p<0.001). The relative risk of Pca metastasis during observation in the intervention compared to control population was: 0.47 (95%CI, 0.35-0.63, p<0.001). The relative risk of Pca specific mortality was also lower in the intervention compared to the control population after a median follow-up of 8.5 years: 0.63 (95%CI, 0.45-0.88, p=0.008); absolute mortality reduction 1.8 deaths per 1000 men.

Conclusions: A relative reduction in Pca metastasis of 53% and Pca mortality of 37% was observed in the intervention population after 8.5 years of observation. The impact of overdiagnosis, quality of life benefits and cost-effectiveness need to be assessed before population-based PSA screening can be recommended.

5.1 INTRODUCTION

Over the past decade, there has been a marked decline in prostate cancer (Pca) mortality, starting in Northern America and later also observed in many European countries where currently mortality rates are lower than in the pre Prostate Specific Antigen (PSA) era.[6, 8, 150]. This decline is likely to be at least in part due to the widespread use of PSA testing, and indeed, the efficacy of PSA screening in lowering Pca mortality has now been established in a randomised controlled trial[102]. The European Randomized Study of Screening for Prostate Cancer (ERSPC) demonstrated a 20% relative reduction in Pca mortality due to PSA screening after comparing the number of men who died from Pca in a screened population with that in a control population where screening was not recommended or performed on a systematic basis[102].

Despite the ERSPC study design recommending no screening in the control population, opportunistic PSA testing occurred in 8-29% of men in the control populations in participating European countries[106-107]. The level of contamination was even higher in the Prostate, Lung, Colorectal and Ovarian (PLCO) trial in the United States, with 52% of men in the control population also undergoing PSA screening[97]. The outcomes of both these studies will have been weakened by the considerable level of Pca screening in the control populations, which may have resulted in an underestimation of the true benefits of Pca screening, and indeed, may have contributed to the lack of difference in Pca mortality between the two arms of the PLCO study.

One of the possible methods of estimating the true effect of PSA screening is applying a secondary analysis using the Cuzick-method: effect of screening in men actually screened[104, 151]. Another method is the selection of a control population with a very low intensity of screening. In Northern Ireland (NI), PSA screening is not recommended and there is a well documented low level of PSA testing (6% of men >50 years old)[152]. Further, men tended not to proceed to prostate biopsy until PSA levels were >10.0ng/ml, with few men with low PSA levels having a prostate biopsy[152]. In the current study, we compared the characteristics and outcomes of a population participating in the ERSPC, section Rotterdam, with men in Northern Ireland (NI) during a time period when asymptomatic PSA screening was infrequent[152].

5.2 MATERIALS AND METHODS

5.2.1 Intervention cohort

Between December 1993 and December 1999, a total of 42,376 men, aged 55-74 years, were randomized in the Rotterdam section of the ERSPC. All men with a prior diagnosis of Pca were excluded. In the current study, men randomized to the intervention arm between 1997 and 1999 were included. This inclusion criterion allowed an equal time of follow-up in the intervention and control population and still more than half of the men included in the intervention arm of the Rotterdam section of the ERSPC since most men were randomized in the year 1997 and 1998. Up to May 1997, men were screened with an interval of four years by PSA measurement, digital rectal examination (DRE) and transrectal ultrasound examination (TRUS). Sextant biopsy was initially offered to men with PSA ≥4.0 ng/ml and/or suspicious finding on DRE and/or TRUS. After November 1997 a biopsy was prompted by PSA ≥3.0 ng/ml only. Treatment decisions were made by local urologists and individual patient preference. Details of the screening methodology were reviewed by Roobol et al[96]. Cancers diagnosed clinically between the two screens or due to opportunistic screening, transurethral resection of the prostate (TURP) for benign disease, and cystoprostatectomy specimens, were identified and included in this cohort as interval cancers. These interval cancers were identified by means of linkage the national cancer registries. Follow-up in this respect was complete until December 31, 2006. Grading of the cancers was done using the Gleason grading system and classified according to the 1992 TNM classification. When an isotope bone scan was not performed, men with stage T1c disease and serum PSA concentration <10.0 ng/ml at diagnosis were classified as M0 and men with serum PSA concentration ≥100.0 ng/ml were classified as M1. In men with a PSA >10 ng/ml and <100 ng/ml at diagnosis, the metastatic status was considered as unknown. Pca mortality was based on the consensus of a Causes of Death Committee (CODC)[153]. This committee reviewed medical records of all men who were deceased with a Pca using a predefined decision tree[153].

5.2.2 Control cohort

Data on 133,287 men, aged 55-74 years between 1 January 1998 and 31 December 1999, in NI were included in the control cohort. Men with a prior diagnosis of any other type of cancer, except non-melanoma skin cancer, were excluded since in the ERSPC almost no men with another type of cancer participated. NI has a stable and homogenous population with little migration (0.7% annually)[152]. For this reason the group of men who were extracted from the NI population register was followed up as a cohort. All men diagnosed with Pca in NI are routinely registered by the Northern Ireland Cancer Registry (NICR). Details of the NICR PSA database and matching process have previously been described[152].Using unique identifiers (name, date of birth, and address), the NICR

links Pca data to their database of all PSA tests performed throughout Northern Ireland and to the Registrar General's Office (Northern Ireland) database of deaths. Causes of death were obtained from official national death certificates, ICD, 9th revision from 1994 until 2000 and 10th revision onwards (World Health Organization, 1992)[154]. Patients were considered to have died from Pca only if Pca (185 or C61) was coded as the primary cause of death on the death certificate. Grading of the cancers was done using the Gleason grading system. Initial treatment data were extracted from medical charts. Cancers were classified according to the 1992 TNM classification with M0 or M1 status based on the result of isotope bone scans. Where bone scans were not performed, men with serum PSA at diagnosis <10.0 ng/ml were classified as M0, men PSA >10 and <100 ng/ml at diagnosis as unknown, whilst men with serum PSA concentration ≥100.0 ng/ml at diagnosis were classified as M1. Diagnostic and mortality data were checked until December 31, 2006.

5.2.3 Validation of cause of death data in Northern Ireland

To validate the cause of death data from death certificates in NI, a random sample of 136 men who had Pca and subsequently died were identified (7.6% of total, median age 77 years). All available information from General Practitioner and hospital charts was extracted and reviewed independently by two authors (PVL, DC). Using the predefined ICOD flowcharts[153], a cause of death was assigned in each case. In cases where there was a discrepancy in the assigned cause of death, the notes were reassessed until a consensus was reached. In 119 men (87.5%), the death certificate data matched that of the performed review. Four men (2.9%) had insufficient information to assign a cause of death. Of the 13 inaccurate recorded causes of death, six were incorrectly recorded as primary cause Pca and seven incorrectly recorded as due to intercurrent disease. This resulted in a death certificate accuracy of 90.4% (SD 2.19).

5.3 STATISTICAL ANALYSIS

The chi-square (χ^2) and the Mann-Whitney *U* tests were used to assess the relationship between categorical and continuous variables, respectively, between the intervention and the control cohort. Pca metastasis and mortality risk ratios between the two populations were estimated using a Poisson regression analysis. For both groups the number of man-years were calculated from the date of their study entry up to their date of death or December 31st, 2006 when still alive. The Nelson-Aalen analysis was used for the graphical estimation of the Pca mortality and Pca metastasis cumulative hazards[155], cumulative survival percentages are presented. In both cohorts the survival time was defined as the time from study entry until Pca death, with censoring at the date of an intercur-

rent death or December 31, 2006. For Pca metastasis the survival time was defined as the time from study entry until Pca metastasis with censoring at date of death if death occurred prior to a metastasis or December 31, 2006. A two-sided *p* value <0.05 was considered to be statistically significant. All analyses were performed with STATA: Data Analysis and Statistical Software, version 10.0.

5.4 RESULTS

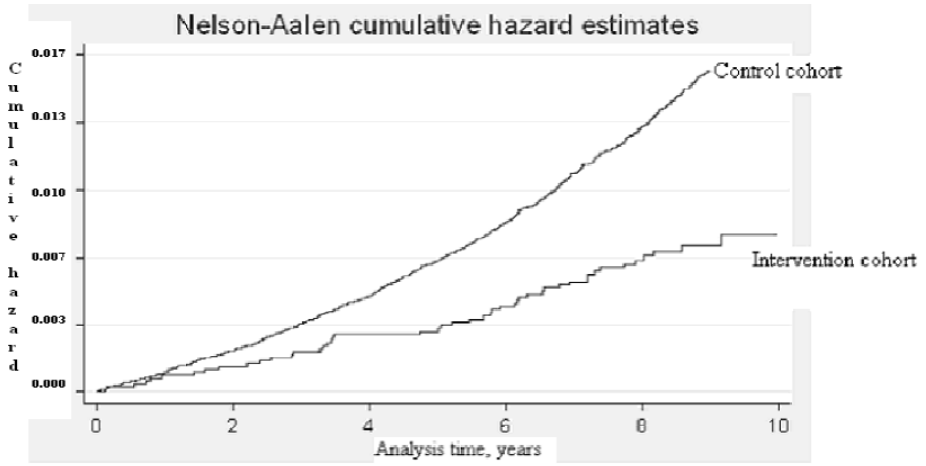
In the intervention cohort, 11,970 men, median age 63 years, were included, with 1,153 (9.6%) of these diagnosed with Pca during the follow-up period. In the control cohort 133,287 men were included, with 3,962 (3.0%) diagnosed with Pca with identical follow-up. Baseline patient characteristics are presented in Table 1. Median age at inclusion was similar, however the age distribution at inclusion was different for both groups (p=0.184 and p<0.001 respectively). Age at diagnosis was higher in the control cohort (median 70 vs. 67 years, *p*<0.001) with a higher median PSA at diagnosis (18.0 vs. 5.1 ng/ml, *p*<0.001). Median follow-up was 8.53 years in the intervention population and 8.72 years in the control population. In the intervention cohort 100% of Pca diagnoses were confirmed histologically by prostate biopsy (99.7%), cystoprostatectomy specimen (0.1%) or TURP (0.2%). In the control cohort 68.2% of men were diagnosed by prostate biopsy; 18.1% and 13.7% were diagnosed by TURP or on the basis of clinical opinion only (no histological confirmation), respectively.

In the intervention cohort 11 men (0.1% of total) and in control cohort 862 men (0.6% of total) had Pca metastasis at diagnosis (*p*<0.001). There was a significant reduction of 53% in Pca metastasis during observation in the intervention population relative to the control population: RR 0.47 (95%CI, 0.35-0.63, *p*<0.001). Pca metastasis cumulative hazard is graphically illustrated in Figure 1. As shown, the cumulative Pca metastasis hazard starts to differ after two years of observation and becomes statistically significant after five years. Furthermore, as demonstrated in Figure 1, this difference in metastatic disease is likely to increase with longer follow-up.

In the intervention cohort 35 (0.29%) men and in the control cohort 627 (0.47%) men died due to Pca or to a Pca intervention related procedure. This equated to a reduction in Pca mortality of 37% in the intervention population relative to the control population: RR 0.63 (95%CI, 0.45-0.88, *p*=0.008). The Pca specific cumulative hazards are graphically illustrated for both cohorts in Figure 2. The difference in Pca specific mortality, expressed by a cumulative hazard, becomes statistically significant six years after the start of observation. After a median follow up of 8.5 years, the absolute rate of Pca mortality was 0.36 per 1,000 person-years in the intervention cohort compared to 0.58

Table 1. Baseline characteristics at study start, diagnosis, and primary treatment modalities
Median follow up 8.5 years

	Control cohort Northern Ireland *N* (% of total)	Intervention cohort ERSSC Rotterdam *N* (% of total)	*P* value
Total participants included	133 287	11 970	
Age (yr), median	63	63	0.184
55 – 60	52 104 (39.1)	4 310 (36.0)	<0.001
61 - 65	33 013 (24.8)	3 153 (26.3)	
66 - 70	23 423 (17.6)	2 723 (22.7)	
71 - 74	24 747 (18.6)	1 784 (14.9)	
Total patients diagnosed,	3962	1153	<0.001
% of total participants	(3.0)	(9.6)	
Age (yr) at diagnosis, median	70	67	<0.001
55 – 60	348 (8.8)	175 (15.2)	<0.001
61 - 65	790 (19.9)	314 (27.2)	
66 - 70	1069 (27.0)	379 (32.9)	
71 - 75	1108 (28.0)	265 (23.0)	
≥ 76	647 (16.3)	20 (1.7)	
PSA at diagnosis (ng/ml), median	18.0	5.1	<0.001
0.0 – 2.9	193 (4.9)	134 (11.6)	<0.001
3.0 – 4.9	168 (4.2)	435 (37.8)	
5.0 – 9.9	740 (18.7)	361 (31.3)	
10.0- 19.9	1 003 (25.3)	126 (10.9)	
≥ 20.0	1 799 (45.4)	89 (7.7)	
Not known or not performed	59 (1.5)	8 (0.7)	
Disease extent			
Not metastasized (M0)	2 718 (68.6)	1 119 (97.0)	<0.001
Metastasized (M1)	862 (21.8)	11 (1.0)	
Not known or not performed	382 (9.6)	23 (2.0)	
Histological differentiation			
Gleason 2 – 6	1 638 (41.3)	790 (68.6)	<0.001
Gleason 7	850 (21.5)	247 (21.4)	
Gleason 8 – 10	932 (23.5)	51 (4.4)	
Not known or not performed	542 (13.7)	65 (5.6)	
Initial treatment			
Radical prostatectomy	277 (7.0)	416 (36.1)	<0.001
Radiotherapy	1 106 (27.9)	442 (38.3)	
Watchful waiting	419 (10.6)	233 (20.2)	
Androgen-deprivation therapy	1 248 (31.5)	48 (4.2)	
Not known or other treatment	912 (23.0)	14 (1.2)	



Years	0	2	4	6	8	10
Intervention cohort						
N at risk	11 970	11 701	11 324	10 912	7315	0
Cumulative N Pca metastasis (survival rate %)	0 (100)	9 (99.9)	20 (99.8)	29 (99.8)	43 (99.7)	47 (99.6)
Control cohort						
N at risk	133 287	127 872	122 567	116 501	104050	0
Cumulative N Pca metastasis (survival rate %)	0 (100)	162 (99.9)	363 (99.7)	631 (99.5)	956 (99.3)	1 126 (99.2)

Figure 1. Incidence of prostate cancer distant metastasis in the intervention and control population.

per 1,000 person-years in the control cohort; the absolute risk difference was 1.8 deaths per 1000 men (1.8 per 1000 men = 627 Pca deaths/133,287 men control population - 35 Pca deaths/11,970 men intervention cohort), which correspond to 555 (1000/1.8) men needed to be screened to save one Pca death[156]. Additional Pca diagnosed by screening resulted in an increase in cumulative incidence with respect to the control population of 67 per 1000 men, i.e. 37 (555/1000*67) cases had to be treated (NNT) in order to prevent one death from Pca[156]. These estimates are all cumulative and therefore interpreted as the probability, or risk, that an individual will have during the 8.5 years of observation[157].

During follow-up, 1,676 men (14.0% of total) died in the intervention cohort, which was significantly lower than the overall mortality of 27,083 men (20.3% of total) in the control cohort: RR 0.70 (95%CI 0.66-0.73 $p<0.001$).

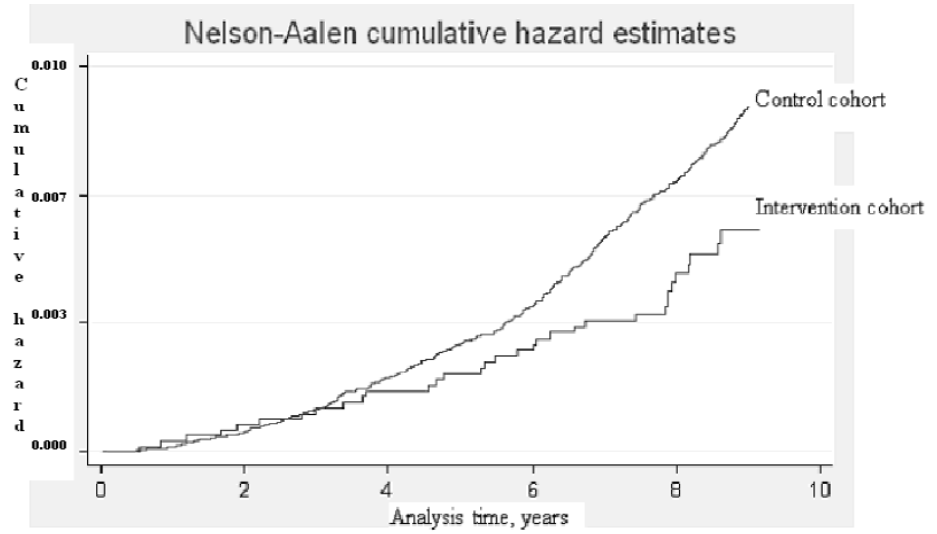
5.5 DISCUSSION

The ERSPC study demonstrated a 20% decrease in Pca mortality due to PSA screening, whilst the PLCO trial did not find any Pca specific mortality reduction[97, 102]. A weakness in both these trials, more so in the PLCO trial, was the level of opportunistic screen-

ing in the control populations. Given the high level of PSA testing and the high rate of screen-detected Pca in many countries throughout the world, any randomized trial will have similar difficulties with contamination of the control population, which may lead to an underestimation of the true value of screening. In the current study, we assessed retrospectively the rate of Pca metastasis and Pca specific mortality in men who did not undergo systematic screening or early investigation, and compared this with men who were prospectively screened for Pca. Our aim was to estimate the true benefits of PSA screening from a Pca screening trial, when there is a low level of contamination in the control population.

The main finding is the absolute mortality reduction of 1.8 deaths per 1000 men in favour of the screened population, which corresponds to a relative risk reduction of 37%, after a median follow-up of 8.5 years. These results compare favorably to the ERSPC study that found an absolute mortality reduction of 0.71 per 1000 men after an average follow-up of 8.8 years, and a relative reduction of 20%[102]. In the present study, 555 men needed to be screened and an additional 37 men needed treatment to prevent one Pca related death, which again, is lower than the ERSPC findings (1410 screened and 48 treated respectively)[102]. The trends in prostate specific mortality in both studies are however similar; there is overlap in the survival curves in the early part of observation, which then diverge with time. In the current study, this divergence happens earlier (4 years vs. 7 years in the ERSPC) and becomes more pronounced over time (Figure 2), leading to a greater overall benefit due to screening, although the mortality difference did not become statistically significant until after six years of observation. By the end of observation, the mortality rate rose more slowly in the intervention population and, given the changes noted in the rates of distant metastasis (Figure 1), this trend is likely to continue with further follow-up. Therefore, there are few benefits of screening in the initial years after PSA testing, but these benefits are likely to increase over time; there is a difference in the rate of distant metastasis in favour of screening after five years, which leads to a disease specific mortality benefit after six years of observation. Hence, screening will only be beneficial in men with a life expectancy from at least another six to eight years.

Although the absolute benefit in terms of deaths prevented by screening might increase after longer follow up, the life-years saved per death prevented might tend to be smaller after more years of follow up. Based on the data of current study we assume an average of 5 years saved per death prevented in current study. Consequently, we expect that the benefit in terms of the number of life years gained will become smaller after longer follow up when the patients become older and will have limited life expectancies. However, for evaluation of the screening studies, these estimates are less used because they



Years	0	2	4	6	8	10
Intervention cohort						
N at risk	11 970	11 706	11 333	10 923	7326	0
Cumulative N Pca specific death (survival rate %)	0 (100)	5 100)	11 (99.9)	19 (99.8)	29 (99.8)	35 (99.7)
Control cohort						
N at risk	133 287	127 972	122 706	116 689	104 226	0
Cumulative N Pca specific death (survival rate %)	0 (100)	39 (100)	145 (99.9)	285 (99.8)	508 (99.6)	627 (99.5)

Figure 2. Prostate cancer specific mortality in the intervention and control population.

are confusing and difficult to understand. Furthermore, the life years gained should be corrected for quality of life, making it quality of life adjusted life years (QALY's) gained.

Another important observation is the screening induced increase in the Pca incidence. The intervention cohort had a 3.2 fold increased rate of Pca diagnosis, with a cumulative incidence increase of 67 per 1000 men. These results compare to 34 per 1000 men in the ERSPC and 12.5 per 1000 men in the PLCO studies, again highlighting the degree of contamination in the control arms of these studies, and the advantages of the control population in the present study[97, 102]. In the current study, the increased incidence rate was much greater than the observed decrease in mortality between the two groups (67 vs. 1.8 per 1000 men respectively), leading to the potential for overdiagnosis (and over-treatment) of the majority of men diagnosed by screening. As the life expectancy of men at end of observation is approximately 10 years, further follow-up will determine how many additional men are diagnosed with Pca and how many die from Pca or concurrent disease, giving more definitive data on overdiagnosis due to PSA screening.

The main limitation is the lack of randomisation, meaning the current study is more akin to studies which have compared rates of prostate cancer diagnosis and prostate specific mortality in areas with high and low rates of PSA screening and radical treatment, such as in Seattle-Connecticut[158] and in the Tyrol region of Austria[159]. The current study has advantages over previous studies as the difference in PSA testing and treatment is more pronounced in the current populations, and will more accurately estimate the potential benefits of PSA screening. Further, as there are individualised data in each population in the current study, survival analyses can also be performed as opposed to simply assessing differences in prostate cancer incidence and mortality rates.

Due to the lack of randomisation, the rates of Pca or Pca mortality may have been different in both groups at the beginning of observation. At inclusion, the incidence of Pca in both countries was very different (86 vs. 62 per 100,000 persons in the Netherlands and NI respectively)[160-161]. This difference was mainly due to the early use of PSA testing in the Netherlands, with Pca rates increasing significantly from a baseline of 62 per 100,000 persons in the end-1980s to over 90 per 100,000 by 2002[160-161]. In NI, there was an equal baseline incidence of 63 per 100,000 persons, but this remained stable until 2000, when the rates started to increase. In contrast, the Pca mortality rates in both countries were remarkably similar, with a slow increase in mortality until its peak in 1995 (34.4 per 100,000 persons in Netherlands vs. 28.5 in NI)[160-161], with a subsequent decrease in both countries. As men in both populations had similar ethnic backgrounds (virtually all white) and had an equivalent median age at inclusion (63 years), they should have a similar baseline risk of Pca. However, the higher level of PSA testing in the Dutch general population before 1995 means that many men have been pre-screened using PSA and will have already been diagnosed with Pca, especially more advanced Pca, and so will not have been offered inclusion into the screen-detected population. This, and the fact that many more men in the Dutch population had underwent PSA testing prior to inclusion in the study population (on average 14% vs. 4%) will bias outcomes in favour of the screened group. Further, it is likely that a small number of men in the NI population may also have undergone PSA screening (9.1% of cancers diagnosed with PSA <5.0ng/ml). There is therefore some degree of contamination in the control population of the present study, although the magnitude of this will be much less than in the ERSPC and PLCO studies[97, 102]. The method of inclusion i.e. men in Rotterdam signing informed consent whilst those in NI being identified retrospectively, also resulted in a healthy screening bias with generally healthier men of higher socio demographic level agreeing to participate in the ERSPC[162]. This resulted in a large difference in overall mortality: RR 0.75 (95%CI 0.73-0.76 $p<0.001$) in favour of the intervention group. As men in the control cohort died sooner, they were more likely to die from a co-morbid cause as opposed to Pca, decreasing the Pca mortality relative to that in the intervention group. Finally, different treatments in both cohorts will have affected outcomes, with men diagnosed and

treated with curative intent at an earlier stage likely to have a better outcome[54, 163-164]. In both groups, following diagnosis, men were free to choose treatment in collaboration with their local urologist. As outlined in Table 1, men in NI had a higher PSA at diagnosis and a higher rate of metastatic disease, they were therefore less likely to undergo prostatectomy and more likely to have androgen deprivation therapy. These differences in treatment are inherent in any study with a wide difference in the intensity of screening.

A number of criteria must be met before population-based screening can be justified[60]. Little is known about the screening risks, side effects of overtreatment and health related quality of life benefits of earlier treatment. Further, given the very large number needed to screen and needed to treat, it seems likely that cost of population-based PSA screening will be considered prohibitive in many countries.

5.6 CONCLUSION

Men undergoing systematic PSA screening had a 3.2 fold increased diagnosis of Pca. After 8.5 years, the rate of Pca metastasis was 53% lower in the intervention population. Further, a significant reduction in Pca specific mortality of 37% was observed in the intervention cohort, however 555 men needed to be screened and 37 men needed treatment to prevent one Pca related death. Longer follow up is likely to demonstrate an increasing mortality benefit in favour of PSA screening, although the impact of overdiagnosis, quality of life benefits and cost-effectiveness need to be assessed before population-based PSA screening can be recommended.


CRUISIN' FOR SCREENING
PROSTATE CANCER

6

Balancing the harms and benefits of early detection of prostate cancer

Cancer. 2010 Sep 13.

Pim J. van Leeuwen, David Connolly, Teuvo L.J. Tammela, Anssi Auvinen, Ries Kranse, Monique J. Roobol, Fritz H. Schröder and Anna Gavin

ABSTRACT

Background: The benefits of prostate cancer (Pca) screening on an individual level remain unevaluated.

Methods: Between 1993 and 1999, a total of 43,987 men, aged 55-74 years, were included in the intervention arm of the European Randomized Study of Screening for Prostate Cancer (ERSPC) section in the Netherlands, Sweden and Finland. A total of 42,503 men, aged 55-74 years, were included in a clinical population in Northern Ireland. Serum PSA (<20.0ng/ml) was measured in all men at study entry. All men were followed for Pca incidence and causes of death until December 31, 2006.

Results: The adjusted absolute difference in Pca specific mortality between the intervention population and the clinical population increased with increasing PSA level at study entry, i.e. 0.05 per 10,000 person-years for men serum PSA 0.0-1.9 ng/ml and 8.8 per 10,000 person-years for men serum PSA 10-19.9 ng/ml. The risks of early detection, i.e. the number needed to investigate (NNI) and number needed to treat (NNT) to save one death from Pca were increased with the decreasing PSA levels at study entry, i.e. the NNI for men serum PSA 0.0-1.9 ng/ml was 24,642 men, and for men serum PSA 10-19.9 ng/ml 133 men; the NNT for men serum PSA 0.0-1.9 ng/ml was 724 men and for men with a serum PSA 10-19.9 ng/ml 60 men.

Conclusions: For men with a low serum PSA, the benefits of aggressive investigation and treatment may be limited since they are associated with a large increase in cumulative incidence and potential overtreatment.

6.1 INTRODUCTION

Prostate cancer (Pca) screening has been subject to much controversy for many years. Recently, the European Randomized Study of Screening for Prostate Cancer (ERSPC) showed a 20% reduction in Pca mortality in the screened population relative to the control population[98]. Secondary analysis of the ERSPC showed that the mortality effect among men was increased to approximately 30% after adjusting for the diluting effect of nonattendance and contamination[105, 165].

The results of the ERSPC did not include patient individual risk stratifications. As blood was not collected at randomization from men in the control group, stratified analysis between the two arms of the ERSPC according to serum PSA on study entry to assess whether the mortality reduction was limited to men with PSA in a particular range was not possible. These analyses are important as the 20% relative reduction in Pca mortality was associated with a considerable increase in the cumulative excess incidence, with screening of 1410 men and treatment of 48 additional cases required to prevent one death from Pca[98].

For this reason, we compared the Pca incidence and Pca specific mortality rates stratified by individual serum PSA level that was measured at study entry in men participating in the intervention arm of the ERSPC and men in Northern Ireland, in whom screening and early detection of Pca was not routinely performed.

6.2 MATERIALS AND METHODS

6.2.1 Intervention population

Between December 1993 and December 1999 a total of 63,153 men, aged 50-74 year, were randomized into the intervention arm of the ERSPC section the Netherlands, Sweden and Finland. For the current study, only men aged 55-74 years, who did not have Pca and were actually screened by PSA, were included. Men with a baseline serum PSA ≥20.0 ng/ml at study entry were excluded since the main focus of this paper was the potential value of early detection. Although legal requirements with respect to randomized trials were different in the Netherlands, Finland and Sweden, written informed consent was required for those who were randomly assigned to the intervention arm of the study[92, 94, 96].

In the Netherlands, men were screened by PSA measurement, digital rectal examination (DRE) and transrectal ultrasound examination (TRUS) between 1993 and 1997. Sextant biopsy was initially offered to men with PSA ≥4.0 ng/ml and/or suspicious finding on DRE and/or TRUS. After May 1997 a biopsy was prompted by PSA ≥3.0 ng/ml only. In Sweden, a sextant biopsy was indicated for men with a level of PSA ≥3.0 ng/

ml. In Finland, men with PSA ≥4.0 ng/ml were defined as screen positive, and men with PSA 3.0-3.9 had an ancillary test (DRE until 1998, free/total PSA ratio with a cut-off ≤ 0.16 from 1999 onwards). In Finland, sextant biopsy was initially offered to the screen positive men, however in 2002 a biopsy procedure with 10-12 biopsy cores was adopted as a general policy. In Sweden men were screened with an interval of two years until the age of 70 years in contrast to men in the Netherlands and Finland who were screened with an interval of four years until the age of 74 years and 71 years respectively. The screening algorithms used in the centers have been extensively described previously[92, 94, 96].

In all centers the treatment decisions were made by a local urologist and based on individual patient preference. Cancers diagnosed between the two screening intervals or after the maximum screening age, clinically or due to opportunistic screening, transurethral resection of the prostate (TURP) for benign disease, and cystoprostatectomy specimens, were considered as well and defined as interval cancers. These interval cancers were routinely identified by means of linkage to the national cancer registry. Cancers were classified according to the 1992 TNM classification. Men with stage T1c disease and serum PSA concentration <10.0 ng/ml were classified as M0 and men with serum PSA concentration ≥100.0 ng/ml were classified as M1, when an isotope bone scan was not performed. In men with a PSA ≥10.0 and <100.0 ng/ml in whom an isotope bonescan was not performed the metastatic status was considered unknown. Pca mortality was based on the consensus of an independent Causes of Death Committee (CODC) in Sweden and the Netherlands and on accurately validated official causes of death certificates in Finland[143, 153]. Diagnostic and mortality data was available until December 31, 2006.

6.2.2 Clinical population

Data on men, aged 55-74 years, who had a first serum PSA between January 1994 and December 1999 in Northern Ireland, were included. Men with a prior diagnosis of Pca or a baseline serum PSA ≥20.0 ng/ml at study entry were excluded. Data were retrospective obtained from a population-based database of the Northern Ireland Cancer Registry (NICR) that included all routinely performed PSA tests since 1994. The NICR maintains this confidential electronic database of all PSA results for Pca surveillance purposes. No personal identifiable information was removed from the database and no patient contact was made during this study. During the years of observation, the clinical population was not systematically screened, because in Northern Ireland early detection and screening was not recommended and the population had a well documented low level of PSA testing (6% of men >50 years old)[152]. Further, men tended not to proceed to prostate biopsy until PSA levels were >10.0ng/ml, with few men with low PSA levels having a prostate biopsy[152]. No individual data on reason for PSA testing was available but recent evidence showed that during the period of 1994-1998 less than 20% of PSA

testing was in asymptomatic men[166]. The NICR registers all Pca cases and links these to their PSA data. Causes of death were obtained from accurately validated official national death certificates, ICD, 9th revision from 1994 until 2000 and 10th revision onwards (World Health Organization, 1992)[105, 154]. Cancers were classified according to the 1992 TNM classification with M0 or M1 based on the result of isotope bone scans. Where bone scans were not performed, men with serum PSA concentration <10.0 ng/ml were classified as M0 whilst men with serum PSA concentration ≥100.0 ng/ml were classified as M1. In men with a PSA ≥10.0 and <100.0 ng/ml at diagnosis in whom an isotope bonescan was not performed, the metastatic status was considered unknown. Diagnostic and mortality data was available until December 31, 2006.

6.3 STATISTICAL ANALYSIS

For both groups the time of follow up was measured from the date of their first PSA test up to their date of death or December 31, 2006. Baseline serum PSA was stratified into four categories, PSA 0.0-1.9, 2.0-3.9, 4.0-9.9 and 10.0-19.9 ng/ml. A multivariate Poisson regression analysis was used with the time of follow-up (divided into 2-years intervals) until either the event of interest or censoring. The following model was used: $Log\ [E(Y)] = log(exp) + \beta_0 + \beta_1 x_1 + ... + \beta_5 x_5$; a generalized linear model with log link function and Poisson distributed errors where $E(Y)$ is expected number of Pca deaths, $log(exp)$ is the logarithm of the follow up time, $(x_1, x_2, ..., x_p)^T$ are the predictive variables, i.e. PSA (categories), age (continue), study population and the time interval since first screening visit (2-years intervals). The βi is the coefficient corresponding to x_i. The term $log(exp)$ was an offset with the parameter estimate constrained to 1 which enables the interpretation of the parameter estimates as rate ratio's. Comparisons between the observed and predicted data of this multivariate model showed the predictions to be accurate. In addition, the number needed to investigate (NNI) and the number needed to treat (NNT) to save one death from Pca were calculated[156], based on the adjusted absolute rate differences. In both populations for baseline serum PSA and age adjusted cumulative hazards were graphically estimated for different PSA categories at study entry using a cumulative hazard method. All analyses were performed with the commercially available STATA package: Data Analysis and Statistical Software, version 10.0; the cumulative hazard figures with Statistical Package for the Social Sciences software, version 16.0.

6.4 RESULTS

6.4.1 Baseline characteristics

A total of 42,503 men were included in the clinical population and a total of 43,987 men were included in the intervention population. Participants had statistically significant differences in their baseline characteristics at study entry (Table 1); the median age and baseline serum PSA were higher in the clinical population. The median follow up time was 8.8 (SD 3.1) and 9.1 (SD 2.2) years for the clinical and intervention population respectively.

Table 1. Characteristics of men at study entry

	Intervention group N (% of total)	Clinical group N (% of total)	P value
Total participants included	43987	42503	
Age (yr), median	61	65	<0.001†
PSA at study entry (ng/ml),			
median	1.18	1.60	<0.001†
0.00– 1.99	32035 (72.8)	25555 (60.1)	<0.001‡
2.00 – 3.99	7467 (17.0)	8703 (20.5)	
4.00 – 9.99	3927 (8.9)	6493 (15.3)	
10.00 – 19.99	558 (1.3)	1752 (4.1)	

† Mann-Whitney U test, ‡ Chi-square test

6.4.2 Prostate Cancer diagnosis

Pca was diagnosed in 1,522 men (3.6%) in the clinical population and in 4,339 men (9.9%) in the intervention population, adjusted rate ratio (RR) 4.61 (95%CI, 4.33-4.91). Patients in the clinical population were diagnosed at an older age and with a higher median serum PSA (Table 2). The number of men with a positive result on an isotope bonescan (or a PSA value of more than 100.0 ng/ml in those without bonescan results) at diagnosis was 6.1 per 1000 men in the clinical group and 1.2 per 1000 men in the intervention group (p<0.001). The median time to a Pca diagnosis was lower in the intervention group than in the clinical group, 4.1 vs. 5.3 years respectively, p<0.001.

The Pca incidence rates increased with increasing baseline PSA level in both study populations (Table 3). The adjusted absolute rate differences on a cancer diagnosis between the intervention group and the clinical group increased with the increasing baseline PSA levels (Table 4).

Table 2. Patient characteristics at diagnosis

	Intervention group *N* (% of total)	Clinical group *N* (% of total)	*P* value
Total patients diagnosed, % of total participants	4339 (9.9)	1522 (3.6)	<0.001‡
Age (yr) at diagnosis, median	66	71	<0.001†
PSA at diagnosis (ng/ml), median	5.0	12.8	<0.001†
Disease extent at diagnosis % of total participants			
Not metastasized (M0)	4285 (9.8)	1129 (2.7)	<0.001‡
Metastasized (M1)	54 (0.1)	261 (0.6)	
Not known	0	124 (0.3)	

† Mann-Whitney U test, ‡ Chi-square test

Table 3. Adjusted rate ratio of prostate cancer incidence for serum PSA at study entry

	Intervention population				Clinical population			
PSA at baseline	N at risk	N diagnosis	Rate ratio (95%CI)	P value	N at risk	N diagnosis	Rate ratio (95%CI)	P value
PSA 0.0 – 1.99	32009	980	*		25555	243	*	
PSA 2.0 – 3.99	7467	1553	6.80(6.27-7.37)	<0.001	8703	313	3.66(3.09-4.33)	<0.001
PSA 4.0 – 9.99	3889	1472	12.62(11.62-13.71)	<0.001	6493	611	9.56(8.22-11.12)	<0.001
PSA 10.0 – 19.99	539	334	21.67(19.09-24.59)	<0.001	1752	355	21.48(18.18-25.38)	<0.001

N: Observed number of men at risk. * Reference group to which other groups are compared. The reference group per definition has a rate ratio of 1.

Table 4. Adjusted absolute difference in prostate cancer incidence between the intervention population and the clinical population

PSA at baseline	Observed Pca incidence Intervention	Observed Pca incidence Clinical	Absolute difference PC incidence
PSA 0.0 – 1.99	23.47	11.52	35.59
PSA 2.0 – 3.99	235.95	43.21	214.32
PSA 4.0 – 9.99	431.29	144.71	377.40
PSA 10.0 – 19.99	709.80	253.04	561.76

* Pca: Prostate cancer. Pca incidence in rates per 10,000 man-years. Absolute difference as rate per 10,000 man-years.

6.4.3 Prostate Cancer mortality

By the end of 2006, the overall mortality was 25.5% in the clinical group and 14.5% in the intervention group, adjusted RR 0.79 (95%CI, 0.77-0.82). In total 236 men (0.6%) died from a Pca related cause of death in the clinical population and 109 men (0.2%) died from a Pca related cause of death in the intervention population. This resulted in

Table 5. Adjusted rate ratio of prostate cancer specific mortality for serum PSA at study entry

	Intervention population				Clinical population			
PSA at baseline	N at risk	N Pca deaths	Relative risk (95%CI)	P value	N at risk	N Pca deaths	Relative risk (95%CI)	P value
PSA 0.0 – 1.99	32009	26	*		25555	29	*	
PSA 2.0 – 3.99	7467	26	3.97(2.29-6.87)	<0.001	8703	44	3.39(2.5-6.29)	<0.001
PSA 4.0 – 9.99	3889	38	10.78(6.46-17.99)	<0.001	6493	89	10.09(6.59-15.43)	<0.001
PSA 10.0 – 19.99	539	19	37.17(20.13-68.62)	<0.001	1752	74	31.05(20.03-48.11)	<0.001

N: Observed number of men at risk. * Reference group to which other groups are compared. The reference group per definition has a RR of 1.

Table 6. Adjusted absolute difference of prostate cancer specific deaths between the clinical population and the intervention population, adjusted NNI and NNT to save one men from prostate cancer death

PSA at baseline	Proportion study population	Median follow-up in years	Observed Pca deaths Intervention	Observed Pca deaths Clinical	Absolute difference PC mortality	NNI	NNT
PSA 0.0 – 1.99	66.6%	8.9	0.92	1.37	0.05	24642	724
PSA 2.0 – 3.99	18.7%	9.0	3.95	6.07	0.47	2393	427
PSA 4.0 – 9.99	12.0%	8.9	11.13	16.71	2.34	492	152
PSA 10.0 – 19.99	2.7%	8.7	40.38	52.75	8.88	133	60

Pca deaths in rates per 10,000 man-years. Absolute difference as rate per 10,000 man-years. NNI: number needed to investigate to save one death from prostate cancer. NNT: number needed to treat to save one death from prostate cancer

for age and baseline serum PSA (continue) adjusted non-significant relative reduction in Pca specific mortality of 20% in the intervention population relative to the clinical population, RR 0.80 (95%CI, 0.63-1.02).

The Pca mortality rates increased with increasing baseline PSA level in both groups (Table 5). Relative to the men with a baseline serum PSA <2.0 ng/ml at study entry, men with a higher baseline serum PSA had a significant increased adjusted RR of dying from Pca in both groups. The absolute difference in Pca specific mortality was 0.05 per 10,000 person years in men with a baseline serum PSA 0-1.9 ng/ml and 8.88 per 10,000 person years in men with a baseline serum PSA 10-19.9 ng/ml, increasing with the increasing baseline PSA level (Table 6).

The age adjusted cumulative hazard graphs for Pca specific death by baseline PSA and study group are presented in Figure 1-4. Relative to the lowest PSA category, the main absolute difference in Pca mortality was observed for the PSA categories 4.0-9.9 ng/ml and 10.0-19.9 ng/ml at study entry. Furthermore, non or a small difference in the cumulative hazard in Pca death was observed for men with a PSA 0.0-1.9 ng/ml and 2.0-3.9 ng/ml at study entry.

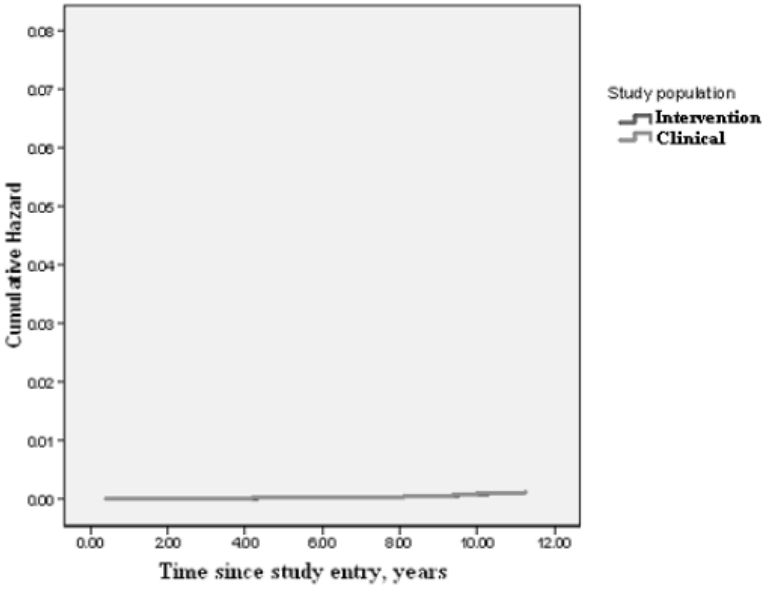


Figure 1. Prostate cancer specific death in men serum PSA 0.0-1.99 ng/ml

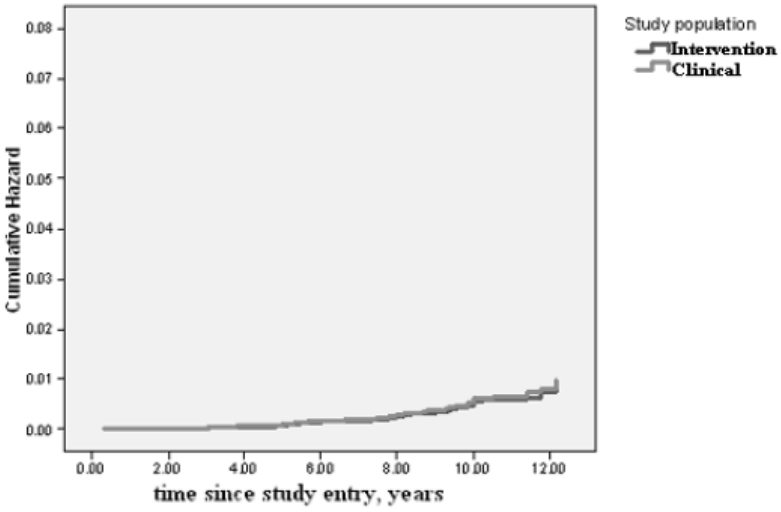


Figure 2. Prostate cancer specific death in men serum PSA 2.0-3.99 ng/ml

6.4.4 Potential harms early detection

Table 6 summarized the adjusted magnitude of early detection and treatment for the baseline PSA categories in terms of the NNI and NNT to save one men from Pca death. NNI and NNT decreased with increasing baseline PSA level. NNI varied from 24,642 men for men with a baseline serum PSA 0.0-1.9 ng/ml to 133 for men with a baseline serum

PSA 10.0-19.9 ng/ml. NNT varied from 724 men for men with a baseline serum PSA 0.0-1.9 ng/ml to 60 for men with a baseline serum PSA 10.0-19.9 ng/ml.

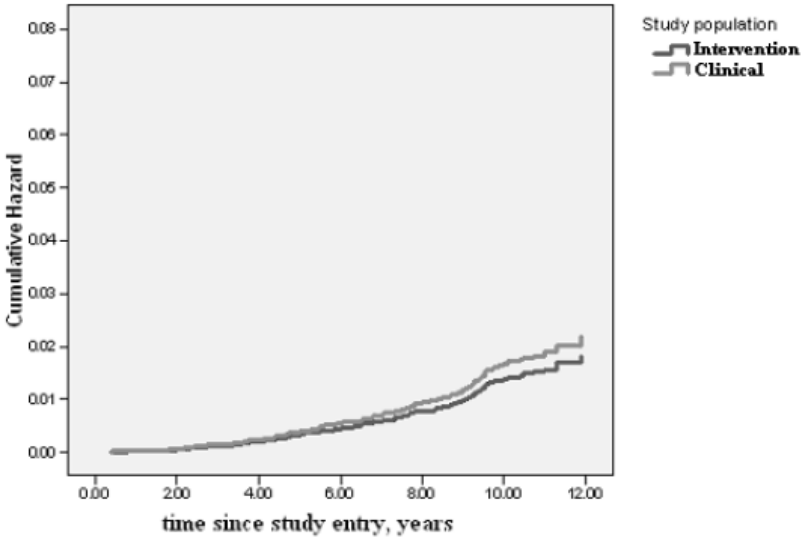


Figure 3. Prostate cancer specific death in men serum PSA 4.0-9.99 ng/ml

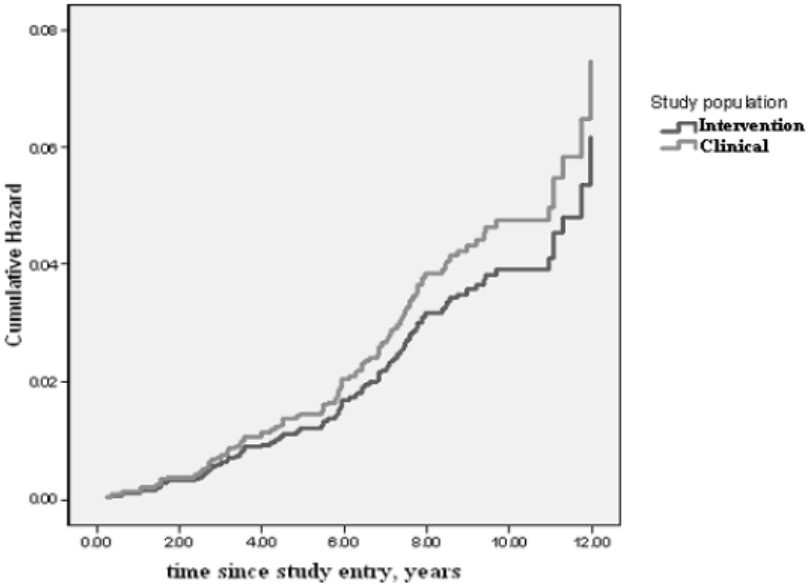


Figure 4. Prostate cancer specific death in men serum PSA 10.0-19.99 ng/ml

6.5 DISCUSSION

Early detection and screening for Pca has potential harms. Pca screening increases the incidence, which might cause needless worry and expense for a lot of men who may be getting treatment for tumours growing too slowly to do any harm[129]. However, screening and early detection has potential benefits as well. These benefits are a reduction in Pca mortality and a decrease in the number of men that suffer from the complications of advanced disease[98, 105]. In current study the potential balance between these harms and benefits is demonstrated using the measures NNI and NNT. The NNI equals one divided by the absolute mortality reduction and the NNT equals 1 divided by the absolute mortality reduction multiplied by the excess Pca incidence. Consequently, the NNI and NNT show, on the one hand the effectiveness of early detection in terms of the reduction in Pca mortality, and on the other hand the harms of early detection in terms of the percentage of men that is diagnosed with a potential overdiagnosed Pca.

This study provides additional information on how the harms and benefits of screening, early detection and treatment are distributed in relation to the baseline PSA levels. It demonstrated that the yield of Pca increased with the increasing baseline serum PSA at study entry. The benefits of early detection might be small for men with a baseline serum PSA 0-3.9 ng/ml at study entry. Despite the short follow-up, especially for men with a baseline serum PSA <2.0 ng/ml at study entry aggressive investigation and treatment yielded little or no mortality reduction while a significant increase in the cumulative incidence of Pca was observed. These observations are in line with studies that show a strong correlation between the lower baseline PSA values and the detection of cancers with potential indolent tumour characteristics[152, 167-169]. These results were confirmed in current study, men with relatively low serum PSA at study entry were more often diagnosed with Pca with a favorable tumour stage and pathological characteristics (data not shown).

The main purpose of this study was to add information to the existing results of the ERSPC by providing a PSA risk stratification that would avoid misuse and maximization of PSA testing. Our results suggest that, assuming that the risk distribution according to the different PSA levels in this study was similar to the ERSPC, most of the absolute reduction in Pca mortality is achieved in men who had a moderately elevated PSA at study entry. In other words, the greatest benefits of early detection programs may be when men, aged 55-74 years, are diagnosed and treated when their serum PSA is in the range 4.0-9.9 ng/ml or 10.0-19.9 ng/ml. Furthermore, following research efforts that recommend more intensive PSA based screening by lowering the PSA cut-off, may greatly increase the number of men that need additional investigations and treatment, whilst having little effect on the reduction of Pca mortality.

Our second main observation is that a large cumulative excess incidence was observed in men with a low baseline serum PSA at study entry in the intervention group. For men with a baseline serum PSA of 0.0-1.99 ng/ml at study entry, the increased risk of being diagnosed with Pca was increased more than 4 fold. In contrast, the Pca specific mortality difference was small, meaning that the potential harms were greater. The NNT to save one man from Pca death was 724 and 427 respectively when the baseline serum PSA was in the range of 0.0-1.9 or 2.0-3.9 ng/ml at study entry, respectively. Consequently, in these men the aggressive investigation and treatment was associated with extensive overtreatment and increase in costs. Furthermore, these observations suggest that for men with lower PSA levels, a screening protocol as currently performed in the ERSPC, may not be a proper tool to reduce Pca mortality. A simultaneous decrease in the quality of life may result. However, since all results are cumulative and the overall mortality is still low, longer follow-up is needed to confirm this early conclusion.

The significant excess incidence rates were mainly a result of repeated systematic screening using a lateralized sextant biopsy technique. A recent review showed that sextant biopsy, either classical or lateralized, will miss 23% or 19% of biopsy detectable Pca with extended biopsy schemes, respectively[170]. Therefore, the excess incidence might be even higher if the currently clinical accepted extended biopsy schemes were used for repeated screening.

The observations of the present study can be compared with the results of the Scandinavian Control Group Prostate (SCGP-4) study since current data included the effect of early detection as well as the effect of early aggressive treatment. The SCGP-4 study showed that radical prostatectomy decreased Pca mortality compared to expected management for men with favourable localised clinically diagnosed disease after a median follow up of 10 years[54]. In the SCGP-4 study, 52% of the patients were diagnosed with Pca having a PSA ≤10 ng/ml. In patients with PSA ≤10 ng/ml at diagnosis the difference in the cumulative incidence of Pca death between radical prostatectomy and watchful waiting was smaller and was observed after more years of follow-up than in the patients diagnosed with a PSA >10 ng/ml[171]. Although the present study is not a randomized controlled trial, the trends in Pca mortality of the SCGP-4 study are similar. In this study the difference in the cumulative risk of death from Pca was observed earlier for men with a baseline serum PSA 10.0-19.9 ng/ml at study entry (Figure 4). Furthermore, for men included in the intervention and clinical cohort, an overlap in the cumulative hazard curves during the first five years was observed for men with baseline serum PSA 4.0-9.9 ng/ml and eight years for men with baseline serum PSA 2.0-3.9 ng/ml at study entry (Figure 2 and 3).

The main limitation of this study is the absence of randomisation, which necessarily results in different patient characteristics at study entry. Statistical adjustment was needed for the difference in age and serum PSA at study entry. Furthermore, the large difference in all cause mortality might have biased the outcomes. Obviously the optimal study design would be the comparative evaluation of the intervention and control arm of the ERSPC. However, since serum PSA was not collected nor PSA measured at study entry in the control arm of the ERSPC, the present study design is an alternative method. The main limitations of the individual study populations and the present study design have previously been described[105, 152].

Furthermore, different treatments in both cohorts might have affected the outcomes, with men diagnosed and treated with curative intent at an earlier stage likely to have a better outcome[54, 163-164]. In both groups, following diagnosis, men were free to choose treatment in collaboration with their local urologist. As outlined in Table 2, men in Northern Ireland had higher PSA levels at diagnosis and a higher rate of metastatic disease, they were therefore less likely to undergo prostatectomy and more likely to have androgen deprivation therapy[105]. These differences in treatment are inherent to any study with a large difference in the intensity of screening and early detection. The distribution of different treatments in the two study groups are published before[105].

Another limitation that might have biased the study is the different PSA assays used. Several assays were used in Northern Ireland which differs from the same PSA assay used in the ERSPC. Cluster analyses that were performed for the different laboratories in Northern Ireland showed that there was no systematic difference in PSA values provided by the different laboratories and that the laboratory of origin did not affect the results. Finally, it could be argued that this study is plagued by a methodological bias, i.e. lead time. Generally, lead time is defined as the time between the detection by screening and the clinical diagnosis if there had been no screening. However, since in current study the observation time is defined as the time difference between the time of death and the time of first PSA measurement, lead time is unlikely to have influenced the outcomes. Nevertheless, it remains unknown if men with a specific age and PSA in a screening population compares equally to men with the same age and PSA in the selected clinical population.

The strong aspect from this study is the risk stratification based on baseline PSA and age. To the best of our knowledge, our study is the first report on a population based study cohort that showed the Pca incidence and mortality in two populations with a different intensity of screening and early detection stratified by baseline age and PSA. Currently, the interpretation of the balance between the risks and benefits is subjective, meaning it is a matter of personal judgement for which PSA level the benefits outweigh the harm. However, the final purpose of current study design is to risk stratify men at baseline

(based on age and PSA) into groups that require no further screening and men that have a higher risk of Pca mortality that should continue screening and early treatment. Currently longer follow-up is needed to give these clinical recommendations.

6.6 CONCLUSIONS

Baseline serum PSA before diagnosis is a strong predictor for Pca mortality in screen and clinical detected Pca. In the absence of standardized early detection programs, PSA might be used for a risk assessment that balances the harms and benefits of early detection in men aged 55-74 years. Current analyses suggest that the significant reduction in disease specific mortality with screening and early detection may be limited to men with baseline elevated PSA levels. In men with a low baseline serum PSA the benefits of continued aggressive investigation and treatment may be limited, whilst they are associated with a large increase in cumulative incidence, overtreatment and costs.


CRUISIN' FOR SCREENING
PROSTATE CANCER

Part III

Prostate Cancer Screening: excess mortality


CRUISIN' FOR SCREENING
PROSTATE CANCER

7

Excess mortality, attendance matters

Submitted

Ries Kranse, Pim J van Leeuwen, Timo Hakulinen, Jonas Hugosson, Teuvo L Tammela, Stefano Ciatto, Monique J Roobol, Marco Zappa, Gunar Aus, Chris H Bangma, Sue M Moss, Anssi Auvinen and Fritz H Schröder

ABSTRACT

Setting: In addition to disease-specific mortality, a randomized controlled cancer screening trial may be evaluated in terms of excess mortality. Excess mortality is calculated in the cancer patients in both study arms and is defined as the difference between the observed and expected mortality. The observed mortality can be estimated accurately by record linkage with the population registry. Thus the method critically hinges upon the assessment of the expected mortality. This issue is of special importance in the case of prostate-cancer screening. Attendees have a lower mortality rate and a higher probability of a prostate cancer diagnosis than non-attendees. Both issues affect the expected mortality in the screening arm patients.

Materials and methods: The effect of accounting for attendance is studied in a prostate cancer screening trial (ERSPC).

Results: Non-attendees had roughly twice the mortality rate of attendees. Approximately twice as many cancers were detected in the screening arm compared to the control arm, primarily in attendees. Unless the difference in mortality rate between attendees and non-attendees is taken into account the expected mortality is overestimated by 0.9 - 3.6 deaths per 1000 person-years . The latter figure corresponds with a high non-attendance proportion.

Conclusions: Attendees have a lower all-cause mortality rate (are healthier) and a higher probability of a prostate cancer diagnosis than non-attendees and than men randomized to the control arm. If attendance is not accounted for, the excess mortality (= observed – expected mortality) is underestimated. The between-arm mortality rate ratio is overestimated (screening is considered more effective than it actually is). These effects may be considerable, notably if non-attendance is common.

INTRODUCTION

A randomized study of screening for cancer typically focuses on the effect of screening on disease-specific mortality. To do so, the cause of death has to be determined. In the European Randomized Study of Screening for Prostate Cancer (ERSPC, an ongoing prostate-cancer screening study in 8 European countries[172]), all deaths in prostate-cancer patients were reviewed by an independent committee. This committee labeled each death in cancer patients as prostate cancer death, possible prostate cancer death, intervention related death or other cause of death[153]. The involvement of human experts at such a critical stage in the evaluation of a study and the potential for error prompted the search for alternatives. An alternative that does not require the human judgment of the cause of death uses the excess mortality rate difference between the screened and the control population[145].

Excess mortality rate is defined as the difference between the actually observed and the expected rate of death. An unbiased estimate of the expected mortality is of crucial importance in a study of excess mortality as is a correct labeling of the cause of death in a study of the disease-specific mortality.

In a previous paper[173] we used identical age-specific mortality rates for calculating the expected mortality for both the screening and the control arm. Also, within the screening arm, no distinction was made in age-specific mortality rates between men who actually attended a screening visit and men who did not.

The above described method of calculating the expected mortality, however, may well be biased. E.g. not attending a screening visit may have a health-related reason. On the other hand, attendees may have a healthier life style (e.g. non- smokers) or have easier access to medical care. Not accounting for these facts may lead to a biased estimate of the expected mortality in the screening arm.

Studying this putative effect is the topic of this paper. We explicitly do not focus on the effectiveness of prostate-cancer screening in terms of a between- arm difference in excess mortality rate, that will be carried out and discussed in a separate paper. We used the material presented in a previous analysis augmented with material from the Finnish[94], the Italian[95] and the Swedish[92] ERSPC centers. These centers are of special interest because the number of non-participants in the screening arm is high. They used pre-consent (i.e. up-front) randomization. This means that the men were enrolled into the study on the basis of the population registry without being aware of this. This results in 100% participation in the control arm and a variable rate of non-attendance in the

screen arm (effectiveness trial, focusing on the effect of screening under real life conditions). In contrast, the participants in the Dutch region of ERSPC were randomized after giving informed consent (efficacy trial, which estimates the maximal effect of screening).

MATERIALS AND METHODS

We studied 149066 participants enrolled in the ERSPC (centers Finland, Italy, Sweden, aged 55-69, and the Netherlands, aged 55-74 at randomization, studied over the period 1993-2006)[68, 172], 66311 men were randomized to the screening arm and 82755 to the control arm. The randomization was 2 to 3 in the Finnish center (more men were randomized to the control arm) and 1 to 1 in the other centers.

In order to study the effect of not accounting for attendance status we have calculated the excess mortality twice, once taking attendance status into account and once ignoring attendance status, using the same baseline data.

The method uses the mortality rate in participants without prostate cancer to calculate the expected mortality in cancer patients. Subsequently, the excess mortality in cancer patients (difference between the expected and observed mortality) is calculated. The excess mortality in participants without prostate cancer is zero, by definition. Hereunder the method is discussed in detail.

For all participants, for all centers, the total follow-up period was subdivided into yearly intervals until either death, emigration or censoring (censoring date for all centers Dec 31, 2006). For each interval the attained age equaled the age in the previous interval for the same individual plus 1 year (age at the initial interval = age at randomization). Each interval in the screening arm was subdivided into an episode of attendance and an episode of non-attendance. Attendance was defined as "being screened" at least once. At the moment of randomization, the status of a participant was defined as non-attendee. It changed to attendee at the moment of first attendance (status retained until death or censoring). For both study arms, each yearly interval was subdivided into a period without and with prostate cancer.

For each attained age, for all episodes in which no cancer was diagnosed, the number of deaths were added separately for each of 5 selections of participants without prostate cancer. These five selections were : attendees, non-attendees, men randomized to the screening arm, men randomized to the control arm and men randomized to either arm of the study. By dividing the respective total number of deaths by the respective num-

ber of person-years, crude estimates of the all- cause mortality rates for each of the five selections were obtained for each attained age.

These crude estimates were smoothed by fitting a Poisson model to the number of deaths with attained age centered at 65 years as sole predictor. The logarithm of the number of person-years at risk was added as an offset to the model constant.

The crude and smoothed mortality rates were compared with the country-specific mortality rates as obtained from the Human Mortality Database.

For each attained age, the total number of person-years in men in whom prostate cancer was diagnosed was calculated for attendees, non-attendees and patients randomized to the control arm, for all study centers.

The number of person-years for each attained age in attendees with prostate cancer was multiplied with the age specific mortality rates in, 1) attendees without prostate cancer, 2) screen arm participants without prostate cancer and 3) participants randomized to either arm without prostate cancer. In this way for each attained age three expected numbers of deaths were calculated. One with attendance status, screening arm and study center accounted for, one with study arm and study center accounted for and one with only the study center accounted for.

The same was done for non-attendee cancer patients. The three multiplication rates used in this case are similar to those used above (but then for non-attendees without prostate cancer). Subsequently the expected number of deaths in cancer patients in the screening arm were calculated by adding the expected number of deaths in attendees and non-attendees.

The expected number of deaths in men randomized to the control arm was calculated by multiplying the person-years in cancer patients for each attained age with the age specific crude mortality rate in control-arm participants without cancer.

Men in whom prostate cancer was found at autopsy were excluded from the calculations of the excess mortality, this was not done in the previous study[173].

For all centers, for all calculated expected deaths, for each attained age the excess number of deaths was calculated as the difference of the observed and expected number of deaths in cancer patients.

For all centers, for all calculated excess number of deaths, for each attained age, the excess mortality rates were calculated by dividing the excess number of deaths by the total number of person-years (i.e. in all participants irrespective of disease status) for that attained age. Additionally, the difference between the expected mortality rate corrected for screening arm only and the expected mortality rate corrected for attendance status was calculated. The same was done for the excess mortality. The mean bias in the expected mortality was calculated for all centers as the sum of the differences between the expected number of deaths over all attained ages divided by the total number of person-years.

For each study center, the details of the calculation of the expected mortality rate (taking and not taking attendance status into account respectively) in a cancer patient with an attained age of 70 years are given. This information is supplied to support the details given in the appendix with actual data.

RESULTS

Table 1 contains a summary of the baseline data and the expected and excess numbers of deaths derived from these baseline data for all study centers.

Note that the excess numbers of deaths calculated directly from the mortality in non cancer participants typically differ from the corresponding reported excess deaths in the table. This illustrates the importance of using attained age- specific mortality rates. E.g. for the Finnish center, for the screening arm, the excess number of deaths derived from the person-years and deaths reported in the table equals 309 - (4695 / 269814) * 11016 = 117.3 (considerably higher than the 56.9 listed).

The proportion of attendance is apparently strongly related to the type of randomization. A much higher proportion (roughly 25%) of non-attendees were present in the Italian, Finnish and Swedish study centers (with up front randomization) . It is about 5 fold higher than in the Dutch center (5% non- attendees). On the other hand, in the Dutch center approximately 50% of the men invited for screening did not give consent[68]. The proportion of cancers detected in the screening arm is much higher than the proportion of cancers detected in the control arm. For all centers accounting for attendance status leads to higher estimates of the excess mortality (compare the excess mortality given in row 1 with the sum of the excess mortalities in the rows labeled A and NA for each center in Table 1).

Table 1: Summary of the baseline data and the expected and excess numbers of deaths derived from these baseline data for all study centers (age ranges at randomization given between brackets). The left panel contains a summary of the data used to calculate the expected mortality in non-cancer participants, the right panel lists the expected and excess mortality if the non-cancer participant mortality rate is applied to cancer patients. Subgroup is coded as follows : S : screening arm, C for control arm, SC screening or the control arm, A for attendee and NA for non-attendee. pYnc is the number of person-years in non- cancer patients, N is the number of men in that specific category, propN is the proportion of men. Note that the denominator of this proportion is the number of men randomized for S and C (proportion is 1 for SC) and the number of men randomized to the screening arm for A and NA. Ndnc is the number of deaths in non-cancer patients, Nc is the number of cancers, propC is the proportion of cancers. The denominator is the number of participants for S, SC and C and the number of cancers for A and NA, pYc is the number of person-years in cancer patients and Ndc is the number of deaths in cancer patients.

Used to derive the expected mortality no cancer patients						Used to derive the expected and excess deaths				
Center (age)	Sub-group	N (propN)	pYnc	Ndnc	Nc(propC)	Sub-group	pYc	Ndc	Expected deaths	Excess deaths
Finland (55-69)	S	31970 (0.40)	269814	4695	2493 (0.08)	S	11016	309	252.1	56.9
	SC	80379	685203	11991	5125 (0.06)	SC	11016	309	255.8	53.2
	A	23774 (0.74)	193586	2300	2097 (0.84)	A	9465	220	152.4	67.6
	NA	8196 (0.26)	76228	2395	396 (0.16)	NA	1552	89	68.3	20.7 (A+NA=88.3)
	C	48409 (0.60)	415389	7296	2632 (0.05)	C	9121	356	223.6	132.4
Italy (55-74)	S	7265 (0.50)	58333	754	280 (0.04)	S	1391	33	23.7	9.3
	SC	14517	117049	1523	413 (0.03)	SC	1391	33	23.3	9.7
	A	5730 (0.79)	41889	475	246 (0.88)	A	1251	21	18.8	2.2
	NA	1535 (0.21)	16444	279	34 (0.12)	NA	140	12	3.3	8.7 (A+NA=10.9)
	C	7252 (0.50)	58716	769	133 (0.02)	C	602	19	10.5	8.5
Rotterdam (55-74)	S	21175 (0.50)	174641	3649	2152 (0.10)	S	13451	428	361.4	66.6
	SC	42318	358520	7399	3054 (0.07)	SC	13451	428	354.7	73.3
	A	19950 (0.94)	163316	3256	2104 (0.98)	A	13246	412	339.8	72.2
	NA	1225 (0.06)	11325	393	48 (0.02)	NA	205	16	9.5	6.5 (A+NA=78.7)
	C	21143 (0.50)	183879	3750	902 (0.04)	C	3645	228	112.8	115.2
Sweden (55-69)	S	5901 (0.50)	60361	1083	697 (0.12)	S	4162	89	92.4	-3.4
	SC	11852	123445	2213	1118 (0.09)	SC	4162	89	91.7	-2.7
	A	4466 (0.76)	39559	564	636 (0.91)	A	3861	69	66.1	2.9
	NA	1435 (0.24)	20802	519	61 (0.09)	NA	300	20	11.4	8.6 (A+NA=11.5)
	C	5951 (0.50)	63084	1130	421 (0.07)	C	1847	71	42.6	28.4

The expected number of deaths in the row labeled SC is much higher than the sum of the expected deaths in the rows labeled A and NA in Table 1. The results labeled with SC are obtained by accounting for the study center only (i.e. the method used in the previous paper[173]). The difference is especially large for the Finnish and the Swedish study centers.

Table 2 lists the excess mortality rates for men with an attained age of 70 for the four centers studied. The bias is the largest for the Swedish center where the "baseline mor-

Table 2. Details for the calculation of the expected mortality rate in men with an attained age of 70 years (see appendix). All rates are given per 1000 person-years.

Center	Mortality rate non-cancer attendees	Mortality rate in non-cancer non-attendees	Proportion attender person-years in non-cancer participants	Proportion non-attender person-years in cancer patients	Expected mortality rate not accounting for attendance status	Expected mortality rate accounting for attendance status	Bias in the expected mortality rate
Finland	0.0200	0.0510	0.75	0.84	27.75	24.96	2.8
Italy	0.0156	0.0281	0.74	0.89	18.85	16.97	1.9
Rotterdam	0.0215	0.0374	0.94	0.99	22.45	21.66	0.8
Sweden	0.0206	0.0515	0.77	0.94	27.71	22.45	5.3

talities" between attendees and non-attendees are quite different combined with quite different proportions of person-years in "non- cancer attendees" and in cancer patients who had attended a screening visit (see appendix for an explanation).

In Figure 1 the bias that results in the expected mortality rate by not accounting for attendance status is plotted for all attained ages for all study centers.

The average difference in the "attendance status accounted for" and "attendance status not accounted for" expected mortality-rates is 3.6, 2.9, 1.1 and 0.9 for Sweden, Finland, Italy and the Netherlands per 1000 person-years respectively.

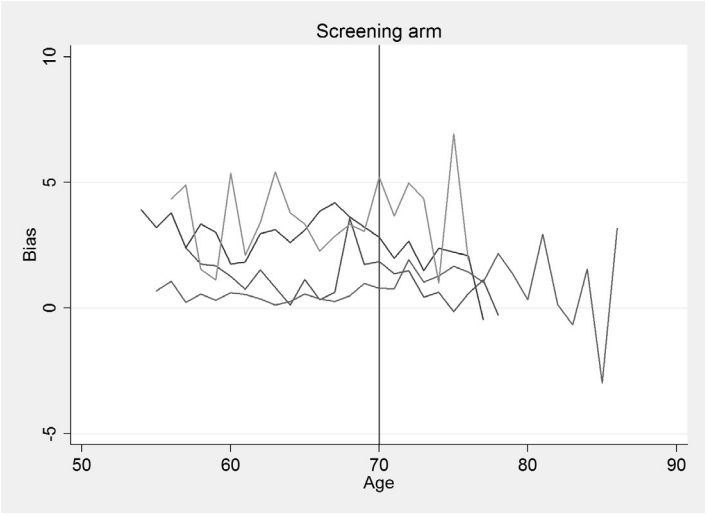


Figure 1. Bias in the expected mortality rate as a function of the attained age for the four centers studied. The bias is caused by not accounting for attendance status in the calculation of the expected mortality rate. The vertical line corresponds with the data given in Table 2. The line with the shortest age-range corresponds with Sweden (yellow), next is Finland (blue), than Italy (red) and the Netherlands (green). All rates are given per 1000 person-years.

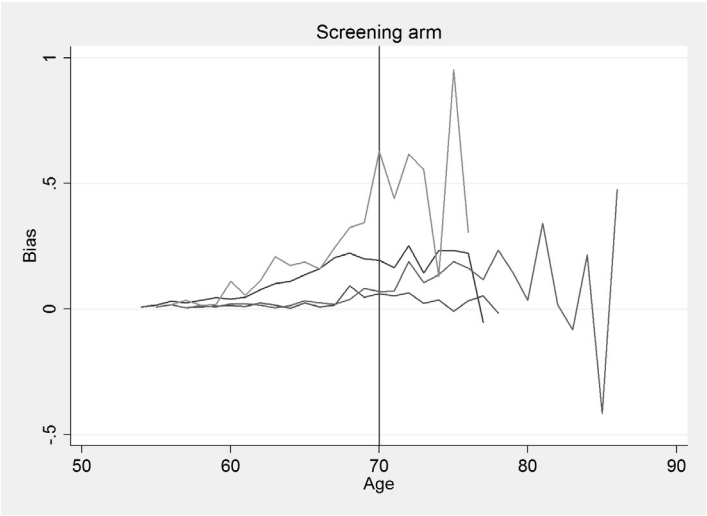


Figure 2. Bias in the excess mortality rate as a function of the attained age for the four centers studied. The line with the shortest age-range corresponds with Sweden (yellow), next is Finland (blue), than Italy (red) and the Netherlands (green). All rates are given per 1000 person-years.

In Figure 2 the bias that results in the excess mortality rate by not accounting for attendance status is plotted, for all attained ages, for all study centers. This plot enables studying the effect of "not accounting for attendance status" on the evaluation of the randomized study by means of a between-arm comparison of the excess mortality rate. The bias increases with increasing age.

The estimates of the expected mortality rates in men without prostate cancer in the different types of participants discerned in this study, as determined by the Poisson regression for the four centers, are given in Table 3. Non-attendees have a roughly twofold higher expected mortality rate than attendees. The expected mortality rates for the screening arm are comparable to the expected mortality rates for the control arm. For Sweden and Finland the expected mortality rates are higher than the mortality rates obtained from the Human Mortality Database (i.e. population-based). For Italy and the Netherlands they are lower.

DISCUSSION

A randomized controlled trial of screening for cancer is typically analyzed in terms of the between-arm difference in the disease-specific mortality rate. Thus the disease-specific mortality rate in the screening arm is compared with the rate in the control arm. This requires accurate information on the cause of death (i.e. human judgment). In addition,

Table 3. Example of the expected mortality rates per center for the four categories of participants discerned in this study for the attained age of 65.In the type column S stands for screening arm, A for attendee, NA for non-attendee and C for control arm.

Center	Type	Mortality rate in a 65 year old man (95%CI)	Mortality rate in the general male population (source Human Mortality Database)
Finland	S	0.0179 (0.0174-0.0185)	0.0167 (2008)
	A	0.0117 (0.0112-0.0122)	
	NA	0.0340 (0.0327-0.0354)	
	C	0.0181 (0.0176-0.0185)	
Italy	S	0.0109 (0.0100-0.0119)	0.0133 (2006)
	A	0.0089 (0.0080-0.0100)	
	NA	0.0157 (0.0138-0.0178)	
	C	0.0115 (0.0106-0.0125)	
Rotterdam	S	0.0130 (0.0124-0.0136)	0.0138 (2008)
	A	0.0122 (0.0116-0.0129)	
	NA	0.0236 (0.0206-0.0271)	
	C	0.0126 (0.0120-0.0132)	
Sweden	S	0.0171 (0.0160-0.0182)	0.0135 (2007)
	A	0.0125 (0.0114-0.0137)	
	NA	0.0264 (0.0242-0.0288)	
	C	0.0170 (0.0160-0.0181)	

other effects of the screening process that affect mortality may be missed (e.g. detection of cardiovascular problems or diabetes). Both problems would be solved by carrying out an all- cause mortality analysis[174]. The reason that is mentioned in the literature for not doing precisely this is that the effect of screening on all-cause mortality is typically small. E.g., for prostate cancer, we assume that a third of the prostate-cancer mortality were avoided by screening. If the life time risk to die of prostate cancer were 3 % (chosen for ease of computation but close to the actual numbers) the effect of screening on all-cause mortality would be 1%. This relative difference is much smaller than the assumed 33% prostate-cancer specific mortality reduction. Showing a 1% effect instead of a 33% effect beyond a reasonable doubt (i.e. with statistical significance) requires prohibitively large numbers of trial participants. The precision of a study increases with the square of the number of events (proportional to the number of participants), thus a study that is twice as precise requires 4 times the number of participants. However, if we assume that a between-arm effect of a randomized controlled trial is confined to the patients in whom cancers are detected, an all-cause mortality analysis can be carried out without needing so many participants[173].

An excess mortality analysis can be applied as an alternative but preferably, at least, as an augmentation of a disease-specific mortality analysis. Given that excess mortality is the difference between the observed and the expected mortality and the possibility to obtain the former by database linkage (i.e. in theory flawlessly), an accurate determination of the expected mortality is of critical importance.

From a comparison of the expected mortality rates with the population mortality rates in Table 3 it is obvious why the latter was not used to estimate the excess mortality. It is lower than the expected mortality rate for Sweden and higher for Italy. A plot for all ages for the Netherlands of the expected mortality rate on the basis of the trial data and the Dutch population mortality rate in males (not shown) shows considerable differences despite the reasonable agreement at age 65 (Table 3). Thus, population-based mortality rates are apparently not specific enough to be used in this study. Therefore we have derived the expected mortality from within the trial, based on the participants in whom no cancer was yet detected.

In screening trials such as the ERSPC, an "intention to screen" based analysis is used. Thus the disease-specific screening arm mortality rate is compared with the corresponding control arm mortality rate, irrespective of actual screening attendance. This approach is conservative. If an effect of screening is detected in an "intent to screen" analysis it is very likely larger in individuals who are actually screened.

Non-attendees have a relatively high mortality rate (health-related,SES or education-related). This can be appreciated from Table 3 (mortality rate in non-attendees is double the mortality rate for attendees in all four centers).

The appendix of this paper shows that not accounting for attendance status in the calculation of the excess mortality may lead to a serious bias. It is affected by two factors. It increases if the ratio attendees/non-attendees in the cancer patients in the screening arm differs more strongly from the ratio attendees/non-attendees in the non-cancer participants. Intuitively this can be understood as follows. Men who show up for screening are healthier (i.e. less likely to die) than men who do not show up. Due to screening, cancers are primarily detected in these men. The bias also increases if the difference in the mortality rates between attendees and non-attendees (both without prostate cancer) increases. If attendance status is not accounted for, the expected number of deaths is overestimated and therefore the excess number of deaths is underestimated. The same holds for the expected and excess mortality rates. Therefore, since this effect occurs in the screening arm, the excess mortality rate difference between the two arms is overestimated. The screening program appears to be more effective than it actually is.

The data presented in this paper illustrate the reality of this problem. For the Swedish study center, the excess mortality in the screening arm is negative if attendance status is not accounted for. This is an unlikely observation since the excess mortality is determined in cancer patients. The largest bias is observed in the three centers with upfront randomization (the related high percentage of non-attendance makes a bias more likely in an intention to screen analysis).

The relation between disease-specific and excess mortality is cancer type dependent. E.g. smoking causes both lung cancer and cardiovascular deaths. The excess mortality is related to both causes and therefore both an excess and a disease-specific mortality analysis should be done. Furthermore, the excess incidence in the screening arm consists of two parts, early diagnosis (cancer is found sooner but would have surfaced anyway) and overdiagnosis (cancer would not have surfaced during a patient's life time). Both add to the expected mortality bias if attendance is not properly accounted for. The amount of overdiagnosis (and thus the bias) strongly depends on the type of cancer. Correcting for attendance is likely more important in prostate cancer screening (relatively high overdiagnosis) than in breast cancer screening with the same follow-up.

There is some opportunistic screening in the control arm of the ERSPC, very likely predominantly in healthy men. Correction for this fact is impossible, it requires information about opportunistic screening at the level of the individual patient (which is not available). Thus the expected mortality for the control arm is estimated too high and the excess mortality too low (using the same logic as given above for the screening arm). This means that if a screening study is evaluated by a between arm study of the excess mortality rates the outcome is conservative (at least if attendance status is accounted for for the screening arm). The actual effect of screening on the excess mortality rate is in reality very likely slightly larger than the calculated value.

Figure 1 shows that the effect of accounting for attendance status on the expected mortality is roughly constant at each attained age. The effect on the excess mortality rate, however, depends on the attained age (Figure 2) because incidence increases with age. The observed mortality is associated with incidence. Thus bias in the excess mortality rate increases with increasing age.

CONCLUSION

Correcting for attendance status is very important in the calculation of the excess mortality rate that can be used in conjunction with a disease-specific mortality analysis in a randomized controlled cancer screening trial.

APPENDIX

Let us look at an age group for a period of time in the screening arm. The expected mortality of the prostate-cancer patients will be based on person-years of those without prostate cancer until the end of the period, death or diagnosis of prostate cancer, whichever comes first. The expected mortality rate without regard to the participation (attendance) status is :

$$r_s = w_{as}\, r_{as} + (1 - w_{as}) r_{ns}, \tag{1}$$

where $w_{as} = py_{as} / py_s$ = proportion of the participants' person-years py_{as} out of all person-years py_s, and r_{as} and r_{ns} are the mortality rates in the participants and non-participants, respectively. When the participation status of the prostate-cancer patients is accounted for, the expected mortality rate is :

$$r_{ps} = w_{ap}\, r_{as} + (1 - w_{ap})\, r_{ns}, \tag{2}$$

where $w_{ap} = py_{ap} / py_p$ is the proportion of the participated prostate-cancer patients' person-years py_{ap} out of all prostate-cancer patients' person-years py_p.

It is easy to see that $r_s = r_{ps}$, if either $w_{as} = w_{ap}$ or $r_{as} = r_{ns}$.

There are no other solutions. Equations (1) and (2) represent lines in the space spanned by r_{as} and r_{ns}.

These lines are equal if they intersect ($r_s = r_{ps}$) or if they are on top of each other ($w_{as} = w_{ap}$).


CRUISIN' FOR SCREENING
PROSTATE CANCER

8

Disease specific mortality may underestimate the total effect of prostate cancer screening

Journal of Medical Screening. 2011 Jan 10

Pim J. van Leeuwen, Ries Kranse, Timo Hakulinen, Monique J. Roobol, Harry J. de Koning, Chris H. Bangma and Fritz H. Schröder

ABSTRACT

Objectives: To study the difference between the disease specific and excess mortality rate in the European Randomized Study of Screening for Prostate Cancer (ERSPC) section Rotterdam.

Methods: A total of 42,376 men were randomized to systematic screening or usual care. The excess number of deaths was defined as the difference between the observed number of deaths in the prostate cancer (PC) patients and the expected number of deaths up to December 31, 2006. The expected number was derived from mortality of all study participants before a possible diagnosis with PC. The disease specific mortality rate was based on the number of men who died from PC. The excess mortality rate based on the arm-specific excess number of deaths and the disease specific mortality rate were compared between the two study arms.

Results: The overall mortality rate was not significantly different between the intervention and the control arms of the study: RR 1.02 (95%CI 0.98-1.07). The disease specific mortality rate was 0.42 men per 1000 person-years in the intervention and 0.48 men per 1000 person-years in the control arm: RR 0.86 (95%CI, 0.64-1.17). The excess mortality rate was 0.40 per 1000 person-years in the intervention arm and 0.61 men per 1000 person-years in the control arm, and the RR for excess mortality was 0.66 (95%CI, 0.39-1.13).

Conclusions: In contrast to the disease specific mortality rates an increased difference in the excess mortality rates was observed between the two arms. This observation may be due to a systematic underestimation of the disease specific deaths, and/or an additional disease related mortality that is measured by an excess mortality analysis but not by a disease specific mortality.

8.1 INTRODUCTION

The European Randomized Study of Screening for Prostate Cancer (ERSPC) has shown a significant reduction in prostate cancer (PC) mortality due to screening[98]. In the final analysis the ERSPC has compared the number of men who died from PC per unit time at risk in both arms of the study, i.e., the disease specific mortality rates. Although this is a generally used endpoint of a randomised controlled screening trial, an alternative to this approach exists. It comprises the estimate of the excess mortality rate in the cancer patients in both arms of the study[145]. This excess mortality rate is based on the actually recorded number of cancer patient deaths in excess of the number expected on the basis of a cohort of cancer free individuals per unit of time.

No differences in the disease specific mortality and excess mortality rates are expected to be found. This is conditional on PC being the only factor affecting differences in mortality, the accurate ascertainment of the disease specific mortality and an accurate estimate of the expected mortality. However, if for example screening for PC has non-PC related effects on the life expectancy of cancer patients, or the cause of death is not accurately assessed, the outcome of the two mortality analyses may differ. Therefore the question whether the disease specific mortality rate differs from the excess mortality rate in men who are systematically screened for PC and in men who were not systematically screened is relevant. The absence of a discrepancy confirms the established effect of screening on PC mortality, the presence of a significant discrepancy warrants further research.

8.2 MATERIALS AND METHODS

8.2.1 Study Population

After signing an informed consent, a total of 42,376 men, aged 55-74 year, were randomized in the Rotterdam section of the ERSPC between 1993 and 2000[68]. Men were allocated to the intervention and the control arm by individual randomisation based on the outcome of a random number generator. Men with a prior diagnosis of PC were excluded. Between December 1993 and May 1997 men in the intervention arm were screened with an interval of four years by PSA measurement, rectal examination (DRE) and transrectal ultrasound examination (TRUS). A sextant biopsy was initially offered to men with PSA ≥4.0 ng/ml and/or a suspicious finding on DRE and/or TRUS. After May 1997 a biopsy was prompted by PSA ≥3.0 ng/ml only. Treatment decisions were taken by the local urologist and at patient's preference. Details of the screening methodology were reviewed by Roobol et al[96]. Cancers diagnosed between the screening intervals or after the age of 74 years clinically or due to opportunistic screening, transurethral

resection of the prostate for benign disease, and cystoprostatectomy specimens, were included as well. These cancers were identified by means of linkage to the national Comprehensive Cancer Registry[160, 175]. Follow-up in this respect was complete up to and including 2006. Men in the control arm were not subject to screening and received usual care if diagnosed with PC. The men with PC in the control arm were identified through the same linkage with the Comprehensive Cancer Registry.

8.2.2 Mortality data

Mortality data of participants who died in the period up to December 31, 2006 were obtained by linking the trial database with the Causes of Death Registry of Statistics Netherlands[176]. Linkage to the Causes of Death Registry of Statistics Netherlands was possible by using the personal administrative number of each participant as a linkage key.

8.2.3 Length of follow-up

Diagnostic and mortality data were available until December 31, 2006. Consequently, the date of censoring was at emigration or December 31, 2006.

8.2.4 Disease specific mortality

For all PC cases in the study who were known to have died, all available information was gathered and anonymised. Subsequently all deceased PC cases were reviewed by a national independent causes of death committee (CODC) using predefined flow charts or by an international committee if no consensus was reached[153]. Patients were determined to have died from PC if they were classified as either "definitely PC death", as "possible PC death" or as "PC intervention related death"[153]. The disease specific mortality was calculated for the both arms of the study.

8.2.5 Excess mortality

The excess mortality was calculated for the both arms of the study and was based on arm-specific excess numbers of deaths of the PC patients. In both arms, the observed number of deaths in PC patients diagnosed during the trial was compared with an expected number of deaths. The expected number of deaths was calculated on the assumption that the patients would have had the same mortality by age at and time since randomization as the study population (intervention and control arms) combined before any diagnosis of PC. The excess number of deaths was defined as the difference between the observed and expected numbers. The excess numbers of deaths were used as numerator data, and the number of person-years per study arm was used as denominator in calculating the excess mortality rates for the both arms of the study.

8.2.6 General population mortality

The age specific mortality rates obtained from the study participants were graphically compared with the age specific mortality rates obtained from nationwide population life tables based on men in the general Dutch population[177]. The mortality rates were expressed in deaths per 1000 person-years at risk per year.

8.3 STATISTICAL ANALYSIS

The excess mortality rates were compared between the study arms as shown in the Appendix 1. The disease specific mortality rates were compared between study arms using a Poisson regression analysis with an indicator of study arm as a predictor and the logarithm of the number of life years as an offset term (predictor with a coefficient of one)[178]. In this paper a two-sided *p* value <0.05 was considered statistically significant. All analyses were performed with the commercially available STATA package: Data Analysis and Statistical Software, version 10.

8.4 RESULTS

8.4.1 Baseline characteristics

A total of 42,317 men of the 42,376 men who signed an informed consent were included in the study (59 men were excluded because of a PC diagnosis before randomization). The median age at randomization was 63 years for the both study arms (Table 1).

Table 1. The study arms by age at randomization

	Intervention arm *N*= 21,175	Control arm *N*= 21,142
Age (yr), median	63.0	63.0
55-59	8,142	8,033
60-64	5,511	5,565
65-69	4,746	4,820
≥ 70	2,776	2,724

The cumulative PC incidence in the intervention arm was 2.4 fold higher than in the control arm, i.e., 2,153 (10.2%) men were diagnosed with PC in the intervention arm and 901 (4.3%) men were diagnosed with PC in the control arm (Table 2). The median age at diagnosis differed significantly between the two groups (p<0.001); men in the intervention arm were diagnosed at an earlier age. Up to the end of 2006, a total of 4,077 (19.2%) men died in the intervention arm (21.67 men per 1000 person years) and a total

Table 2. Prostate cancer patients by age at diagnosis

	Intervention arm N= 2,153 (10.2% of men randomized)	Control arm *N* = 901 (4.3% of men randomized)	*P* value
Age (yr), median	68	71	< 0.001 †
55-59	165 (7.7)	23 (2.6)	< 0.001 ‡
60-64	458 (21.3)	122 (13.5)	
65-69	716 (33.2)	236 (26.2)	
70-74	650 (30.2)	274 (30.4)	
≥ 75	164 (7.6)	246 (27.3)	

† Mann-Whitney U test, ‡ Chi-square test

of 3,977 (18.8%) died in the control arm (21.20 men per 1000 person years). This resulted in a cumulative all cause mortality that was not significantly different between the two arms: RR 1.02 (95%CI 0.98-1.07), Table 3.

Table 3. Mortality outcomes in the intervention and control arm ERSPC-Rotterdam

Randomization arm	Numb. of men	Numb. PC	All cause mortality rate	RR (95%CI)	PC death rate	RR (95%CI)	Excess mortality rate	RR (95%CI)
Intervention	21,175	2,153	21.67	1.02 (0.98-1.07)	0.42	0.86 (0.64-1.17)	0.40	0.66 (0.39-1.13)
Control	21,142	901	21.20		0.48		0.61	

PC: prostate cancer, Rates as the ratio of deaths to 1000 person-years of exposure, RR: rate ratio

8.4.2 Disease specific mortality rate

Up to the end of 2006, in the intervention arm 0.42 men per 1000 person years and in the control arm 0.48 men per 1000 person years died from a PC related cause of death according to the CODC. This resulted in a non-significant reduction in PC mortality rate of 14% in the intervention population relative to the control arm: RR 0.86 (95%CI, 0.64-1.17), Table 3.

8.4.3 Excess mortality

Up to the end of 2006, the excess mortality rate was 0.40 men per 1000 person-years in the intervention arm and 0.61 men per 1000 person-years in the control arm. This resulted in a non-significant reduction in the excess mortality rate of 34% in the intervention population relative to the control population: RR 0.66 (95%CI, 0.39-1.13), Table 3.

8.4.4 Disease specific mortality versus excess mortality

The excess mortality rate was slightly lower than the disease specific rate in the intervention arm of the study. The absolute risk difference was 0.02 men per 1000 person years. The excess mortality rate was considerable higher than the disease specific rate in the

Table 4. Mortality outcomes according to age at randomization

Age	No. of men I arm	No. of men C arm	PC death rate I arm	PC death rate C arm	Excess Mortality rate I arm	Excess Mortality rate C arm
55-64	13,653	13,598	0.19	0.25	0.18	0.27
65-74	7,522	7,544	0.83	0.89	0.80	1.19

PC: prostate cancer, I: Intervention, C: Control, Rates as the ratio of deaths to 1000 person-years of exposure, RR: rate ratio

control arm of this study. The absolute risk difference was 0.13 men per 1000 person years. The difference between disease specific and excess mortality rate in the control arm of the study was most evident for the older ages at randomization (Table 4).

8.4.5 Mortality rates study population versus general population

The age-specific mortality rates among study participants were favourable compared to men in the general population (Figure 1). The mortality rates increased with increasing attained age in both populations. The mortality ratio was about 1.4 fold higher among men in the general population for all ages.

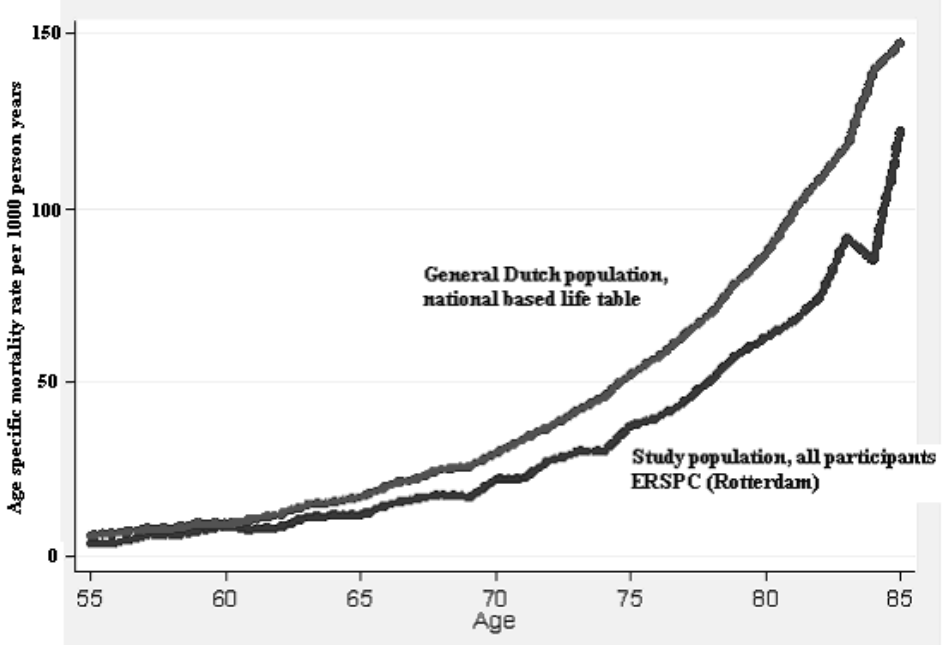


Figure 1. The age specific mortality rate per 1000 person-years, for men in the general Dutch population and all men participating in the ERSPC (Rotterdam).

8.6 DISCUSSION

The ERSPC was initiated to detect the effect of population-based PC screening on PC specific mortality. The ERSPC (all centers included) was designed with a power of 0.86 to detect a 25% intervention effect in men actually screened if the contamination remained limited and the follow-up was complete up to the end of 2008[179]. As a result, a 20% relative reduction in PC mortality with a significance level of p=0.04 was reported by the ERSPC based on the data up to the end of 2006; the relative reduction in the men actually screened was 27%[98]. The focus in this first report was the difference in PC specific mortality rates[98]. In the current report we have studied the excess mortality in addition to the disease specific mortality in the Rotterdam section of the ERSPC. The excess mortality rate was notably higher than the disease specific mortality rate in the control arm. As a result, in comparison to the disease specific mortality rates, an increased difference in the excess mortality rates was observed between the two arms of the ERSPC Rotterdam. Neither of the relative reductions were statistical significant. Therefore results after longer follow-up and complementary data of the total ERSPC are needed to confirm these observations.

Any difference between the excess and disease specific mortality rate could be due to a violation of the appropriateness of the basic assumptions (accurate estimation of expected mortality in case of excess mortality, accurate classification of cause-of-death in the case of disease specific mortality, and the absence of an intervention related effect on the life expectancy that is unrelated to PC in case of excess mortality)[145]. Since a difference was observed, this difference will be discussed based on the basis of the above three possible explanations.

8.6.1 Accurate estimation of expected mortality

Typically, in relative survival analysis, the expected mortality is estimated from nationwide population life tables stratified by age, sex and period[180]. In this screening setting, the general population may not be suitable as a reference population since the people who participate might well be healthier and of higher socioeconomic classes, i.e., subject to healthy screenee bias[181-182]. For this reason, in the present study, it seemed appropriate to obtain the expected mortality from all participants before a possible diagnosis of PC. A comparison of the mortality rates from the Dutch national statistics [177] with the rates derived according to the procedure outlined above, indicated that the mortality rates were considerably lower among study participants than among men in the general population (Figure 1). These results are consistent with a study on socioeconomic status patterns among study participants, which showed that men who participated in the study had higher income levels and lower poverty or deprivation status[162]. Therefore,

the national population is not an appropriate reference group for the estimation of the excess mortality in a PC screening trial. Additionally, an excess mortality could be overestimated if PC patients would have characteristics which are related to PC but also carry an increased risk of death other than from PC. For example, patients diagnosed with smoking-related cancers will experience excess mortality, compared to the reference population, due to both the cancer and other smoking related conditions. Such a relation is not known for PC, however, it could be that there is an unknown factor related to clinical diagnosed PC in particular that is carrying a higher risk of death from other causes. In principal, such a confounding effect could be of influence on the observed difference between the disease specific mortality and excess mortality rates in the current study. Therefore, this has to be considered as an explanation for the observed difference.

8.6.2 Random variation excess mortality

The confidence interval of the rate ratio of the two excess mortality rates between the two study arms was estimated by the delta method, Appendix 1[183-184]. To examine the appropriateness of the assumptions made in this approach, the confidence interval was also estimated by means of a computer simulation (bootstrap procedure)[185]. Using a bootstrap procedure has the benefit of not making assumptions about the distribution of the coefficients. The procedures consist of a repeated sampling with replacement from the dataset under study and carrying out the modelling for each sample thus obtained. The rate ratio of the two excess mortality rates between the two study arms obtained from the bootstrap equaled 0.68 (95%CI 0.28-1.08).

8.6.3 Accurate classification of cause-of-death

In the present study the causes of death of every PC patient is ascertained by a committee[153]. Such committees are based on the experience in breast cancer screening trials where independent committees have reviewed causes of death to ensure a correct interpretation of the mortality results published[133]. It is assumed to be the most optimal method, especially in PC where death on the death certificates differs often from the clinical picture[135, 137]. However, it is in practice impossible to blind the arm of the study entirely before the committee reviews the records of deceased men. The records are for example not blinded for stage at diagnosis although this is a very suggestive parameter for the arm of the study[186]. It remains possible that the lack of adequate blinding of the records with respect to study arm has resulted in a bias in the cause of death ascertainment. Efforts have been made to optimise the accuracy of the assessment of the cause of death, e.g. all PC deaths are reviewed by three persons independently without exchanging opinions[153]. However, despite these efforts, it might be possible that the reviewers have underestimated the PC mortality by being too cautious in stating that a patient has died from PC in the control arm of the study, in order not to overestimate

the study outcomes. Furthermore, it might be that due to the study protocol more information could be gathered from patients diagnosed with PC in the intervention arm. Consequently, the CODC made a more valid decision for the deceased men who were participating in the intervention arm. Therefore, this potential biases might be an explanation for the observed difference in both mortality rates in the control population. Finally, any algorithm used for the review of causes of deaths is bound to be based on some arbitrary definitions which are needed to ascertain comparability of decisions but which may lead to discrepancies with the excess mortality.

8.6.4 Absence of an intervention related effect on life expectancy that is unrelated to PC.

Assuming that the observed difference in outcomes between the two mortality rates is not due to the two previously described explanations, the difference may be due to the theoretical difference of the two mortality analyses. Disease specific mortality is included in excess mortality but is not identical to it, i.e. disease specific mortality refers to mortality due to only one particular cause of death such as disease progression and treatment related mortality and excess mortality measures both the direct and indirect mortality due to the cancer of interest, e.g. including cachexia, uraemia, suicide depression and loss of interest in life. For this reason, a screening study may reduce the disease specific mortality but side effects of the screening procedure (anxiety) may cause an increased death of other causes. In such a case the excess mortality is higher than the disease specific mortality. On the other hand, the screening procedure may detect health problems unrelated to the disease under study which may lead to effective medical treatment. In this case the excess mortality will be less than the disease specific mortality. These latter considerations may thus explain the differences between the mortality rates in the intervention arm of the study (Table 3).

Based on the observation that the excess mortality rate exceeds the disease specific mortality rate in the control arm of the study, the life expectancy of more men may be negatively influenced than estimated on the basis of a disease specific mortality analysis. Possibly, in a number of cases, PC specific mortality has failed to esthablish to which extent PC is a contributary factor for death. This is a known phenomenon since only 50-85% of men with advanced PC die from their disease, depending on age and the extent of the disease at diagnosis[187]. As a result, in these men with advanced disease who did not die from PC, the advanced disease could have shortened the patient's life by making the patient more susceptible for other causes of death. For example, the death of patients who are physically affected by PC but eventually die from a secondary infection, will be classified as an intercurrent cause of death. In such a case the question remains if the patient would have died when the disease of interest was absent. Based

on our results this may, amongst others, explain the observed discrepancy between the two types of mortality rates, especially in the control arm since the control arm contains more men with a clinically advanced disease.

The main difference between the excess mortality and disease specific mortality rate in the control arm was observed for the older age group at randomization (Table 4). This observation is in good agreement with the previous excess mortality studies on mammography screening for breast cancer that also observed a more pronounced discrepancy between the excess mortality and disease specific mortality analysis for the older ages[188-189]. In line with suggestions made in the previous publications, this could be the result of a more uncertain individual cause of death determination in these age groups since the overall mortality in these older ages was higher.

8.6.5 Overall mortality

This study and the total ERSPC were not designed to detect a statistically significant effect in overall mortality[98, 179]. We estimated that if the effect of screening were the only between arm difference in the mortality pattern a 30% disease specific mortality reduction would translate into a 0.7 % overall mortality reduction. In order to show this effect of screening on overall mortality at the two-sided 5% level of statistical significance, more than 3 million study participants have to be randomized (assuming a median of 9 year follow-up and a 15% overall mortality over the 9 year period)[190]. Therefore, if PC screening has an impact on the overall mortality, this will be very small.

8.6.6 Limitations

This study was performed by a single centre of the ERSPC. However, this centre was not powered to analyze the effect of a significant difference in disease specific mortality alone[179]. Consequently, final conclusions can only be made after confirmative results based on data of the total ERSPC.

8.7 CONCLUSIONS

For the evaluation of PC screening, a study of the excess mortality analysis is of independent additional value to a disease specific analysis since it measures the effect of screening in the presence of all competing risks. The between arm difference in the excess mortality rates was more than twofold the between arm difference in the disease specific mortality rates. Two possible explanations for the observed discrepancy outlined in this paper are: 1) a systematic underestimation of deaths from PC by the CODC in the control arm of the study; 2) an additional disease related mortality that is measured by an excess

mortality analysis but not with a disease specific mortality. Furthermore, although the effect in terms of excess mortality was larger than the disease specific mortality, PC screening might have no or a very small effect on the overall mortality.

In any randomised study with a disease specific mortality reduction as an endpoint, an additional excess mortality study should be mandatory as it may reveal additional important information. In the present study it may imply that the effects of PC screening on mortality is considerably larger than reported in the initial analysis that focussed on disease specific mortality only.

APPENDIX 1.

Statistical comparison between the two excess mortality rates of the two study arms

The two excess mortality rates are presented as $m_i = e_i/n_i$, where i = 1, 2 is the study arm (i= 1: intervention arm and i= 2: control arm), e_i =arm-specific excess number of deaths and n_i = arm-specific number of observed person-years. In this study, the quantities n_1 and n_2 are regarded to be constants since the total numbers of death in the arms are large.

The e_i are not regarded as constants, since e_i depends on N_i = number of prostate cancer cases per study arm and the chances of patient survival (a source of extra random variation). Furthermore, each e_i is also a difference between two numbers of deaths among the N_i patients:

$$e_i = d_i - d_i^*,$$

where d_i = the observed number of deaths per study arm and d_i^* = the expected number of deaths per study arm. Neither one of them is regarded to be constant as they depend on N_i, a Poisson-distributed random variable.

In this calculation, the d_i and d_i^* are mathematically related to the N_i as:

$$d_i = N_i q_i \text{ and } d_i^* = N_i q_i^*,$$

where the q_i and q_i^* are death proportions, observed and expected respectively. The q_i^* is regarded to be fixed as it is based on the large total numbers of deaths in the study population. The q_i is assumed to be the parameter of a binomial distribution with radix N_i.

In this comparison the quantity of interest is:

$$\text{Rate ratio (RR)} = m_1/m_2 \text{ and, } \ln(RR) = \ln(m_1) - \ln(m_2).$$

The variance of the RR is estimated by the delta method[183-184] as:

$$\text{var}(\ln(RR)) = \text{var}(m_1)/m_1^2 + \text{var}(m_2)/m_2^2.$$

The 95% confidence limits for ln(RR) based on the delta method are calculated as ln(RR) +1.96 sqrt (var (ln(RR))) as the upper bound and ln(RR) − 1.96 sqrt (var (ln(RR))) as the lower bound , where sqrt = square root. Consequently, the confidence limits for the RR are obtained by an exponentiation of these upper and lower bounds.

The var (m_1) and var (m_2) are calculated as:

$$m_i = e_i/n_i = (d_i - d_i^*)/n_i = N_i (q_i - q_i^*)/n_i.$$

var (m_i) = var (numerator)/n_i^2, with var (numerator) = var (N_i) $(q_i - q_i^*)^2 + N_i^2$ var $(q_i - q_i^*)$

$$= N_i (q_i - q_i^*)^2 + N_i^2 \text{ var } (q_i).$$

Now the Poisson assumption of N_i, constancy assumption of q_i^* and the delta method for variance of a product have been made use of. A binomial assumption for q_i further gives:

$$\text{var } (q_i) = q_i (1 - q_i)/N_i$$

and

$$\text{var (numerator)} = N_i (q_i - q_i^*)^2 + N_i q_i (1 - q_i) = N_i ((q_i - q_i^*)^2 + q_i (1 - q_i)).$$

This can be inserted into the formula of var (m_i), i=1,2, in order to give the confidence interval for the RR.


CRUISIN' FOR SCREENING
PROSTATE CANCER

9

Prostate cancer screening reduces the excess mortality

Submitted.

Pim J van Leeuwen, Ries Kranse, Timo Hakulinen, Jonas Hugosson, Teuvo L Tammela, Stefano Ciatto, Monique J Roobol, Marco Zappa, Harry J de Koning, Chris H Bangma, Sue M Moss, Anssi Auvinen, and Fritz H Schröder

ABSTRACT

Objectives: To study the excess mortality rate in the European Randomized Study of Screening for Prostate Cancer (ERSPC).

Methods: A total of 141578 men, age 55-69 years, were randomized to systematic screening or usual care in the ERSPC sections Finland, Italy, the Netherlands and Sweden. The excess number of deaths was defined as the difference between the observed number of deaths in the prostate cancer (PC) patients and the expected number of deaths up to December 31, 2006. The expected number was derived from mortality of all study participants before a diagnosis with PC adjusted for study centre, study arm and study attendance. The excess mortality rates were compared between the two study arms.

Results: The PC incidence rate was 9.25 per 1000 person-years in the intervention arm and 5.49 per 1000 person-years in the control arm, RR 1.69 (95%CI 1.62-1.76). The excess mortality rate was 0.29 per 1000 person-years in the intervention arm and 0.37 men per 1000 person-years in the control arm; the RR for excess mortality was 0.77 (95%CI, 0.55-1.08). The absolute risk reduction in the excess mortality was 0.08 per 1000 person-years. The overall mortality rate was not significantly different between the intervention and the control arms of the study: RR 0.99 (95%CI 0.96-1.01).

Conclusions: Although the reduction in excess mortality was not statistically significant, the between-arm reduction in excess mortality rate was in excellent agreement with the previously report 20% reduction in the disease-specific mortality. This finding corroborates that PSA-based screening reduces the rate of excess deaths from PC. Estimation of excess mortality can be used in studies to evaluate the effects of PC screening on mortality.

INTRODUCTION

Randomized studies have shown that Prostate Specific Antigen (PSA)-based screening for prostate cancer (PC) causes a significant reduction in PC specific mortality in men aged 55-69 years at invitation for screening[1, 2]. The effect was studied by comparing the number of deaths from PC in the screened and unscreened population of the trials. The cause of death of all men who died and were diagnosed with PC was labeled as either "death from" or "death with" PC by independent committees in each of the participating countries[3]. No effect of screening was found on the all-cause mortality.

There have been debates on the issue of whether disease-specific death should be the endpoint of a cancer screening trial[4-6]. All-cause mortality is potentially a more valid outcome than disease-specific mortality since all-cause mortality analyses are superior in discovering the ineffectiveness and, in some cases, the side-effects of an intervention. For example, a screening program can reduce the disease-specific mortality whereas the side effects of the screening procedure (anxiety, depression or even the treatment for the disease being screened for itself) cause an increased risk of death of other causes. In such a case the disease- specific mortality decreases while the overall mortality increases. In clinical trials, such effects are known from fibrates for cholesterol reduction[7], antiarrhythmics following myocardial infarction[8, 9], and liberal red cell transfusion in the critically ill[10].

However, for evaluation of a screening trial, disease-specific mortality remains the best surrogate endpoint since screening trials would require millions of subjects to have the statistical power to detect a reduction in all-cause mortality. In a recent study we have presented an alternative to assess the total effect of screening on the mortality in a randomized trial[11]. It comprises a comparison of the estimates of the excess mortality rate in the cancer patients in both arms of the study. These excess mortality rates are based on the actually recorded number of cancer patient deaths in excess of the number expected on the basis of a cohort of cancer free individuals per unit of time. To study the use of excess mortality in prostate cancer screening further this study is based on data of the four largest centres of the European Randomized Study of Screening for Prostate Cancer (ERSPC).

STUDY DESIGN

The ERSPC was designed as a randomized, multicenter trial of screening for PC, with the rate of death from PC as the primary outcome[12]. As complementary approach to the primary outcome, it was planned to calculate the 'excess mortality' in PC patients (i.e. patients diagnosed after randomization)[13].

In the present study, data from the four largest ERSPC centres are analysed to have robust estimates of the excess mortality. The procedure of recruitment and randomization differed among centres. In Finland, Italy, and Sweden men underwent randomization before written informed consent was provided by those allocated to the screening arm (population-based effectiveness trial). In the Netherlands, men provided informed consent before randomization (efficacy trial). In Finland a total of 80377 men aged 55-69 were randomized, in Italy a total of 14517 men, in Sweden a total of 11852 men, and in the Netherlands a total of 34832 men. In Italy, Sweden and the Netherlands the study was performed based on randomization in a 1:1 ratio to the screening group or the control group. In Finland, the size of the screening group to the control group was approximately 1:1.5.

Screening tests and indications for biopsy

In Sweden a PSA cut-off value of 3.0 ng/ml was used as an indication for biopsy[14]. In Finland, a PSA value of 4.0 ng/ml or more was defined as positive test by which men were referred for biopsy; those with a value of 3.0 to 3.9 ng/ml underwent an ancillary test (digital rectal examination until 1998 and calculation of the ratio of the free PSA value to the total PSA value (biopsy indication if FT-ratio < 0.16) starting in 1999) and were referred for biopsy if either of the two tests was positive[15]. In Italy, a PSA value of 4.0 ng/ml or more was defined as positive, but men with a PSA value of 2.5 to 3.9 ng/ml also underwent ancillary tests (digital rectal examination and transrectal ultrasonography)[16]. In the Netherlands, up to February 1997, a combination of digital rectal examination, transrectal ultrasonography, and PSA testing (with a cutoff value of 4.0 ng per milliliter) was used for screening; in 1997, this combination was replaced by PSA testing only (biopsy indication if PSA >= 3 ng/ml)[17]. Centres used sextant biopsies guided by transrectal ultrasonography. In June 1996, lateralized sextant biopsies were recommended. In Italy, transperineal sextant biopsies were used. In Finland, a biopsy procedure with 10 to 12 biopsy cores was adopted in 2002 as a general policy for the two study groups. The screening interval in Finland, Italy and the Netherlands was 4 years; Sweden used a 2-year interval. Treatment of PC was performed according to local policies and guidelines. The distribution of treatments that were applied to the screening group and the control group was comparable[18].

Follow up data

Cancer incidence data were obtained by linking the trial database with the regional (Italy, Sweden) or national (the Netherlands, Finland) Cancer Registry. Mortality data of participants who died in the period up to December 31, 2006 were obtained by linking the trial database with the National Causes of Death Registry. Linkage was possible by using the personal administrative number of each participant as a linkage key. Causes

of death were evaluated in a blinded fashion and according to a standard algorithm or, after validation, on the basis of official causes of death[3]. Patients were determined to have died from PC if they were classified as either "definitely PC death", as "probable PC death" or as "PC intervention related death"

STATISTICAL ANALYSIS

Follow-up for diagnosis and mortality analyses began at randomization and ended at death, emigration, or date of censoring (December 31, 2006), whichever came first. Men who were randomized to the intervention arm were classified as attendees or non-attendees using the following definition: after randomization all men in the intervention arm were considered as non-attendees until the date that they participated for the first time in the study by a PSA test, DRE or TRUS. At this date their status switched from non-attendee to attendee. Men who were not screened during the study were considered as non-attendees during the complete follow-up.

The age-specific expected number of deaths for PC patients in the screening arm who attended a screening visit was calculated on the basis of the age-specific observed mortality rate in study participants randomized to the screening arm who attended a screening visit (and before any PC diagnosis.). The age-specific expected number of deaths for PC patients in the screening arm who did not attend a screening visit was calculated on the basis of the age-specific observed mortality rate in study participants randomized to the screening arm who did not attend a screening visit (and before any PC diagnosis). The age-specific expected number of deaths for PC patients in the control arm was calculated on the basis of the age-specific observed mortality rate in study participants who were randomized to the control arm of the study before any PC diagnosis. The excess mortality for the three groups of PC patients mentioned above was calculated by subtracting the appropriate expected number of deaths from the observed number of deaths for the group. Thus the expected number of deaths was calculated based on the assumption that the PC patients would have had the same age-specific mortality as the study population taking into account attendance status (attendees and non-attendees in the intervention arm and all men in the control arm before any diagnosis of PC).

For each study arm, the sum of the number of life years of all participants from the moment of randomisation until either censoring or death was calculated. This sum was divided by the number of men randomized to that arm which yields the average number of life-years until either censoring or death per study arm per participant. Subtracting the number of life years per participant in the control arm from the number of life-years per participant in the screening arm yields the life-time gained per participant.

Poisson regressions (using the stata generalized linear model function glm) were used to calculate all rate ratios, confidence intervals of rate ratios and their associated p-values[19]. In this approach the observations (deaths or number of PC cases) are assumed to be distributed according to a Poisson distribution. The natural logarithm of the expected value of this Poisson distribution is assumed to be a linear function of the study arm (coded as an indicator value, 1 for the screening arm, 0 for the control arm) with an added constant consisting of a fixed part (model constant) and the natural logarithm of the number of person-years. In this parameterization the exponentiated parameter associated with the study arm indicator yields the ratio of the hazard in the screening arm and the control arm (the hazard rate ratio) and the exponentiated model constant yields the hazard in the control arm. However, the excess mortality is modelled with a different link function. In the case of excess mortality the expected value of the Poisson distribution does not equal the natural logarithm of the number of deaths but the natural logarithm of the excess number of deaths (i.e. the difference of the observed and expected number of deaths). The 95% confidence interval of the excess mortality hazard rate ratio and its associated p-value were based on the delta method (accounting for attendance in the intervention arm and difference in randomization scheme in Finland) shown in the Appendix. A two-sided *p* value <0.05 was considered statistically significant. All analyses were performed with the commercially available STATA package: Data Analysis and Statistical Software, version 10.

RESULTS

Baseline characteristics

A total of 141578 men, age 55-69 years at randomization, were included in the study. Of these men 62578 were assigned to the intervention group and 79000 to the control group. The PC incidence rate in the intervention arm was 1.69 fold higher than in the control arm, i.e., 5206 (8.3% of total) men were diagnosed with PC in the intervention arm and 3871 (4.9% of total) men were diagnosed with PC in the control arm (Table 1). Up to the end of 2006, 9673 (15.5% of total) men died in the intervention arm and 12309 (15.6% of total) died in the control arm. This resulted in an all cause mortality that was not significantly different between the two arms: RR 0.99 (95%CI 0.96-1.01, p=0.274), Table 2.

Disease-specific mortality rate

Up to the end of 2006, in the intervention arm 188 men and in the control arm 296 men died from a PC-related cause of death according to the causes of death committee (CODC). The corresponding disease-specific mortality rate was 0.33 men per 1000

Table 1. Number of subjects, deaths and results of screening according to study group, attendees and non-attendees.

	Intervention	Attendees	Non-attendees	Controls	All
Number of men	62578	50242	12336	79000	141578
Age at randomization, median	59	60	59	59	59
mean	59.8	59.9	59.6	59.7	59.8
Person-years	562521	437709	124813	705264	1267786
Number of PC	5206	4681	525	3871	9077
PC rates	9.25	10.70	4.21	5.49	7.16
Number of deaths	9673	6100	3573	12309	21982
Death rates	17.20	13.94	28.63	17.45	17.34
Number of PC deaths	188	126	62	296	484
PC death rate	0.33	0.29	0.50	0.42	0.38

Attendees: men randomized to intervention arm who attended screening, Non- attendees: men randomized to intervention arm who never attended screening, PC: prostate cancer, Deaths: Number of all cause deaths, Rates are expressed as number per 1000 person-years

Table 2. Effect of screening on prostate cancer incidence, overall mortality, prostate cancer specific mortality and excess mortality

	Intervention*	Controls*	Relative risk	95% CI	P value
PC rates	9.25	5.49	1.69	1.62 – 1.76	< 0.001
Death rates	17.2	17.5	0.99	0.96 – 1.01	0.274
PC death rate	0.33	0.42	0.79	0.66 – 0.96	0.014
Excess mort rate	0.29	0.37	0.77	0.55 – 1.08	0.132

*All rates are expressed as number per 1000 person-years

person-years in the intervention arm and 0.42 men per 1000 person-years in the control arm. This resulted in a statistically significant reduction in PC mortality rate of 21% in the intervention population relative to the control arm: RR 0.79 (95%CI, 0.66-0.95, p=0.014), Table 2.

Excess mortality

Up to the end of 2006, the excess mortality rate was 0.29 men per 1000 person-years in the intervention arm and 0.37 men per 1000 person-years in the control arm. This resulted in a non-significant reduction in the excess mortality rate of 23% in the intervention population relative to the control population: RR 0.77 (95%CI, 0.55-1.08, p=0.132), Table 2. This equaled with an absolute risk reduction in the excess mortality of 0.075 per 1000 person-years. The life time gained per participant was approximately 23 days after a median follow-up of nine years.

DISCUSSION

Since the latest publications by the ERSPC trial[1, 2], there is growing consensus that PSA-based screening for PC reduces mortality from the disease. Some issues remain unresolved, most notably those regarding the cost-effectiveness of screening, as well the concerns about the quality of life related to screening. Nearly no concerns consist on the validity of PC-specific mortality as endpoint of a PC screening trial. Nevertheless, in the ERSPC trial, screening might have reduced the number of deaths from PC while side effects of the screening procedure (anxiety, depression) might have caused an increase in the death from other causes. Up to the present study these screening effects remained unknown. The present study showed a reduction in the rate in excess mortality as an effect of screening that is almost identical to the reduction in the rate in disease-specific mortality. This is an important observation since by definition disease-specific mortality is included in excess mortality but is not identical to it. Disease-specific mortality refers to mortality due to only one particular cause of death such as disease progression and treatment related mortality and excess mortality measures both the direct and indirect mortality due to the cancer of interest, e.g. including cachexia, uraemia, suicide, depression and loss of interest in life. Although in this study the RR of excess mortality did not reach the formal limits of statistical significance, the present study strongly corroborates the finding that PSA-based screening reduces the direct and indirect mortality related to PC and strongly supports the fact that PSA-based screening reduces the rate of death from PC. The support provided by this analysis is obtained in an objective way since it does not require human expertise to determine a cause of death.

After a median of nine years follow up, the ERSPC as a whole showed an alpha spending adjusted RR for death from PC in the intervention group of 0.80 (95%CI, 0.65-0.98;p=0.04) based on data including 162243 men randomized in seven different countries[1]. Data in the present study showed a reduction in the PC specific mortality rate of 21%, RR 0.79 (95%CI, 0.66-0.95;p=0.014). The PC related cause of death was based on the consensus of the CODC that evaluated all deceased cases in a blinded fashion according to a standard algorithm[3]. Such committees are assumed to be the most optimal method for cause of death ascertainment, especially in PC where death on the death certificates differs often from the clinical picture. Nevertheless, in practice it was impossible to blind the arm of the study entirely before the committee reviewed the records of deceased men. The records were for example not blinded for stage at diagnosis although this is a very suggestive variable for the arm of the study. Therefore it remained possible that the lack of adequate blinding of the records with respect to study arm had resulted in a bias in the cause of death ascertainment. Based on the findings in the present study this potential bias did not influence the study outcomes.

The results are not completely in line with a previous study that compared the effect of screening on the disease-specific and excess mortality in the ERSPC section Rotterdam only[11]. The previous study showed in contrast to the disease-specific mortality rates an increased difference in the excess mortality rates between the two study arms. The main difference between the excess mortality and disease-specific mortality rate was observed for the older age group at randomization, i.e. 70-74 years, men that were not included in the present study. The results for men age 55-64 years at randomization in the present study were in line with the results presented before[11]. The effect for age is in good agreement with excess mortality studies on mammography screening for breast cancer that also observed a more pronounced discrepancy between the excess mortality and disease-specific mortality analysis for the older ages[19, 20]. As a result, the excess mortality analyses should be incorporated in the evaluation of the screening trial when longer follow up is included. In line with suggestions made in the previous publication, it could be that there is an increase in the uncertainty of individual cause of death ascertainment in older age groups when the overall mortality in these older ages is increased.

Excess mortality analyses are often limited by the validity of the expected mortality. Typically, in an excess mortality analysis, the expected mortality is estimated from nationwide population life tables stratified by age, sex and period. However, in a screening setting, the general population does not seem to be suitable as a reference population since the people who participate are healthier and of higher socioeconomic classes, i.e., subject to healthy screenee bias[11, 21]. For this reason, in the present study, the expected mortality was based on all participants before a diagnosis of PC and adjusted for study arm and study participation. These adjustments were needed. This is shown in detail by Kranse et al[22]. This study showed that if the expected mortality were to be estimated without a correction for attendance status, an excess mortality analysis on PC screening would over-estimate the effect of screening since men that participate and attend in a screening trial have been shown to be healthier and have a decreased risk of death from other causes compared to men who do not attend. Additional adjustment by estimating the expected mortality with correction for noncompliance among men participating in the intervention arm with a biopsy indication might have optimized the excess mortality since men in the intervention arm who had a raised PSA but were not biopsied were likely to have higher all cause mortality than those with biopsy. Although this seems to have a very small effect on the final outcomes, it might be incorporated in future excess mortality analyses.

Population-based screening entails an intervention in a healthy population and, therefore, potentially should have a favorable harm-benefit trade-off, which at the moment is not yet proven. Therefore, although this study corroborates that PSA-based screening reduces the mortality, population-based PC screening cannot be recommended at the

moment. The PC incidence in the intervention arm was 1.69 fold higher than in the control arm. However, since PC is a leading chronic condition affecting men and the incidence of overdiagnosed PC is increasing rapidly, there is an urgent need to resolve these issues. Therefore additional studies are needed. The present study showed that the estimation of excess mortality can be used in future studies to evaluate the effects of screening on PC mortality. This is useful since disease-specific mortality requires reliably coded information on cause of death and assumes that cancer mortality is independent of competing risk mortality, an assumption which is especially for PC only approximately true.

In conclusion, the PSA-based screening seems to have benefit on the excess mortality. The estimated excess mortality reduction of 23% is in excellent agreement with our earlier results.

APPENDIX

Assessment of the 95% confidence interval of the excess mortality rate ratio for a single study centre by accounting for participation.

The calculation of the confidence interval of the excess mortality rate ratio without taking attendance in the screening arm of the study into account has been described in detail in Appendix of van Leeuwen et al [200].

The two excess mortality rates (xsmr) are presented as $m_i = e_i/n_i$, where i = 1, 2 is the study arm (i= 1: intervention arm and i= 2: control arm), e_i =arm-specific excess number of deaths and n_i = arm-specific number of observed person-years. N_i = number of prostate cancer cases per study arm. q_i and q_i* are death proportions in prostate cancer patients, observed and expected respectively.

Having introduced the basic method [200] and notation we can introduce a first refinement, accounting for attendance in the screening arm. Since the attendance issue only affects the screening arm the formula for index i =2 are unaffected. In the following formula the subscript 1p denotes attendees and 1n denotes non-attendees.

The excess mortality rate ratio (xsmrR), now accounting for participation (p) and non-participation (n) in the screening arm becomes:

$$xsmrR = \frac{N_{1p}(q_{1p} - q_{1p}^*) + N_{1n}(q_{1n} - q_{1n}^*)}{N_2(q_2 - q_2^*)} \quad \frac{n_2}{n_1} =$$

$$\frac{num}{den} \quad \frac{n_2}{n_1}, \qquad (1)$$

Using the delta method the variance of the logarithm of (1) can be calculated as in [11]

$$var(\ln(xsmrR)) = \left(\frac{n_2}{n_1}\right)^2 \left(\frac{var(num)}{(num)^2} + \frac{var(den)}{(den)^2}\right).$$

The var(num) and var(den) are calculated as:

$$var(num) = N_{1p}((q_{1p} - q_{1p}^*)^2 + q_{1p}(1 - q_{1p})) + N_{1n}((q_{1n} - q_{1n}^*)^2 + q_{1n}(1 - q_{1n}))$$

$$var(den) = N_2((q_2 - q_2^*)^2 + q_2(1 - q_2)).$$

Calculating the pooled excess mortality rate ratio over all study centers and its confidence interval.

The above section deals with the calculation of the confidence interval of the xsmrR, accounting for attendance in the intervention arm in one study center. In the present study we consider the four study centers combined. In three study centers the randomization is 1 to 1, in one center (the Finnish center) it is 2 to 3, for every two men randomized to the intervention arm, 3 are randomized to the control arm. Thus if we calculate the pooled xsmrR we have to account for the different randomization scheme used in Finland. Were this not done, the contribution of the Finnish center to the xsmr in the control arm would be different than the contribution of the Finnish center to the xsmr in the screening arm. In that case, an observed difference in the xsmrR can be attributed to either this imbalance or to a true difference between the intervention arm and the control arm excess-mortality rate (or any combination of the two). Thus correction for this imbalance is necessary to enable an unambiguous interpretation of the xsmrR.

Let us denote the research centers by r = 1,2,3 and 4. Then we get:

$$xsmrR = \frac{N_{1pr}(q_{1pr} - q_{1pr}^*) + N_{1nr}(q_{1nr} - q_{1nr}^*)}{N_{2r}(q_{2r} - q_{2r}^*)(n_{1r} / n_{2r})}$$

$$= \frac{num_r}{den_r (n_{1r} / n_{2r})}.$$

Overall,

$$xsmrR = \frac{\sum_{r=1}^{4} num_r}{\sum_{r=1}^{4} den_r (n_{1r} / n_{2r})} = \frac{NUM}{DEN} .$$

Then,

$$var(\ln(xsmrR)) = var(NUM) / (NUM)^2 + var(DEN) / (DEN)^2,$$

in which,

$$var(NUM) = \sum_{r=1}^{4} var(num_r)$$

and,

$$var(DEN) = \sum_{r=1}^{4} (n_{1r} / n_{2r})^2 \, var(den_r).$$

The quantities var(numr) and var(denr) are given in the previous attachment as var(num) and var(den). Note that in practice to a very good approximation for this specific study $n_{1r} / n_{2r} = 2/3$ for the Finnish center and $n_{1r} / n_{2r} = 1$ for the other centers.


CRUISIN' FOR SCREENING
PROSTATE CANCER

10

Increased non-prostate cancer death risk in clinically diagnosed prostate cancer

BJU Int. 2012 Jan 30

Pim J. van Leeuwen, Suzie J. Otto, Ries Kranse, Monique J. Roobol, Meelan Bul, Xiaoye Zhu, Harry de Koning and Fritz H. Schröder

ABSTRACT

Objectives: To assess the cause specific mortality unrelated to prostate cancer (PC) itself in men with screen and clinically diagnosed PC.

Methods: Study among participants of the European Randomized Study of Screening for Prostate Cancer. Based on consensus of causes of death committee, all men who died from PC were excluded. In the intervention arm, cases were men with a screen detected PC, aged 55-74 years, between 1993-2001. These cases were matched to 2 controls in whom no cancer was found after biopsy, and 2 controls in whom no cancer was suspected screening. In control arm, cases were men with clinically diagnosed PC, aged 55-74 years, between 1993-2001. These cases were matched to 4 controls without PC. Matching was done with respect to date of birth, -screening and/or -diagnosis. Men were followed up to December 31, 2007.

Results: No statistically significant difference in overall mortality between cases and controls in the intervention arm was observed: RR 1.26 (95%CI0.96-1.65; p=0.102) and RR 1.13 (95%CI0.86-1.47; p=0.381). In the control arm, the overall mortality was statistically significantly higher in cases relative to controls: RR 1.43 (95%CI1.03-2.00; p=0.033). This difference was due to an increased risk of dying from neoplasms and disease of the circulatory or respiratory system among cases: RR 1.61 (95%CI1.12-2.29; p=0.009). The study was limited by the relatively small size.

Conclusions: Increased mortality unrelated to PC itself was observed in men with clinically diagnosed PC, but not in screen detected PC. The excess mortality in men with clinically diagnosed PC seems due to significantly increased risk of dying from neoplasm and disease of the circulatory or respiratory system. Results have to be studied more thoroughly in further clinical trials.

10.1 INTRODUCTION

Prostate cancer (PC) has become the most common non-cutaneous diagnosed cancer in men in Europe and the United States[5]. Currently, only a small percentage of men with predominantly localized PC die from a PC related cause of death. Cardiovascular disease (CVD) is the primary or secondary cause of death in most PC patients[144].

Recent evidence has suggested that there might be a positive correlation between increased cardiovascular risk and the PC incidence, progression and treatment for PC[191-196]. Furthermore, some methods to reduce the risk for CVD seem to be similar to methods to reduce the risk for high risk PC[197-199]. In addition, recent evidence has suggested that the excess mortality is lower among men who were diagnosed with screen detected PC in comparison to men with clinically diagnosed PC, possibly due to the use of medications for CVD and/or the change to a healthier lifestyle of men with a screen detected PC[200]. The current study was designed to quantify the excess mortality measured in the study by van Leeuwen et al.[200].

Previously we have demonstrated that CVD mortality is not increased among men diagnosed with prostate cancer in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC) compared to the general Dutch population[201]. However, no distinction was made between screen detected and symptomatically detected patients, and to socioeconomic class which is known to be higher in men participating in the ERSPC in comparison to the general population. Making use of the detailed information on each individual participant in the ERSPC Rotterdam, we compared the incidence of overall mortality and non-prostate cancer causes of death among men without PC and men with screen-detected PC on the one, and unscreened men with symptomatically diagnosed PC on the other hand. With this analysis we have the aim to assess whether men with screen detected and symptomatically diagnosed PC are at an increased risk of death and of which particular causes. The comparison group for men with screen detected PC consisted of men who were screened for PC but in whom no PC was detected. Unscreened men in the control arm with symptomatically PC were compared with men participating in the control population of the ERSPC Rotterdam who were not diagnosed with PC.

10.2 MATERIALS AND METHODS

10.2.1 Study Population

All men were participants of the Rotterdam section of the ERSPC. All men signed an informed consent before randomization to systematic screening (intervention arm) and usual care (control arm).

Men in the intervention arm were screened with an interval of four years by PSA measurement, rectal examination (DRE) and transrectal ultrasound examination (TRUS) between December 1993 and May 1997. A sextant biopsy was initially offered to men with PSA ≥4.0 ng/ml and/or a suspicious finding on DRE and/or TRUS. After May 1997 a biopsy was prompted by PSA ≥3.0 ng/ml only. All cancers were classified according to the 1992 TNM classification. Men who were diagnosed with PC were treated by the local urologist. Details of the screening methodology were reviewed by Roobol et al[96].

10.2.2 Design

The current study is designed as a prospective cohort study including cases and controls where case subjects were men diagnosed with PC (screen-detected or symptomatically) and control subjects were men who were not diagnosed with PC until death or the cut-off date December 31, 2007. Matching was performed to ensure equal risks of non-prostate cancer death among the case and control subjects.

10.2.3 Selection of screen-detected cases and matched controls

All males, aged 55-74 year, in the intervention arm of ERSPC Rotterdam diagnosed with a screen detected PC between January 1, 1995 and December 31, 2004 were eligible as case subject (case intervention arm) and were actually selected if the tumour was diagnosed at screening and was localized (defined as stage T1C, N0, M0 and serum PSA < 20.0 ng/ml). Cases were excluded if they were determined as PC related cause of death before December 31, 2007. These inclusion criteria were made to select only screen detected PCs. Only T1c cancers were included since T1c PC is the disease diagnosed in a patient at screening without any clinical sign they have the disease. The date of diagnosis of the case was the index date. The control subjects were randomly selected from the intervention arm, alive at the date of diagnosis of the case and were matched to the case subjects by month and year of birth, self reported health status (i.e., good, moderate or poor) and month and year of screening. Each time a case subject was included in the study, we randomly selected four controls; two control subjects in whom no cancer was found after a prostate biopsy and two control subjects who had a normal PSA value (PSA < 3.0 ng/ml) and did not undergo prostate biopsies or were not suspicious for having cancer. If there was no match in respect to the previous mentioned variables, then a control subject was selected who was born 1 month before or after the case subject. If that still did not lead to a match, then a control subject was selected born 2 months after the case subject and then, if necessary, 2 months before the case subject was born, and so forth, until a match was found with a maximum difference of 6 months. Two most optimal groups were chosen since both might be criticisable for a reason. The aim of the current study design was to indentify controls who were in the absence of PC. We selected men with a PSA < 3.0 ng/ml at a screening visit although it is known that there

is no PSA value where below PC is not detectable[72], and men with no cancer found at a sextant biopsy although it is known that a sextant prostate biopsy misses a substantial percentage of PC[170].

10.2.4 Selection of unscreened symptomatically diagnosed cases and matched controls

These case subjects, aged 55-74 year, were all men randomized to the control arm of the ERSPC, section Rotterdam, who were clinically diagnosed with PC, defined as stage >T1c, N0/N1, M0/M1 and serum PSA > 0 ng/ml, between January 1, 1995 and December 31, 2004 (case control arm). Cases were excluded if they were determined as having died from a PC related cause before December 31, 2007. T1c cancers were excluded since T1c PC is the disease diagnosed in a patient at screening without any clinical sign they have the disease. The controls were randomly selected from the control arm, alive at the date of diagnosis of the case, and matched to the case subjects on month and year of diagnosis with respect to month and year of birth, self reported health status (i.e., good, moderate or poor) and month and year of randomization. Each time a case subject was included in the study, we randomly selected 4 controls. Control subjects were men who were randomized to the control arm of the ERSPC, section Rotterdam, alive at the time of diagnosis of the case, and not diagnosed with PC. If there was no match in respect to the previous mentioned variables, the same complementary matching process was performed as described before.

10.2.5 Mortality Data

Mortality data of case and control subjects in both the intervention and the control arm that died in the period up to December 31, 2007 were obtained by linking the trial database with the Causes of Death Registry of Statistics Netherlands. Causes of death were based on the national certificates coded according to the International Classification of Diseases (9th revision [ICD9]; 1995) and ICD10 (from 1996 onward), and the causes of death grouping was based on the tabulation list for main primary causes of death of Statistics Netherlands[154]. For all case subjects in the study who were known to have died, all available information was gathered and anonymised. Subsequently, cases classified as having died from PC related death by the independent causes of death committee (CODC) of the ERSPC Rotterdam were excluded from the study. The CODC reviews all deceased PC cases using predefined flow charts. Patients were determined to have died from PC if they were classified as either "definitely PC death", "possible PC death" or as "PC intervention related death"[153].

10.3 STATISTICAL ANALYSES

The mortality rates were compared between cases and controls using a Poisson regression analysis with an indicator of study arm as a predictor and the logarithm of the number of person years as an offset term (predictor with a coefficient of one)[178]. For the cases and controls, the time of follow-up was measured from the date of randomization up to date of death or December 31, 2007. *P*-values <0.05 were considered statistically significant. All analyses were performed with the commercially available STATA package: Data Analysis and Statistical Software, version 11.

10.4 RESULTS

10.4.1 Baseline characteristics

A total of 372 cases and 1488 controls participating in the intervention arm of the ERSPC section Rotterdam were included in this study, and a total of 221 cases and 884 controls participating in the control arm of the ERSPC section Rotterdam. The age at diagnosis and tumor characteristics of the cases are presented in Table 1. The median age at diagnosis differed significantly between the two groups ($p<0.001$); cases in the intervention arm were diagnosed at an earlier age.

10.4.2 Case and Control Subjects Intervention Arm

The median follow up from time of diagnosis up to either death or end of follow up was 8.9 years for the cases. Up to the end of 2007, a total of 82 (22.0%) cases died (26.2 men per 1000 person years). A total of 139 (18.7%) controls died in whom no cancer was found after biopsy, and a total of 155 (20.8%) controls in whom no cancer was suspected (Table 2). This resulted in an all cause mortality that was not significantly different between cases and controls: RR 1.26 (95%CI 0.96-1.65; $p=0.102$) for the cases relative to the controls in whom no cancer was found after biopsy; and RR 1.13 (95%CI 0.86-1.47; $p=0.381$) for the cases relative to the controls in whom no cancer was suspected. In addition no significant difference was found between the cases and all controls in the intervention arm; RR 1.19 (95%CI 0.93-1.52; $p=0.168$).

With respect to CVD, a total of 26 (7.0%) cases died from CVD (8.3 men per 1000 person years) and a total of 47 (6.3%) controls in whom no cancer was found after biopsy, and a total of 53 (7.1%) controls in whom no cancer was suspected. Consequently, no statistically significant difference in CVD was observed between the cases and controls; RR 1.18 (95%CI 0.73-1.90; $p=0.503$) for the cases relative to the controls in whom no cancer was found after biopsy; and RR 1.04 (95%CI 0.65-1.67; $p=0.854$) for the cases relative to the

Table 1. Baseline characteristics at diagnosis

	Cases Intervention arm, *N* (% of total)	Cases Control arm, *N* (% of total)	*P* value
Total participants included	372	221	
Age (yr), median	66.6	68.5	<0.001
55 – 60	68 (18.3)	19 (8.6)	<0.001
61 - 65	107 (28.8)	56 (25.3)	
66 - 70	115 (30.9)	79 (35.8)	
71 – 74	82 (22.0)	67 (30.3)	
PSA at diagnosis (ng/ml), median	4.7	11.6	<0.001
Tumour Stage			
T1c	372 (100.0)	-	
T2	-	141 (63.8)	
T3	-	76 (34.4)	
T4	-	4 (1.8)	
Histological differentiation			
Gleason 2 – 6	293 (78.8)	90 (40.7)	<0.001
Gleason 7	70 (18.8)	77 (34.8)	
Gleason 8 – 10	9 (2.4)	30 (13.6)	
Not known or not performed	-	24 (10.9)	
Primary Treatment			
Surgery	156 (41.9)	59 (26.7)	<0.001
Radiotherapy	131 (35.3)	116 (52.5)	
Watchful waiting	83 (22.3)	13 ((5.9)	
Hormone	-	30 (13.6)	
Other	2 (0.5)	3 (1.3)	
Adjuvant Treatment			
Hormone	4 (1.1)	28 (12.7)	<0.001

controls in whom no cancer was suspected. The RR on CVD was 1.11 (95%CI 0.72-1.70; p=0.643) for all cases relative to all the controls.

10.4.3 Causes of death intervention arm

The six most frequent causes of death among the cases and controls in the intervention arm are presented in Table 2. Neoplasms were the most frequent cause of death, followed by CVD. Despite all PC related deaths based on the consensus of the CODC were excluded, in 3 cases the primary cause of death was PC (3.6% of all deaths). The mortality from neoplasms, CVD and diseases of the respiratory system together was not statistically significant different in the cases relative to the controls; RR 1.20 (95%CI 0.92-1.58; p=0.180).

Table 2. Causes of death in the intervention arm stratified by study group

Study Group	Cases N (%)	Controls: All N (%)	RR (95%CI)	Controls: No PC after biopsy N (%)	RR (95%CI)	Controls: No PC suspected N (%)	RR (95%CI)
Total included	372 (100)	1488 (100)		744 (100)		744 (100)	
Total no. of deaths	82 (22.0)	294 (19.8)	1.19 (0.93-1.52)	139 (18.7)	1.26 (0.96-1.65)	155 (20.8)	1.13 (0.86-1.47)
Total CVD deaths	26 (7.0)	100 (6.7)	1.11 (0.72-1.70)	47 (6.3)	1.18 (0.73-1.90)	53 (7.1)	1.04 (0.65-1.67)
Neoplasms	40 (10.7)	118 (7.9)		54(7.2)		64 (8.6)	
Respiratory system	1 (0.2)	20 (1.3)		7 (0.9)		13 (0.5)	
Digestive system	2 (0.5)	7 (0.5)		4 (0.5)		3 (0.4)	
Endocrine and metabolic	1 (0.3)	10 (0.7)		5 (0.7)		5 (0.7)	
Abnormal clin or lab findings	3 (0.8)	11 (0.7)		7 (0.9)		4 (0.5)	

Only the six most frequently recorded primary causes of death of Statistics Netherlands are presented; clin: clinical; lab: laboratory; RR: Relative Risks using a Poisson regression analysis with an indicator of study arm as a predictor and the logarithm of the number of person years as an offset term (predictor with a coefficient of one)

10.4.4 Case and Control Subjects Control Arm

For the cases the median follow up from time of diagnosis up to either death or end of follow up was 6.1 years. A total of 47 (21.3%) cases died (34.2 men per 1000 person years) and a total of 134 (15.2%) controls (Table 3). This resulted in an all cause mortality that was statistically significantly different between cases and controls: RR 1.43 (95%CI 1.03-2.00; p=0.033). The CVD was not statistically significantly different between the cases and controls (Table 3). A total of 11 (5.0%) cases died from CVD (8.0 men per 1000 person years) and a total of 30 (3.4%) controls died from CVD; RR 1.50 (95%CI 0.75-2.99; p=0.250) for cases relative to the controls.

10.4.5 Causes of death Control Arm

The six most frequent causes of death among the cases and controls in the control arm are presented in Table 3. Neoplasms were the most frequent cause of death followed by CVD and diseases of the respiratory system. Based on data of the Causes of Death Registry of Statistics Netherlands, a total of 6 death cases were recorded as primary cause PC, although all PC related deaths based on the consensus of the CODC were excluded (12.8% of all deaths). The mortality from neoplasms, CVD and diseases of the respiratory system together was statistically significantly higher among the cases relative to the controls; RR 1.61 (95%CI 1.12-2.29; p=0.009).

Table 3. Causes of death in the control arm stratified by study group

Study Group	Cases N (%)	Controls: N (%)	RR (95%CI)
Total included	221 (100)	884 (100)	
Total no. of deaths	47 (21.3)	134 (15.2)	1.43 (1.03-2.00)
Total CVD deaths	11 (5.0)	30 (3.4)	1.50 (0.75-2.99)
Neoplasms	26 (11.8)	62 (7.0)	
Respiratory system	5 (2.3)	15 (1.7)	
Digestive system	2 (0.9)	4 (0.4)	
Endocrine and metabolic	1 (0.4)	2 (0.2)	
Abnormal clin or lab findings	0 (0.0)	6 (0.7)	

Only the six most frequently recorded primary causes of death of Statistics Netherlands are presented; clin: clinical; lab: laboratory; RR: Relative Risks using a Poisson regression analysis with an indicator of study arm as a predictor and the logarithm of the number of person years as an offset term (predictor with a coefficient of one)

10.6 DISCUSSION

In this study, men with clinically diagnosed PC were at an increased risk of dying from a cause unrelated to PC itself compared to men without PC, RR: 1.43 (95%CI 1.03-2.00; p=0.033) if all PC deaths based on the consensus of the CODC were excluded. This difference was based on an increased risk of dying from neoplasms and diseases of the circulatory and respiratory system, together: RR 1.61 (95%CI 1.12-2.29; p=0.009). Results are limited by a relatively small sample size. The observations are in line with previous findings of the ERSPC section Rotterdam in which a secondary mortality analysis showed that the excess mortality was higher than the disease specific mortality in the control arm of the ERSPC Rotterdam. This previous study suggested a possible underestimation of deaths from PC by the CODC in the control arm of the study, or an additional disease related mortality that was measured by an excess mortality analysis but not with a disease specific mortality analysis[200]. The current study confirms that both these suggestions have a contribution to the observed excess mortality in the control arm of the study.

All cases in this study that were determined by the CODC to have died from a PC related cause of death were excluded from the study. For this reason, in theory, the overall mortality in cases and controls was expected to be similar. However, this was not the case in the control arm of the study. Based on data of the death certificates, 3 cases with a screen detected PC (0.81% of all cases) and 6 cases with a clinically diagnosed PC (2.71% of all cases) were recorded with PC being the primary cause of death. This observation, in spite of being based on small numbers, suggests that the accuracy of the

cause of death determination is higher in screen detected than in clinically diagnosed PC. As suggested previously, it might be possible that the reviewers have underestimated the PC mortality by being too cautious in stating that a patient has died from PC in the study and that the information to make a judgement is more limited in the control arm of the study. Furthermore, it might be that due to the study protocol more information could be gathered from patients diagnosed with PC in the intervention arm resulting in a higher accuracy of the death certificates.

What are the reasons why men with PC have an increased risk of dying from causes unrelated to PC itself. First, recently, Hemelrijck et al. observed an increased risk of nonfatal and fatal CVD in patients with PC receiving curative treatment, surveillance or especially androgen-deprivation therapy (ADT). Findings that are in line with previous studies; a systematic review of 4 studies showed that men who underwent ADT had a significantly increased risk of CVD (summary hazard ratio, 1.17; 95%CI, 1.07-1.29)[202]. ADT is also associated with osteoporosis and fractures, diabetes and with an increased risk of colorectal cancer[203-205]. Given the widespread use of ADT to treat especially locally advanced PC, treatment with ADT might result in an increased risk of death in patients with clinically diagnosed PC in particular. In the current study, 26.6% of the cases in the control arm received primary and/or adjuvant treatment with ADT in contrast to 1.1% of the cases in the intervention arm (Table 1). Secondly, the median age at diagnosis in men participating in the control arm was higher than men in the screening arm. Consequently, the increased risk of dying from causes unrelated to PC itself might be associated to the increasing age at diagnosis. However, no effect of age at diagnosis could be measured among men diagnosed with PC in the screening arm.

The risk of excess mortality in screen detected PC was suggested to be lower due to changes in medical regimen, medication and lifestyle unrelated to PC itself[200]. It is known that the incidence of clinically insignificant PC is large among men participating in a screening program[129]. Nevertheless a large percentage of these men obtain invasive treatment with curative intent. Consequently, after a PC diagnosis, the first contact with the urologist and/or radiation oncologist often involves screening of patients general health including the measures of the vital signs that lead to other interventions that might be associated with the decrease in non PC specific mortality. It has been proven that the use of these medications is increased in men and women diagnosed with PC or breast cancer subsequent to screening[206-209]. Further, the change in medical treatments is studied in detail by a study among 180 men with screen detected PC. This study showed among patients diagnosed with a screen detected PC a significant change in medical treatments and prescriptions unrelated to PC itself subsequent to their diagnosis[209]. In total 72% of men had a change in medical regimen after diagnosis, 61% had a change in medication and 29% received a new medical diagnosis. In total 24 (14%) had a multiple gated acquisition scan, treadmill or persantine thallium test, and 23 (13%)

had another treatment performed, most frequently a cardiac catheterization. Among the men with changes to their medical regimen 36 (59%) began a new or different medication, which in 67% was a beta-blocker[209]. Meta analyses including twenty-three randomized controlled trials showed a statistically significantly lower CVD and all cause mortality (relative risk 0.76, 95% CI: 0.68, 0.84) in the groups treated with beta-blockers than in the control groups after a mean of 1 year[210]. Other frequently new started medication were statins. A meta-analyses including nine studies showed that the use of statins for secondary prevention in elderly patients with documented coronary heart disease reduced all-cause mortality 22% and reduced coronary heart disease mortality 30% median use of 3 years[211-213]. Also relatively frequently new started medications were inhaled corticosteroid and long-acting beta-2 agonists for treatment for chronic obstructive pulmonary disease. Using the combination of inhaled corticosteroid and long-acting beta-2 agonist therapy more than 6 months have shown to be associated with a 20% reduced total mortality[214]. Although these medication effects are shown in large randomized controlled trials, we might suggest that they have at least contributed to the observed effect among men with screen-detected cancers in the present study.
Since there is emerging evidence that lifestyle factors can alter the rate of progression of indolent PC, many men frequently take on the responsibility of improving their general health by making lifestyle and dietary changes[215-216]. A small randomized clinical trial showed that intervention participants had significantly improved their lifestyle compared with controls at 12 months[217]. Recommendations given to prevent progression of indolent PC include for example, choose nutritious foods, focus on fruits and vegetables, drink green tea, eat omega-3 rich foods, choose healthy fats, manage stress and take supplements. Consequently these recommendations might also have an effect on the non-PC related mortality, incidence and mortality rates of other neoplasms. Urologists might, therefore, be aware of positive effects of life-style modification and might give more attention to co-morbidities. Further urologists might develop programs in conjunction with for example cardiologists to counsel asymptomatic men before initiating early PC treatment, in order to identify those who are at an increased risk of non-PC related mortality. However, the results presented in the present study have a low level of evidence, which should not automatically support the changes in medical treatment in men with screen detected PC. Therefore, this issue should be studied more thoroughly in further analyses or other clinical trials. For example, the effect of the change in medical treatments might be studied in detail to assess the specific effect of the various medications on the outcomes in men with screen-detected PC.

There are limitations in this study: The number of deaths was relatively small. Consequently the CIs for the estimates of the cause specific death groups were rather wide. No data was available on smoking, BMI and serum lipid levels, although these are strong

risk factors for CVD. Although this study reached the optimal match between cases and controls that ensured an equal risk of non-prostate cancer death, it remains possible that especially the cases in the control arm were less healthy as part of their late diagnosed clinically significant PC (possibly these men refused the health system in general). In addition, the potential role of the use of medication that decreases the risk of CVD was not assessed in this study. Finally, data from this study after longer follow up is needed to confirm current observations. The strength of this study includes the validity of the control cases. In general, control cases are obtained from the general population although PC patients often have shown to be a selection of men considering general health and social economic class. Our study also has strong aspects. Participants were part of a randomized controlled trial and followed by regular matching with national cancer registries. Cases and controls were part of a homogeneous study population.

10.7 CONCLUSIONS

We concluded that men with clinically diagnosed PC have an increased risk of death unrelated to PC itself. This excess mortality was likely due to a significantly increased risk of dying from neoplasm and disease of the circulatory or respiratory system. No increased risk in all cause mortality unrelated to PC itself was observed among men diagnosed with PC subsequent to screening. Many effects might have influenced these observations, however, the relatively increased use of ADT in clinically diagnosed PC and the change of medical regimes and medication among men with screen detected PC may have had the most influence. These results should be studied more thoroughly in further clinical trials. The present study is limited by the relatively small sample size and small number of events. Finally, if changes in the medical regimens really do affect all cause mortality in men with PC, uro-oncologists should look carefully at the management of abnormal parameters of the circulatory and respiratory system, and should encourage PC patients to make lifestyle modifications.


CRUISIN' FOR SCREENING
PROSTATE CANCER

Part IV

How to screen for prostate cancer: risk assessment


CRUISIN' FOR SCREENING
PROSTATE CANCER

11

Towards an optimal interval for prostate cancer screening

Eur. Urol. 2012 Jan 61 (1): 171-6.

Pim J. van Leeuwen, Monique J. Roobol, Ries Kranse, Marco Zappa, Sigrid Carlsson, Meelan Bul[1], Xiaoye Zhu, Chris H. Bangma, Fritz H. Schröder and Jonas Hugosson

ABSTRACT

Background: The rate of decrease in advanced cancers is an estimate for determining the prostate cancer (PCa) screening program effectiveness.

Objectives: To assess the effectiveness using a 2-or 4-years screening interval.

Design, Setting, and Participants: Men, aged 55-64 years, were participants of two centres of European Randomized Study of Screening for Prostate Cancer (ERSPC): Gothenburg (2-year screening interval, n=4202) and Rotterdam (4-year screening interval, n=13301). We followed participants until date of PCa, date of death, last-follow up December 31, 2008, or up to a maximum of 12-years after initial screening. Potentially life-threatening (advanced) cancer was defined as cancer with at least one of following characteristics: clinical stage ≥T3a, M1 or N1, PSA >20.0ng/ml or Gleason score ≥8.

Interventions: We compared proportional total (advanced) cancer incidence (screen detected and interval cases), defined as the ratio of the observed number of (advanced) cancers to the expected numbers of (advanced) cancers based on the control arm of the study.

Measurements: The proportional cancer incidence from second screening round until end of observation was compared using a 2-years or 4-years screening interval.

Results and limitations: From screening round two until end of observation, proportional cancer incidence was 3.64 in Gothenburg and 3.08 in Rotterdam; RR1.18 (95%CI 1.04-1.33;p=0.009). Proportional advanced cancer incidence was 0.40 in Gothenburg and 0.69 in Rotterdam; RR0.57 (95%CI 0.33-0.99;p=0.048); the RR for detection of low risk PCa was 1.46 (95%CI 1.25-1.71;p<0.001). This study was limited by the assumption that PSA testing in the control arm was similar in both two centres.

Conclusion: A 2-year screening interval significantly reduced the incidence of advanced PCa, however increased the overall risk of being diagnosed with (low risk) PCa compared to a 4-year interval in men aged 55 to 64 years. Individualized screening algorithms have to be improved to provide the strategy for this issue.

11.1 INTRODUCTION

The European Randomized Study of Screening for Prostate Cancer (ERSPC) has shown that screening reduces mortality from prostate cancer (PCa)[98, 165, 218]. Experience from cervical and breast cancer screening programmes has demonstrated that the programme effectiveness can differ between centres and countries[219-220]. One of the main criteria used in evaluating and comparing the effectiveness of a screening is the change in the rate of disease diagnosed at an advanced stage[221]. This is reflected in the incidence and tumour stage distribution of interval cancers among those screened in the years following a negative screening test and in the tumour stage distribution of cases diagnosed at second and subsequent screening rounds[222].

Within the ERSPC, the screening intervals differed between the participating centres. Most centres used a 4-year screening interval while in Gothenburg, Sweden, a 2-year interval was used[98]. A previous study showed no difference in the rate of aggressive interval cancers between the Gothenburg (Sweden) and Rotterdam (the Netherlands) study[223]. However, the results might be inconclusive, since breast cancer screening studies have indicated that caution is required when comparing the absolute interval cancer rates in different countries[220]. Therefore, the proportionate interval cancer incidence is defined as the measure of choice for between-programme comparisons since it allows adjustment for geographical differences in cancer incidence[221, 224].

11.2 MATERIAL AND METHODS

11.2.1 Characterization of the Study Population

Between December 1993 and December 1999 a total of 62,322 men, aged 50-74 year, were randomized, 42,376 men, age 55-74, in Rotterdam and 9,973 men, age 50-64, in Gothenburg. To obtain a similar age distribution between the two centers, only men who were aged 55-64 years at start screening in both study arms, were included in this study. The procedure of recruitment and randomization differed among the centres. In Gothenburg, men underwent randomization before written informed consent was provided (population-based effectiveness trial), in Rotterdam men gave informed consent before randomization (efficacy trial).

11.2.2 Screening

The screening protocol is described in detail in Schröder et al[172]. During the whole study period, in Gothenburg men were screened with an interval of two years until an upper age limit of 67-71 (mean 69) years, in contrast to men in Rotterdam who were screened with an interval of four years until an upper age limit of 71-74 (mean 73). Interval cancers

were defined as any case of PCa diagnosed up to 24 months (Gothenburg) or 48 months (Rotterdam) after participation in the first or a subsequent screening round. Cancers diagnosed in men after they refused a scheduled screening round or a biopsywere not considered as interval cancers. All cancers were classified according to the 1992 TNM classification. A potentially life-threatening advanced cancer was defined as a PCa with one of the following characteristics: clinical stage ≥T3a, N1 or M1, or a serum PSA >20.0ng/ml or Gleason score ≥8 at biopsy. A low risk PCa was defined as cancer with clinical stage T1c, Gleason score ≤6 and PSA ≤ 10.0 ng/ml at diagnosis.

11.2.3 The underlying rate of prostate cancer incidence

The underlying rate of PCa incidence was obtained from the population participating in the control arm of the study centre. This allowed an age, time period, and tumour stage specific estimate for the underlying incidence.

11.3 STATISTICAL ANALYSIS

The time of follow up was measured from the date of first screening up to either the date of diagnosis, date of death or December 31, 2008. To achieve a comparable length of observation in the two study centers, the maximum length of follow-up was 12 years for every individual. Men with a longer follow-up were censored at 12 years after the date of first screening. The definitions of terms used in this study are summarized in Table 1. The following outcomes were calculated from the prevalence screen and subsequent screening rounds: the number of (advanced) detected cancers (including both screen-detected and interval cancers) and the proportional incidence of (advanced) cancers. The proportional (advanced) cancer incidence was defined as the ratio of the observed number of cancers (including both screen-detected and interval cancers) to the expected number of cancers among survivors in the screening group during the period of observation. The number of expected (advanced) cancer was estimated by multiplying age and calendar period specific (advanced) PCa rates of the control population to the age and calendar period specific person-years generated in the screening group. This allowed an estimate of the baseline PCa risk in Sweden and the Netherlands and thus adjustment for the difference in the baseline risk between the two countries[225]. Contamination (i.e. PSA-testing in the control population) was assumed to be similar in the two study centres.

Confidence intervals of proportional incidence ratio's were calculated assuming that the observed number of events followed a Poisson distribution with the logarithm of the observed number of life years as offset term[226]. A Poisson regression model with the logarithm of the expected number of cancers as offset term was used to estimate rela-

Table 1. Definitions of terms used in this study

Term	Definition
Advanced cancer	PC with one of the following characteristics: ≥T3a, N1, M1, PSA >20 ng/ml, or Gleason score ≥8
Interval cancer	PC diagnosed in a man who had a screening test with or without further assessment, which was negative for PC, either before the next invitation to screening, or within a time period equal to the screening interval in case the man had reached the upper age limit for screening.
Prevalence screen-detected cancer	PC diagnosed in a man who had at first invitation a screening test with further assessment which was positive for PC.
Proportional (advanced) cancer incidence	The ratio of the observed number of (advanced) cancers and the for age and country specific expected number of (advanced) cancers in the absence of screening.
Screen detected cancer	PC diagnosed in a man who had at second or subsequent round a screening test with further assessment which was positive for PC.

tive risks (RRs) of the centre of Rotterdam versus the centre of Gothenburg by modality of cancer occurrence and aggressiveness. All statistical tests were two sided. *P* values less than 0.05 were considered statistically significant. Analyses were performed with STATA package: Data Analysis and Statistical Software, version 11.0.

11.4 RESULTS

After age selection, i.e., of men who were aged 55–64 years at the time of first screening, 4202 men from the intervention arm in Gothenburg and 13301 men from the intervention arm in Rotterdam were included. During the study period (up to December 31, 2008), men in Gothenburg had a maximum of six screenings with a median follow-up of 12.0 years. Men in Rotterdam had a maximum of three screening visits with a median follow-up time of 11.2 years. A total of 5950 men, age 55-64 years at randomization, were included in the control arm in Gothenburg and 13966 men, age 55-64 years at randomization, were included in the control arm in Rotterdam. Based on these data, the expected incidence rate for (advanced) PC was calculated for each year after randomization.

Outcomes of the proportional advanced cancer and low risk cancer incidence are shown in Table 2,3 and 4. The data relate to the intervention (observed data) and control (expected data) population of the ERSPC Gothenburg and Rotterdam. During round 2-6 the proportional screen detected cancer incidence was RR 3.64 (95%CI 2.92-4.53) in Gothenburg and RR 3.08 (95%CI 2.67-3.55) in Rotterdam. During round 2-6 the proportional advanced cancer incidence was RR 0.40 (95%CI 0.22-0.71) in Gothenburg and RR 0.69 (95%CI 0.50-0.96) in Rotterdam. In Table 5 the differences in program effectiveness are presented comparing the Gothenburg and Rotterdam centre. A statistically

Table 2. Number of men screened (NS), number cancers detected (DNC), observed person years (PY), Expected number cancers (ENC), proportional number detected cancers (RR).

	NS	DNC	PY	ENC	RR	95% CI
Gothenburg, screen round						
Screen 1	4202	130	8102	20	6.50	4.06-10.41
Screen 2 – 6	9487	371	18115	102	3.64	2.92-4.53
Rotterdam, screen round						
Screen 1	13301	544	50418	142	3.83	3.18-4.61
Screen 2 – 3	17540	780	66138	253	3.08	2.67-3.55

Table 3. Number of men screened (NS), number advanced detected cancers (DNC), observed person years (PY), Expected number advanced cancers (ENC), proportional number advanced detected cancers (RR).

	NS	DNC	PY	ENC	RR	95% CI
Gothenburg, screen round						
Screen 1	4202	24	8102	11	2.14	1.07-4.45
Screen 2 - 6	9487	16	18115	40	0.40	0.22-0.71
Rotterdam, screen round						
Screen 1	13301	165	50418	54	3.05	2.25-4.15
Screen 2 - 3	17540	64	66138	92	0.69	0.50-0.96

Table 4. Number of men screened (NS), number low risk detected cancers (DNC), observed person years (PY), Expected number low risk cancers (ENC), proportional low risk detected cancers (RR).

	NS	DNC	PY	ENC	RR	95% CI
Gothenburg, screen round						
Screen 1	4202	66	8102	5	13.20	5.32-32.76
Screen 2 - 6	9487	236	18115	25	9.44	6.25-14.26
Rotterdam, screen round						
Screen 1	13301	270	50418	41	6.58	4.74-9.15
Screen 2 - 3	17540	472	66138	73	6.47	5.05-8.27

Table 5. Relative risk and 95% confidence intervals (CI) for the proportional (advanced) cancers in the Gothenburg (Goth) study relative to the Rotterdam (Rott) study

Cancers, round	Relative Risk (RR)	95% CI	P- value
	Goth. vs. Rott.		
Total cancer incidence, round 2 – 6	1.18	1.04-1.33	0.009
Advanced cancer incidence, round 2 – 6	0.57	0.33-0.99	0.048
Low risk cancer incidence, round 2 – 6	1.46	1.25-1.71	< 0.001

significant reduction was observed in the proportional advanced cancer incidence after the first screening episode (round and interval 1) in favour of using a 2-year screening interval, RR 0.57 (95%CI 0.33-0.99; p=0.048). The likelihood for a PC to be diagnosed at repeated screening was increased using a 2-year relative to a 4-year interval, RR 1.18 (95%CI 1.04-1.33;p=0.009); for the detection of a low risk PCa the RR was 1.46 (95%CI 1.25-1.71; p<0.001).

11.5 DISCUSSION

In an intension to screen analysis, data on seven centres of the ERSPC showed that screening for PC produced a 20% relative reduction in PCa mortality, using a 4-year screening interval in 86% of the participants, nine years after the onset of screening[98]. The Gothenburg screening study, which is part of the ERSPC, showed that the relative reduction in PC mortality increased up to 44% fourteen years after onset of screening men every second year[218]. Based on the present study, the use of a shorter screening interval might have contributed to the relatively high reduction in PCa mortality. This study showed that screening with a 2-year relative to a 4-year interval significantly reduced the incidence of advanced cancer with 43% in men aged 55 to 64 years. Any difference in mortality between the two centers should be studied more thoroughly in further analyses when mortality data is available by an appropriate length in follow-up. At present, we can conclude that screening biannually decreases the risk of being diagnosed with advanced PCa, however, that the effect on PCa mortality remains uncertain.

Despite these observations, the optimal interval for PCa screening remained undefined. First, the effects of screening need also to be evaluated by means of an appropriate economic analysis. Biannually screening is likely to increase costs associated to the screening programme. In addition, simply shortening the screening interval would probably lead to more overdiagnosis. The likelihood for PCa diagnosed at repeated screening was increased using a 2-year compared to a 4-year interval, RR 1.18 (95%CI 1.04-1.33;p=0.009). Moreover, the proportional low risk PCa incidence was significant higher in the program using biannually screening, RR 1.46 (95%CI 1.25-1.71;p<0.001). Since overdiagnosis remains the key limitation in the implementation of population based PC screening, other strategies are needed to define the optimal screening interval. This is the first study that showed that some men might benefit from more frequent screening in terms of a reduced rate of advanced disease with potentially improved PC mortality outcomes. As a result, we should try to indentify the men who need a more frequent screening program, which make individual screening intervals based on patient's individual risks an urgent need. Studies have shown that a more individualized program,

that includes a combination of other diagnostics, is sensitive for the estimation of the future risk of PC[227-228]. Based on these studies, and studies that have related the interval to the initial PSA level, a candidate's individual screening interval might be defined[229-230]. For example, Roobol et al. concluded that the screening interval of men aged 55-65 years with a PSA of 1.0 ng/ml could be as long as 8 years with a minimal risk of missing aggressive PC at a curable stage[231]. Two more recent studies, that determined the potential benefit of screening stratified by the participant's initial PSA, concluded that men with an initial PSA ≤1.0 and <2.0 ng/ml are unlikely to benefit from repeated screening since the cancer in these men is not likely to become life threatening if they remain unscreened[232-233]. This issue should be studied more thoroughly in further clinical trials.

The present study has limitations. Due to the design of the study, the time of observation from screening round 2 until the end of observation differed between the two study centers, i.e., 10 years in Gothenburg and 8 years in Rotterdam. In addition, the screening tests differed among the two centers in part of the initial screening round. Men in Gothenburg were screened by PSA only whereas men in Rotterdam were screened by a combination of PSA, DRE and TRUS up to 1997[172]. Consequently, this difference might have influenced the outcomes. In Rotterdam a more intense screening strategy at first screening round was used, resulting in a PCa detection rate at first screening round of 3.9% in Rotterdam versus 2.8% in Gothenburg (p=0.0006). To assess the potential consequence of the difference in study design, we also studied the proportional cancer incidence during the first 4 years and last 8 years of observation in Gothenburg and Rotterdam. This study showed a proportional cancer incidence of 7.65 in Gothenburg and 3.83 in Rotterdam; RR 2.00 (95%CI 1.71-2.33;p<0.001), and a proportional advanced cancer incidence of 1.70 in Gothenburg and 3.05 in Rotterdam; RR 0.56 (95%CI 0.38-0.83;p=0.004) during the first four years after randomization. During the following eight years of observation, the proportional cancer incidence was 3.00 in Gothenburg and 3.08 in Rotterdam; RR 0.97 (95%CI 0.84-1.11;p=0.696) and the proportional advanced cancer incidence was 0.32 in Gothenburg and 0.69 in Rotterdam; RR 0.46 (95%CI 0.24-0.88;p=0.019). This alternative analysis would not have changed our conclusions. One additional key observation was made. After two screening rounds no increase in the overall detection of PCa was found comparing the 2-year with the 4 year screening interval. The second limitation of our study is the difference in randomization procedure in the two centres. In Gothenburg, consent was obtained after randomization, with a response and participation rate at first screening of 64%. In Rotterdam, consent was obtained before randomization, with a response rate of approximately 55%, but participation rate of men randomized was 95%. Thus noncompliant men were excluded before randomization in Rotterdam but not in the Gothenburg study. This so called healthy screenee bias, might have resulted in a relatively more favourable PCa distribution in the control population of the Rotterdam

study. Since in this study the proportional incidence of advanced interval cancers was used (a rate ratio related to the control population of the study) this potential bias could have overestimated the proportional incidence in favour of the Gothenburg study. Finally, contamination by PSA testing in the control population, and the percentage of men that had undergone PSA tests before randomization was assumed to be similar although no data was available to test this assumption.

11.6 CONCLUSIONS

This study compared the interval cancers in a population with a 2-year and a 4-year screening interval using the best possible control for sources of artefactual variation across the different populations. Both screening with a 2-year and a 4-year interval significantly reduced the risk of being diagnosed with advanced cancer. A 2-year screening interval significantly reduced the incidence of advanced PCa up to 43% relative to a program that uses a 4-year interval in men aged 55 to 64 years. However, the screening programs that use more frequent screening showed to be associated with increased numbers of investigations and with an increase in the incidence of low risk PCa up to 46%. Therefore the identification of men that can benefit from more frequent screening is needed to define the optimal patient-individual screening interval. Finally, the effects of screening need also to be evaluated by means of an appropriate economic analysis, not merely on evidence of effectiveness in terms of stage distribution and PC specific mortality, and needs to take into account the quality of life aspects of the screening program.


CRUISIN' FOR SCREENING
PROSTATE CANCER

12

Screening: should more biopsies be taken in larger prostates?

BJU Int. 2009 Oct;104(7):919-24.

Pim van Leeuwen, Roderick van den Bergh, Tineke Wolters, Fritz Schröder and Monique Roobol

ABSTRACT

Objectives: To assess the number of missed prostate cancers and the frequency of aggressive disease when taking lateralized sextant prostate biopsies, irrespective of the total prostate volume (Pvol), during screening for prostate cancer.

Subjects and methods: Men participating in the European Randomized Study of Screening for Prostate Cancer, Rotterdam section, aged 55–74 years, with a prostate-specific antigen (PSA) level of ≥ 3.0 ng/mL, and a negative sextant biopsy result at the initial screening round, were followed for 8 years. Cases of prostate cancer detected during the follow-up by screening, or detected clinically as interval cancers, were assessed. Pvol at the initial screening round was related to the number of cancers found during the follow-up. Furthermore, the frequency of aggressive cancer (N1 or M1, PSA > 20.0 ng/mL, Gleason > 7 as evaluated using multivariate logistic regression analysis, including age, PSA level and Pvol.

Results: In the total of 1305 men, 152 prostate cancers were detected during 8 years of follow-up (11.6%); 23 were classified as aggressive (15.1%), and 50 (32.9%) were detected as interval cancers. There was a significant relation between a larger Pvol at the initial screening round and fewer cancers (odds ratio 0.1, P <0.001). In multivariate logistic regression, the initial PSA level (odds ratio 3.21, 95% confidence interval, CI 1.2–8.3) and smaller Pvol (0.08, 95% CI 0.03–0.26) were statistically significant predictors for all cancers and aggressive cancers (PSA odds ratio 70.37, 95% CI 13.5–366.2; Pvol odds ratio 0.03, 95% CI 0.01–0.35).

Conclusions: Men with a smaller Pvol and an initially high PSA level were at greater risk of cancer detection and of an aggressive cancer during the follow-up. The use in clinical practice of volume-adjusted biopsy schemes should not be implemented automatically in screening programmes with repeated screening.

12.1 INTRODUCTION

PSA testing in combination with an efficient and accurate method of prostate biopsy represents the standard method for the early detection of prostate cancer. Systematic sextant biopsy was proposed and popularized by Hodge et al.[29] and has been the standard protocol for many years. Later studies applying extended biopsy protocols showed that the sextant biopsy misses 10–30% of biopsy-detectable cancers [234-235]. Prostate volume (Pvol) can affect the cancer detection rate, as a sextant biopsy might under-sample larger prostates[236-237].

Although the effectiveness of prostate cancer screening is not confirmed by level I evidence, screening is common and even recommended. Randomized multicentre studies are ongoing, including the European Randomized Study of Screening for Prostate Cancer (ERSPC), that is designed to study the effect of population-based screening for prostate cancer on prostate cancer mortality rates and quality of life and, if proven effective, to establish the optimum screening programme [238]. In the ERSPC section Rotterdam, men are screened with an interval of 4 years by PSA testing and, when the PSA level is ≥ 3 ng/mL, with a lateralized sextant biopsy, irrespective of the prostate volume. Obviously, due to this study protocol, biopsies are taken in a considerable number of men with an elevated PSA level produced by benign prostate enlargement. Furthermore, probably due to the standard use of the lateralized sextant biopsy scheme, a significant percentage of cancers remain undetected in men with large prostates. Consequently, these cancers might surface at repeated screening rounds, or as interval cancers with less favourable cancer characteristics.

We analysed whether men with larger prostates, who had a biopsy indication but with negative results at the initial screening round, were diagnosed with cancer more frequently as interval cancers or at subsequent screening rounds, and whether these had less favourable tumour characteristics. We then considered whether more biopsy cores should be taken in men with larger prostates, for adequate prostate cancer screening, or that other baseline variables might be more predictive for detecting aggressive prostate cancers.

12.2 SUBJECTS AND METHODS

The ERSPC was designed to study the effect of population-based screening for prostate cancer on prostate cancer mortality rates and quality of life. Between December 1993 and December 1999, in the Dutch centre of the ERSPC, 42 376 men aged 55–75 years

were randomized after providing informed consent. In all, 21 210 men were randomized to the intervention arm. Men with a previous diagnosis of prostate cancer were excluded. For the current analysis only those participants were included who were screened by the use of an exclusively PSA- based algorithm, between May 1997 and December 1999. In these men the PSA level as used as the only biopsy indication, applying a threshold of 3.0 ng/mL, irrespective of the findings on a DRE and TRUS. A detailed description of the methods of the screening trial is provided by Roobol et al.[96]. At 4 and 8 years after the initial screening, men were invited to undergo repeated screening, when aged 74 years. Cancers diagnosed clinically between the screening intervals or by opportunistic screening, TURP for benign disease, and cystoprostatectomy specimens (i.e. 'interval cancers') were also considered. The serum PSA level was measured using the Hybritech Tandem-E assay (Beckman-Coulter, San

Diego, CA, USA). Pvol and transition zone (TZ) volume (TZvol) were measured by TRUS

(model 1846 mainframe, Bruel & Kjaer, Glostrup, Denmark) and a 7-MHz biplanar endorectal transducer. The lateralized sextant biopsy consisted of right- and left- sided lateral cores from the base, mid and apex of the prostate. An additional core biopsy was taken from any suspicious area on TRUS.

All cancers were classified according to the primary TNM classification of 1992, and graded using the Gleason grading system. Men with T1c and a serum PSA level < 10 ng/mL were classified as M0, and men with a serum PSA level of ≥ 100 ng/mL were classified as M1, if an isotope bone scan was not taken. Aggressive cancer was defined as cancer that had at least one of the following characteristics at diagnosis: stage N1 or M1, serum PSA level > 20.0 ng/mL, or a Gleason score > 7.

For statistical analysis the Pvol at the initial screening round was stratified in groups (<40, 40–60, >60 mL) and related to the number of prostate cancers found at the second and third screening round, by screening or clinically, during the two intervals (0–4 and 4–8 years). These results were analysed using the chi-square test and two-sided Fisher's exact test. For additional statistical comparisons Pvol, TZvol and PSA values were logarithmically transformed to obtain normal distributions. Subsequently, the total number of prostate cancers detected at, or before, the second and third screening round were statistically analysed using univariate logistic regression, to assess the correlation with the continuum of Pvol and TZvol at the initial screening round. In addition, multivariate logistic regression analyses were used to assess all screen and clinically detected cancers during the 8 years of follow up, including the initial screening round variables of log PSA value, age and log Pvol, using a backward stepwise method with variables rejected at a $P < 0.05$. For all tests, a two- sided $P < 0.05$ was considered to indicate statistical significance.

12.3 RESULTS

During the first screening round, 10 754 men were invited for screening using a PSA level of ≥3.0 ng/mL as a biopsy indication, between May 1997 and December 1999; 10 191 (94.8%) men actually had a PSA test. In all 2146 (21.1%) men had a PSA level of ≥3.0 ng/mL, of whom 1849 (86.2%) were biopsied. Prostate cancer was detected in 541 (29.3%) men. Consequently, 1305 men, who all had a benign biopsy result and in whom the Pvol was measured, were included in the present study for further evaluation at repeated screening rounds. The baseline patient characteristics of these 1305 men are shown in Table 1. Figure 1 presents a flow diagram of the study including the following 8 years.

Table 1 The baseline characteristics of the 1305 subjects

Variable	Median (range) or n (%)
Age, years	65 (55-74)
55-59	227 (17.4)
60-64	358 (27.4)
65-69	396 (30.3)
70-75	324 (24.8)
PSA, ng/mL	4.0 (3.0-46.7)
Pvol, mL	47.4 (13.1-207.7)
00.0 – 40.0	408 (31.3)
40.1 – 60.0	565 (43.3)
> 60.0	332 (25.4)
Tzvol, mL	29.1 (4.5-182.9)
00.0 – 20.0	305 (23.4)
20.1 – 30.0	372 (28.5)
> 30.0	618 (47.4)
Unknown/not done	10 (0.1)

At the second screening round after 4 years, 976 (74.8%) participants were invited between July 2001 and October 2004; 734 (75.2%) men had a PSA test, of whom 628 (85.6%) had a serum PSA level of ≥3.0 ng/mL. Overall, 581 (92.5%) men were biopsied, of whom 62 (10.7%) had prostate cancer. In addition, 26 interval cancers were detected between the first and second screening round. Of all 88 cancers the median Pvol was 43.4 mL at the initial screening round, by contrast with an initial median Pvol of 47.4 mL in all men with a benign biopsy result at the second screening round. Of 88 cancers, 12 (13.6%) were aggressive; the total number of cancers and number of aggressive cancers in relation to Pvol the initial screening round for 1246 patients (who were alive at the second screening) are shown in Table 2. Statistically more cancers were detected in smaller prostates, but there was no significant difference for the total number of ag-

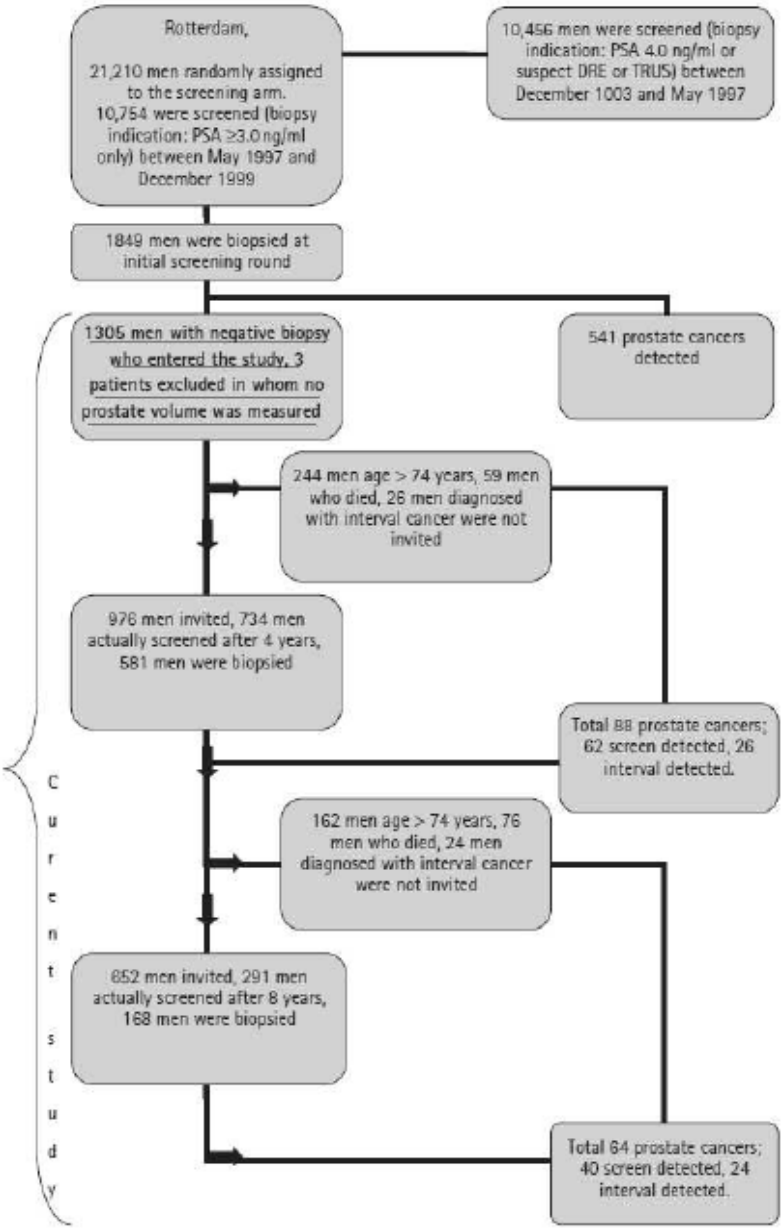


Figure 1. A flow diagram of the study.

gressive cancers and Pvol stratified in groups of < 40 mL and > 40 mL (Fisher's exact two-sided, P = 0.545).

The total number of detected prostate cancers in relation to the total Pvol and TZvol (continues variables) was analysed using univariate logistic regression. This showed a

Table 2 Screen- and interval-detected prostate cancer during the First and second 4 years of follow-up

	Pvol, ml				
	0-40	40.1-60	>40	>60	Total
First 4 years					
Cancer, n (%)	37 (9.5)	36 (6.7)	-	15 (4.7)	88 (7.1)
No cancer, n (%)	351 (90.5	505 (93.3)	-	301 (95.3)	1157 (92.9)
P	<0.005				
Aggressive cancer, n (%)	4 (1.0)	-	8 (0.9)	-	12 (1.0)
No aggressive cancer / no cancer	384 (99.0)	-	849 (99.1)	-	1233 (99.0)
P	0.545				
Second 4 years					
Cancer, n (%)	31 (9.4)	23 (4.9)	-	10 (3.5)	64 (5.9)
No cancer, n (%)	300 (90.6)	443 (95.1)	-	275 (96.5)	1018 (94.1)
P	<0.005				
Aggressive cancer, n (%)	7 (2.1)	-	4 (0.5)	-	11 (1.0)
No aggressive cancer / no cancer	324 (97.9)	-	747 (99.5)	-	1071 (99.0)
P	<0.005				

Agressive prostate cancer; one of the following characteristics at diagnosis: stage M1 or N1, serum PSA level >20.0 ng/ml, or a Gleason score >7

significant relation for the total number of prostate cancers with the smaller Pvol (odds ratio 0.217, P = 0.02) and with TZvol (0.327, P = 0.02) at the initial screening round. There was no relation for the total number of aggressive prostate cancers with Pvol or TZvol at the initial screening round (P = 0.915 and 0.891, respectively).

At third screening round (4 years later), 652 (71.3%) men were invited for a PSA testbetween July 2005 and March 2008, and 291 (44.6%) were screened by PSA testing, of whom 229 (78.7%) had a serum PSA level of ≥3.0 ng/mL. A biopsy was taken in 168 (73.4%) men, of whom 40 were diagnosed with prostate cancer (23.8%). In 24 men an interval cancer was detected during the second 4 years of follow-up. Of all 64 cancers the median Pvol was 40.0 mL at the initial screening round, by contrast with a median Pvol of 47.6 mL in men with a benign biopsy result at the third screening round. Of all 64 cancers, 11 (17%) were aggressive. After stratifying Pvol as in the previously defined groups, there was a significant relation between the total number of prostate cancers and smaller Pvol at the initial screening round (Table 2). Subsequently, there was a significant relation for the total number of aggressive cancers and Pvol (Table 2). The predictive value of Pvol and TZvol (continuous variables) for detecting prostate cancer was assessed using univariate logistic regression. Men with, at the initial screening round, a small Pvol and/or small Tzvol were more often diagnosed with prostate cancer during the follow-up than men with high Pvol (odds ratio 0.05, P < 0.001) and TZvol (0.10, P < 0.001). However, there was no significant relation for aggressive prostate

Table 3 Predictive value of different predictors of prostate cancer and aggressive prostate cancer in multivariate analyses

	Multivariate analyses, odds ratio (95% CI), *P*	
Variable	All cancers	Aggressive cancers
Log Pvol	0.085 (0.028-0.259), <0.001	0.032 (0.003-0.345), 0.004
Log initial PSA	3.213 (1.242-8.313, 0.016	70.367 (13.522-366.191), <0.001
Age	0.957 (0.925-0.990), 0.011	1.010 (0.930-1.096), 0.817

cancer with the continuum of Pvol and TZvol at the initial screening round (P = 0.066 and 0.093, respectively).

Table 3 shows the results of the multivariate logistic regression analysis, with serum PSA level, Pvol and age at the initial screening round included. All 152 prostate cancers, including 23 aggressive prostate cancers, all detected up to 8 years after the initial screening round, were included. A small Pvol, lower age and an increased baseline PSA level at the initial screening round were significant predictors for cancer (Table 3). In addition, a multivariate logistic regression analysis, with initial PSA, Pvol and age included, showed that an increased baseline PSA level and smaller Pvol were significantly predictive for detecting aggressive prostate cancer during the 8-year follow-up (Table 3).

12.4 DISCUSSION

In this study we showed that men with a smaller Pvol, who had, at the initial screening round, an indication for biopsy and a negative biopsy result, were at greater risk of being diagnosed with prostate cancer, and of aggressive prostate cancer, during the 8 years of follow-up. This was in contrast with our intuitive assumption which implied that using lateralized sextant biopsies irrespective of Pvol would result in more missed cancers in larger prostates, due to under-sampling. Our initial theory supported the need to obtain more biopsies because taking only six biopsy cores in men with large prostates might miss many aggressive cancers, and possibly affect the prostate cancer mortality rate. In line with this, the ideal endpoint for the present study would be the prostate cancer mortality rate. However, as no men in the current study died from prostate cancer until December 2006, the number of detected aggressive prostate cancers for different Pvol replaced the prostate cancer mortality rate as a surrogate endpoint.

In all, 23 aggressive prostate cancers were screen- or clinically detected during the 8 years of follow-up. Stratifying the initial Pvol by groups with < 40 and > 40 mL resulted in 11 (2.7%) of 408 men and 12 (1.3%) of 897 men who were detected with an aggres-

sive prostate cancer, respectively. This resulted in an increased risk for men with smaller prostates of developing a clinically significant cancer with an unfavourable perspective. Nevertheless, this relative risk stratification could be influenced by the unequal Pvol distribution over the different volume groups, as 68.7% of the men had a Pvol of > 40 mL at the initial screening round. Therefore, the percentage of aggressive cancers detected in larger prostates might be lower than in small glands because larger prostates are more often biopsied in previous screening rounds secondary to the increased serum PSA level that might be produced by benign prostate tissue. Consequently, men with smaller prostates and a PSA level of ≥3.0 ng/mL at the previous round were a priori at increased risk of cancer, as their elevated PSA level might not be produced by benign prostate tissue. However, even when the number of patients was normally distributed over Pvol groups of < 40 and > 40 mL, and the total number of detected aggressive prostate cancers was unaffected, there would be no significantly higher risk for larger Pvol, as in all, 11 and 12 aggressive prostate cancers were detected in men with a Pvol of < 40 and > 40 mL.

In ine with our initial hypothesis, it might be possible that better sampling of smaller prostates influenced the number of detected cancers. In addition, at the second and third screening round, lateralized sextant biopsies were taken irrespective of Pvol. For this reason we made a subanalysis including only the 'clinically' detected interval cancers during the 8 years of follow-up. In this analysis, eight (2.0%) of 408 men with an initial Pvol of < 40 mL, and 10 (1.1%) of 897 with a Pvol of > 40 mL had an aggressive prostate cancer. Consequently, this subanalysis showed no increased relative risk of an aggressive prostate cancer diagnosis for men with larger prostates.

By contrast with our initial theory, the current results have no implication for changing the study protocol by taking more biopsy cores in men with larger prostates. After 8 years of follow-up, more aggressive tumours were detected in small prostates than in larger prostates. These results were confirmed in multivariate analyses including age, baseline PSA level and Pvol. In addition, in multivariate analyses the baseline serum PSA level was a significant predictor for detecting aggressive prostate cancer. This implies that if there should be an adjustment of the study protocol, patients with smaller prostates and a significantly increased baseline serum PSA level with a benign biopsy result at the initial screening should be re-screened after a shorter interval. However, the current study design does not have the power to introduce this adjustment to the study protocol. Also, these results suggest that volume-adjusted biopsy schemes should not be implemented automatically in screening programmes with repeated screening rounds. Because men with larger prostates are at greater risk of a benign elevated serum PSA level, the introduction of a volume-based biopsy scheme validated in the clinical setting

into a prostate cancer screening programme will result in increasing over-diagnosis. Consequently, as over- diagnosis is one of the major concerns of prostate cancer screening, an increase in over-diagnosis could interfere with the effectiveness and public acceptance of a screening programme. For this reason there is a need for an adjusted prostate biopsy scheme particularly developed for a systematic screening situation, which might differ from clinical schemes used routinely.

The present study results are in line with those from studies that have focused on the relation of tumour grade and Pvol [239-241]. These studies were published after the results of the Prostate Cancer Prevention Trial, showing that men with a smaller Pvol, treated with finasteride, were 25% less likely to be diagnosed with prostate cancer but 67% more likely to have high-grade disease (Gleason ≥7)[72]. Although these observations might be attributable to a biopsy artefact of gland size, whereby small glands are relatively oversampled, further studies showed that gland size not only affects the detection rate, but is also associated with the distribution of Gleason score. Three studies taken together all indicate that prostate cancer detection is greater in small glands and at increased risk of clinically significant upgrading after radical prostatectomy[239-241].

The present study included a follow-up of 8 years. Possibly more (aggressive) prostate cancers will be detected after a longer follow-up. Consequently this could change the conclusions obtained from present study results in the future. However, the results show that during the more recent years of observation, the interval-detected aggressive prostate cancers were mostly diagnosed in small prostates. This makes it unlikely that a larger number of aggressive prostate cancers in larger prostates might be expected in the following years. Moreover, based on the estimated lead-time, most missed aggressive prostate cancers at the initial screening round would be clinically detected during 8 years of follow-up[129, 242].

The number of biopsy cores taken affects the cancer detection rate. Several groups concluded that the lateralized sextant biopsy method under-samples prostates and consequently might fail to detect a significant proportion of clinically important tumours [243-244]. Recently, Schröder et al. [170] reviewed seven studies that compared the percentages of cancers detected in the clinical setting by extended and sextant biopsy schemes. This review reported that classical sextant and lateralized sextant biopsies would have missed 23% and 19% of the detectable cancers. It was also confirmed that a volume-adjusted, increased-core regimen significantly increases the positive biopsy rate, with no significant increase in morbidity of the procedure [245]. The most comprehensive data of the European Prostate Cancer Detection Study resulted in the Vienna nomogram, and it was concluded that 8–18 biopsy cores should be taken, based on Pvol and age, to en-

sure a 90% certainty of cancer detection [246]. However, the optimum number of cores that should be taken as an initial strategy in a screening setting remains a controversial topic, especially for prostate cancer screening, as over-diagnosis is one of the major concerns of screening. The present study does not aim to provide an answer to this question.

In conclusion, currently Pvol-based biopsy schemes are routinely used in clinical practice to assure adequate sampling. However, the present study does not support the implementation of these same volume-based schemes for a screening setting. We found that when applying a volume-independent lateralized sextant biopsy scheme in a prostate cancer screening programme with repeated screenings, large prostates are not undersampled with respect to the number and aggressiveness of tumours found, compared to small prostates, during the 8 years of follow-up. The total number of cancers and number of aggressive cancers was higher in small prostates. For prostate cancer screening, the question of the validity of a sextant biopsy regimen for the balance between missed aggressive cancers and overdiagnosis remains.


CRUISIN' FOR SCREENING
PROSTATE CANCER

13

Prostate cancer mortality in men participating in screening: risk analysis of men managed by radical surgery for clinically localized prostate cancer

Submitted

Pim J. van Leeuwen, Roderick C.N. van den Bergh, Teuvo L.J. Tammela, M. Kwiatkowski, L. Denis, Chris H. Bangma, S. Ciatto, Jonas Hugosson, Fritz H. Schröder.

ABSTRACT

Background: Competing causes of deaths in men with screen detected prostate cancer (PCa) are poor defined. Results of clinical diagnosed cases are inaccurate for predicting the outcomes since the survival of screen detected cancers are in principal lengthened due to an earlier diagnosis.

Objectives: Estimation of the survival based on a competing risk analysis stratified by age, stage, serum Prostate Specific Antigen (PSA) and histological findings for men diagnosed with clinically localized PCa who were managed by radical surgery.
Design, Settings and Participants: Total of 2234 participants of the European Randomized Study of Screening for Prostate Cancer (ERSPC), aged 55 to 74 years, were diagnosed due to screening between 1993 and 2000 with PCa (staged <T2c, not N1, not M1 and PSA <20.0 ng/ml). 990 of these men were treated with radical surgery and followed until the end of 2008 for overall and disease specific mortality.

Measurements: Estimates of the probability of dying from PCa or other causes of death.
Results and limitations: Median follow-up was 9.9 years. 168 men died during follow-up, of which 23 due to PCa. Men with tumours with Gleason scores of ≤6, 7 and ≥8, face a 2.1%, 4.2% or 16.1% probability, respectively, of dying from PCa within 15 years of diagnosis. The main limitations of this study were the small number of PCa deaths. The risk of dying from Pca relative to the risk of dying from other causes was higher for the younger ages and advanced Gleason scores.

Conclusions: Men with a clinically localised PCa (PSA <20.0 ng/ml), treated by radical surgery, have a very heterogeneous risk of dying either from PCa or other causes depending on their Gleason score, stage and age. These risk estimates might be used as preoperative information and counselling for patients diagnosed with a localized PCa due to screening.

13.1 INTRODUCTION

Prostate cancer (PCa) is the most common form of cancer affecting men in the Western world[126, 247]. The International Agency for Research on Cancer projected that in 2008 there were 382,251 new cases of PCa (22% of all adult male cancers) and 89,319 deaths (9% of all male cancer deaths) from PCa, resulting in an incidence/mortality ratio of more than 4:1 [126]. These numbers highlight the difference between the number of men who are diagnosed with PCa and the number of men who actually die from PCa. This is mainly due to the increased detection by Prostate Specific Antigen (PSA) testing, which started in the early nineties. PSA testing has resulted in PCa being diagnosed at an earlier age, when it is at a relatively earlier stage, and lower tumour grade and serum PSA. Many of these patients are treated with radical surgery; in the United States and Western European countries between 45% and 50% of patients with localized PCa are currently treated by radical prostatectomy[248-249]. Overall, these patients have a competing risk of dying either from PCa or other causes, depending on the aggressiveness and stage of the disease, co-morbidity and age.

This study was designed to estimate survival for men diagnosed with screen-detected clinically localized PCa and treated with radical surgery. The primary objective of the analysis was to estimate the probability of dying from PCa or other causes after radical treatment given a patient's tumour histology, stage, PSA and age at diagnosis in the PSA era.

13.2 METHODS

All men included in this study were participants of the European Randomized Study of Screening for Prostate Cancer (ERSPC). The study was designed to study the effect of population-based screening for PCa on mortality and quality of life[68]. The study started in 1993 and randomized a total of 267,994 men in the age range of 50–74 years.

The inclusion criteria for the current study were: (1) randomization to the intervention arm; (2) diagnosis with screen detected clinically localised PCa between 1993 and 2000, aged 55-74 years; (3) treated by radical surgery. Patients undergoing a combination of radical surgery with radiation therapy, or patients initially followed by an active surveillance program or treated initially with endocrine treatment were excluded to evaluate the outcomes among men treated by radical prostatectomy that were highly comparable among each other.

13.2.1 Clinical information

All cancers were classified according to the TNM classification of 1992. All men with a clinical PCa ≥cT2c and/or cN1 and/or cM1 disease or with a PSA ≥20 ng/ml at diagnosis were excluded. We defined these inclusion criteria to estimate the survival of men with PCa after screening in a low and favourable clinical stage, presuming that these cancers are currently most often diagnosed in communities with a high frequency of PSA testing[248]. Men with T1c disease and serum PSA concentration less than 10.0 ng/ml were classified as M0, even through a bone scan was not performed. Men with PSA 10-20 ng/ml and stage NX or MX were not excluded.

13.2.2 Pathological Evaluation

The pathological information was based on the prostate biopsy specimens. Due to the ERSPC screening protocol, most information was based on lateralized sextant biopsies. Grading of the cancers was done using the Gleason grading system. Standardization of pathology procedures was coordinated and achieved by the work of the international pathology committee[98].

13.2.3 Treatment procedure

For all patients the treatment of PCa was performed according to the local policies and guidelines, in high and low volume centres. Radical surgery was performed by an open or a (robotic assisted) laparoscopic procedure. Comparing outcomes of different surgical techniques was not the aim of this study.

13.2.4 Mortality data

Mortality data of participants who died in the period up to December 31, 2008 were obtained by linking the trial database with the National Causes of Death Registry. Linkage to the National Causes of Death Registry was possible by using the personal administrative number of each participant as a linkage key. Causes of death were evaluated in a blinded fashion and according to a standard algorithm or, after validation, on the basis of official causes of death[153]. Patients were determined to have died from PCa if they were classified as either "definitely PCa death", as "probable PCa death" or as "PCa intervention related death"

13.3 STATISTICAL METHODS

Follow-up for mortality analyses began at diagnosis and ended at death, emigration, or a uniform censoring date (December 31, 2008). The primary outcomes for this study were estimates of the probability of dying from PCa or other competing causes given a

patient's age, PSA and tumour histology grade at diagnosis. The counts and cumulative survival estimates of men with each of the 3 outcomes of interest (alive, deceased from PCa, and deceased from other causes) are presented. To estimate the proportions of men who died from PCa, died from competing medical hazards, or were still alive 15 years following diagnosis, we applied the 2 fitted rates to the proportion of men still alive at the beginning of each successive follow-up interval. Therefore, the following model was used: $Log\ [E(Y)] = log(exp) + \beta_0 + \beta_1 x_1 + ... + \beta_5 x_5$; a generalized linear model with log link function and Poisson distributed errors where $E(Y)$ is expected number of PCa deaths, $log(exp)$ is the logarithm of the follow up time, $(x_1, x_2, ..., x_p)^T$ are the predictive variables, i.e. age, Gleason score, serum PSA and the time since diagnosis. The βi is the coefficient corresponding to x_i. The term $log(exp)$ was an offset with the parameter estimate constrained to 1. All analyses were performed with STATA: Data Analysis and Statistical Software, version 11.0.

13.4 RESULTS

13.4.1 Study population

Between 1993 and 2000, a total of 2234 men were diagnosed with a histological proven localized PCa (clinical staged ≤T2c, NO, M0 and PSA < 20 ng/ml) in the intervention arm of the ERSPC. From all patients complete PSA, stage and pathology information was available at diagnosis. The primary treatment modalities are shown in Table 1. A total of 990 men were treated with radical surgery. A detailed description of these men is presented in Table 1. Median follow-up was 9.9 years (inter quartile range = 8.9-11.2 years). A total of 23 men died from a PCa related cause of death, from which 4 cases were assessed as intervention related death. Men of older ages were at increased risk of an intervention related cause of death.

Table 1. Initial treatment modalities prostate cancer <T2c, not N1, not M1 and PSA <20.0 ng/ml)

Total number patients diagnosed	2234
Surgery	990 (44.3)
Radiotherapy	535 (24.0)
Active surveillance	506 (22.6)
Endocrine	59 (2.7)
Surgery and Radiotherapy	3 (0.1)
Surgery + Endocrine treatment	41(1.8)
Radiotherapy + Endocrine	75 (3.4)
Other or Unknown	25 (1.1)

13.4.2 Relative risk prostate cancer mortality

In multivariate analysis the relative risk of PCa specific death for patients with Gleason score ≥8 was statistically significant increased relative to patients with a Gleason score ≤6 tumours; RR 7.22 (95%CI 2.72- 19.15, p <0.001) (Table 2). No statistically significant difference was observed in the RR of dying from PCa for men with PSA 10-20.0 ng/ml relative to men with PSA < 10.0 ng/ml at diagnosis, RR 0.77 (95%CI 0.23- 2.61, p=0.676).

Table 2. Patient characteristics at diagnosis

Total participants included	990
Age (yr), median	63.0
55 - 64	616 (62.2)
65 – 74	374 (37.8)
PSA at diagnosis (ng/ml), median	5.3
0.0 - 9.9	856 (86.5)
10.0 - 20.0	134 (13.5)
Clinical stage	
T1	512 (51.7)
T2	478 (48.3)
N0	583 (54.4)
NX	452 (45.6)
M0	971 (98.1)
MX	19 (1.9)
Histological differentiation	
Gleason 2 – 6	751 (75.9)
Gleason 7	171 (17.2)
Gleason 8 – 10	68 (6.9)

13.4.3 Mortality Outcomes by age and Gleason score

The distribution and outcomes of 990 patients treated by radical surgery for localized PCa are presented in Table 2. The patients are stratified by age and the histology of the biopsy specimen classified according to the Gleason score. Estimates of the cumulated probability of dying from either PCa or other causes are presented as a 15-year outcome in Table 3 and as function of time since diagnosis in Figure 1.

For men aged 55-64 years at diagnosis the PCa specific 10-years cumulated mortality estimates were 1.1%, 2.3% and 13.0% for Gleason scores ≤6, 7 and ≥8 respectively. The 10-years PCa specific mortality estimates for men aged 65-74 years were 2.1%, 4.2% and 7.7% for Gleason scores ≤6, 7 and ≥8 respectively.

Table 3 and Figure 1 demonstrate that few men with Gleason score ≤6 tumours identified by prostate biopsy had progression leading to death from PCa; 1.2% and

Table 3. Multivariate relative risks prostate cancer specific mortality.

Characteristic	Relative Risk	95% CI	*P* - value
Age at diagnosis	1.01	0.91 – 1.12	0.818
PSA at diagnosis	1.00	0.90 – 1.11	0.993
Gleason score ≤ 6	*		
Gleason score 7	2.01	0.68 – 5.79	0.199
Gleason score ≥ 8	7.22	2.72 – 19.15	< 0.001

* Reference group to which compared. The reference group per definition has a relative risk of 1.

2.7% of men aged 55-64 and 65-74 years, respectively, died from PCa within 15 year of diagnosis (median 9.9 years). From all men aged 55-64 years at diagnosis that died within 15 years after diagnosis, 7.7%, 17.6% and 55.3% died from PCa if they harboured a disease with Gleason scores of ≤6, 7 and ≥8, respectively. For men aged 65-74 at diagnosis these percentages were 9.0%, 13.5% and 41.7% for Gleason scores ≤6, 7 and ≥8, respectively.

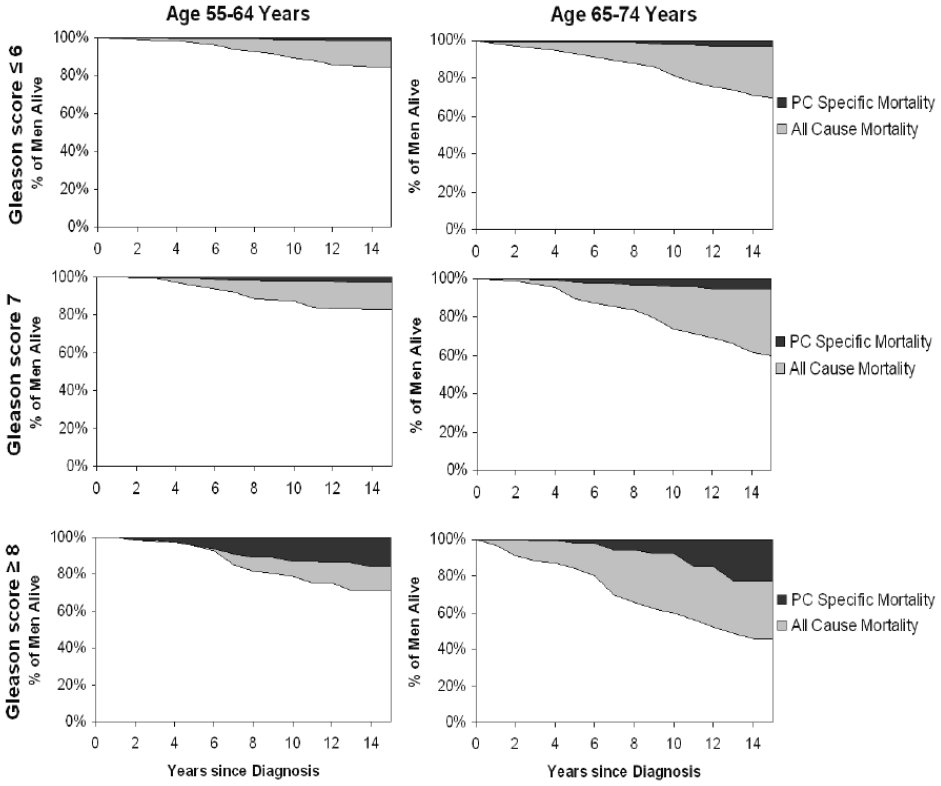


Figure 1. Estimated mortality from prostate cancer and other causes of death.

Table 4. Observed counts and estimated cumulative percentage of men deceased from either prostate cancer or an intercurrent cause of death by age and Gleason score up to 15 years after diagnosis (median 9.9 years).

Age 55-64 years at diagnosis					
Gleason Score	No. at risk (%)	No. Alive (%)	No. Deceased From Other Causes (%)	No. Deceased From Prostate Cancer (%)	% Prostate Cancer Deaths of all Deaths
≤ 6	495 (100)	437 (84.5)	52 (14.3)	6 (1.2)	7.7 %
7	88 (100)	76 (83.0)	10 (14.0)	2 (3.0)	17.6 %
8-10	33 (100)	25 (71.2)	4 (14.9)	4 (15.9)	55.5 %
All	616 (100)	538 (83.3)	66 (14.5)	12 (2.9)	16.7 %
Age 65-74 years at diagnosis					
Gleason Score	No. at risk (%)	No. Alive (%)	No. Deceased From Other Causes (%)	No. Deceased From Prostate Cancer (%)	
≤ 6	256 (100)	206 (70.0)	45 (27.3)	5 (2.7)	9.0 %
7	83 (100)	60 (60.0)	20 (34.5)	3 (5.5)	13.5 %
8-10	35 (100)	18 (46.0)	14 (31.5)	3 (22.5)	41.7 %
All	374	284 (58.3)	79 (36.2)	11 (5.5)	13.2 %

13.4 DISCUSSION

Accurate predictions of outcomes in cancer specific and overall survival are useful for patients and clinicians before choosing treatment. Currently, there is limited survival data that is appropriate for the men diagnosed with localized PCa as a result of screening. The reported outcomes are primarily of the patients diagnosed with PCa in the pre PSA era[55, 250-251]. These outcomes are inaccurate for predicting the survival of patients with a screen detected PCa since the survival outcomes of these cancers are in principal lengthened due to a diagnosis at an earlier age[252]. This is called the lead-time which for screen detected PCa is estimated between the 6 and 12 years depending on age at diagnosis[129]. Considering treatment of men with a localized screen detected PCa, we realize that more information is needed than the outcomes of one single treatment modality. In our opinion, only prospective randomized data can solve the controversy supporting surgery, radiation or observation. No randomized controlled trials have been performed to solve this dilemma for men with a screen detected localized PCa. These studies are ongoing, however results are not expected soon [253-254].

This study provides risk stratifications for men diagnosed with a screen detected localised PCa who consider treatment with radical surgery. It demonstrates a large heterogeneity among patients with PCa clinical stage <T2c and PSA <20.0 ng/ml. Men with a localised disease and a Gleason score ≤6 in their biopsy specimen harbour a 2.1% PCa mortality risk while this risk in men with a Gleason score ≥8 disease was estimated to be 16.1% after 15 years of observation. Our study showed that PCa remains a major cause

of death in particular in younger men with high-grade disease who were screened and treated with radical prostatectomy. Additionally, the relative contribution of the death unrelated to PCa itself to the overall mortality increased with increasing age at diagnosis and decreased with the Gleason scores.

More than an exponential increase in the risk of dying from PCa was observed for the increasing Gleason scores. This was reflected in the more than 2 fold higher RR for a man with a Gleason score ≥8 relative to a Gleason score 7, and the RR of 2.01 for a men with a Gleason score 7 relative to Gleason score ≤6, (Table 2). These observations are in line with studies showing that patients with a Gleason score 8-10 typically have a markedly higher PCa death rate compared with the lower scores[251, 255-256]. However, the difference in risk between the high and low Gleason scores might be larger in our population due to two effects.

First, screening detects a considerable part of insignificant cancers that would never have surfaced during the lifetime of a patient. In the current study these patients are included in the most favourable risk group, i.e., Gleason scores ≤6. As a result, the survival in this group is improved by an increase in the number of patients who would not have died of the disease in the first place, which highlights the possibility for active surveillance strategies (i.e., deferred curative treatment for low-risk prostate cancer until the perceived disease progression). Observational single centre and retrospective population-based cohort studies of active surveillance for localized prostate cancer have reported favourable outcomes, however studies that show longer follow up are needed[41-42, 257-258]. Nevertheless, for Gleason scores ≤6 PCa, we might conclude that cancer specific survival is very favourable, but that is in part due to a significant percentage of overtreatment.

Second, the increased PCa mortality risks for Gleason score 8 relative to Gleason score 6 diseases is probably due to the limitation in the clinical staging of men with high risk Gleason scores. The clinical stage in most of our patients was based on the serum PSA, the findings by DRE, transrectal ultrasound (TRUS), and an isotope bonescan when indicated. As a result, 41% of our patients with Gleason 8-10 disease and clinical stage <T2c, NO, MO and PSA < 20 ng/ml had extraprostatic disease (i.e., ≥pT3) in their radical prostatectomy specimen (data not shown). This percentage is line with the outcomes of the Partin tables, a predictive model that has been developed to estimate the probability of extraprostatic extension, seminal vesicle invasion and lymph node invasion[259]. After we entered the clinical data of our patients with a Gleason score 8-10 disease in the Partin tables, the estimated probability of extraprostatic extension was 38% (95%CI 30-47). In conclusion, a large percentage of patients with high Gleason scores might not harbour a localised disease, which we cannot assess due to the limitations of the currently used staging modalities.

The competing risks of dying from PCa in men with favourable Gleason scores are very low, i.e. 8% and 9% of all deaths for ages 55-64 and 65-74 years, respectively.

In total 4 of the 990 men (0.4%) who were treated with radical surgery were determined as intervention related death. This is in line with 30-day mortality after radical prostatectomy percentages presented in earlier series varying between 0.5% and 0.7%[260-261]. We observed no statistically significant association with age, although earlier series showed that radical prostatectomy complication rates were associated with increasing age and co-morbidity. Relative to other major surgeries in older ages[262], intervention related mortality after radical prostatectomy is low and should be low considering the risk of overdiagnosis and overtreatment that is associated PCa. Overdiagnosis is estimated between 23-66% in screen-detected PCa. Therefore, as there is no evident reason to assume a different risk of intervention related deaths in overdiagnosed or non-overdiagnosed cancers, roughly 1 or 2 deaths in our group might be ascribed to overtreatment. Consequently, the risk of death associated to overtreatment is roughly estimated between 0.1 and 0.2% in this patient cohort. Although these risks seem to be low, especially in older people these risks might be worth to discuss when considering a radical prostatectomy in a patient with favourable characteristics.

Most studies that reported outcomes after radical prostatectomy are limited by the inclusion of patients diagnosed before the introduction of PSA, or by the use of surrogate endpoints (e.g. biochemical recurrence) rather than PCa specific mortality[55, 250, 263-264]. In 2,578 men with localized PCa diagnosed and treated before the wide introduction of PSA testing, the 10-year cancer specific mortality was 6%, 20% and 23% for well, moderate and poorly differentiated cancer, respectively[265]. In 751 men diagnosed with clinically nonmetastatic PCa between 1971 and 1984 the 15-year cancer specific death was 4% to 5%, 8% to 9%, 12% to 13%, 23% to 26% and 30% to 32% for Gleason scores 2 to 4, 5, 6, 7 and 8 to 10, respectively[251]. So, the cancer specific mortality estimates in the present study compare favourable to earlier series across all Gleason scores. Corresponding 10- and 15-year data in the current study was 1.1% and 2.1%, 2.3% and 4.2%, 13.0% and 16.1% for the Gleason scores ≤6, 7 and ≥8, respectively. In comparison to the earlier series the percentage of cancer specific deaths was roughly 2 fold decreased for men with Gleason scores 8-10, 5 fold decreased for Gleason scores 7, and 6 fold decreased for Gleason scores ≤6. These results confirm that earlier studies are inaccurate for predicting the survival of patients with a screen detected PCa, that the survival of screen-detected cancers across all Gleason scores, especially lower scores, are lengthened due to a diagnosis at an earlier time, and that the survival is likely improved by detecting insignificant cancers that would not have been diagnosed in the absence of screening. To which extent the survival outcomes are diluted by these effects are impossible to establish in this study.

There are studies reporting the outcomes after radical prostatectomy in patients diagnosed after PSA testing. In a study by Stephenson et al. the overall 15-year PCa specific mortality was 12% for patients treated in the PSA era, and the 15-year PCa specific mortality for patients with a PSA <10 ng/ml, Gleason 6, T1c or T2a disease was 2%[256]. D'amico et al. reported the PCa specific mortality after radical prostatectomy for patients with clinically localised PCa diagnosed in the PSA era. The 10-year cumulative PCa specific mortality was < 1% for patients with a PSA < 10 ng/ml, Gleason score ≤ 6 and T1c or T2a disease, and 4% for patients with a PSA < 10-20 ng/ml, Gleason score 7 and T2b disease[255]. The additional findings were in line with ours. The relative contribution of deaths unrelated to PCa itself increased with advancing age, and high-grade PCa remained a major cause for patients who were treated with radical prostatectomy of all ages, but in particular the younger ages.

In comparison to studies [256] and [255], strength of our study is the prospective data collection and the PCa specific mortality that was based on the consensus of a cause of death committee. In addition, the current study did not include only patients treated at high-volume centres, and thus may presents an estimate for those patients who are diagnosed in the PSA era and are treated in the community setting. Finally, all men were diagnosed in a defined screening program.

Our study has limitations. The number of PCa specific deaths was small, and thus the survival estimates have to be interpreted with caution. Additionally, no difference in outcome between men treated by either an open or (robot assisted) laparoscopic radical prostatectomy was obtained. The histological information was most often based on sextant biopsies. However, sextant biopsies are related to a significant under grading. Therefore, more extended biopsy schemes are performed in current clinical practise. No data on co-morbidity was included although co-morbidity is a known strong determinant of survival among men with PCa. Despite these limitations, the results of this large prospective study of patients with localised PCa that were treated with radical prostatectomy provide useful information for patients and clinicians.

13.5 CONCLUSION

Men with a screen-detected clinically localised PCa, treated by radical surgery, have a very heterogeneous risk of dying either from PCa or other causes depending on their age and histology of the biopsy specimen. Men, aged 55- 74 years at diagnosis, with tumours that have Gleason scores of ≤6, 7 and ≥8, face a 2.1%, 4.2% or 16.1% probability, respectively, of dying from PCa within 15 years of diagnosis. The relative contribution of deaths unrelated to PCa itself increased with advancing age and decreasing Gleason scores. For men with high-grade PCa, Pca remained a major cause of death despite early

detection by screening followed by radical prostatectomy and screening. The present estimates on cancer specific mortality compare favourable to earlier series across all Gleason scores indicating that disease specific survival of screen-detected cancer is lengthened due diagnoses at an earlier point in time, and likely is increased by the treatment of men that likely would never have surfaced clinical symptoms of PCa in the absence of screening.


CRUISIN' FOR SCREENING
PROSTATE CANCER

Part V

General Discussion


CRUISIN' FOR SCREENING
PROSTATE CANCER

14

General discussion

Population Prostate Cancer Screening - Less Might Be More

Pim J. van Leeuwen

14.1 INTRODUCTION

Prostate cancer (PC) is a disease that most frequently occurs in elderly men, and there is a progressive increase in PC detection and mortality rates with age[7, 152, 266]. Opportunistic PSA testing has resulted in increasing PC incidence with patients diagnosed at an earlier age[8, 149]. Mortality from PC remains the most common in patients 70 years of age or older. In 2008, the median age at the time of death from PC in the Netherlands was 80 years[160]. Approximately, 3.5% of patients die before reaching the age of 60, 28.4% die between 60 and 74 years of age, and 68.1% die after the age of 75 years.

Demographic projections indicate that the number of men aged 65 years and older will increase from the current 2.5 million men to 4.5 million in the Netherlands during the next 30 years[176]. The factors contributing to this change include the low and high rates of fertility during certain eras (e.g., during the depression and after World War II), advancements in medicine and sanitation, and reduced mortality due to elimination of many infectious diseases across all age groups[267-268]. The aging of the Dutch population during the next 30 years will be driven by the "baby boom" cohort born between 1946 and 1964. In 2011, the oldest individuals in this cohort will turn 65. By 2040, the percentage of persons in the age group 65 years and older will increase from the current 15% to 26% of the population.

As PC is a disease that is associated with high mortality rates in elderly men, there is reason for concern regarding the estimated increase in the number of men aged 65 years and older. An increase of at least 51% in the number of new cases of PC is estimated by the year 2035. The total mortality rate from PC is estimated to increase by 131%, from 2050 to 5855 PC deaths (Figure 1)[269]. As a result, PC will become a major health problem during the upcoming decades. Screening for PC might be a strategy to prevent

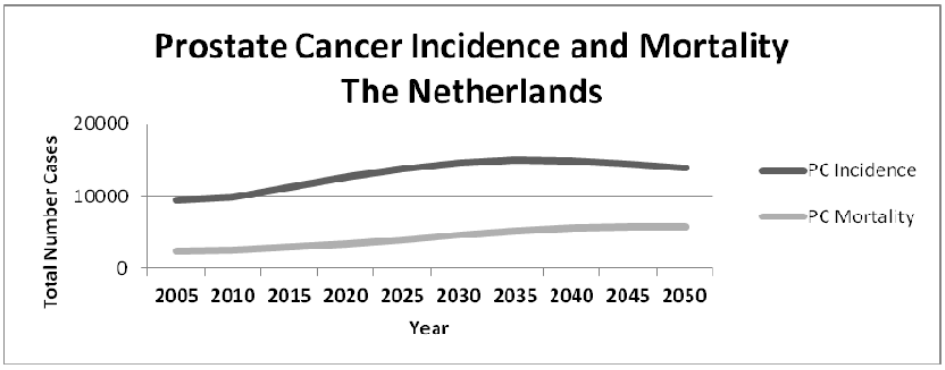


Figure 1. Total prostate cancer incidence and mortality estimates in the Netherlands between 2000 and 2050, obtained from [274].

the increase in PC morbidity and mortality. This, and the key findings presented in parts I-IV, will be the subject of the general discussion of this thesis.

14.2 PARTS I AND II

The results presented in chapters 5 and 6 will be discussed with respect to two large randomised controlled trials based on screening for PC. The European Randomised Study of Screening for Prostate Cancer (ERSPC) reported that screening for PC has the potential to reduce disease-specific mortality[172]. The Prostate, Lung, Colorectal and Ovarian (PLCO) trial in the United States did not show a reduction in PC mortality resulting from screening[270]. The mortality outcomes for both of these studies, especially the PLCO trial, were biased by neglecting the considerable levels of PSA testing in the control populations[271]. The data discussed in chapter 5 showed that the adjusted benefit of screening on a population-wide basis was higher than the 20% relative reduction in PC mortality that was reported by the ERSPC after a median follow up of 9 years. The study in this thesis estimated a relative reduction in PC mortality of 37%[105]. Individual patient mortality results stratified by PSA were not obtained in the randomised controlled trials because no blood from the men in the control arm of the trials was collected and stored for PSA measurement at the end of the trials[172, 270]. Therefore, in chapter 6, the PSA-specific data from men participating in the intervention arm of the ERSPC were compared with the PSA-specific data men in Northern Ireland (NI). The data presented in Chapter 6 showed that baseline PSA was useful for developing a risk stratification scheme with the potential to increase the balance between the harm and benefit of PC screening[232].

The data presented in chapters 5 and 6 and the data from the ERSPC as a whole were used to quantify the harm-benefit tradeoffs that were associated with PSA screening. The harm-benefit tradeoffs were demonstrated by the NNS/NNI* and NNT* to save one man from PC-related death. The NNS/NNI was calculated as (1/absolute reduction in PC mortality) and the NNT as ((1/absolute reduction in PC mortality) * excess incidence). As a result, diluted estimates of the absolute mortality reduction and excess incidence due to PSA testing in the control population were reflected by a modification of the harm-benefit tradeoffs. Furthermore, both the absolute mortality reduction and the excess incidence depend on the duration of the follow-up. This makes the harm-benefit tradeoffs a sub-optimal measure under the condition of a limited follow-up. Due to the reduction in absolute mortality, the increases in the excess incidence and the duration of follow-up were biased and limited by the data from the randomised controlled trails. In chapters 5 and 6, the harm-benefit tradeoffs will be discussed in more detail.

---

* Appendix 1 page 246

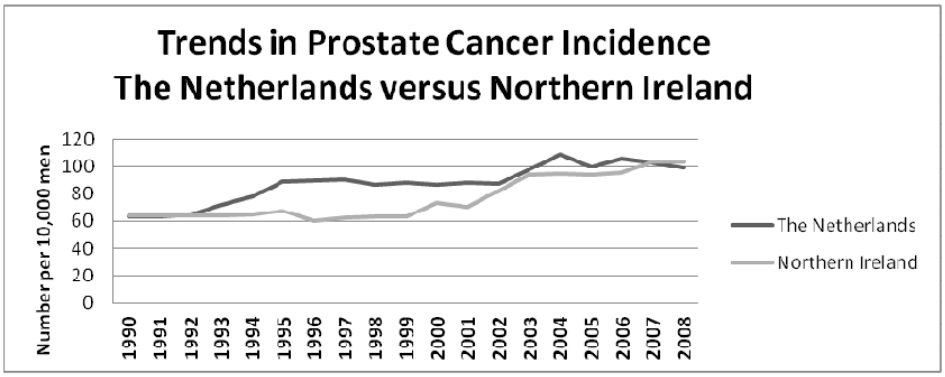


Figure 2. Prostate Cancer Incidence in the Netherlands and Northern Ireland 1990 – 2008. Obtained from[164,278].

The results described in chapter 5 showed that the excess PC incidence was largely underestimated by the ERSPC and PLCO trials. After a median of 9 years of follow-up, 9.6% of the men were diagnosed with PC in the intervention arm of the ERSPC-Rotterdam, 4.3% of the men were diagnosed with PC in the control arm of the ERSPC-Rotterdam, and 3.0% of the men were diagnosed with PC in NI[272]. In total, 6.0% of the men were diagnosed with PC in the control arm of the PLCO after a median follow-up of seven years[97]. It is unlikely that the difference in incidence between the control arm of the ERSPC and the NI is reflected by a difference in the natural risk of developing the disease (chapter 4). This difference reflects the change in the frequency of PSA testing in the general population of the two study populations[152]. As shown in Figure 2, the PC incidence was equal in the Netherlands and NI in 1990; it started to increase in the Netherlands in 1992, but it remained stable in NI until 1999. Comparable trends in PC incidence were shown in Finland, Sweden, Italy, Spain and Switzerland, showing an increase in the PC incidence starting in the early 1990s[9]. The impact of PSA testing on incidence was first reported in the United States, where it was associated with a doubling in the rates of prostate cancer from 1986 to 1992[147]. PSA testing in the Netherlands, which has been available since the mid-1980s, remained limited until around 1993, perhaps due to recommendations that advised against its use in asymptomatic men. When we compared the incidence rates in the Netherlands and NI and the number of PSA tests performed in those countries, a close relationship between the use of PSA screening and cancer incidence was revealed. In 2000, the rate of PSA tests per 1000 men over 45 years of age was almost 140 in the Netherlands and around 50 in NI, which is consistent with the more rapid increase in incidence observed in the Netherlands during the 1990s.

The ERSPC estimated that serum PSA was measured in 20% of men from the control population of the study[165]. In total 52% of men in the control arm of the PLCO were

tested for PC by measuring the PSA[270]. PSA testing has resulted in an increase of the PC incidence in the control arm of the PLCO and a decrease of the excess incidence after seven years of follow-up[270]. A more robust estimate of the excess incidence was made in chapter 5, showing an increase in the cumulative incidence in the intervention population relative to the control population of 67 per 1000 men after a median of nine years of follow-up. The excess incidence was 34 per 1000 in the ERSPC as a whole and 13 per 1000 in the PLCO[172, 270]. Taking into account the effect of PSA contamination on the excess incidence, the ERSPC underestimated the number of men (48 in total) that would receive additional treatment to prevent one man from PC after nine years of follow-up. Similar conclusions can be drawn from the data of the Gothenburg screening study, which presented a NNT of 12 after 14 years of follow-up[218].

The results presented in chapter 5 show that the reduction in PC mortality was underestimated by the randomised controlled trials. For clear reasons, including PSA contamination in the control arm and intense pre-screening prior to randomisation in both arms, no effect on the mortality rate was found in the PLCO[271, 273]. A secondary analysis of the ERSPC showed that the relative reduction that was observed after 9 years of follow-up increased after adjustment for contamination and nonattendance from 20% to 29-31%[165]. In this study, contamination was based on tumour stage specific data, assuming that 23.9% of the T1C cases, 11.2% of the T2 cases, 12% of the T3 cases and 9% of the T4 cases in the control arm were detected as a result of PSA testing in asymptomatic men (Table 1)[165]. After a median of 9 years, the stage distribution of the cancers diagnosed in the control population of the ERSPC was 6.7% T1a/b cases, 42.7% T1C cases, 30.7% T2 cases, 16.4% T3 cases and 3.5% T4 cases (Table 1)[274]. In contrast, of all of the cases of PC diagnosed in the Netherlands in 1993, in men aged 65-74 years, 12.4% had T1a/b-stage

Table 1. Stage distribution of the control population of the ERSPC adjusted for contamination in comparison to the stage distribution of all cases diagnosed in the Netherlands in 1993. Data obtained from [169,271,280].

	A	B	C	D	E
	Control ERSPC	Control ERSPC	Control ERSPC, adjusted for contamination	The Netherlands 1993	The Netherlands 1999
	Stage distribution, %	Assumed contamination, %	Stage distribution, %	Stage distribution, %	Stage distribution, %
T1a/b	6.7	0	8.1	12.4	6.2
T1c	42.7	23.9	39.1	3.9	16.0
T2	30.7	11.2	31.7	38.9	43.3
T3	16.4	12.0	17.2	12.9	12.3
T4	3.5	9.0	3.9	32.2	22.2

disease, 3.9% had T1C, 38.7% had T2, 12.9% had T3 and 32.2% had a T4-stage disease[266]. A comparison is presented in Table 1, and the stage distribution of all of the cases diagnosed in the control arm of the ERSPC is presented in column A. Next, the estimated percentage of men diagnosed as a result of opportunistic PSA testing in the control arm is presented (column B), and the stage distribution of all cases diagnosed in the control population of the ERSPC after adjustment for contamination is given in column C. Finally, the stage distribution of all cancers diagnosed in the general population of the Netherlands in 1993 and 1999 is presented in the last columns (D and E). Comparing the adjusted stage distribution of the control population of the ERSPC with the stage distribution of the cases diagnosed in 1993-1999, the adjusted stage distribution seems to be more favourable. This suggests that the effect of contamination by the ERSPC as a whole is underestimated. Consequently, the relative effect of screening on the PC rate of mortality may be higher than 31% after 9 years of follow-up. In chapter 5, we estimated a relative reduction of 37% after 9 years of follow-up. Considering the other limitations described in chapter 5, the true relative reduction is likely to be between 31 and 37% after a median follow-up of nine years.

The data described in chapter 6 demonstrate that risk stratification based on baseline PSA values can be used to optimise the harm-benefit trade-off in a PC screening programme. Men with low initial PSA values are unlikely to benefit from early detection. This observation has clear clinical consequences and allows for specific individualised risk stratifications after measuring men's PSA baseline. As a result, men at high risk can be informed about their more favourable harm-benefit trade off with respect to the overall NNS and NNT presented by the randomised controlled trials. In contrast, men with a serum PSA <1.0 ng/ml (36% of all men) or men with PSA <2.0 ng/ml (67% of all men) can be reassured that even if they received a biopsy revealing detectable cancer, it is unlikely to become life threatening during their lifetime. Such risk stratification measures may lead to an increased acceptance of screening among men and might increase the compliance among those who are at high risk, if they are informed of their risk status and their individualised harm-benefit tradeoffs. The results presented chapter 6 are in line with a recent study by Vickers et al. who determined the relationship between concentrations of PSA at age 60 and subsequent diagnosis of clinically relevant PC in an unscreened population[233]. They showed that 60-year-old men with PSA concentrations below the median (≤1 ng/ml) were unlikely to have clinically relevant PC (0.5% risk of metastasis by the age 85 and 0.2% risk of death from PC). The risk of dying from PC for men with PSA levels lower than 1.0 ng/ml in Northern Ireland after 9 years of follow-up was 0.1%[233]. Because population-based screening is not a reasonable option for reducing a 0.2% risk of cancer-specific death after 25 years, systematic repeated screening should not be applied to men with low baseline serum PSA values.

The data presented in chapter 6 illustrate the potential power of using risk stratification in population-based screening programmes. The reductions in the relative and absolute mortality increased with the increasing serum PSA values at the start of the study. This influenced the NNS and NNT because the absolute mortality reduction was used as the denominator in the calculations of the NNS and NNT. As a result, the harm-benefit tradeoffs associated with PSA-screening were more favourable for men with higher PSA values at the start of the study. In other words, the benefits increased with the increase in PC mortality. As a result, the harm-benefit trade-offs are likely to differ between populations in Europe. The PC mortality rates in the Nordic European countries (Norway, Sweden, Denmark, Iceland, and Estonia) are five times higher than those reported in several Central and Eastern European countries. Based on the data presented in chapter 6 and in Vickers et al[233], harm-benefit trade-offs are likely to differ between different European countries, with more favourable harm-benefit trade-offs in countries with higher PC mortality rates, such as Norway, Sweden, Denmark, Iceland and Estonia. These differences must be taken into account when applying the results of the randomised controlled trials to a specific population in Europe.

The results on mortality and excess incidence discussed in chapter 5 and 6 are limited by the short duration of the follow-up. With PC being a major cause of death in elderly men (the median age of PC-related death is 80 years) and a median age of 69 years at the end of the follow-up in the study described in chapter 6, most men were at risk of dying from PC during the years after termination of the follow-up period. Therefore, even if the relative reduction in PC mortality remains stable during the next years, the absolute reduction in the rates of mortality from PC will increase with longer follow up. The association between age and the relative and absolute reduction is illustrated in Figure 3. Applying a 20% relative reduction to the age-specific PC mortality rate shows

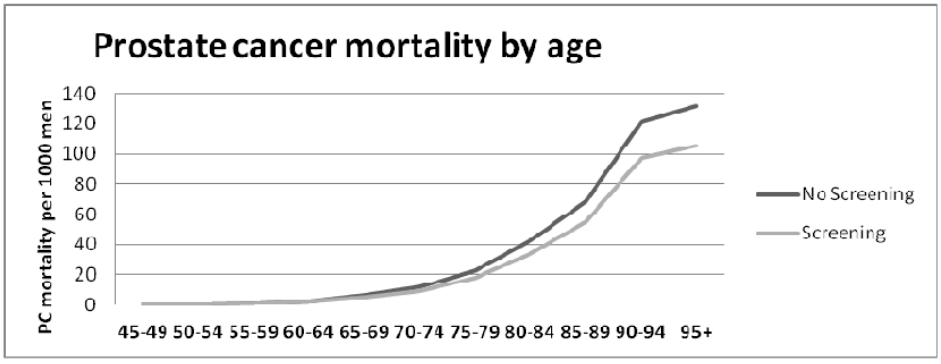


Figure 3. Prostate Cancer (PC) mortality stratified by age at the time of death. The figure illustrates that with a constant relative reduction in PC mortality of 20%, the absolute reduction in the PC mortality rate will increase with the age at the time of death from PC.

that the absolute mortality reduction will increase 6.4-fold from 0.13 per 1000 men aged 65-69 years to 0.83 per 1000 men aged 80-84 years. Consequently, using the absolute PC mortality reduction as the denominator for the calculation of the harm-benefit tradeoff, this tradeoff will change favourably with a longer follow-up in an ageing population. In addition, as the excess incidence of screening depends on the duration of follow-up, the harm-benefit trade-off is also limited by the short-term estimate of the excess incidence. Following the principle of screening, the excess incidence is expected to decrease because the incidence in screened men is expected to drop to a much lower level than in unscreened men after men in the screening population have reached the optimal age for cessation screening[275]. Therefore, it would be more realistic to replace the numerator of the NNT calculation with an estimate of the incidence of overdiagnosis. Draisma et al. estimated the frequency of overdiagnosis in the Rotterdam section of the trial and found it to be as high as 44-55% of the screen-detected cases[129]. The rate of overdiagnosis of screen-detected cases in the US population of men aged 50-84 years between 1985-2000 was estimated to be 23-42%[276]. Assuming a 20% relative PC mortality reduction and an overdiagnosis frequency between 23 and 55%, the NNT to save one man from PC death is estimated to be between 19 and 5 men.

Finally, the relative reduction in PC mortality is expected to increase with longer follow-up. This prediction is based on the cumulative incidence in mortality in both study arms of the ERSPC as a whole, the Gothenburg screening study and the results described in chapter 5. Further, as shown in chapters 5 and 6 and by the ERSPC as a whole, it takes several years for the benefits of PC screening to become evident. Therefore, the steady state for the reduction of the relative mortality rate, which is the final effect of screening, remained undefined in these studies. An estimate of the steady state for the reduction of the relative mortality rate was made by Hanley et al. who showed in a time-specific re-analysis of the ERSPC that the relative mortality reduction is expected to increase. The time-specific re-analysis of the PC-related deaths that occurred in the first 12 years of follow-up suggested that if screening were to be carried out for several years, and if the follow-up were pursued into the window where the reduction in mortality becomes manifest, the relative mortality reduction would be 50–60%[277].

14.3 PART III

A randomised controlled trial is the most reliable method for determining the effects of screening because participants are assigned to one of the two study arms on the basis of chance[278]. The effect of screening is evaluated by comparing the number of cancer-specific deaths in the intervention and control group per unit time[279]. Disease-specific

mortality is considered to be the optimal surrogate endpoint, but actually all-cause mortality is a superior endpoint because it also has the potential to measure the unexpected lethal side effects of screening. Whilst all-cause mortality might be the ideal endpoint, in practice, the relatively small contribution of disease-specific mortality to all-cause mortality means that no screening trial can feasibly be powered to show an effect on all-cause mortality. Assuming a 30% relative reduction in disease-specific mortality would require that more than 3 million men be randomised to show a statistically significant 0.7% relative reduction in all-cause mortality (assuming a median nine-year follow-up and a 15% overall mortality over the 9-year period). In other words, disease-specific mortality is defined as the endpoint of a screening trial because it is numerically impossible to show a benefit in the all-cause mortality. Nevertheless, it would be preferable to have a measure that can capture the unexpected lethal side effects of the intervention. In chapters 7 through 10, we described an alternative surrogate endpoint, i.e., excess mortality. Excess mortality is defined as the difference between the actually observed and the expected number of deaths. As a result, an excess mortality analysis measures the difference between the mortality in a group of patients and the mortality in that group that would be expected in the absence of cancer.

In a randomised controlled cancer screening trial, the cause of death is based on the most likely primary cause of death. This is recorded on the death certificate or is determined by a cause-of-death committee that reviews patients' medical records using predetermined clinical algorithms. Participants that are diagnosed with cancer are considered to die from either the cancer of interest or an intercurrent cause of death. Participants that were not diagnosed with cancer during their lifetime are considered to have died from an intercurrent cause. Cause-of-death committees are required in randomised controlled trials to ensure a correct interpretation of the mortality results[133, 280]. Using these committees, the cause of death is based on the consensus of independent reviewers who have gathered all of the available information from hospital charts, outpatient visit letters, laboratory results and reports from pathology and radiology[153]. Ideally, the review is carried out in a 'blinded' fashion, where the reviewers are unaware of the trial arm to which the subject belongs. In practise, however, it may be difficult to achieve this blinding because information about screening is likely to be included in the notes, e.g., the clinical stage at the time diagnosis and information provided during the patient's consultation with a general practitioner or urologist is illustrative for determining the patient's screening status. Cancer screening aims to provide an earlier diagnosis. As a result, the cancer incidence is typically significantly higher in the intervention group in randomised controlled trials. Consequently, the causes of death in men that are participating in the intervention group are more frequently based on the death certificate or on the consensus of the cause-of-death committee. Because more subjects are reviewed

in the intervention group, it is assumed that there is no cause of death ascertainment-bias to ensure the principal of randomisation. All potential misclassifications that can occur during the process of ascertaining the cause of death are assumed to be equally distributed. If not, any such bias would be larger in the intervention group due to the imbalance in the cause-of-death evaluations. This may result in a net bias. For many types of cancer, it can be difficult to determine whether the patient died entirely due to the cancer in question or due to completely unrelated causes. Especially for patients who die at relatively older ages and have a considerable number of co-morbidities, the dichotomy seems to be difficult to assess. Studies that have questioned the validity of death-certificate-assigned causes of death in cancer and studies that have reviewed the agreement between death certificates and the consensus of the cause-of-death committee have revealed inaccurate coding, with increased levels of inaccuracy for individuals who die at an older age and have an advanced clinical stage of disease at the time of diagnosis[133, 137, 281-283]. Therefore, cause-specific mortality as an endpoint of randomised controlled trials for cancer screening might be biased due to the determination of the end point. Consequently, it would be desirable to have an additional surrogate endpoint to verify this potential bias. In chapters 7 through 10, we demonstrate that excess mortality can be used to verify the validity of the disease-specific mortality outcomes in a randomised controlled trial.

We concluded that in any randomised controlled trial with disease-specific mortality as an endpoint, an additional excess mortality study is mandatory as it may reveal additional important information. However, we did not advise the replacement of disease-specific mortality with excess mortality as the optimal surrogate endpoint of a randomised controlled trial of cancer screening. Hence, excess mortality is a complementary approach to disease-specific mortality because it measures a different parameter. In a PC screening trial, disease-specific mortality refers to the mortality rate that is directly related to PC due to causes such as disease progression and treatment-related mortality. Excess mortality measures both the direct and indirect mortality due to PC, e.g., including cachexia, uraemia, suicidal depression and the loss of interest in life. In addition, an excess mortality analysis may not be superior because it is also subject to uncertainty and potential bias. Most of these sources of uncertainty and bias are presented and discussed in chapter 7.

A valid estimation of the expected mortality is the key assumption of an excess mortality analysis, and this is the main limitation of such an analysis. To perform an accurate excess mortality analysis, the expected mortality in a group of men with PC is assumed to be equal to the expected mortality in the non-cancer group if the group with cancer had no cancer. In PC analyses, this assumption is often difficult to make. In general, men who are diagnosed with a screen-detected cancer are healthier and of higher socioeconomic classes than men in the general population[162, 181]. Therefore, as shown in

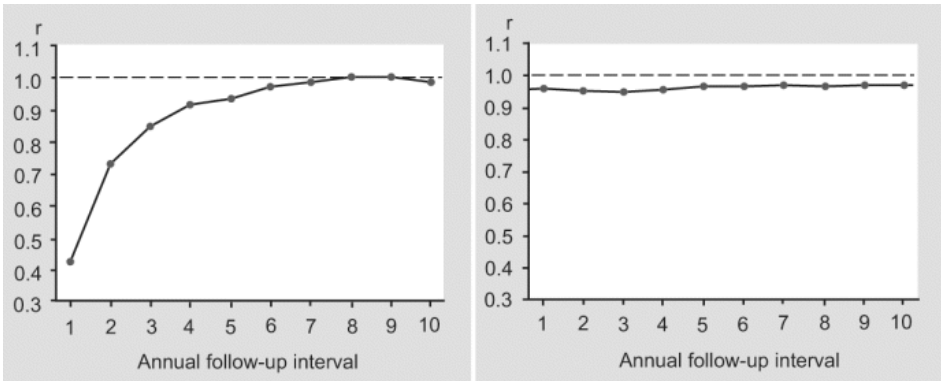


Figure 4. Interval-specific (conditional) estimates of relative survival for patients diagnosed with cancers of the stomach (left graph) and female breast (right graph) in Finland, obtained with permission from [145].

chapters 7 and 8, the national population is not an appropriate reference group for the estimation of the expected mortality in a PC screening trial, and adjusted estimates are needed. A major limitation of an excess mortality analysis in general is that no test is available to test the key assumption of the analysis, i.e., the correctness of the estimate of the expected mortality. One method for validation is to estimate the interval-specific excess mortality ratios, suggested by Dickman et al[252]. This method might be useful for excess mortality estimations in cancers such as those of the stomach and pancreas, with no excess mortality 6-8 years after diagnosis[284]. For these cancers, the interval-specific excess mortality ratio is expected to be approximately 0 at 8 years after diagnosis. Consequently, an excess mortality ratio at the interval 8 years after diagnosis would be representative for a valid estimation of the expected mortality (Figure 4). However, this method is not useful for cancers such as those of the prostate or breast, where excess mortality remains at a relatively constant level for many years following diagnosis (Figure 4)[284]. Therefore, an alternative validation method is needed. Meanwhile, as long as a standardised method to validate the expected mortality is lacking, the interpretation and value of the excess mortality analyses remain questionable.

We applied a method to validate the data presented in chapter 8 and 9, which was not described in the method sections of the chapters. This verification method contains a test that measures the accuracy of the expected mortality. The method is possible in studies that have based the expected mortality on the non-cancer-diagnosed participants in a randomised controlled trial for cancer screening (the method performed in chapters 8 and 9). The validation process begins with taking a random sample of approximately 1000 participants without cancer from the study population. These participants are roughly similar (e.g., regarding screening attendance, age at diagnosis, date randomisation) to the cases that were diagnosed with cancer; the only exception is that the random sample cases were not diagnosed with cancer. In addition, the excess number of deaths

is defined as the difference between the observed number of deaths in the random sample and the expected number of deaths determined using a bootstrap method. The outcome of this estimation is expected to be approximately 0. If this is the case, it means that the expected mortality is estimated with good confidence. Our validation test is feasible and does not require difficult routines in statistical packages. Most importantly, it allows for quantification of the main assumption and, if the assumption is proven, it supports the excess mortality results. Therefore, this validation test should become a standardised method in all future excess mortality estimations.

Both the two studies presented in chapters 8 and 9 are limited by the lack of a statistically significant difference between the excess mortality rates. In chapters 8 and 9, excess mortality rate ratios were compared with the disease-specific mortality rate ratios. The confidence intervals of the disease-specific mortality rate ratios were narrower than the confidence intervals of the excess mortality rate ratios. The 95% confidence interval of the excess-mortality rate ratio was assessed by means of the delta method (chapters 8 and 9, appendix 1) with the observed number of deaths per study arm in cancer patients and the expected number of deaths per study arm in cancer patients not assumed to be constants because they depends on the number of PC cases per study arm and the chances of patient survival (a source of extra random variation). In contrast, the ratio of disease-specific mortality in the intervention group to the mortality in the control group was estimated by means of a Poisson regression analysis with the observed number of deaths per study arm in cancer patients regarded as constant. Consequently, it is evident that in some cases, it may be difficult to know the cause of death in cases of PC. For this reason, in the ERSPC study, efforts have been made to optimise the accuracy of the assessment of the cause of death in PC patients using a committee. Despite this, it remains impossible to confirm cases of PC-specific death without any uncertainty. For example, the cause of death often differed after the case was reviewed by three independent persons who did not exchange opinions about the case. Consequently, we might account for these uncertainties when estimating the 95% confidence interval of the disease-specific mortality rate ratio. For example, if the confidence interval for the disease-specific mortality rate ratio accounted for 5% of 10% uncertainty in cases of PC-related mortality, the RR with a 95%CI was not RR 0.79 (95%CI, 0.66-0.95), but rather RR 0.79 (95%CI, 0.62-0.99) and RR 0.79 (95%CI, 0.59-1.02), respectively, for 5% and 10% uncertainty in the number of PC-specific deaths in both the two study arms.

The results described in chapter 8 suggest that the effect of PC screening on mortality is larger than reported in the analyses that focussed on disease-specific mortality only, especially in men who were 70-74 years of age at the time of randomisation. This contains an additional segment of disease-related mortality that is measured with an excess mortality analysis, but not with a disease-specific mortality analysis. As discussed in chapters 7 through 10, disease-specific mortality may miss important benefits of

cancer screening because of misclassification of the causes of death. On the other hand, the difference might be explained by the observed deviance in non-PC mortality in screen- and clinically diagnosed PC (chapter 10). Men with clinically diagnosed PC were found to be at an increased risk of death due to causes unrelated to PC itself (a significantly increased risk of dying from neoplasms and diseases of the circulatory or respiratory system was detected). These findings should be subject of further research. To date, there is limited evidence available that men with PC have an increased risk of having a non-PC related causes of death. Based on our findings, components of metabolic syndrome might be associated with an increased risk for PC progression. These components, including obesity, an abdominal fat distribution and hyperinsulinaemia, should be investigated further[285-286]. It is known that metabolic syndrome is a major risk factor for cardiovascular disease and adult-onset or type 2 diabetes mellitus[287]. The findings described in chapters 8 and 10 suggest that lifestyle interventions that were related to a diagnosis with screen-detected cancer have had a major impact in decreasing the risk of non-PC-specific mortality. Apart from changes in lifestyle, other health problems unrelated to PC may be detected and treated earlier, which may lead to improvements in the life expectancy. For this reasons, studies are needed to assess the role of changes in lifestyle, the use of medications and medical interventions after PC screening. In addition, studies that assess the mechanisms of the preventive effects of medications, such as statins, on the natural history of PC are of interest with respect to the findings presented in chapters 8 and 10[197-198].

A 24% relative effect of screening was measured by use of the excess mortality analysis presented in chapter 9. However, more arguments than just a reduction in mortality are needed to support the introduction of population-based screening. Firm evidence of the positive effects is needed to outweigh the costs, both economic and in terms of quality of life, of adoption of screening in the general population (chapter 2). A clear consequence of the excess mortality analyses is that screening reduces rate of PC-related mortality, while side effects of the screening procedure (e.g., anxiety, depression and related to a PC diagnosis) are unlikely to cause an increased risk of death.

14.3 PART IV

Part IV of this thesis aims to improve the currently used screening, early detection and treatment strategies. It includes an optimisation of the screening algorithm to potentially increase the PC mortality reduction, and particularly, the prevention of unnecessary biopsies as well as the selective detection of clinically significant PCs to reduce overdiagnosis. Finally, it discusses the difficulties that men have who are diagnosed with localised PC and consider curative treatment.

The screening programme currently used by the ERSPC might need improvement in the following areas. Today, the most significant negative effects of a PC screening programme are overdiagnosis and the healthy life years that are lost due to an earlier PC diagnosis as a result of screening. Consequently, to improve the health benefit trade-off, the overdiagnosis has to be reduced. Furthermore, the diagnosis of a screen-detected cancer might be postponed while having the same positive effect of reducing the PC mortality rate. Furthermore, the number of biopsies has to be reduced, even as a significant side effect of invasive treatment of PC.

PC screening using a serum PSA-based threshold as a sole indicator for prostate biopsy lacks specificity. This results in large numbers of unnecessary biopsies and detection of indolent PCs, but it also misses significant cancer diagnoses in men with PSA levels below the chosen PSA cut-off value[72, 74]. Future screening strategies should aim to reduce the number of potentially unnecessary biopsies and potentially indolent PC diagnoses using risk-profile-based cut-off values as indicators for prostate biopsy. However, despite numerous attempts to assemble a set of risk factors for PC, no individualised screening strategy utilising such stratifications has been introduced to date. Only hypothetical exercises have emerged. These show that an individualised screening algorithm using other available prebiopsy information (e.g., family history, DRE, TRUS findings, prostate volume) in addition to the PSA level can result in a considerable reduction of unnecessary biopsies and detection of insignificant PC[227-228]. In addition, new PC markers in serum and urine, in conjunction with these models, have been shown to improve their sensitivity and specificity[288]. The hypothetical models have clear implications for the improvement of screening strategies, making a first step toward the definition of a patient's individual risk. The models could be the basis for future individual risk-based screening protocols, with factors such as screening interval, age of cessation of screening and the number of biopsies taken based on the patient's individual risk. Multivariate risk stratifications seem to be the proper future approach because novel 'perfect markers', with high sensitivity and specificity, are far from being introduced for clinical use and screening programmes[289]. Most recent discoveries, such as the gene-fusion TMPRSS2:ERG or prostate cancer antigen 3 (PCA3), are not able to replace PSA while improving its performance[290-291]. Therefore, while basic research remains the main priority, improvement of PC screening seems to be restricted to screening based on patients individual age, comorbidity, family history, PSA, DRE and prostate volume, with incorporation of new makers such as TMPRSS2:ERG and PCA3. Finally, these risk strategies should be able to differentiate between 'low-risk' cancers that are not likely to become harmful during a patient's life and 'high-risk' cancers that require interventions.

To date, the most commonly recommended screening interval is annual[292-293]. In the ERSPC, screening intervals of 2 and 4 years were used. In chapter 11, the proportional interval of cancer incidence was studied comparing a 2-year with a 4-year screening

interval. We concluded that the effectiveness of screening might be increased with a shorter screening interval. However, the data in chapter 11 show that a fixed biannual screening interval will still be too long for some cancers but is too short for most other cancers. Since overdiagnosis remains the major limitation of PC screening, it is unjustified to recommend a fixed screening interval in future screening algorithms. Nevertheless, the outcomes of the work described in chapter 11 have clear clinical implications. This study showed for the first time that some men can benefit from more frequent screenings. Consequently, individual risk factors might be defined to select subgroups that are likely to benefit from shorter screening intervals. In chapter 6 and in [263, 285-286], it is shown that individualised screening intervals can be defined based on baseline PSA measurements. This supports the use of close surveillance with a more frequent screening interval for some and an extended screening interval for others. This approach will prevent unnecessary biopsies, tests and visits.

Using an individual risk-based strategy to define the optimal screening interval is an approach that is currently available to reduce overdiagnosis. Consequently, more aspects of the screening algorithm should be based on the patient's individual risk. There is limited research addressing the mortality reduction due to screening among older men, with the optimal age of PC screening cessation remaining uncertain. It is known that overdiagnosis increases with age at the time of screening[129]. However, the optimal age of cessation of PC screening should be defined by more factors than age alone. There is great variation in the number and severity of comorbities, suggesting that a screening algorithm based solely on chronologic age will be incomplete. Life expectancy decreases not only with increasing age, but also with the number and severity of chronic diseases and the level of functional decline[294]. Health, functional status, and comorbidity are the most appropriate indicators of expected life span, compared with chronological age alone[295-296]. In addition, smoking, obesity and being overweight are associated with large decreases in life expectancy and increases in early mortality[297]. Predictions of life expectancy based on a multivariate set of co-morbidities are needed. These predictions must be incorporated in the decision of whether to continue or discontinue the screening procedure, e.g., whether to perform a biopsy or not.

To date, some PC guidelines recommend starting PSA testing at an age of 40 years[292] because PSA measurements before the age of 50 might help to risk-stratify men based on the frequency and/or type of later PC screening. Two studies have associated the serum PSA level with the long term risk of PC. Loeb et al. examined men in their 40s and showed that the subsequent PC diagnosis was 14.6-fold higher for men with a baseline PSA level between 0.7 and 2.5 ng/ml compared to men with PSA <0.7 ng/ml[169]. Lilja et al. assessed the PC risk among men younger than 50 years and showed that the PSA level at age 44–50 years was strongly associated with the likelihood of developing PC up to 25 years later[298]. The odds ratio for a PC diagnosis at a PSA value of 0.51–1.0 ng/ml was 2.51

compared to PSA ≤0.50 ng/ml, which roughly corresponded to the population average. The odds ratio increased to 7.02 for a PSA of 1.0-1.5 ng/ml, and it increased further up to 19.01 for a PSA of 2.01-3.0 ng/ml compared to a PSA ≤0.50 ng/ml. Although these risk strategies are promising and in line with the results that are presented in chapter 6, it currently seems to be unjustified to lower the age of onset for PSA testing. Without a clear guideline for how to interpret the early PSA measurements and acceptable biopsy indications, lowering the age of screening onset will result in earlier diagnoses and in an undefined increase in overdiagnosis. Consequently, this will increase the number of years lived with cancer and increase the 62% negative effect estimated by Heijnsdijk et al. As a result, other changes to the screening algorithm seem to have a higher priority.

The number biopsies taken might also be based on a patient's individual risk. In the ERSPC, sextant biopsies are used, although sextant biopsies are considered to be obsolete in current clinical practice. In clinical practice, for an initial biopsy, a minimum of 10 but not >18 systematic cores are recommended, with 14-18 cores in glands ≥ 50 cm^3. Further biopsy sets, either as extended repeats or as a saturation biopsies (≥20 cores), including the transition zone, are warranted in young and fit men with a persistent suspicion of PC[299]. Sextant biopsies are considered obsolete because extended biopsy schemes have been shown to increase the cancer detection rate[170, 300]. However, a gain in the cancer yield is only beneficial to the patients or participants when more clinically significant cancers are diagnosed without further increasing the detection of clinically insignificant PC. As the prostate volume is inversely related to cancer detection, this parameter is generally applied as an indicator for the number of cores per biopsy. Subsequently, this hypothesis could also be obtained in a PC screening programme[301-302]. The hypothesis was tested in chapter 12. We found that, although a sextant biopsy was applied without considering prostate volume, men with smaller prostate gland volumes were at greater risk of aggressive cancer during the follow-up. In contrast, based on the hypothesis, an increased number of aggressive cancers was expected in the larger prostate gland volumes. We concluded that the volume-adjusted biopsy schemes should not be implemented automatically in a screening programme with repeated screening because this will further increase the rate of overdiagnosis. Lateralised sextant biopsies miss approximately 19% of detectable cancers compared to a volume-adjusted biopsy protocol[170, 299]. However, the number of potentially missed cancers with a poor outcome in terms of progression-free survival and deaths from PC has been shown to be low[170]. The use of extended biopsy schemes is likely to detect more potentially indolent PCs and thus to increase the overdiagnosis and overtreatment. Considering the dilemma of overdiagnosis and overtreatment, the usefulness of more extended biopsy schemes within this setting must be questioned. On the other hand, a percentage of potential clinically relevant cancers are clearly missed using

sextant biopsies. Therefore, extended biopsies might be introduced in future screening strategies. However, verification of the individual risk is needed first to decrease the number of biopsy indications and the detection of clinically insignificant cancers. Finally, the number of biopsies might not be solely based on the prostate volume, but also on the risk of 'significant' PC.

Another strategy to reduce overdiagnosis could be the primary prevention of PC. Multiple randomised trials of prevention strategies, including pharmaceutical, dietary, and supplementation, have been performed[303]. The most recent trials are the Selenium and Vitamine E Cancer Prevention Trial (SELECT), the Prostate Cancer Prevention Trial (PCPT) and REduction by DUtasteride of prostate Cancer Events (REDUCE trial). The SELECT showed no benefit of chemoprevention with vitamin E and selenium supplementation for reducing the incidence of PC[304]. The PCPT and REDUCE trial, which evaluated the use of 5α-reductase inhibitors (5ARI) with PC prevention as a primary endpoint, showed a successful result. In the PCPT, the period prevalence of PC was reduced by 24.8% in the 5ARI (Finasteride) group compared with the placebo group (18.4% vs. 24.4%, P <.001)[305]. The REDUCE trial showed that the PC incidence decreased by 23% in men who received the 5ARI dutasteride in comparison to those who received a placebo after four years[306]. However, in the PCPT trial, the prevalence of high-grade tumours (Gleason score 7-10) detected through biopsy increased in the 5AR group (37.0% of all tumours) compared with the placebo group (22.2%; P <.001)[305]. These unfavourable outcomes in the PCPT seem to be based on study biases, i.e., (the detection was better in men with small prostate gland volumes). The REDUCE study has demonstrated that use of 5ARI is safe and effective for reducing the risk of cancer, regardless of the risk stratum[307-309]. If the incidence of PC with a Gleason score of 5-6 significantly can be reduced, and at the same time no increase in high-grade tumours with 5ARI use occurs, chemoprevention might have the potential to reduce the detection of favourable indolent disease, making screening more beneficial. However, more research is needed. Both trials were based on a study design in which men in both study arms (5ARI and placebo) were biopsied at a fixed interval without a clinical indication for biopsy. This study design makes the clinical implication difficult to determine. In patients with benign prostatic hyperplasia (BPH), 5-ARIs suppress serum PSA levels by about 50%, and when using 5-ARIs, the PSA kinetics are poor defined. To date, it is clear that a rise in PSA levels, but not their stabilisation, indicates a biopsy and the possible presence of Gleason 7–10 cancer[310-311]. However, additional information is needed to answer the question of what should be done if the PSA is not rising. Therefore, further research is needed to define which men might be treated with a 5ARI, and more importantly, what should be the indicator for a biopsy when using 5ARI. Until these results are available, primary prevention with 5ARI should not be recommended to decrease the risk of PC. This is in line with the recently published FDA policies.

To date, population-based PC screening has not been introduced, and there is a wide variation in the detection strategies favoured among clinicians. In some countries, testing for PC in asymptomatic men is permitted after men are informed of the potential benefits and costs[293, 312-313]. This seems to be ethically correct, however, the decision to screen for PC is not based solely on the facts about screening in clinical practice. The patient's emotional state and the physician's preference are of major influence. One of the arguments against screening is overdiagnosis; however, this is difficult to translate into men's individual risk. When defining a patient's individual risk, the potential benefits of early detection are often considered more important than the potential harms. Nevertheless, it is the aim of physicians to provide the facts, allowing men to decide for or against screening. Therefore, physicians should use decision aids with PSA testing. There is strong, consistent evidence that patient decision aids improve a patients' understanding of PSA testing[314]. Most studies have indicated that decisional conflict is reduced among men who use a decision aid[315]. Secondly, physicians should be made aware of the risks of diagnosing PC based only on age and serum PSA. Nomograms and risk calculators have shown better accuracy than PSA in predicting the detection of PC through biopsy[228]. These tools, such as the SWOP/ERSPC risk calculator and the Prostate Cancer Prevention Trial (PCPT) prostate cancer risk calculator, assess other factors, including digital rectal examination (DRE), family history for prostate cancer, a prior biopsy, race, and age. These tools have been shown to result in a considerable reduction in the number of unnecessary biopsies[228, 316]. A preliminary assessment of the risk of high-grade PC revealed by prostate biopsy can be made based on the PSA level, an abnormal digital rectal examination, and a family history of PC[317-318]. Both of these online risk calculators are user friendly[319-320].

The risk calculators aim to allow men to understand the risks of PC in an individual setting. However, it remains questionable whether it will be possible for men to distinguish between the risks. Currently, the outcome of the risk calculators is the chance of a biopsy-detectable cancer, i.e., a score in the range of 0-100%[321]. Unfortunately, there is no threshold value defined for the overall cancer risk or for the high-grade cancer risk that should prompt a biopsy. Currently, the risk level at which a biopsy should be considered is based on factors related to a man's life expectancy (age, comorbidities) as well as his level of concern regarding his risk. It is unclear whether patients are able to appreciate the difference between a 10% and 20% overall PC risk. Therefore, men's life expectancy, based on age and comorbidity, might be added to the nomogram after which the nomogram should result in a positive or negative recommendation for biopsy without presenting the actual risk. Moreover, because high-grade tumours appear to pose the greatest risk and because there is a greater need to diagnose these high-grade cancers early, the risk of high-grade disease should be the primary determinant of the need for a prostate biopsy.

The next challenging step in PC care occurs after screening and early diagnosis, i.e., the definition of the optimal treatment for men with localised PC. It can be a challenge to distinguish between patients who are at a higher risk of death from PC and those who are more likely to die from other causes. To date, there is no universally accepted definition of clinically significant or insignificant disease. Studies have shown that the cancer volume, clinical stage and tumour histology are important predictors of the long-term outcome[21, 322-323]. These characteristics are used to define which cancers should be treated and which can be managed expectantly. Based on these results, methods to reduce the overtreatment of indolent tumours have been implemented with active surveillance programmes for small, low-risk tumours. These programmes are urgently needed in conjunction with screening programmes to limit the increase in economic costs and to improve the overall quality of life outcomes. Sanda et al. reported on quality of life and satisfaction with the outcome following either surgery or radiation. Long-term incontinence persists in 2-22% of patients, and erectile dysfunction affects 20-90% of men depending on factors such as the baseline sexual function, the age of the patient and surgical technique used[109]. In chapter 13, we have provided the disease outcomes for 990 men with localised screen-detected PC who were treated with a radical prostatectomy. The results contribute to the understanding of treatment effectiveness among men with screen-detected localised PC stratified by the patients' clinical characteristics. The patient age, tumour stage and Gleason score provided the most relevant data concerning the likelihood of dying from PC. According to the data presented in chapter 13, men who harbour tumours with a Gleason score larger than 7 are at a significant risk of death from PC. These men should consider aggressive treatment, such as radical surgery. By contrast, a large group of men with low-risk PC might have had similar outcomes with active surveillance. Several studies have shown outcomes from large, observational cohorts of men who have selected active surveillance as a treatment alternative[41-42, 257, 324]. These studies suggested that active surveillance is a viable treatment option for men with low-grade, early-stage, and favourable-risk PC with limited life expectancy. The benefits of active surveillance are clear: This approach avoids the side effects associated with other treatment options, such as erectile dysfunction or incontinence. However, cancer spread during surveillance remains a great risk, and it requires lifetime of regular testing with PSA, DRE and biopsies. Therefore, it is questionable whether active surveillance is an adequate solution to the problem of overdiagnosis associated with screening. Recently, a retrospective analysis by Stattin et al. showed that among those with low-risk disease, the calculated cumulative 10-year PC-specific mortality was 2.4% in the surveillance group and 0.7% in the group that underwent surgery and radiation therapy[258]. These results seem to be in favour of curative treatment; however, a careful interpretation of the results reveals that 1427 men required curative treatment to save one man from PC death over a period of 10 years. The data highlight the present and future dilemma. An

upcoming RCT will possibly show a small benefit of surgery or radiation therapy over active surveillance; however, a large number of men would require curative treatment to prevent one PC-specific death. This leads to the question of how patients should be informed and what information should be given. Consolation with a man diagnosed with localised, low-risk PC will be extremely difficult. Emotions will frequently stimulate the patient to choose the smaller benefit and thus opt for the curative treatment. Based on data on overdiagnosis and observational active surveillance, active surveillance can be considered as a treatment option. However, until better biomarkers for indolent disease are identified, we should bear in mind that the risk of disease progression persists no matter how small or well differentiated the disease seems to be. Therefore, more effort is needed to reduce the diagnosis of the low-risk cancers because the diagnosis of these cancers is too likely to result in curative treatment. Diagnostic tools that specify which cancers to detect and which cancers not to detect would be preferable.

14.5 CONCLUSIONS

To date, convincing evidence is lacking that PC screening is effective in PC mortality reduction or that it produces a beneficial net effect, i.e., the benefits exceed the adverse effects, such as overtreatment and overdiagnosis. Moreover, it is uncertain if the lack of a net effect will be solved by the expected increase in PC mortality reduction after a longer follow-up of the on-going randomised controlled trials of PC screening. A current problem is the inability to target men at risk of developing or harbouring not just histological PC, but also a significant disease that requires intervention. Strategies that aim to reduce overtreatment by not treating low-risk cases of PC and keeping them in the window of curability are still in the experimental phase. Establishing population-based screening for PC as a health policy will depend on lowering the rate of overdiagnosis while focusing on methods for more selective screening and achieving an acceptable risk-benefit ratio. This thesis contains significant information addressing this dilemma and recommends adjustments in the screening strategy. The following conclusions are drawn:

- The effect of PC screening on the PC mortality and metastases are underestimated by the on-going randomised controlled trials due to the effect of PSA testing (contamination) on the control populations in the trials.
- Overdiagnosis, and thus the NNT, are underestimated by the on-going randomised controlled trials due to the effect of PSA testing (contamination) in the control populations of the trials.

- The baseline serum PSA level before diagnosis is a strong predictor of PC mortality in screen-detected and clinically detected PC.
- For men with a low serum PSA, the benefits of aggressive testing and treatment seem to be limited because they are associated with a large increase in the cumulative incidence and potential overtreatment.
- The estimation of excess mortality can be used in future studies to evaluate the effects of screening on PC mortality.
- The excess mortality analysis further strengthens previous reports on a beneficial effect of PSA screening.
- In any randomised study with disease-specific mortality reduction as an endpoint, an additional excess mortality study should be mandatory, as it may reveal additional important information regarding the effect of screening.
- Screening reduces the PC mortality while side effects of the screening procedure (e.g., anxiety and depression related to a PC diagnosis) do not increase the risks of all-cause death.
- Men with clinically diagnosed PC are at an increased risk of death unrelated to PC itself.
- A programme with a 2-year, as opposed to a 4-year, screening interval leads to a significant reduction in the incidence of advanced cancer but is associated with an increase in the incidence of "low-risk" screen-detected PC.
- A patient-specific screening interval seems to be the most efficient approach, as some men benefit from more frequent screenings, but most do not.
- Men with a smaller prostate gland volume are at an increased risk of harbouring tumour cells and developing an aggressive cancer.
- For men with high-grade PC, PC remained a major cause of death despite early detection through screening followed by radical prostatectomy and screening.
- The disease-specific survival of localised screen-detected cancers that are treated by radical surgery is lengthened due to diagnosis at an earlier point in time and the treatment of men who probably would never have presented clinical symptoms of PC in the absence of screening.

14.5 FUTURE PERSPECTIVE

If PC screening is implemented as a strategy to significantly decrease the mortality from PC, while considering and improving the overall quality of life aspects, the overall benefits of screening will increase during the next decades. Due to the ageing of the general population, the total rate of PC-specific mortality is expected to increase by 131% in the Netherlands between 2010 and 2050, from 2524 PC-related deaths to 5855 PC-related

deaths[269]. Assuming a 20% relative reduction in PC mortality as an effect of screening in the general Dutch population, screening would prevent 505 PC deaths in the total Dutch population in 2010 and 1711 PC deaths in 2050. Meanwhile, ageing of the national population will have a large impact on the incidence PC. The total incidence of PC is expected to increase from 9905 new cases in 2010 to 14967 new cases in 2035, assuming no increase in the frequency of PSA testing (Figure 1). Assuming an age-specific 50% overdiagnosis as a result of introducing population-based PC screening, the incidence would increase by 126% to 22450 new cases in 2035.

These future estimates indicate that we are faced with a significant dilemma related to PC. Therefore, we need to make decisions regarding how to continue our research on population-based PC screening. PC screening has proven to be effective for reducing the PC-specific mortality by providing level 1a evidence. However, it has been shown to be associated with a disproportionate increase in PC incidence. To improve the outcomes of population-based screening for PC, two future directions are possible.

1. Do more: Optimise PSA-based screening by lowering the PSA cut-off, shortening the screening interval and decreasing the age at which screening begins. This might increase the reduction in PC-specific mortality and the incidence of metastatic disease. Furthermore, it might positively affect the rate of all-cause mortality among men with screen-detected PC due to an early change in their medical regimens for co-morbidities. However, the more intensive screening strategy will certainly significantly increase the harm associated with screening.
2. Do less: Introduce screening based on individualised risks. Make the screening intervals as long as possible, discontinue repeated screening in men as soon as possible, and define biopsy cut-offs based on a multivariate set of predictions that are based on the prediction of "significant PC". This will require ways to predict which individuals are likely to develop high-risk cancers. Studies with this focus might result in fewer biopsies and diagnoses and thus in harm-benefit tradeoffs that justify introduction of population-based screening in the future.

Finally, we have to realise that population-based screening has the potential to reduce mortality from PC but not to completely eliminate PC mortality from the general population. There will always be cancers that will escape detection despite screening. To save or cure the men with these exceptional cancers, other prevention strategies or treatments are indicated. Wilson and Junger stated in 1968: "the condition sought should be an important health problem". This means that population-based screening is not the optimal strategy to prevent death or progression in these few cancers that escape detection through the current programme.

14.6 APPENDIX

Explanation of the statistical measures used in the general discussion.

The Number Needed to Screen (NNS) or Number Needed to Investigate (NNI) are presented as,

NNS = NNI = (1/absolute reduction in PC mortality),

where the absolute mortality reduction is based on the absolute difference between the mortality rate in two populations. In the general discussion, the absolute difference in mortality is defined as the number of deaths per 10,000 person-years in the control population minus the number of deaths per 10,000 person-years in the screening population.

The absolute difference in mortality = ((number deaths/10,000 person-years) control group) - ((number deaths/10,000 person-years) screening group).

The Number Needed to Treat (NNT) is presented as,

NNT = ((1/absolute reduction in PC mortality) * excess incidence),

NNT = ((NNS) * excess incidence),

Where the excess incidence is reduction is based on the absolute difference between the incidence rate in two populations. In the general discussion, the excess incidence is defined as the number of cancers diagnosed per 10,000 person-years in the screening population minus the number of cancers diagnosed per 10,000 person-years in the control population,

Excess incidence= ((the number of cancers diagnosed/10,000 person-years) screening group) – (the number of cancers diagnosed/10,000 person-years) control group).


CRUISIN' FOR SCREENING
PROSTATE CANCER

Part VI

Appendices

Summary

Prostate cancer is the most common cancer among men in the Netherlands. A total of 9500 men were diagnosed with prostate cancer, and 2300 died from prostate cancer-related causes in the Netherlands in 2006.

During the last three decades, many researchers have tried to assess the effectiveness of screening for prostate cancer. The objective of early detection (screening) is to detect a disease at a stage in its natural history where treatment can be applied to prevent death or suffering. By use of digital rectal examination (DRE) and prostate specific antigen (PSA), it is possible to diagnose prostate cancer at an earlier clinical stage. The European Randomised Study of Screening for Prostate Cancer (ERSPC) was begun in the early 1990s to study the effect of population-based screening for prostate cancer on the prostate cancer mortality rates and the quality of life. Today, it is known that screening has the potential to decrease the rate of prostate cancer-specific mortality significantly. However, prostate cancer screening is also associated with significant negative effects. Due to screening, many prostate cancers are diagnosed that would otherwise remain clinically unrecognised until the individual died from other causes. Currently, 50% of all diagnosed prostate cancers are slow-growing tumours that will not result in clinical signs or symptoms during the patient's lifetime. The diagnosis of these cancers is referred to as overdiagnosis. Overdiagnosis appears to be especially problematic when it results in invasive treatment of tumours that are unlikely to be harmful. This thesis, entitled "prostate cancer screening, the effect on mortality and incidence", includes studies on the effectiveness of prostate cancer screening and studies that contribute to the improvement of the screening programme.

Chapter 1 contains a general introduction on the different aspects of prostate cancer diagnosis and treatment. Chapter 2 is a review of the prostate cancer screening studies up to 2011. Based on the results presented in chapter 2, we concluded that screening for prostate cancer reduces the disease-specific mortality significantly. Nevertheless, there is no scientific basis for the introduction of a population-based prostate cancer screening programme because a number of essential criteria have not been met. For example, no robust evidence is available regarding the economical and quality-of-life implications of screening for prostate cancer. Chapter 3 comprises an overview of the value of measuring the trends in prostate cancer incidence, survival and mortality. The interpretations of these trends have shown to be confusing without proper knowledge of the natural history of prostate cancer and the statistical measures. We showed that trends in prostate cancer incidence are not reliable for measuring the increases in environmental risk factors, as the practices in screening and early detection have changed

more than necessary during the last two decades. Trends in prostate cancer survival are suspected to be unreliable because they are based on the same series of patients as incidence rates, and any inflation of incidence due to the inclusion of less malignant or nonmalignant diseases creates a spurious increase in survival rates. As a result, prostate cancer survival rates should be interpreted with caution and with the knowledge of the effects of lead time, length time and overdiagnosis on survival outcomes. Trends in prostate cancer mortality are the principal measures through which the success of the different interventions on a population basis can be determined. Because the mortality rate measures of the number of deaths (in general, or due to a specific cause) in some populations are scaled to the size of those populations per unit time, the measure is not biased by the number of cancers diagnosed.

In chapters 5 and 6, we studied the effect of prostate cancer screening on the prostate cancer mortality rate and incidence. In chapter 5, we compared two populations, 11,970 men in the Netherlands and 133,287 men Northern Ireland. Men in the Netherlands received systematic screening, and men in Northern Ireland receive usual care. During follow-up (median 9 year), 1,153 men (9.6%) were diagnosed with prostate cancer in the screening group (the Netherlands), and 3,962 men (3.0%) were diagnosed in the control group (Northern Ireland). In the screening group, 35 (0.29%) men, and in the control cohort 627 (0.47%) men, died due to prostate cancer. This equated to a reduction in prostate cancer mortality of 37% in the intervention population relative to the control population: RR 0.63 (95%CI, 0.45-0.88, p=0.008). Therefore, 555 men needed to be screened to prevent one prostate cancer death. In chapter 6, we tried to identify men, stratified by serum PSA, who would potentially gain the most benefit from prostate cancer screening. We showed that the potential harms of repeated screening exceeded the potential benefits in men aged 55-74 years with a serum PSA 0.1 to 1.9 ng/ml. The main conclusion was that in men with a low baseline serum PSA, the benefits of continued aggressive investigation and treatment may be limited, whilst they are associated with a large increase in cumulative incidence, overtreatment and costs.

In part 3 (chapters 7 through 10), of this thesis we aimed to assess the effect of prostate cancer screening on the excess mortality. Excess mortality is defined as the difference between the total (all-cause) mortality of the patients and the mortality that would be expected in the absence of cancer. In chapter 7, we described a method for performing a robust excess mortality analysis in a randomised control cancer screening trial. Chapters 8 through 10 present the results of the excess mortality analysis. Chapter 8 is based on the analysis of 42,376 men that were randomised to systematic screening or usual care in the ERSPC the Netherlands. After a median follow-up, the disease-specific mortality rate was 0.42 men per 1000 person-years in the intervention and 0.48 men per 1000 person-years in the control arm: RR 0.86 (95%CI, 0.64-1.17). The excess mortality rate was 0.40 per 1000 person-years in the intervention arm and 0.61 men per 1000 person-

years in the control arm, and the RR for excess mortality was 0.66 (95%CI, 0.39-1.13). The difference was mainly among men aged 70-74 year at randomisation. In other words, we found a between-arm difference in the excess mortality rates that was more than twofold the between-arm difference in the disease-specific mortality rates. Two possible explanations for the observed discrepancy were found for this observation: 1) a systematic underestimation of the number of deaths from prostate cancer by the cause-of-death committee of the study; 2) additional disease-related mortality that is measured by an excess mortality analysis but not with a disease-specific mortality. The main results found in chapter 8 were studied in more detail in chapter 10. This additional analysis confirmed that men with clinically diagnosed prostate cancer were at an increased risk of death unrelated to prostate cancer itself (increased excess mortality relative to the disease-specific mortality). The excess mortality was due to a significantly increased risk of dying from neoplasm and diseases of the circulatory or respiratory system. Many effects might have influenced these observations; however, the relatively increased use of hormone treatment in clinically diagnosed prostate cancer and the change of medical regimes and medication among men with screen-detected prostate cancer may have had the most influence. As a result, we concluded that if changes in the medical regimens do affect all-cause mortality in men with prostate cancer, uro-oncologists should look carefully at the management of abnormal parameters of the circulatory and respiratory system, and they should encourage prostate cancer patients to make lifestyle modifications. In chapter 9, we studied the excess mortality in prostate cancer screening based on data from the four largest centres of the European Randomised Study of Screening for Prostate Cancer (ERSPC). For this analysis, a total of 141578 men, aged 55-69 years, were randomised into groups that received either systematic screening or usual care in the ERSPC sections of Finland, Italy, the Netherlands and Sweden. After a mean follow-up of nine years, the prostate cancer incidence rate was 9.24 per 1000 person-years in the intervention arm and 5.49 per 1000 person-years in the control arm, RR 1.68 (95%CI 1.61-1.75). The excess mortality rate was 0.29 per 1000 person-years in the intervention arm and 0.36 men per 1000 person-years in the control arm; the RR for excess mortality was 0.76 (95%CI, 0.54-1.07). We concluded that PSA-based screening had a clear benefit on the excess mortality and that the estimated excess mortality reduction of 24% was in excellent agreement with the disease-specific mortality that was reported in a previous report by our group, implying that PSA-based screening is effective in reducing the mortality from prostate cancer.

Chapter 11 aimed to assess the optimal interval (time between two screening rounds) for prostate cancer screening. The rate of decrease in advanced cancers was estimated for a programme using a 2-year and one using a 4-year interval to determine the cancer screening programme effectiveness. Men aged 55-64 years from two centres of the ERSPC- Gothenburg (2-year screening interval, n=4202) and Rotterdam (4-year screening

interval, n=13,301) were included. The results showed that screening with a 2-year or a 4-year interval both significantly reduced the risk of being diagnosed with advanced cancer. A 2-year screening interval significantly reduced the incidence of advanced prostate cancer up to 43% relative to a programme that used a 4-year interval. The screening programmes that used more frequent screening were found to be associated with increased numbers of investigations and overall cancer incidence. Chapter 12 contains an evaluation of the use of sextant biopsies in men with larger prostate glands who participated in a screening programme with a 4-year interval. Men with smaller prostate volumes and initially high serum PSA levels were at a greater risk of cancer detection and an aggressive cancer during the follow-up. We concluded that the use in clinical practice of volume-adjusted biopsy schemes should not be implemented automatically in screening programmes with repeated screening. In chapter 13, a competing risk stratification (dying from prostate cancer or another cause of death) was made for men with localised prostate cancer who consider treatment with radical surgery. We found that men with a screen-detected clinically localised prostate cancer treated by radical surgery have a heterogeneous risk of dying either from prostate or other causes depending on their age and the histological characteristics of the biopsy specimen. Men aged 55-74 years at diagnosis with tumours that have Gleason scores of ≤6, 7 and ≥8, face a 2.1%, 4.2% or 16.1% probability, respectively, of dying from prostate cancer within 12 years of diagnosis. The relative contribution of deaths unrelated to prostate itself increased with advancing age and decreasing Gleason scores. We concluded that for men with high-grade prostate cancer, prostate cancer remained a major cause of death despite early detection by screening followed by radical prostatectomy and screening.

The results of this thesis described in parts 1 through 4 are discussed in part 5. In this part, the results described in this thesis are discussed in relation to the most recent literature and suggestions for further research are mentioned.

Samenvatting (Dutch)

Prostaatkanker is de meest voorkomende maligniteit onder mannen in Nederland. In het jaar 2006 werden er 9500 mannen gediagnosticeerd met prostaatkanker en overleden in totaal 2300 mannen aan de gevolgen van prostaatkanker. Hiermee is prostaatkanker één van de belangrijkste gezondheidsproblemen voor mannen in Nederland. In 1993 heeft dit geresulteerd in de start van een uitgebreid onderzoek naar de effectiviteit van de vroege opsporing van prostaatkanker in acht Europese landen.

Na de ontdekking van het prostaat specifiek antigeen (PSA), en het bewijs dat PSA als een serum marker voor de detectie van prostaatkanker kon worden gebruikt, kan prostaatkanker in een vroeg stadium worden gediagnosticeerd. Screening van prostaatkanker betreft het onderzoeken van een in principe gezonde populatie om de nog asymptomatische gevallen van prostaatkanker op het spoor te komen, in de veronderstelling dat de prostaatkanker in een vroeg stadium beter te behandelen is en hiermee de sterfte aan prostaatkanker kan worden verlaagd. De European Randomized Study of Screening for Prostate Cancer (ERSCP) is begin jaren negentig gestart om meer inzicht in te krijgen in de effectiviteit van screening naar prostaatkanker. Inmiddels is bekend dat door middel van screening de sterfte aan prostaatkanker kan worden verminderd. Screening op prostaatkanker brengt echter ook nadelen met zich mee. Bij een deel van de mannen ontwikkelt de prostaatkanker zich namelijk zo langzaam, dat mannen niet als gevolg van, maar wel met prostaatkanker overlijden. Bij sommige mannen blijft de tumor zo klein, dat ze er tijdens hun leven zelfs geen klachten van ondervinden. Wanneer bij deze mannen de tumor door middel van screening toch vroeg wordt opgespoord, worden ze wellicht onnodig behandeld middels een ingrijpende operatie of bestraling. Het probleem van screening is dat steeds meer van deze onschuldige prostaatkanker worden gediagnosticeerd. Dit proefschrift, getiteld “prostate cancer screening, the effect on prostate cancer mortality and incidence” geeft inzicht in het effect van prostaatkanker screening en geeft antwoord op enkele specifieke vragen betreffende de verbetering van het screeningsprogramma.

Hoofdstuk 1 is een algemene introductie waarin de diagnostiek en behandeling van prostaatkanker uiteen wordt gezet. Hoofdstuk 2 omvat een literatuuroverzicht van de gepubliceerde studies naar het effect van prostaatkanker screening tot medio 2011. Op basis van het overzicht in hoofdstuk 2 wordt geconcludeerd dat de door screening bereikte daling in prostaatkanker sterfte onvoldoende is om screening naar prostaatkanker te adviseren. Zodoende kan niet worden geadviseerd om prostaatkanker screening op bevolkingsniveau in te voeren. In hoofdstuk 3 worden de voor en nadelen van het meten van trends in prostaatkanker incidentie, overleving en sterfte besproken.

Geïllustreerd wordt dat de trend in prostaatkanker incidentie geen goede uitkomstmaat is om de mogelijke invloed van verschillende omgevingsfactoren op het krijgen van prostaatkanker te meten. Geïllustreerd wordt dat de verandering in de prostaatkanker incidentie volledig beïnvloed wordt door het toenemend gebruik van de PSA test. Dit maakt de mogelijke invloed van andere factoren op verandering in incidentie onmeetbaar. Op basis van de resultaten in hoofdstuk 2 zijn de trends in prostaatkanker overleving een slechte maat om de verandering in de agressiviteit en de behandeling van prostaatkanker te meten. Zoals geïllustreerd in hoofdstuk 3 zijn de trends in prostaatkanker sterfte de enige uitkomstmaat waarmee het effect van de veranderingen in preventieve maatregelen en de behandelingen van prostaatkanker op de bevolking kan worden gemeten.

Hoofdstuk 5 en 6 geeft inzicht in het effect van prostaatkanker screening op de prostaatkanker incidentie en sterfte. In hoofdstuk 5 worden twee populaties vergeleken, 133287 mannen in Noord Ierland en 11970 mannen in Nederland. De mannen in Nederland kregen om de 4 jaar een PSA-bepaling aangeboden (screening), mannen in Noord Ierland niet. Tijdens de follow-up (gemiddeld 9 jaar) werd er bij 1153 (9.6%) mannen prostaatkanker gevonden in de screeningsgroep (Nederland), versus 3962 (3.6%) in de controlegroep (Noord Ierland). Tijdens de onderzoeksperiode overleden 35 (0.29%) mannen in de screeningsgroep aan prostaatkanker, 627 (0.47%) mannen overleden aan prostaatkanker in de controlegroep. Het relatieve risico om te overlijden aan prostaatkanker in de screeningsgroep ten opzichte van dat in de controlegroep was RR: 0,63 (95%CI 0,45-0,88), een reductie van 37% in prostaatkanker specifieke sterfte als een gevolg van screening. Hoofdstuk 6 is een aanvulling op hoofdstuk 5. In tegenstelling tot de resultaten in hoofdstuk 5 is het hier gelukt om een groep mannen te identificeren die relatief het meeste voordeel hebben van prostaatkanker screening. Daarnaast is er een groep mannen geïdentificeerd bij wie de nadelen van screening zo groot zijn dat verdere screening moet worden afgeraden. Zo wordt in hoofdstuk 6 geillustreerd dat bij mannen van 55-74 jaar die een lage PSA-waarde hebben (tussen de 0.1 en 1.9 ng/ml) de nadelen van herhaald screening aanzienlijk groter zijn dan de voordelen. Op basis van deze resultaten van hoofdstuk 6 kan worden geconcludeerd dat als het initiele PSA (eerste meting) zou worden gehanteerd als criteria voor herhaald screenen, er een beter evenwicht kan ontstaan tussen de voor- en nadelen van prostaatkanker screening. Op die manier zouden we kunnen voorkomen dat mannen een behandeling ondergaan die ze eigenlijk niet nodig hebben.

In deel 3 (hoofdstuk 7-10) worden vier studies gepresenteerd die het effect van prostaatkanker screening op de “excess mortality” onderzoeken. Excess mortality is gedefinieerd als het verschil tussen de waargenomen totale sterfte in patiënten met prostaatkanker en de verwachte sterfte op basis van een gelijke groep in de bevolking naar leeftijd en geslacht. In hoofdstuk 7 wordt aangetoond dat het mogelijk is om een

betrouwbare "excess mortality" analyse te maken binnen een gerandomiseerde studie naar het effect van prostaatkanker screening. De resultaten betreffende de effectiviteit van prostaatkanker screening op de excess mortality volgen in hoofdstuk 8-10. Hoofdstuk 8 betreft de analyse van totaal 42376 mannen, leeftijd 55-74 jaar, gerandomiseerd (1:1) in de ERSPC-Nederland. De resultaten tonen een ziekte specifieke sterfte van 0.42 man per 1000 persoonsjaren in de screening groep en 0.48 man per 1000 persoonsjaren in de controle groep na gemiddelde follow-up van 9 jaar, RR: 0.86 (95%CI, 0.64-1.17), een reductie van 14%. De excess mortality was 0.40 per 1000 persoonsjaren in de screening groep en 0.61 man per 1000 persoonsjaren in de controle groep: RR 0.66 (95%CI, 0.39-1.13), een reductie van 34%. Hiermee was het effect van screening meer dan twee keer zo groot op de excess mortality ten opzichte van de ziekte specifieke sterfte. De conclusies en mogelijke verklaringen gegeven in hoofdstuk 8 waren: ten eerste, er is sprake van een systematische onderschatting van de prostaatkanker specifieke sterfte door de "cause of death committee" van de studie; ten tweede, er is sprake van een prostaatkanker gerelateerde sterfte die gemeten kan worden met een excess mortality analyse maar niet met de ziekte specifieke mortaliteit analyse. Deze observaties en suggesties zijn aanvullend bestudeerd in hoofdstuk 10. In deze case-control studie is de mogelijke aanvullende prostaatkanker gerelateerde sterfte onderzocht. Op grond van de resultaten in hoofdstuk 10 blijkt dat mannen met klinisch gediagnosticeerd prostaatkanker, naast een verhoogde kans om te overlijden aan prostaatkanker, een verhoogde kans hebben om te overlijden aan andere doodsoorzaken ten opzichte van met mannen die zijn gediagnosticeerd met prostaatkanker als een gevolg van screening. In hoofdstuk 10 hadden mannen met klinisch gediagnosticeerd prostaatkanker een verhoogd risico om te overlijden aan een aan andere vorm van kanker, aan hart-en-vaat ziektes en longziektes. Een verklaring voor deze bevinding is het gebruik van de hormonale therapie bij mannen met klinisch gediagnosticeerd prostaatkanker. Daarnaast wordt er mogelijk een overlevingsvoordeel gezien bij mannen met een screen-detected prostaatkanker doordat zij op relatief jonge leeftijd (na vroege diagnose prostaatkanker) een verandering maken in hun gezondheid (bijkomend positief effect diagnose prostaatkanker). In hoofdstuk 9 omvat de excess mortality analyse in de vier grootste centra (Finland, Italië, Zweden en Nederland) van de ERSPC. Hierbij zijn in totaal 141578 mannen, leeftijd 55-69 jaar, gerandomiseerd tussen screening en de reguliere gezondheidszorg. Na een gemiddelde follow-up van negen jaar was de prostaatkanker incidentie 9.25 per 1000 persoonsjaren in de screening groep en 5.49 per 1000 persoonsjaren in de controle groep, RR 1.69 (95%CI 1.62-1.76). De excess mortality rate was 0.29 per 1000 persoonsjaren in de screening groep en 0.37 man per 1000 persoonsjaren in de controle groep; de RR voor excess mortality was 0.77 (95%CI, 0.54-1.08), reductie 23%. Op grond van deze resultaten is geconcludeerd dat het verschil in excess mortality tussen de twee studiearmen goed overeenkomt met de eerder aangetoonde 20% reductie in de ziekte

specifieke sterfte. De bevindingen bevestigen dat PSA screening de sterfte aan prostaatkanker kan reduceren. Verder toont het aan dat een excess mortality analyse, mits op de door in hoofdstuk 7-9 beschreven methode, uitgevoerd kan worden gebruikt om het effect van screening op kanker te onderzoeken.

In deel 4 van het proefschrift worden er drie belangrijke deelvragen binnen de effectiviteit van prostaatkanker screening beantwoord. Allereerst is er in hoofdstuk 11 gekeken naar de waarde van het screeningsinterval (de tijd tussen twee screenings onderzoeken). Vergeleken is het aantal gediagnosticeerde agressieve prostaatkanker (kankers relatief te laat gediagnosticeerd met een slechte prognose) wanneer een screeningsinterval van 2 of 4 jaar wordt gebruikt. Resultaten in hoofdstuk 11 laten zien dat met beide screeningsintervallen een significante reductie in het aantal agressieve prostaatkanker kan worden bereikt. Echter, met een programma dat een 2-jaars screenings interval hanteert, daalt de incidentie van het aantal agressieve kankers met nog eens met 43% ten opzichte van een programma met een 4-jaars interval. Het nadeel is wel dat het programma met het kortere interval in totaal meer kankers diagnosticeert en daarmee de overdiagnose verder laat stijgen.

In hoofdstuk 12 is de waarde van het aantal prostaatbiopten bestudeerd in mannen met een groter prostaatvolume. De hypothese was dat in mannen met een groter prostaatvolume meer biopten zouden moeten worden genomen omdat anders kankers in metname grotere prostaten gemist zouden worden. Resultaten laten zien dat mannen met een initieel verhoogde PSA en mannen met een relatief kleiner prostaatvolume een verhoogd risico hebben op de diagnose van een agressief prostaatkanker gedurende de follow-up. Deze bevindingen tonen aan dat in een screeningsprogramma het aantal prostaatbiopten niet moet worden aangepast naar het totale prostaatvolume.

In hoofdstuk 13 is een risico stratificatie gemaakt bedoeld voor mannen met prostaatkanker die geïnformeerd willen worden over hun prognose indien gekozen wordt voor behandeling middels chirurgie (radicale prostatectomie). In totaal werden 1043 mannen, in de leeftijd van 55 tot 75 jaar, gescreend op prostaatkanker en gediagnosticeerd met een gelokaliseerde vorm van prostaatkanker. De resultaten laten zien dat mannen die behandeld zijn met een radicale prostatectomie voor een bij screening gediagnosticeerd gelokaliseerd prostaatkanker met een Gleason score ≤6, 7 en ≥8, respectievelijk een 1.3%, 2.7% of 10.3% kans hebben om te overlijden aan prostaatkanker in een periode van 12 jaar na diagnose. Het relatieve risico om te overlijden aan prostaatkanker, ten opzichte van het risico om te overlijden aan een andere doodsoorzaak, is het hoogst voor de relatief jonge man (55-60 jaar) met een relatef hoge Gleason score (Gleason score ≥8). Geconcludeerd is dat mannen met een bij screening gediagnosticeerd gelokaliseerd prostaatkanker Gleason score ≤6 die behandeld zijn met een radicale prostatectomie een minimale kans hebben om te overlijden aan prostaatkanker in een

periode van 12 jaar na diagnose. Echter, van de mannen met gelokaliseerd prostaatkanker Gleason score ≥8 die overlijden in de periode van 12 jaar na diagnose, overlijden 1 op de 2.2 mannen alsnog aan de gevolgen van prostaatkanker ondanks de screening en behandeling met radicale chirurgie.

De resultaten uit deel 1-4 zijn besproken in deel 5. Hierbij zijn meerdere vergelijkingen gemaakt met de recente literatuur. Dit heeft geresulteerd in verscheidende conclusies evenals suggesties voor toekomstig onderzoek.

List of authors

G. Aus (Gunnar)
Associate Professor of Urology
Department of Urology
Sahlgrenska University Hospital
Gothenburg, Sweden

A. Auvinen (Ansi)
Professor of Epidemiology
Deparment of Public Health
University of Tampere
Tampere, Finland

C.H. Bangma (Chris)
Professor of Urology
Department of Urology
Erasmus University Medical Center
Rotterdam, the Netherlands

R.C.N. van den Bergh (Roderick)
PhD researcher
Department of Urology and Public Health
Erasmus University Medical Center
Rotterdam, the Netherlands

A. Black (Amanda)
Epidemiologist
National Cancer Institute
Rockville, United States

M. Bul (Meelan)
PhD researcher
Department of Urology
Erasmus University Medical Center
Rotterdam, the Netherlands

S. Carlsson (Sigrid)
PhD researcher
Department of Urology
Sahlgrenska University Hospital
Gothenburg, Sweden

S. Ciatto (Stefano)
Professor Radiological Oncology
Department of DMI
Institute for Cancer Prevention
Florence, Italy

D. Connolly (David)
Consultant of Urology
Department of Urology
Belfast City Hospital
Belfast, Northern Ireland

L. Denis (Louis)
Professor of Urology
Department of Urology
Oncology Centre Antwerp
Antwerp, Belgium

A. Gavin (Anna)
Director Northern Ireland Cancer Registry
Northern Ireland Cancer Registry
Belfast, Northern Ireland

T. Hakulinen (Timo)
Professor of Epidemiology
Finnish Cancer Registry
Helsinki, Finland

J. Hugosson (Jonas)
Professor of Urology
Department of Urology
Sahlgrenska University Hospital
Gothenburg, Sweden

H.J. de Koning (Harry)
Professor of Public Health
Department of Public Health
Erasmus University Medical Center
Rotterdam, the Netherlands

R. Kranse (Ries)
Statistician Engineer
Rotterdam Cancer Registry
Rotterdam, the Netherland

P.J. van Leeuwen (Pim)
PhD researcher
Department of Urology
Erasmus University Medical Center
Rotterdam, the Netherlands

S.M. Moss (Sue)
Epidemiologist
Cancer Screening Evaluation Unit
The Institute of Cancer Research
London, United Kingdom

M. Kwiatkowski
Professor of Urology
Department of Urology
Urologische Klinik Aarau
Aurau, Switzerland

S.J. Otto (Suzie)
Postdoc researcher
Department of Public Health
Erasmus University Medical Center
Rotterdam, the Netherlands

M.J. Roobol (Monique)
Epidemiologist
Department of Urology
Erasmus University Medical Center
Rotterdam, the Netherlands

F.H. Schröder (Fritz)
Professor of Urology
Department of Urology
Erasmus University Medical Center
Rotterdam, the Netherlands

E.W. Steyerberg (Ewout)
Professor of Medical Decision Making
Department of Public Health
Erasmus University Medical Center
Rotterdam, the Netherlands

T.L. Tammela (Teuvo)
Professor of Urology
Department of Urology
Tampere University Hospital
Tampere, Finland

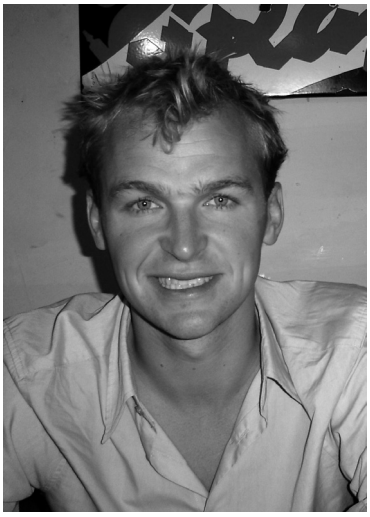
H.A. van Vugt (Heidi)
PhD researcher
Department of Urology and Public Health
Erasmus University Medical Center
Rotterdam, the Netherlands

T. Wolters (Tineke)
PhD researcher
Department of Urology and Pathology
Erasmus University Medical Center
Rotterdam, the Netherlands

M. Zappa (Marco)
Professor of Epidemiology
Unit of Epidemiology
Institute for Cancer Prevention
Florence, Italy

Y.Zhu (Xiaoye)
PhD researcher
Department of Urology
Erasmus University Medical Center
Rotterdam, the Netherlands

Curriculum Vitae



Pim Johannes van Leeuwen, the author of this thesis, was born on the 18th of November 1980 in Heemstede, the Netherlands. After he attended the Vrije School the Hague, he finished pre-university education in 2000. He continued to study Medicine at the Erasmus University Rotterdam the same year. In 2004, he finished his master after completing his master thesis at the Department Pediatric Surgery of the Sophia Children Hospital Rotterdam. In 2005 he started his internships. His final internship he attended at the department Urology of the VUMC, Amsterdam. After he completed his medical degree he worked as a resident at the Department of Urology under supervision of Prof.dr. C.H. Bangma. In April 2008 he met Prof. dr. F.H. Schröder which was the start of his PhD research project on the value of early detection of prostate cancer. In January 2011 he started his residency Urology by first attending two years general surgery in the Maasstad hospital, Rotterdam. In 2013 he will continue his residency at the Department of Urology of the Sint Franciscus Gasthuis Rotterdam, followed by two years at the Department of Urology of the Erasmus Medical Centre Rotterdam. When he is not working, Pim likes to spend time with his friends and family. He likes sport; he plays soccer, often goes running, cycling and likes to spend time on a surf or snowboard. Further, he visits often the theatre, is acting in plays, and likes to cook a special diner.

List of Publications

1. Wat zijn de langetermijneffecten van botulinen toxine-A bij overactieve blaas? van Leeuwen PJ, Blok BFM. Urologen Vademecum. 2008 Apr; Huisartsen Vademecum. 2008 Aug.
2. Pharmacokinetics and analgesic effects of intravenous propacetamol vs rectal paracetamol in children after major craniofacial surgery. Prins SA, Van Dijk M, van Leeuwen PJ, Searle S, Anderson BJ, Tibboel D, Mathot RA. Paediatr Anaesth. 2008 Jul.
3. The Prostate Cancer Prevention Trial and European Randomized Study of Screening for Prostate Cancer risk calculators indicating a positive prostate biopsy: a comparison. van den Bergh RC, Roobol MJ, Wolters T, van Leeuwen PJ, Schröder FH. BJU Int. 2008 Nov.
4. Screening and prostate-cancer mortality in a randomized European study. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Määttänen L, Bangma CH, Aus G, Villers A, Rebillard X, van der Kwast T, Blijenberg BG, Moss SM, de Koning HJ, Auvinen A; ERSPC Investigators. N Engl J Med. 2009 Mar.
5. Screening: should more biopsies be taken in larger prostates? van Leeuwen PJ, van den Bergh R, Wolters T, Schröder F, Roobol M. BJU Int. 2009 May.
6. Should Pathologists Routinely Report Prostate Tumour Volume? The Prognostic Value of Tumour Volume in Prostate Cancer. Wolters T, Roobol MJ, van Leeuwen PJ, van den Bergh RC, Hoedemaeker RF, van Leenders GJ, Schröder FH, van der Kwast TH. Eur Urol. 2009 Jul.
7. Prostate cancer mortality in screen and clinically detected prostate cancer: Estimating the screening benefit. van Leeuwen PJ, Connolly D, Gavin A, Roobol MJ, Black A, Bangma CH, Schröder FH. Eur J Cancer. 2009 Oct.
8. Metastasis-free Survival in Screen and Clinically Detected Prostate Cancer: A Comparison between the European Randomized Study of Screening for Prostate Cancer and Northern Ireland. van Leeuwen PJ, Connolly D, Gavin A, Napolitano G, Roobol MJ and Schröder FH. AUAnews. 2009 Nov.
9. Eleven-Year Outcome of Patients with Prostate Cancers Diagnosed During Screening After Initial Negative Sextant Biopsies. Schröder FH, van den Bergh RC, Wolters T, van Leeuwen PJ, Bangma CH, van der Kwast TH, Roobol MJ. Eur Urol. 2009 Nov.
10. Is delayed radical prostatectomy in men with low-risk screen-detected prostate cancer associated with a higher risk of unfavorable outcomes? van den Bergh RC, Steyerberg EW, Khatami A, Aus G, Pihl CG, Wolters T, van Leeuwen PJ, Roobol MJ, Schröder FH, Hugosson J. Cancer. 2010 Jan.
11. Early detection and screening for prostate cancer. van Leeuwen PJ, Roobol M, Schröder FH. Evidence Based Urology. 2010 Jan.

12. Prostaatkanker screening: Huidige inzichten en aanbevelingen voor de toekomst. van Leeuwen PJ, Roobol MJ. NTVU. 2010 Feb.
13. Words of wisdom. Re: Comparative Effectiveness of Minimally Invasive vs Open Radical Prostatectomy. van Leeuwen PJ. Eur Urol. 2010 Mar.
14. Impact of PSA testing and prostatic biopsy on cancer incidence and mortality: comparative study between the Republic of Ireland and Northern Ireland. Carsin AE, Drummond FJ, Black A, van Leeuwen PJ, Sharp L, Murray LJ, Connolly D, Egevad L, Boniol M, Autier P, Comber H, Gavin A. Cancer Causes Control. 2010 Jun.
15. Risk factors: Increased risk of suicide after prostate cancer diagnosis. van Leeuwen PJ, Schröder FH. Nat Rev Urol. 2010 Jul.
16. Performance of Prostate Cancer Antigen 3 (PCA3) and Prostate-Specific Antigen in Prescreened Men: Reproducibility and Detection Characteristics for Prostate Cancer Patients with High PCA3 Scores (≥100). Roobol MJ, Schröder FH, van Leenders GL, Hessels D, van den Bergh RC, Wolters T, van Leeuwen PJ. Eur Urol. 2010 Sep.
17. The implementation of screening for prostate cancer. van Leeuwen PJ, van Vugt HA, Bangma CH. Prostate Cancer Prostatic Dis. 2010 Sep.
18. Daily, monthly and seasonal variation in PSA levels and the association with weather parameters. Connolly D, van Leeuwen PJ, Bailie J, Black A, Murray LJ, Keane PF, Gavin A. Prostate Cancer Prostatic Dis. 2010 Oct.
19. Performance of the prostate cancer antigen 3 (PCA3) gene and prostate-specific antigen in prescreened men: exploring the value of PCA3 for a first-line diagnostic test. Roobol MJ, Schröder FH, van Leeuwen PJ, Wolters T, van den Bergh RC, van Leenders GJ, Hessels D. Eur Urol. 2010 Oct.
20. Balancing the harms and benefits of early detection of prostate cancer. van Leeuwen PJ, Connolly D, Tammela TL, Auvinen A, Kranse R, Roobol MJ, Schroder FH, Gavin A. Cancer. 2010 Oct.
21. A Critical Analysis of the Tumor Volume Threshold for Clinically Insignificant Prostate Cancer Using a Data Set of a Randomized Screening Trial. Wolters T, Roobol MJ, van Leeuwen PJ, Bergh RC, Hoedemaeker RF, Leenders GJ, Schröder FH, Kwast TH. J Urol. 2010 Nov.
22. Blinded and uniform causes of death verification in cancer screening: a major influence on the outcome of a prostate cancer screening trial? Otto SJ, van Leeuwen PJ, Hoekstra JW, Merckelbach JW, Blom JH, Schröder FH, Roobol MJ, de Koning HJ. Eur J Cancer. 2010 Nov.
23. Prostate Cancer Detection and Dutasteride: Utility and Limitations of Prostate-Specific Antigen in Men with Previous Negative Biopsies. van Leeuwen PJ, Kölble K, Huland H, Hambrock T, Barentsz J, Schröder FH. Eur Urol. 2010 Dec.
24. Disease specific mortality may underestimate the total effect of prostate cancer screening. van Leeuwen PJ, Kranse R, Hakulinen T, de Koning H, Roobol MJ, Bangma CH, Schröder FH. J Med Screen. 2010 Dec.
25. Identifying and characterizing "escapes"-men who develop metastases or die from prostate cancer despite screening (ERSPC, section Rotterdam). Zhu X, van Leeuwen PJ, Bul M, Bangma CH, Roobol MJ, Schröder FH. Int J Cancer. 2011 Jan.

26. Prostate Cancer Incidence and Disease-Specific Survival of Men with Initial Prostate-Specific Antigen Less Than 3.0ng/ml Who Are Participating in ERSPC Rotterdam. Bul M, van Leeuwen PJ, Zhu X, Schröder FH, Roobol MJ. Eur Urol. 2011 Jan.
27. Reply to Balancing the harms and benefits of early detection of prostate cancer. van Leeuwen PJ, Connolly D. Cancer. 2011 Feb.
28. Updating the prostate cancer risk indicator for contemporary biopsy schemes. Bul M, Delongchamps NB, Steyerberg EW, de la Roza G, van Leeuwen PJ, Zhu X, van Vugt HA, Haas GP, Schroder FH, Roobol MJ. Can J Urol. 2011 Apr.
29. 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. Zhu X, van Leeuwen PJ, Bul M, Otto SJ, de Koning HJ, Bangma CH, Schröder FH, Roobol MJ. Eur Urol. 2011 Aug.
30. Critical assessment of prebiopsy parameters for predicting prostate cancer metastasis and mortality. van Leeuwen PJ, Kranse R, van den Bergh RCN, Wolters T, Schröder FH, Roobol MJ. Can J Urol. 2011 Dec.
31. Comparison of incidentally detected prostate cancer with screen-detected prostate cancer treated by prostatectomy. Wolters T, Montironi R, Mazzucchelli R, Scarpelli M, Roobol MJ, van den Bergh RC, van Leeuwen PJ, Hoedemaeker RF, van Leenders GJ, Schröder FH, van der Kwast TH. Prostate. 2012 Jan.
32. Increased non-prostate cancer death risk in clinically diagnosed prostate cancer. van Leeuwen PJ, Otto SJ, Kranse R, Roobol MJ, Bul M, Zhu Y, de Koning H and Schröder FH. BJU Int. 2012 Jan.
33. Towards an optimal interval for prostate cancer screening. van Leeuwen PJ, Roobol MJ, Kranse R, Zappa M, Carlsson S, Bangma CH, Schröder FH and Hugosson J. Eur Urol. 2012 Jan.
34. Prostate cancer screening has no effect on prostate cancer specific mortality over 20 years of follow-up of Swedish men. van Leeuwen PJ. Evid Based Med. 2012 Feb.
35. Efficacy vs. effectiveness trial: consequences in prostate cancer incidence, overall mortality and disease-specific mortality within a multicenter randomized controlled screening trial. Zhu Y, van Leeuwen PJ, Holmberg E, Bul M, Carlson S, Schröder FH, Roobol MJ, Hugosson J. Submitted.
36. Excess mortality, attendance matters. Kranse R, van Leeuwen PJ, Hakulinen T, Hugosson J, Tammela TL, Ciatto S, Roobol MJ, Zappa M, Aus G, de Koning HJ, Bangma CH, Moss SM, Auvinen A, and Schröder FH. Submitted.
37. Prostate cancer screening reduces the excess mortality. van Leeuwen PJ, Kranse R, Hakulinen T, Hugosson J, Tammela TL, Ciatto S, Roobol MJ, Zappa M, Aus G, de Koning HJ, Bangma CH, Moss SM, Auvinen A, and Schröder FH. Submitted.
38. Prostate Cancer Mortality after Screening: Risk Analysis of Men Managed by Radical Surgery for Clinically Localized Prostate Cancer. van Leeuwen PJ, Kranse R, Hakulinen T, Hugosson J, Tammela TL, Ciatto S, Roobol MJ, Zappa M, Aus G, de Koning HJ, Bangma CH, Moss SM, Auvinen A, and Schröder FH. Submitted.

Acknowledgements

Looking back, I am surprised and at the same time very grateful for all I have received throughout these years. A thesis like this you never perform on your own. First and foremost, I would like to thank the more than 162,000 participants of the European Randomized Study of Screening for Prostate Cancer. Furthermore, there are a number of people in particular that I would like to thank for providing invaluable help in my research work.

I would like to express my sincere gratitude to Professor Chris Bangma for giving me the opportunity to carry out this PhD research project. Thank you for your support and sharing your vast knowledge and experience in the broad field of prostate cancer research. Besides being a talented and internationally recognized professor, you have been a dedicated supervisor who is distinctive in his gratitude and sincerity.

Next, I am extremely grateful to Professor Fritz Schröder without whom this thesis would certainly never have been completed, nor even started. With enthusiasm, energy and the most inspiring clear-sighted and inquiring mind, you made me derive enormous pleasure from my research during these years. It is known that you are a talented urologist and researcher; you are equally talented in cooperating smoothly with the new generation of researchers. It has been a privilege to have had you as supervisor. After I have defended my thesis, I would like to invite you for a traditional Kazakhstan drink.

I am most thankful to Ries Kranse, my statistical father, coach and supervisor, from whom I have learned most, if not all, of what I know today about the statistics of clinical trials. Ries, I sincerely enjoyed (almost always) working together with you and I will always remain thankful for the knowledge and insights you have provided me. To make any difference in medical research, we need more people like you who have the talent to think out of the box or against the grain. I look forward to collaborating with you again in the near future.

Next I am very thankful to Monique Roobol, one of my supervisors and the study coordinator of the ERSPC section Rotterdam. You are a keen scientist, I admire your perseverance. You have shown to me that it is possible to be both a hard working associate professor and a careful mother to your own children as well as your 'PhD-kids'. It is only a matter of time before you will become professor in the early detection of prostate cancer.

I would like to take this opportunity to also thank Professor Timo Hakulinen for his support and advice on the excess mortality analyses. I admire your critical view; the finest combination between precision and kindness. Thank you for sharing your knowledge and experience with me.

I would like to mention a special thanks to Professor Jonas Hugosson. You are a Professor that has the skills to give the right intellectual challenge to young ambitious people. Thank you for the fruitful discussions, analytical thinking, and sharing your broad experience in prostate cancer research with me. Also many thanks to the other members of the ERSPC study group. I am honored to be part of this unique European research group. Special thanks for Sue Moss, Marco Zappa, Teuvo Tammella, Annsi Auvinen and Sigrid Carlsson, it has been a pleasure to work together with you.

A part of this thesis was performed and written in Belfast, Northern Ireland. My stay and all the work performed over there were invaluable. I am most thankful to Anna Gavin who gave me the opportunity to become part of the Northern Ireland Cancer Registry. Next, many thanks to David Connolly without whom the work in Northern Ireland would not even have started. David, thanks for sharing your knowledge and experience with me, and for the linguistic editing of the papers we made. Finally, I would like to thank all co-workers of the Northern Ireland Cancer Registry. Special thanks go to Collin Fox, Richard Middleton, Gulio Napolitano, Amanda Black, Liam Murray, David Donnelly, Finian Bannon, Susan McGookin, Jackie Kelly, Donna Floyd and Wendy Hamill.

The main part of this thesis was performed and written at the ERSPC office in Rotterdam, the Netherlands. A great support for this thesis was performed by the employees of the screening office. Marlies, Conja, Meavis, Maaike, Lakhsmi and Heleen, I would like to thank you all for the vast support. It would be impossible to finish this thesis without your important work in running the study and collecting the data. During my period as a PhD student, several fellow students worked at the screening office. Meelan, Xiaoye, Heidi, Lionne, Roderick and Tineke, thanks for your friendship and support during work. Especially Tineke and Roderick thanks for the warm welcome in the office, for inspiring me to the research field, and for the fun during our trips around the world with ERSPC-airlines! Other young urologists, especially Rogier, Joost, Niels, Stijn, Jos and Susanne, it is a pleasure to become "urological-colleagues".

I would also like to thank the department of public health Erasmus MC, Rotterdam, with a special mention for Professor Harry de Koning who has had a very important contribution to the ERSPC study. For me, it would be a privilege to get the opportunity to do more research in your department in the coming years. Suzie Otto, Eveline Heijnsdijk and all others of the department of public health Erasmus MC, thanks for the kind cooperation and for always being available for support and advice.

Finishing a thesis is a long journey which makes some personal support essential. Thanks to my friends who have relieved this journey. Special thanks for my close friends and the boys I know as the 'Krijgertjes', 'the Lords', 'the pretty boys of VOC 12', 'the KGB's', and boys who like to visit 'the Wappie's.

Special thanks to my sisters and little brother. Meike, Rosalie and Ewoud, I'm so grateful for being part of a close family. Although we are all going our own way, our affinity has shown to be vast in sharing the essential things in life. Thanks to my parents, because you have been the first and most loyal supporters. Mama, at an event like this it remains even more a grief to have you not in our midst, especially you deserved to be part of this celebration. Daddy, you are doing a wonderful job in the way you are taking care of your children.

Many thanks to the members of the committee for defending my thesis, Professor Bangma, Professor Hakulinen, Professor Kiemeney, Professor de Koning, Professor Kuipers, Professor de Reijcke, Professor Schröder, and Doctor Roobol.

Ries and Ewoud it feels trustworthy that you fellows will support and back me up during the defense of my thesis. I'm looking forward; it will be a remarkable day!

References

1. Walsh PC RA, Vaughan ED, Wein AJ. Diagnosis and Staging of Prostate Cancer. Cambell's Urology. 2003;8:3055-79.
2. Thorpe A, Neal D. Benign prostatic hyperplasia. Lancet. 2003 Apr 19;361(9366):1359-67.
3. Verhamme KM, Dieleman JP, Bleumink GS, van der Lei J, Sturkenboom MC, Artibani W, et al. Incidence and prevalence of lower urinary tract symptoms suggestive of benign prostatic hyperplasia in primary care--the Triumph project. Eur Urol. 2002 Oct;42(4):323-8.
4. Bosch JL, Hop WC, Kirkels WJ, Schroder FH. The International Prostate Symptom Score in a community-based sample of men between 55 and 74 years of age: prevalence and correlation of symptoms with age, prostate volume, flow rate and residual urine volume. Br J Urol. 1995 May;75(5):622-30.
5. Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol. 2007 Mar;18(3):581-92.
6. Bouchardy C, Fioretta G, Rapiti E, Verkooijen HM, Rapin CH, Schmidlin F, et al. Recent trends in prostate cancer mortality show a continuous decrease in several countries. Int J Cancer. 2008 Jul 15;123(2):421-9.
7. SEER. SEER Cancer Statistics Review 1975-2006. http://seercancergov/csr/1975_2006/results_merged/topic_lifetime_riskpdf. 2006.
8. Collin SM, Martin RM, Metcalfe C, Gunnell D, Albertsen PC, Neal D, et al. Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study. Lancet Oncol. 2008 May;9(5):445-52.
9. Bray F, Lortet-Tieulent J, Ferlay J, Forman D, Auvinen A. Prostate cancer incidence and mortality trends in 37 European countries: An overview. Eur J Cancer. 2010 Nov;46(17):3040-52.
10. Dixon H, Scully M, Wakefield M, Murphy M. The prostate cancer screening debate: public reaction to medical controversy in the media. Public Underst Sci. 2009 Jan;18(1):115-28.
11. McNeal JE, Redwine EA, Freiha FS, Stamey TA. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. Am J Surg Pathol. 1988 Dec;12(12):897-906.
12. Collin SM, Metcalfe C, Donovan J, Lane JA, Davis M, Neal D, et al. Associations of lower urinary tract symptoms with prostate-specific antigen levels, and screen-detected localized and advanced prostate cancer: a case-control study nested within the UK population-based ProtecT (Prostate testing for cancer and Treatment) study. BJU Int. 2008 Nov;102(10):1400-6.
13. Osborn JL, Getzenberg RH, Trump DL. Spinal cord compression in prostate cancer. J Neurooncol. 1995;23(2):135-47.

14. Tazi H, Manunta A, Rodriguez A, Patard JJ, Lobel B, Guille F. Spinal cord compression in metastatic prostate cancer. Eur Urol. 2003 Nov;44(5):527-32.
15. Duran I, Tannock IF. Disseminated intravascular coagulation as the presenting sign of metastatic prostate cancer. J Gen Intern Med. 2006 Nov;21(11):C6-8.
16. Lin DW, Porter M, Montgomery B. Treatment and survival outcomes in young men diagnosed with prostate cancer: a Population-based Cohort Study. Cancer. 2009 Jul 1;115(13):2863-71.
17. 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 Nov 15;119(10):2366-71.
18. Jewett H. Significance of the palpable prostatic nodule. J Am Med Assoc. 1956;Mar 10(160(10)):838-9.
19. Smith DS, Catalona WJ. Interexaminer variability of digital rectal examination in detecting prostate cancer. Urology. 1995 Jan;45(1):70-4.
20. Gosselaar C, Kranse R, Roobol MJ, Roemeling S, Schroder FH. The interobserver variability of digital rectal examination in a large randomized trial for the screening of prostate cancer. Prostate. 2008 Apr 11.
21. Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA. 1994 Feb 2;271(5):368-74.
22. Thompson IM, Rounder JB, Teague JL, Peek M, Spence CR. Impact of routine screening for adenocarcinoma of the prostate on stage distribution. J Urol. 1987 Mar;137(3):424-6.
23. Lilja H. Biology of prostate-specific antigen. Urology. 2003 Nov;62(5 Suppl 1):27-33.
24. 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 Apr 25;324(17):1156-61.
25. Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. N Engl J Med. 1987 Oct 8;317(15):909-16.
26. Stamey TA, Kabalin JN, McNeal JE, Johnstone IM, Freiha F, Redwine EA, et al. Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. II. Radical prostatectomy treated patients. J Urol. 1989 May;141(5):1076-83.
27. Benson MC, Whang IS, Pantuck A, Ring K, Kaplan SA, Olsson CA, et al. Prostate specific antigen density: a means of distinguishing benign prostatic hypertrophy and prostate cancer. J Urol. 1992 Mar;147(3 Pt 2):815-6.
28. Benson MC, Whang IS, Olsson CA, McMahon DJ, Cooner WH. The use of prostate specific antigen density to enhance the predictive value of intermediate levels of serum prostate specific antigen. J Urol. 1992 Mar;147(3 Pt 2):817-21.
29. Hodge KK, McNeal JE, Terris MK, Stamey TA. Random systematic versus directed ultrasound guided transrectal core biopsies of the prostate. J Urol. 1989 Jul;142(1):71-4; discussion 4-5.

30. Routh JC, Leibovich BC. Adenocarcinoma of the prostate: epidemiological trends, screening, diagnosis, and surgical management of localized disease. Mayo Clin Proc. 2005 Jul;80(7):899-907.
31. Epstein JI, Lieberman PH. Mucinous adenocarcinoma of the prostate gland. Am J Surg Pathol. 1985 Apr;9(4):299-308.
32. Epstein JI, Woodruff JM. Adenocarcinoma of the prostate with endometrioid features. A light microscopic and immunohistochemical study of ten cases. Cancer. 1986 Jan 1;57(1):111-9.
33. Sawczuk I, Tannenbaum M, Olsson CA, deVere White R. Primary transitional cell carcinoma of prostatic periurethral ducts. Urology. 1985 Apr;25(4):339-43.
34. Tetu B, Ro JY, Ayala AG, Johnson DE, Logothetis CJ, Ordonez NG. Small cell carcinoma of the prostate. Part I. A clinicopathologic study of 20 cases. Cancer. 1987 May 15;59(10):1803-9.
35. Little NA, Wiener JS, Walther PJ, Paulson DF, Anderson EE. Squamous cell carcinoma of the prostate: 2 cases of a rare malignancy and review of the literature. J Urol. 1993 Jan;149(1):137-9.
36. Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol. 1974 Jan;111(1):58-64.
37. Epstein JI. An update of the Gleason grading system. J Urol. 2010 Feb;183(2):433-40.
38. Helpap B, Egevad L. The significance of modified Gleason grading of prostatic carcinoma in biopsy and radical prostatectomy specimens. Virchows Arch. 2006 Dec;449(6):622-7.
39. Gofrit ON, Zorn KC, Steinberg GD, Zagaja GP, Shalhav AL. The Will Rogers phenomenon in urological oncology. J Urol. 2008 Jan;179(1):28-33.
40. Parker C. Active surveillance: towards a new paradigm in the management of early prostate cancer. Lancet Oncol. 2004 Feb;5(2):101-6.
41. van den Bergh RC, Roemeling S, Roobol MJ, Aus G, Hugosson J, Rannikko AS, et al. Outcomes of Men with Screen-Detected Prostate Cancer Eligible for Active Surveillance Who Were Managed Expectantly. Eur Urol. 2008 Sep 17.
42. 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 Jan 1;28(1):126-31.
43. Morgan PB, Hanlon AL, Horwitz EM, Buyyounouski MK, Uzzo RG, Pollack A. Radiation dose and late failures in prostate cancer. Int J Radiat Oncol Biol Phys. 2007 Mar 15;67(4):1074-81.
44. Pollack A, Zagars GK, Starkschall G, Antolak JA, Lee JJ, Huang E, et al. Prostate cancer radiation dose response: results of the M. D. Anderson phase III randomized trial. Int J Radiat Oncol Biol Phys. 2002 Aug 1;53(5):1097-105.
45. Zietman AL, DeSilvio ML, Slater JD, Rossi CJ, Jr., Miller DW, Adams JA, et al. Comparison of conventional-dose vs high-dose conformal radiation therapy in clinically localized adenocarcinoma of the prostate: a randomized controlled trial. JAMA. 2005 Sep 14;294(10):1233-9.

46. Litwin MS, Gore JL, Kwan L, Brandeis JM, Lee SP, Withers HR, et al. Quality of life after surgery, external beam irradiation, or brachytherapy for early-stage prostate cancer. Cancer. 2007 Jun 1;109(11):2239-47.
47. Zelefsky MJ, Kuban DA, Levy LB, Potters L, Beyer DC, Blasko JC, et al. Multi-institutional analysis of long-term outcome for stages T1-T2 prostate cancer treated with permanent seed implantation. Int J Radiat Oncol Biol Phys. 2007 Feb 1;67(2):327-33.
48. Zelefsky MJ, Chan H, Hunt M, Yamada Y, Shippy AM, Amols H. Long-term outcome of high dose intensity modulated radiation therapy for patients with clinically localized prostate cancer. J Urol. 2006 Oct;176(4 Pt 1):1415-9.
49. Nelson B, Kaufman M, Broughton G, Cookson MS, Chang SS, Herrell SD, et al. Comparison of length of hospital stay between radical retropubic prostatectomy and robotic assisted laparoscopic prostatectomy. J Urol. 2007 Mar;177(3):929-31.
50. Kundu SD, Roehl KA, Eggener SE, Antenor JA, Han M, Catalona WJ. Potency, continence and complications in 3,477 consecutive radical retropubic prostatectomies. J Urol. 2004 Dec;172(6 Pt 1):2227-31.
51. Saranchuk JW, Kattan MW, Elkin E, Touijer AK, Scardino PT, Eastham JA. Achieving optimal outcomes after radical prostatectomy. J Clin Oncol. 2005 Jun 20;23(18):4146-51.
52. Penson DF, McLerran D, Feng Z, Li L, Albertsen PC, Gilliland FD, et al. 5-year urinary and sexual outcomes after radical prostatectomy: results from the Prostate Cancer Outcomes Study. J Urol. 2008 May;179(5 Suppl):S40-4.
53. Penson DF, McLerran D, Feng Z, Li L, Albertsen PC, Gilliland FD, et al. 5-year urinary and sexual outcomes after radical prostatectomy: results from the prostate cancer outcomes study. J Urol. 2005 May;173(5):1701-5.
54. Bill-Axelson A, Holmberg L, Ruutu M, Haggman M, Andersson SO, Bratell S, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med. 2005 May 12;352(19):1977-84.
55. Bill-Axelson A, Holmberg L, Filen F, Ruutu M, Garmo H, Busch C, et al. Radical prostatectomy versus watchful waiting in localized prostate cancer: the Scandinavian prostate cancer group-4 randomized trial. J Natl Cancer Inst. 2008 Aug 20;100(16):1144-54.
56. Samson DJ, Seidenfeld J, Schmitt B, Hasselblad V, Albertsen PC, Bennett CL, et al. Systematic review and meta-analysis of monotherapy compared with combined androgen blockade for patients with advanced prostate carcinoma. Cancer. 2002 Jul 15;95(2):361-76.
57. Tannock IF, de Wit R, Berry WR, Horti J, Pluzanska A, Chi KN, et al. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. N Engl J Med. 2004 Oct 7;351(15):1502-12.
58. Habbema JD, van Oortmarssen GJ, van der Maas PJ. Mass screening for cancer: the interpretation of findings and the prediction of effects on morbidity and mortality. Clinics in laboratory medicine. 1982 Sep;2(3):627-38.
59. Pienta KJ. Critical appraisal of prostate-specific antigen in prostate cancer screening: 20 years later. Urology. 2009 May;73(5 Suppl):S11-20.
60. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. Bol Oficina Sanit Panam. 1968 Oct;65(4):281-393.

61. Population cancer screening in Canada: strategic and priorities. Canadian Strategy for Cancer Control. 2002.
62. Lindegren ML, Kobrynski L, Rasmussen SA, Moore CA, Grosse SD, Vanderford ML, et al. Applying public health strategies to primary immunodeficiency diseases: a potential approach to genetic disorders. MMWR Recomm Rep. 2004 Jan 16;53(RR-1):1-29.
63. Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. Bull World Health Organ. 2008 Apr;86(4):317-9.
64. Davidson P GJ. Should mass screening for prostate cancer be introduced at the national level? Copenhagen, WHO Regional Office for Europe. Available at http://wweurowhoint/Document/E82958pdf. 2004.
65. Union E. Council recommendation on cancer screening. Official Journal of the European Union. 2003;L 327/34.
66. Committee UNS. Second report of the UK National Screening Committee. Available at wwwnscnhsuk 2004. 2004.
67. Regelgeving W-e. Wet op het bevolkingsonderzoek. http://wettenoverheidnl. 2009.
68. Schroder FH, Denis LJ, Roobol M, Nelen V, Auvinen A, Tammela T, et al. The story of the European Randomized Study of Screening for Prostate Cancer. BJU Int. 2003 Dec;92 Suppl 2:1-13.
69. Sindhwani P, Wilson CM. Prostatitis and serum prostate-specific antigen. Curr Urol Rep. 2005 Jul;6(4):307-12.
70. Rao AR, Motiwala HG, Karim OM. The discovery of prostate-specific antigen. BJU Int. 2008 Jan;101(1):5-10.
71. 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 Jul 6;294(1):66-70.
72. 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 May 27;350(22):2239-46.
73. Thompson IM, Ankerst DP. Prostate-specific antigen in the early detection of prostate cancer. Cmaj. 2007 Jun 19;176(13):1853-8.
74. Schroder FH, Carter HB, Wolters T, van den Bergh RC, Gosselaar C, Bangma CH, et al. Early detection of prostate cancer in 2007. Part 1: PSA and PSA kinetics. Eur Urol. 2008 Mar;53(3):468-77.
75. Postma R, Schroder FH, van Leenders GJ, Hoedemaeker RF, Vis AN, Roobol MJ, et al. Cancer detection and cancer characteristics in the European Randomized Study of Screening for Prostate Cancer (ERSPC)--Section Rotterdam. A comparison of two rounds of screening. Eur Urol. 2007 Jul;52(1):89-97.
76. Krumholtz JS, Carvalhal GF, Ramos CG, Smith DS, Thorson P, Yan Y, et al. Prostate-specific antigen cutoff of 2.6 ng/mL for prostate cancer screening is associated with favorable pathologic tumor features. Urology. 2002 Sep;60(3):469-73; discussion 73-4.

77. Gohagan JK, Prorok PC, Hayes RB, Kramer BS. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. Control Clin Trials. 2000 Dec;21(6 Suppl):251S-72S.
78. Loeb S, Catalona WJ. What is the role of digital rectal examination in men undergoing serial screening of serum PSA levels? Nat Clin Pract Urol. 2009 Feb;6(2):68-9.
79. Crawford ED, DeAntoni EP, Etzioni R, Schaefer VC, Olson RM, Ross CA. Serum prostate-specific antigen and digital rectal examination for early detection of prostate cancer in a national community-based program. The Prostate Cancer Education Council. Urology. 1996 Jun;47(6):863-9.
80. Schroder FH, van der Maas P, Beemsterboer P, Kruger AB, Hoedemaeker R, Rietbergen J, et al. Evaluation of the digital rectal examination as a screening test for prostate cancer. Rotterdam section of the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst. 1998 Dec 2;90(23):1817-23.
81. Yamamoto T, Ito K, Ohi M, Kubota Y, Suzuki K, Fukabori Y, et al. Diagnostic significance of digital rectal examination and transrectal ultrasonography in men with prostate-specific antigen levels of 4 NG/ML or less. Urology. 2001 Dec;58(6):994-8.
82. Bozeman CB, Carver BS, Caldito G, Venable DD, Eastham JA. Prostate cancer in patients with an abnormal digital rectal examination and serum prostate-specific antigen less than 4.0 ng/mL. Urology. 2005 Oct;66(4):803-7.
83. Andriole GL, Levin DL, Crawford ED, Gelmann EP, Pinsky PF, Chia D, et al. Prostate Cancer Screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: findings from the initial screening round of a randomized trial. J Natl Cancer Inst. 2005 Mar 16;97(6):433-8.
84. Borden LS, Jr., Wright JL, Kim J, Latchamsetty K, Porter CR. An abnormal digital rectal examination is an independent predictor of Gleason > or =7 prostate cancer in men undergoing initial prostate biopsy: a prospective study of 790 men. BJU Int. 2007 Mar;99(3):559-63.
85. Ghavamian R, Blute ML, Bergstralh EJ, Slezak J, Zincke H. Comparison of clinically nonpalpable prostate-specific antigen-detected (cT1c) versus palpable (cT2) prostate cancers in patients undergoing radical retropubic prostatectomy. Urology. 1999 Jul;54(1):105-10.
86. Okotie OT, Roehl KA, Han M, Loeb S, Gashti SN, Catalona WJ. Characteristics of prostate cancer detected by digital rectal examination only. Urology. 2007 Dec;70(6):1117-20.
87. Gosselaar C, Roobol MJ, Roemeling S, de Vries SH, Cruijsen-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 May 1;66(6):625-31.
88. Rodriguez LV, Terris MK. Risks and complications of transrectal ultrasound. Curr Opin Urol. 2000 Mar;10(2):111-6.
89. Djavan B, Waldert M, Zlotta A, Dobronski P, Seitz C, Remzi M, et al. Safety and morbidity of first and repeat transrectal ultrasound guided prostate needle biopsies: results of a prospective European prostate cancer detection study. J Urol. 2001 Sep;166(3):856-60.

90. Villers A, Malavaud B, Rebillard X, Bataille V, Iborra F. ERSPC: features and preliminary results of France. BJU international. 2003 Dec;92 Suppl 2:27-9.
91. Berenguer A, Lujan M, Paez A, Santonja C, Pascual T. The Spanish contribution to the European Randomized Study of Screening for Prostate Cancer. BJU Int. 2003 Dec;92 Suppl 2:33-8.
92. Hugosson J, Aus G, Bergdahl S, Fernlund P, Frosing R, Lodding P, et al. Population-based screening for prostate cancer by measuring free and total serum prostate-specific antigen in Sweden. BJU Int. 2003 Dec;92 Suppl 2:39-43.
93. Kwiatkowski M, Huber A, Stamm B, Lehmann K, Wernli M, Hafeli A, et al. Features and preliminary results of prostate cancer screening in Canton Aargau, Switzerland. BJU Int. 2003 Dec;92 Suppl 2:44-7.
94. Finne P, Stenman UH, Maattanen L, Makinen T, Tammela TL, Martikainen P, et al. The Finnish trial of prostate cancer screening: where are we now? BJU Int. 2003 Dec;92 Suppl 2:22-6.
95. Ciatto S, Gervasi G, Frullini P, Zendron P, Zappa M. Specific features of the Italian section of the ERSPC. BJU Int. 2003 Dec;92 Suppl 2:30-2.
96. Roobol MJ, Schroder FH. European Randomized Study of Screening for Prostate Cancer: achievements and presentation. BJU Int. 2003 Dec;92 Suppl 2:117-22.
97. Andriole GL, Grubb RL, 3rd, Buys SS, Chia D, Church TR, Fouad MN, et al. Mortality Results from a Randomized Prostate-Cancer Screening Trial. N Engl J Med. 2009 Mar 18.
98. 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 Mar 26;360(13):1320-8.
99. van der Cruijsen-Koeter IW, Vis AN, Roobol MJ, Wildhagen MF, de Koning HJ, van der Kwast TH, et al. Comparison of screen detected and clinically diagnosed prostate cancer in the European randomized study of screening for prostate cancer, section rotterdam. J Urol. 2005 Jul;174(1):121-5.
100. Grubb RL, 3rd, Pinsky PF, Greenlee RT, Izmirlian G, Miller AB, Hickey TP, et al. Prostate cancer screening in the Prostate, Lung, Colorectal and Ovarian cancer screening trial: update on findings from the initial four rounds of screening in a randomized trial. BJU Int. 2008 Dec;102(11):1524-30.
101. van der Cruijsen-Koeter IW, Roobol MJ, Wildhagen MF, van der Kwast TH, Kirkels WJ, Schroder FH. Tumor characteristics and prognostic factors in two subsequent screening rounds with four-year interval within prostate cancer screening trial, ERSPC Rotterdam. Urology. 2006 Sep;68(3):615-20.
102. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and Prostate-Cancer Mortality in a Randomized European Study. The New England journal of medicine. 2009 Mar 18.
103. 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 a large screening trial (ERSPC). Int J Cancer. 2009 Sep 8.

104. 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). European urology. 2009 Jul 28.
105. van Leeuwen PJ, Connolly D, Gavin A, Roobol MJ, Black A, Bangma CH, et al. Prostate cancer mortality in screen and clinically detected prostate cancer: Estimating the screening benefit. Eur J Cancer. 2009 Oct 3.
106. Otto SJ, van der Cruijsen 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 Jun 20;105(3):394-9.
107. Ciatto S, Zappa M, Villers A, Paez A, Otto SJ, Auvinen A. Contamination by opportunistic screening in the European Randomized Study of Prostate Cancer Screening. BJU Int. 2003 Dec;92 Suppl 2:97-100.
108. White WM, Sadetsky N, Waters WB, Carroll PR, Litwin MS. Quality of life in men with locally advanced adenocarcinoma of the prostate: an exploratory analysis using data from the CaPSURE database. J Urol. 2008 Dec;180(6):2409-13; discussion 14.
109. 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 Mar 20;358(12):1250-61.
110. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane database of systematic reviews (Online). 2006(4):CD001877.
111. Poplack SP, Tosteson AN, Grove MR, Wells WA, Carney PA. Mammography in 53,803 women from the New Hampshire mammography network. Radiology. 2000 Dec;217(3):832-40.
112. Vejborg I, Olsen AH, Jensen MB, Rank F, Tange UB, Lynge E. Early outcome of mammography screening in Copenhagen 1991-99. J Med Screen. 2002;9(3):115-9.
113. Wang H, Karesen R, Hervik A, Thoresen SO. Mammography screening in Norway: results from the first screening round in four counties and cost-effectiveness of a modeled nationwide screening. Cancer Causes Control. 2001 Jan;12(1):39-45.
114. Institute. NC. Inventory of colorectal cancer screening activities in ICSN countries, June 2006. Available at http://appliedresearchcancergov/icsn/colorectal/screeninghtml. 2006.
115. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database Syst Rev. 2007(1):CD001216.
116. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol. 2008 Jun;103(6):1541-9.
117. Grazzini G, Castiglione G, Ciabattoni C, Franceschini F, Giorgi D, Gozzi S, et al. Colorectal cancer screening programme by faecal occult blood test in Tuscany: first round results. Eur J Cancer Prev. 2004 Feb;13(1):19-26.

118. Pignone M, Saha S, Hoerger T, Mandelblatt J. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2002 Jul 16;137(2):96-104.
119. Goldie SJ, Gaffikin L, Goldhaber-Fiebert JD, Gordillo-Tobar A, Levin C, Mahe C, et al. Cost-effectiveness of cervical-cancer screening in five developing countries. N Engl J Med. 2005 Nov 17;353(20):2158-68.
120. Manser RL, Irving LB, Stone C, Byrnes G, Abramson M, Campbell D. Screening for lung cancer. Cochrane Database Syst Rev. 2004(1):CD001991.
121. Gohagan J, Marcus P, Fagerstrom R, Pinsky P, Kramer B, Prorok P. Baseline findings of a randomized feasibility trial of lung cancer screening with spiral CT scan vs chest radiograph: the Lung Screening Study of the National Cancer Institute. Chest. 2004 Jul;126(1):114-21.
122. Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, Galen B, et al. The National Lung Screening Trial: overview and study design. Radiology. 2011 Jan;258(1):243-53.
123. Laara E, Day NE, Hakama M. Trends in mortality from cervical cancer in the Nordic countries: association with organised screening programmes. Lancet. 1987 May 30;1(8544):1247-9.
124. Sankaranarayanan R, Nene BM, Dinshaw KA, Mahe C, Jayant K, Shastri SS, et al. A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. Int J Cancer. 2005 Sep 10;116(4):617-23.
125. Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N. Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. Health Technol Assess. 2004 May;8(20):iii, 1-78.
126. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. Eur J Cancer. 2010 Mar;46(4):765-81.
127. Bullard J, Coleman MP, Robinson D, Lutz JM, Bell J, Peto J. Completeness of cancer registration: a new method for routine use. Br J Cancer. 2000 Mar;82(5):1111-6.
128. Wells WA, Carney PA, Eliassen MS, Tosteson AN, Greenberg ER. Statewide study of diagnostic agreement in breast pathology. J Natl Cancer Inst. 1998 Jan 21;90(2):142-5.
129. Draisma G, Boer R, Otto SJ, van der Cruijsen IW, Damhuis RA, Schroder FH, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst. 2003 Jun 18;95(12):868-78.
130. Quinn M, Babb P. Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons. BJU Int. 2002 Jul;90(2):162-73.
131. Kleinbaum DG KM. Survival Analysis: A Self-Learning Text. Statistics for Biology and Health. 2005.
132. Klein DG KM. Survival analysis: a self learning text. Springers Publishers, New York. 2005(2).
133. Nystrom L, Larsson LG, Rutqvist LE, Lindgren A, Lindqvist M, Ryden S, et al. Determination of cause of death among breast cancer cases in the Swedish randomized mammography screening trials. A comparison between official statistics and validation by an endpoint committee. Acta Oncol. 1995;34(2):145-52.

134. Parkin DM. The evolution of the population-based cancer registry. Nat Rev Cancer. 2006 Aug;6(8):603-12.
135. Albertsen PC, Walters S, Hanley JA. A comparison of cause of death determination in men previously diagnosed with prostate cancer who died in 1985 or 1995. J Urol. 2000 Feb;163(2):519-23.
136. Percy C, Stanek E, 3rd, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health. 1981 Mar;71(3):242-50.
137. Penson DF, Albertsen PC, Nelson PS, Barry M, Stanford JL. Determining cause of death in prostate cancer: are death certificates valid? J Natl Cancer Inst. 2001 Dec 5;93(23):1822-3.
138. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. Natl Cancer Inst Monogr. 1961 Sep;6:101-21.
139. Berkson J, Gage RP. Calculation of survival rates for cancer. Proc Staff Meet Mayo Clin. 1950 May 24;25(11):270-86.
140. Dickman PW, Hakulinen T. Population-based cancer survival analysis. 2006.
141. Weiss NS, Rossing MA. Healthy screened bias in epidemiologic studies of cancer incidence. Epidemiology. 1996 May;7(3):319-22.
142. Prorok PC, Connor RJ, Baker SG. Statistical considerations in cancer screening programs. Urol Clin North Am. 1990 Nov;17(4):699-708.
143. Makinen T, Karhunen P, Aro J, Lahtela J, Maattanen L, Auvinen A. Assessment of causes of death in a prostate cancer screening trial. Int J Cancer. 2008 Jan 15;122(2):413-7.
144. Newschaffer CJ, Otani K, McDonald MK, Penberthy LT. Causes of death in elderly prostate cancer patients and in a comparison nonprostate cancer cohort. J Natl Cancer Inst. 2000 Apr 19;92(8):613-21.
145. Lenner P. The excess mortality rate. A useful concept in cancer epidemiology. Acta Oncol. 1990;29(5):573-6.
146. Feuer EJ, Merrill RM, Hankey BF. Cancer surveillance series: interpreting trends in prostate cancer--part II: Cause of death misclassification and the recent rise and fall in prostate cancer mortality. J Natl Cancer Inst. 1999 Jun 16;91(12):1025-32.
147. Etzioni R, Legler JM, Feuer EJ, Merrill RM, Cronin KA, Hankey BF. Cancer surveillance series: interpreting trends in prostate cancer--part III: Quantifying the link between population prostate-specific antigen testing and recent declines in prostate cancer mortality. J Natl Cancer Inst. 1999 Jun 16;91(12):1033-9.
148. Barry MJ. Screening for prostate cancer--the controversy that refuses to die. N Engl J Med. 2009 Mar 26;360(13):1351-4.
149. Etzioni R, Tsodikov A, Mariotto A, Szabo A, Falcon S, Wegelin J, et al. Quantifying the role of PSA screening in the US prostate cancer mortality decline. Cancer Causes Control. 2008 Mar;19(2):175-81.
150. Kvale R, Auvinen A, Adami HO, Klint A, Hernes E, Moller B, et al. Interpreting trends in prostate cancer incidence and mortality in the five Nordic countries. J Natl Cancer Inst. 2007 Dec 19;99(24):1881-7.
151. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. Stat Med. 1997 May 15;16(9):1017-29.

152. Connolly D, Black A, Gavin A, Keane PF, Murray LJ. Baseline prostate-specific antigen level and risk of prostate cancer and prostate-specific mortality: diagnosis is dependent on the intensity of investigation. Cancer Epidemiol Biomarkers Prev. 2008 Feb;17(2):271-8.
153. 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 Int. 2003 Dec;92 Suppl 2:71-8.
154. WHO. International statistical classification of diseases and related health problems. Tenth revision. Geneva: WHO. 1992;1.
155. Aalen. Non parametric inference for a family of counting processes. Ann Statist. 1978;6:701-726.
156. Rembold CM. Number needed to screen: development of a statistic for disease screening. Bmj. 1998 Aug 1;317(7154):307-12.
157. Day NE. Cancer Incidence in Five Continents. Cumulative rate and cumulative risk. IARC Sci Publ. 1992(120):862-4.
158. Lu-Yao G, Albertsen PC, Stanford JL, Stukel TA, Walker-Corkery ES, Barry MJ. Natural experiment examining impact of aggressive screening and treatment on prostate cancer mortality in two fixed cohorts from Seattle area and Connecticut. Bmj. 2002 Oct 5;325(7367):740.
159. Bartsch G, Horninger W, Klocker H, Pelzer A, Bektic J, Oberaigner W, et al. Tyrol Prostate Cancer Demonstration Project: early detection, treatment, outcome, incidence and mortality. BJU Int. 2008 Apr;101(7):809-16.
160. Integraalkankercentrumrotterdam. IKCnet. http://wwwikcnetnl/IKR/indexphp. 2009.
161. WHO. Data and Statistics. wwwwhoint. 2008.
162. Otto SJ, Schroder FH, de Koning HJ. Low all-cause mortality in the volunteer-based Rotterdam section of the European randomised study of screening for prostate cancer: self-selection bias? Journal of medical screening. 2004;11(2):89-92.
163. Messing EM, Manola J, Yao J, Kiernan M, Crawford D, Wilding G, et al. Immediate versus deferred androgen deprivation treatment in patients with node-positive prostate cancer after radical prostatectomy and pelvic lymphadenectomy. Lancet Oncol. 2006 Jun;7(6):472-9.
164. 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 Jul 13;360(9327):103-6.
165. 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 Oct;56(4):584-91.
166. Connolly D. Long term follow-up of men with elevated prostate specific antigen levels: a population based study. . MD Thesis Belfast: Queen's University Belfast. 2007.
167. Gann PH, Hennekens CH, Stampfer MJ. A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. JAMA. 1995 Jan 25;273(4):289-94.

168. 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 Sep;58(3):411-6.
169. Loeb S, Roehl KA, Antenor JA, Catalona WJ, Suarez BK, Nadler RB. Baseline prostate-specific antigen compared with median prostate-specific antigen for age group as predictor of prostate cancer risk in men younger than 60 years old. Urology. 2006 Feb;67(2):316-20.
170. Schroder 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 Feb;57(2):256-66.
171. Holmberg L, Bill-Axelson A, Garmo H, Palmgren J, Norlen BJ, Adami HO, et al. Prognostic markers under watchful waiting and radical prostatectomy. Hematol Oncol Clin North Am. 2006 Aug;20(4):845-55.
172. 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 Mar 26;360(13):1320-8.
173. van Leeuwen PJ, Kranse R, Hakulinen T, Roobol MJ, de Koning HJ, Bangma CH, et al. Disease-specific mortality may underestimate the total effect of prostate cancer screening. J Med Screen. 2010;17(4):204-10.
174. Black WC. Overdiagnosis: An underrecognized cause of confusion and harm in cancer screening. J Natl Cancer Inst. 2000 Aug 16;92(16):1280-2.
175. van der Sanden GA, Coebergh JW, Schouten LJ, Visser O, van Leeuwen FE. Cancer incidence in The Netherlands in 1989 and 1990: first results of the nationwide Netherlands cancer registry. Coordinating Committee for Regional Cancer Registries. Eur J Cancer. 1995 Oct;31A(11):1822-9.
176. CBS. http://statline.cbs.nl/StatWeb/publication. CBS. 2009;1.
177. http://www.cbs.nl/nl-NL/menu/cijfers/kerncijfers/default.htm. sterftecijfers Centraal Bureau voor de Statistiek. 2008.
178. Rabe-Hesketh S EB. A handbook of statistical analyses using Stata. Boca Raton: Chapman & Hall. 1998.
179. de Koning HJ, Liem MK, Baan CA, Boer R, Schroder FH, Alexander FE. Prostate cancer mortality reduction by screening: power and time frame with complete enrollment in the European Randomised Screening for Prostate Cancer (ERSPC) trial. Int J Cancer. 2002 Mar 10;98(2):268-73.
180. Buckley JD. Additive and multiplicative models for relative survival rates. Biometrics. 1984 Mar;40(1):51-62.
181. Nijs HG, Essink-Bot ML, DeKoning HJ, Kirkels WJ, Schroder FH. Why do men refuse or attend population-based screening for prostate cancer? J Public Health Med. 2000 Sep;22(3):312-6.
182. Nijs HG, Tordoir DM, Schuurman JH, Kirkels WJ, Schroder FH. Randomised trial of prostate cancer screening in The Netherlands: assessment of acceptance and motives for attendance. J Med Screen. 1997;4(2):102-6.
183. Davison AC. Statistical Models. 2003;pp. 33-35.

184. Oehlert GW. A note on the delta method. The American Statistician. 1992;46, 27–29.
185. Hesterberg T MD, Monaghan S, Clipson A, Epstein R. Bootstrap Methods and Permutation Tests. Statistics. 2005.
186. Postma R, van Leenders AG, Roobol MJ, Schroder FH, van der Kwast TH. Tumour features in the control and screening arm of a randomized trial of prostate cancer. Eur Urol. 2006 Jul;50(1):70-5.
187. Beemsterboer PM, de Koning HJ, Birnie E, van der Maas PJ, Schroder FH. Advanced prostate cancer: course, care, and cost implications. Prostate. 1999 Jul 1;40(2):97-104.
188. Larsson LG, Nystrom L, Wall S, Rutqvist L, Andersson I, Bjurstam N, et al. The Swedish randomised mammography screening trials: analysis of their effect on the breast cancer related excess mortality. J Med Screen. 1996;3(3):129-32.
189. Jonsson H, Tornberg S, Nystrom L, Lenner P. Service screening with mammography of women aged 70-74 years in Sweden. Effects on breast cancer mortality. Cancer Detect Prev. 2003;27(5):360-9.
190. Dubben HH. Trials of prostate-cancer screening are not worthwhile. Lancet Oncol. 2009 Mar;10(3):294-8.
191. Van Hemelrijck M, Garmo H, Holmberg L, Ingelsson E, Bratt O, Bill-Axelson A, et al. Absolute and Relative Risk of Cardiovascular Disease in Men With Prostate Cancer: Results From the Population-Based PCBaSe Sweden. J Clin Oncol. 2010 Jun 21.
192. Fall K, Fang F, Mucci LA, Ye W, Andren O, Johansson JE, et al. Immediate risk for cardiovascular events and suicide following a prostate cancer diagnosis: prospective cohort study. PLoS Med. 2009 Dec;6(12):e1000197.
193. Fang F, Keating NL, Mucci LA, Adami HO, Stampfer MJ, Valdimarsdottir U, et al. Immediate risk of suicide and cardiovascular death after a prostate cancer diagnosis: cohort study in the United States. J Natl Cancer Inst. 2010 Mar 3;102(5):307-14.
194. Thompson MM, Garland C, Barrett-Connor E, Khaw KT, Friedlander NJ, Wingard DL. Heart disease risk factors, diabetes, and prostatic cancer in an adult community. Am J Epidemiol. 1989 Mar;129(3):511-7.
195. Hager M, Mikuz G, Bartsch G, Kolbitsch C, Moser PL. The association between local atherosclerosis and prostate cancer. BJU Int. 2007 Jan;99(1):46-8.
196. Grundmark B, Garmo H, Loda M, Busch C, Holmberg L, Zethelius B. The metabolic syndrome and the risk of prostate cancer under competing risks of death from other causes. Cancer Epidemiol Biomarkers Prev. 2010 Aug;19(8):2088-96.
197. Platz EA, Clinton SK, Giovannucci E. Association between plasma cholesterol and prostate cancer in the PSA era. Int J Cancer. 2008 Oct 1;123(7):1693-8.
198. Platz EA, Leitzmann MF, Visvanathan K, Rimm EB, Stampfer MJ, Willett WC, et al. Statin drugs and risk of advanced prostate cancer. J Natl Cancer Inst. 2006 Dec 20;98(24):1819-25.
199. Buschemeyer WC, 3rd, Freedland SJ. Obesity and prostate cancer: epidemiology and clinical implications. Eur Urol. 2007 Aug;52(2):331-43.
200. van Leeuwen P.J. KR, Hakulinen T, Roobol M.J., de Koning H, Bangma C.H., Schröder F.H. Disease specific mortality may underestimate the total effect of prostate cancer screening. J Med Screen. 2010;in press.

201. Otto SJ, Schroder FH, de Koning HJ. Risk of cardiovascular mortality in prostate cancer patients in the Rotterdam randomized screening trial. J Clin Oncol. 2006 Sep 1;24(25):4184-9.
202. Taylor LG, Canfield SE, Du XL. Review of major adverse effects of androgen-deprivation therapy in men with prostate cancer. Cancer. 2009 Jun 1;115(11):2388-99.
203. Keating NL, O'Malley AJ, Smith MR. Diabetes and cardiovascular disease during androgen deprivation therapy for prostate cancer. J Clin Oncol. 2006 Sep 20;24(27):4448-56.
204. Shahinian VB, Kuo YF, Freeman JL, Goodwin JS. Risk of fracture after androgen deprivation for prostate cancer. N Engl J Med. 2005 Jan 13;352(2):154-64.
205. Gillessen S, Templeton A, Marra G, Kuo YF, Valtorta E, Shahinian VB. Risk of colorectal cancer in men on long-term androgen deprivation therapy for prostate cancer. J Natl Cancer Inst. 2010 Nov 10.
206. Sprague BL, Trentham-Dietz A, Nichols HB, Hampton JM, Newcomb PA. Change in lifestyle behaviors and medication use after a diagnosis of ductal carcinoma in situ. Breast Cancer Res Treat. 2010 Nov;124(2):487-95.
207. Thomson CA, Flatt SW, Rock CL, Ritenbaugh C, Newman V, Pierce JP. Increased fruit, vegetable and fiber intake and lower fat intake reported among women previously treated for invasive breast cancer. J Am Diet Assoc. 2002 Jun;102(6):801-8.
208. Ligibel JA, Partridge A, Giobbie-Hurder A, Golshan M, Emmons K, Winer EP. Physical activity behaviors in women with newly diagnosed ductal carcinoma-in-situ. Ann Surg Oncol. 2009 Jan;16(1):106-12.
209. Walsh RM, Thompson IM. Prostate cancer screening and disease management: how screening may have an unintended effect on survival and mortality-the camel's nose effect. J Urol. 2007 Apr;177(4):1303-6.
210. McAlister FA, Wiebe N, Ezekowitz JA, Leung AA, Armstrong PW. Meta-analysis: beta-blocker dose, heart rate reduction, and death in patients with heart failure. Annals of internal medicine. 2009 Jun 2;150(11):784-94.
211. Leizorovicz A, Lechat P, Cucherat M, Bugnard F. Bisoprolol for the treatment of chronic heart failure: a meta-analysis on individual data of two placebo-controlled studies--CIBIS and CIBIS II. Cardiac Insufficiency Bisoprolol Study. Am Heart J. 2002 Feb;143(2):301-7.
212. Hjalmarson A, Goldstein S, Fagerberg B, Wedel H, Waagstein F, Kjekshus J, et al. Effects of controlled-release metoprolol on total mortality, hospitalizations, and well-being in patients with heart failure: the Metoprolol CR/XL Randomized Intervention Trial in congestive heart failure (MERIT-HF). MERIT-HF Study Group. JAMA. 2000 Mar 8;283(10):1295-302.
213. Afilalo J, Duque G, Steele R, Jukema JW, de Craen AJ, Eisenberg MJ. Statins for secondary prevention in elderly patients: a hierarchical bayesian meta-analysis. Journal of the American College of Cardiology. 2008 Jan 1;51(1):37-45.
214. Kliber A, Lynd LD, Sin DD. The effects of long-acting bronchodilators on total mortality in patients with stable chronic obstructive pulmonary disease. Respir Res. 2010;11:56.

215. Parsons JK, Newman VA, Mohler JL, Pierce JP, Flatt S, Marshall J. Dietary modification in patients with prostate cancer on active surveillance: a randomized, multicentre feasibility study. BJU Int. 2008 May;101(10):1227-31.
216. Hadley CW, Miller EC, Schwartz SJ, Clinton SK. Tomatoes, lycopene, and prostate cancer: progress and promise. Exp Biol Med (Maywood). 2002 Nov;227(10):869-80.
217. Daubenmier JJ, Weidner G, Marlin R, Crutchfield L, Dunn-Emke S, Chi C, et al. Lifestyle and health-related quality of life of men with prostate cancer managed with active surveillance. Urology. 2006 Jan;67(1):125-30.
218. 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 Aug;11(8):725-32.
219. Taylor R, Page A, Bampton D, Estoesta J, Rickard M. Age-specific interval breast cancers in New South Wales and meta-analysis of studies of women aged 40-49 years. J Med Screen. 2004;11(4):199-206.
220. Taylor R, Supramaniam R, Rickard M, Estoesta J, Moreira C. Interval breast cancers in New South Wales, Australia, and comparisons with trials and other mammographic screening programmes. J Med Screen. 2002;9(1):20-5.
221. Day NE, Williams DR, Khaw KT. Breast cancer screening programmes: the development of a monitoring and evaluation system. Br J Cancer. 1989 Jun;59(6):954-8.
222. van der Cruijsen-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 Oct 1;95(19):1462-6.
223. Roobol MJ, Grenabo A, Schroder FH, Hugosson J. Interval cancers in prostate cancer screening: comparing 2- and 4-year screening intervals in the European Randomized Study of Screening for Prostate Cancer, Gothenburg and Rotterdam. J Natl Cancer Inst. 2007 Sep 5;99(17):1296-303.
224. Bulliard JL, Sasieni P, Klabunde C, De Landtsheer JP, Yankaskas BC, Fracheboud J. Methodological issues in international comparison of interval breast cancers. Int J Cancer. 2006 Sep 1;119(5):1158-63.
225. Bray F, Sankila R, Ferlay J, Parkin DM. Estimates of cancer incidence and mortality in Europe in 1995. Eur J Cancer. 2002 Jan;38(1):99-166.
226. Breslow NE DN. Statistical methods in cancer research. IARC Sci Publ. 1987;2(The design and analysis of cohort studies):1-406.
227. Roobol MJ, Schroder FH, Crawford ED, Freedland SJ, Sartor AO, Fleshner N, et al. A Framework for the Identification of Men at Increased Risk for Prostate Cancer. J Urol. 2009 Sep 14.
228. 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. 2009 Sep 1;57(1):79-85.
229. Carter HB, Epstein JI, Chan DW, Fozard JL, Pearson JD. Recommended prostate-specific antigen testing intervals for the detection of curable prostate cancer. Jama. 1997 May 14;277(18):1456-60.

230. Ito K, Yamamoto T, Ohi M, Takechi H, Kurokawa K, Suzuki K, et al. Possibility of re-screening intervals of more than one year in men with PSA levels of 4.0 ng/ml or less. Prostate. 2003 Sep 15;57(1):8-13.
231. Roobol MJ, Roobol DW, Schroder FH. Is additional testing necessary in men with prostate-specific antigen levels of 1.0 ng/mL or less in a population-based screening setting? (ERSPC, section Rotterdam). Urology. 2005 Feb;65(2):343-6.
232. van Leeuwen PJ, Connolly D, Tammela TL, Auvinen A, Kranse R, Roobol MJ, et al. Balancing the harms and benefits of early detection of prostate cancer. Cancer. 2010 Sep 13.
233. Vickers AJ, Cronin AM, Bjork T, Manjer J, Nilsson PM, Dahlin A, et al. Prostate specific antigen concentration at age 60 and death or metastasis from prostate cancer: case-control study. BMJ. 2010;341:c4521.
234. Presti JC, Jr., Chang JJ, Bhargava V, Shinohara K. The optimal systematic prostate biopsy scheme should include 8 rather than 6 biopsies: results of a prospective clinical trial. J Urol. 2000 Jan;163(1):163-6; discussion 6-7.
235. Gore JL, Shariat SF, Miles BJ, Kadmon D, Jiang N, Wheeler TM, et al. Optimal combinations of systematic sextant and laterally directed biopsies for the detection of prostate cancer. J Urol. 2001 May;165(5):1554-9.
236. Karakiewicz PI, Aprikian AG, Meshref AW, Bazinet M. Computer-assisted comparative analysis of four-sector and six-sector biopsies of the prostate. Urology. 1996 Nov;48(5):747-50.
237. Karakiewicz PI, Hanley JA, Bazinet M. Three-dimensional computer-assisted analysis of sector biopsy of the prostate. Urology. 1998 Aug;52(2):208-12.
238. Schroder FH, Denis LJ, Roobol M, Nelen V, Auvinen A, Tammela T, et al. The story of the European Randomized Study of Screening for Prostate Cancer. BJU Int. 2003 Dec;92 Suppl 2:1-13.
239. Briganti A, Chun FK, Suardi N, Gallina A, Walz J, Graefen M, et al. Prostate volume and adverse prostate cancer features: fact not artifact. Eur J Cancer. 2007 Dec;43(18):2669-77.
240. Freedland SJ, Isaacs WB, Platz EA, Terris MK, Aronson WJ, Amling CL, et al. Prostate size and risk of high-grade, advanced prostate cancer and biochemical progression after radical prostatectomy: a search database study. J Clin Oncol. 2005 Oct 20;23(30):7546-54.
241. Kassouf W, Nakanishi H, Ochiai A, Babaian KN, Troncoso P, Babaian RJ. Effect of prostate volume on tumor grade in patients undergoing radical prostatectomy in the era of extended prostatic biopsies. J Urol. 2007 Jul;178(1):111-4.
242. Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. Biometrics. 2008 Mar;64(1):10-9.
243. Remzi M, Fong YK, Dobrovits M, Anagnostou T, Seitz C, Waldert M, et al. The Vienna nomogram: validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. J Urol. 2005 Oct;174(4 Pt 1):1256-60; discussion 60-1; author reply 61.

244. Durkan GC, Sheikh N, Johnson P, Hildreth AJ, Greene DR. Improving prostate cancer detection with an extended-core transrectal ultrasonography-guided prostate biopsy protocol. BJU Int. 2002 Jan;89(1):33-9.
245. Paul R, Scholer S, van Randenborgh H, Kubler H, Alschibaja M, Busch R, et al. Morbidity of prostatic biopsy for different biopsy strategies: is there a relation to core number and sampling region? Eur Urol. 2004 Apr;45(4):450-5; discussion 6.
246. Vashi AR, Wojno KJ, Gillespie B, Oesterling JE. A model for the number of cores per prostate biopsy based on patient age and prostate gland volume. J Urol. 1998 Mar;159(3):920-4.
247. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin. 2009 Jul-Aug;59(4):225-49.
248. Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate cancer. J Clin Oncol. 2010 Mar 1;28(7):1117-23.
249. Hamilton AS, Albertsen PC, Johnson TK, Hoffman R, Morrell D, Deapen D, et al. Trends in the treatment of localized prostate cancer using supplemented cancer registry data. BJU Int. 2010 Aug 24.
250. Han M, Partin AW, Pound CR, Epstein JI, Walsh PC. Long-term biochemical disease-free and cancer-specific survival following anatomic radical retropubic prostatectomy. The 15-year Johns Hopkins experience. Urol Clin North Am. 2001 Aug;28(3):555-65.
251. Sweat SD, Bergstralh EJ, Slezak J, Blute ML, Zincke H. Competing risk analysis after radical prostatectomy for clinically nonmetastatic prostate adenocarcinoma according to clinical Gleason score and patient age. J Urol. 2002 Aug;168(2):525-9.
252. Dickman PW, Adami HO. Interpreting trends in cancer patient survival. J Intern Med. 2006 Aug;260(2):103-17.
253. Donovan J, Mills N, Smith M, Brindle L, Jacoby A, Peters T, et al. Quality improvement report: Improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. Commentary: presenting unbiased information to patients can be difficult. BMJ. 2002 Oct 5;325(7367):766-70.
254. Wilt TJ, Brawer MK, Barry MJ, Jones KM, Kwon Y, Gingrich JR, et al. The Prostate cancer Intervention Versus Observation Trial:VA/NCI/AHRQ Cooperative Studies Program #407 (PIVOT): design and baseline results of a randomized controlled trial comparing radical prostatectomy to watchful waiting for men with clinically localized prostate cancer. Contemp Clin Trials. 2009 Jan;30(1):81-7.
255. D'Amico AV, Moul J, Carroll PR, Sun L, Lubeck D, Chen MH. Cancer-specific mortality after surgery or radiation for patients with clinically localized prostate cancer managed during the prostate-specific antigen era. J Clin Oncol. 2003 Jun 1;21(11):2163-72.
256. Stephenson AJ, Kattan MW, Eastham JA, Bianco FJ, Jr., Yossepowitch O, Vickers AJ, et al. Prostate cancer-specific mortality after radical prostatectomy for patients treated in the prostate-specific antigen era. J Clin Oncol. 2009 Sep 10;27(26):4300-5.
257. 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. J Urol. 2007 Dec;178(6):2359-64; discussion 64-5.

258. Stattin P, Holmberg E, Johansson JE, Holmberg L, Adolfsson J, Hugosson J, et al. Outcomes in localized prostate cancer: National Prostate Cancer Register of Sweden follow-up study. J Natl Cancer Inst. 2010 Jul 7;102(13):950-8.
259. Makarov DV, Trock BJ, Humphreys EB, Mangold LA, Walsh PC, Epstein JI, et al. Updated nomogram to predict pathologic stage of prostate cancer given prostate-specific antigen level, clinical stage, and biopsy Gleason score (Partin tables) based on cases from 2000 to 2005. Urology. 2007 Jun;69(6):1095-101.
260. Begg CB, Riedel ER, Bach PB, Kattan MW, Schrag D, Warren JL, et al. Variations in morbidity after radical prostatectomy. N Engl J Med. 2002 Apr 11;346(15):1138-44.
261. Alibhai SM, Leach M, Tomlinson G, Krahn MD, Fleshner N, Holowaty E, et al. 30-day mortality and major complications after radical prostatectomy: influence of age and comorbidity. J Natl Cancer Inst. 2005 Oct 19;97(20):1525-32.
262. Polanczyk CA, Marcantonio E, Goldman L, Rohde LE, Orav J, Mangione CM, et al. Impact of age on perioperative complications and length of stay in patients undergoing noncardiac surgery. Ann Intern Med. 2001 Apr 17;134(8):637-43.
263. Han M, Partin AW, Zahurak M, Piantadosi S, Epstein JI, Walsh PC. Biochemical (prostate specific antigen) recurrence probability following radical prostatectomy for clinically localized prostate cancer. J Urol. 2003 Feb;169(2):517-23.
264. Stephenson AJ, Scardino PT, Eastham JA, Bianco FJ, Jr., Dotan ZA, Fearn PA, et al. Preoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. J Natl Cancer Inst. 2006 May 17;98(10):715-7.
265. Gerber GS, Thisted RA, Scardino PT, Frohmuller HG, Schroeder FH, Paulson DF, et al. Results of radical prostatectomy in men with clinically localized prostate cancer. JAMA. 1996 Aug 28;276(8):615-9.
266. 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 Jul;46(11):2077-87.
267. Lutz W, Sanderson W, Scherbov S. The end of world population growth. Nature. 2001 Aug 2;412(6846):543-5.
268. Lutz W, Sanderson W, Scherbov S. The coming acceleration of global population ageing. Nature. 2008 Feb 7;451(7179):716-9.
269. Leeuwen PJ ZX, Bul M, Schröder FH, Roobol MJ, Bangma CH. Prostaatkankerincidentie en mortaliteit de komende jaren - zijn we voorbereid? . NTGV. 2010;6(Oktober):162.
270. Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med. 2009 Mar 26;360(13):1310-9.
271. Schroder FH, Roobol MJ. ERSPC and PLCO prostate cancer screening studies: what are the differences? Eur Urol. 2010 Jul;58(1):46-52.
272. van Leeuwen PJ, Connolly D, Gavin A, Roobol MJ, Black A, Bangma CH, et al. Prostate cancer mortality in screen and clinically detected prostate cancer: estimating the screening benefit. Eur J Cancer. 2010 Jan;46(2):377-83.
273. van Leeuwen PJ, van Vugt HA, Bangma CH. The implementation of screening for prostate cancer. Prostate Cancer Prostatic Dis. 2010 Sep;13(3):218-27.

274. 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 Nov;46(17):3082-9.
275. van Oortmarssen GJ, Habbema JD, van der Maas PJ, de Koning HJ, Collette HJ, Verbeek AL, et al. A model for breast cancer screening. Cancer. 1990 Oct 1;66(7):1601-12.
276. Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst. 2002 Jul 3;94(13):981-90.
277. Hanley JA. Mortality reductions produced by sustained prostate cancer screening have been underestimated. J Med Screen. 2010;17(3):147-51.
278. Cuervo LG, Aronson JK. The road to health care. BMJ. 2004 Jul 3;329(7456):1-2.
279. Black WC. Randomized clinical trials for cancer screening: rationale and design considerations for imaging tests. J Clin Oncol. 2006 Jul 10;24(20):3252-60.
280. Moss S. Design issues in cancer screening trials. Stat Methods Med Res. 2010 Oct;19(5):451-61.
281. Fall K, Stromberg F, Rosell J, Andren O, Varenhorst E, South-East Region Prostate Cancer G. Reliability of death certificates in prostate cancer patients. Scand J Urol Nephrol. 2008;42(4):352-7.
282. Goldoni CA, Bonora K, Ciatto S, Giovannetti L, Patriarca S, Sapino A, et al. Misclassification of breast cancer as cause of death in a service screening area. Cancer Causes Control. 2009 Jul;20(5):533-8.
283. Otto SJ, van Leeuwen PJ, Hoekstra JW, Merckelbach JW, Blom JH, Schroder FH, et al. Blinded and uniform causes of death verification in cancer screening: a major influence on the outcome of a prostate cancer screening trial? Eur J Cancer. 2010 Nov;46(17):3061-7.
284. Klint A, Engholm G, Storm HH, Tryggvadottir L, Gislum M, Hakulinen T, et al. Trends in survival of patients diagnosed with cancer of the digestive organs in the Nordic countries 1964-2003 followed up to the end of 2006. Acta Oncol. 2010 Jun;49(5):578-607.
285. Nobes JP, Langley SE, Laing RW. Metabolic syndrome and prostate cancer: a review. Clin Oncol (R Coll Radiol). 2009 Apr;21(3):183-91.
286. Laukkanen JA, Laaksonen DE, Niskanen L, Pukkala E, Hakkarainen A, Salonen JT. Metabolic syndrome and the risk of prostate cancer in Finnish men: a population-based study. Cancer Epidemiol Biomarkers Prev. 2004 Oct;13(10):1646-50.
287. Mottillo S, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, et al. The metabolic syndrome and cardiovascular risk a systematic review and meta-analysis. J Am Coll Cardiol. 2010 Sep 28;56(14):1113-32.
288. Perdona S, Cavadas V, Di Lorenzo G, Damiano R, Chiappetta G, Del Prete P, et al. Prostate Cancer Detection in the "Grey Area" of Prostate-Specific Antigen Below 10 ng/ml: Head-to-Head Comparison of the Updated PCPT Calculator and Chun's Nomogram, Two Risk Estimators Incorporating Prostate Cancer Antigen 3. Eur Urol. 2010 Oct 12.

289. Jansen FH, Roobol M, Jenster G, Schroder FH, Bangma CH. Screening for prostate cancer in 2008 II: the importance of molecular subforms of prostate-specific antigen and tissue kallikreins. Eur Urol. 2009 Mar;55(3):563-74.
290. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. Clin Cancer Res. 2007 Sep 1;13(17):5103-8.
291. Salagierski M, Verhaegh GW, Jannink SA, Smit FP, Hessels D, Schalken JA. Differential expression of PCA3 and its overlapping PRUNE2 transcript in prostate cancer. Prostate. 2010 Jan 1;70(1):70-8.
292. Carroll PR, Albertsen PC, Greene K. Prostate-Specific Antigen Best Practice Statement: 2009 Update. American Urological association. 2009.
293. Network NCC. NCCN Guidelines for Detection, Prevention, & Risk Reduction, Prostate Cancer Early Detection http://wwwnccnorg. 2010.
294. Inouye SK, Peduzzi PN, Robison JT, Hughes JS, Horwitz RI, Concato J. Importance of functional measures in predicting mortality among older hospitalized patients. JAMA. 1998 Apr 15;279(15):1187-93.
295. Walter LC, Bertenthal D, Lindquist K, Konety BR. PSA screening among elderly men with limited life expectancies. JAMA. 2006 Nov 15;296(19):2336-42.
296. Lu-Yao G, Stukel TA, Yao SL. Prostate-specific antigen screening in elderly men. J Natl Cancer Inst. 2003 Dec 3;95(23):1792-7.
297. Peto R, Whitlock G, Jha P. Effects of obesity and smoking on U.S. life expectancy. N Engl J Med. 2010 Mar 4;362(9):855-6; author reply 6-7.
298. Lilja H, Ulmert D, Bjork T, Becker C, Serio AM, Nilsson JA, et al. Long-term prediction of prostate cancer up to 25 years before diagnosis of prostate cancer using prostate kallikreins measured at age 44 to 50 years. J Clin Oncol. 2007 Feb 1;25(4):431-6.
299. 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 Sep 4.
300. Eichler K, Hempel S, Wilby J, Myers L, Bachmann LM, Kleijnen J. Diagnostic value of systematic biopsy methods in the investigation of prostate cancer: a systematic review. J Urol. 2006 May;175(5):1605-12.
301. Eskew LA, Bare RL, McCullough DL. Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate. J Urol. 1997 Jan;157(1):199-202; discussion -3.
302. de la Taille A, Antiphon P, Salomon L, Cherfan M, Porcher R, Hoznek A, et al. Prospective evaluation of a 21-sample needle biopsy procedure designed to improve the prostate cancer detection rate. Urology. 2003 Jun;61(6):1181-6.
303. Rittmaster RS, Fleshner NE, Thompson IM. Pharmacological approaches to reducing the risk of prostate cancer. Eur Urol. 2009 May;55(5):1064-73.
304. Lippman SM, Klein EA, Goodman PJ, Lucia MS, Thompson IM, Ford LG, et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). JAMA. 2009 Jan 7;301(1):39-51.

305. Thompson IM, Goodman PJ, Tangen CM, Lucia MS, Miller GJ, Ford LG, et al. The influence of finasteride on the development of prostate cancer. N Engl J Med. 2003 Jul 17;349(3):215-24.
306. Andriole GL. Headline Results of the REDUCE Trial: The Effect of Dutasteride on Prostate Cancer Risk Reduction. American Urological Association 104th Annual Scientific Meeting. 2009;Late Breaking Abstract 1.
307. Thompson IM, Tangen CM, Parnes HL, Lippman SM, Coltman CA, Jr. Does the level of prostate cancer risk affect cancer prevention with finasteride? Urology. 2008 May;71(5):854-7.
308. Thompson IM, Tangen CM, Goodman PJ, Lucia MS, Parnes HL, Lippman SM, et al. Finasteride improves the sensitivity of digital rectal examination for prostate cancer detection. J Urol. 2007 May;177(5):1749-52.
309. Thompson IM, Chi C, Ankerst DP, Goodman PJ, Tangen CM, Lippman SM, et al. Effect of finasteride on the sensitivity of PSA for detecting prostate cancer. J Natl Cancer Inst. 2006 Aug 16;98(16):1128-33.
310. van Leeuwen PJ, Kolble K, Huland H, Hambrock T, Barentsz J, Schroder FH. Prostate Cancer Detection and Dutasteride: Utility and Limitations of Prostate-Specific Antigen in Men with Previous Negative Biopsies. Eur Urol. 2010 Dec 4.
311. Andriole GL, Bostwick D, Brawley OW, Gomella L, Marberger M, Montorsi F, et al. The Effect of Dutasteride on the Usefulness of Prostate Specific Antigen for the Diagnosis of High Grade and Clinically Relevant Prostate Cancer in Men With a Previous Negative Biopsy: Results From the REDUCE Study. J Urol. 2011 Jan;185(1):126-31.
312. PCCNC. 2009 Prostate Cancer Screening Guidelines http://wwwpccncorg/pdf/ProstateScreeningpdf. 2009.
313. Heidenreich A, Bellmunt J, Bolla M, Joniau S, Mason M, Matveev V, et al. EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Treatment of Clinically Localised Disease. Eur Urol. 2010 Oct 28.
314. O'Connor AM, Bennett CL, Stacey D, Barry M, Col NF, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2009(3):CD001431.
315. Volk RJ, Hawley ST, Kneuper S, Holden EW, Stroud LA, Cooper CP, et al. Trials of decision aids for prostate cancer screening: a systematic review. Am J Prev Med. 2007 Nov;33(5):428-34.
316. 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 Jan;57(1):79-85.
317. Kaplan DJ, Boorjian SA, Ruth K, Egleston BL, Chen DY, Viterbo R, et al. Evaluation of the Prostate Cancer Prevention Trial Risk calculator in a high-risk screening population. BJU Int. 2010 Feb;105(3):334-7.
318. Roobol MJ, Schröder FH, Trottier G, Fernandes KA, Kranse R, Fleshner NE. Development of the European Randomized study of Screening for Prostate Cancer (ERSPC) risk calculator for high grade PCa and assessing and comparing its performance with the PCPT risk calculator in a Canadian cohort. abstract Annual EAU Congress Vienna. 2011.

319. Parekh DJ, Ankerst DP, Higgins BA, Hernandez J, Canby-Hagino E, Brand T, et al. External validation of the Prostate Cancer Prevention Trial risk calculator in a screened population. Urology. 2006 Dec;68(6):1152-5.
320. van Vugt HA, Roobol MJ, Kranse 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 Apr;47(6):903-9.
321. Kranse R, Roobol M, Schroder FH. A graphical device to represent the outcomes of a logistic regression analysis. Prostate. 2008 Nov 1;68(15):1674-80.
322. Stamey TA, McNeal JE, Yemoto CM, Sigal BM, Johnstone IM. Biological determinants of cancer progression in men with prostate cancer. JAMA. 1999 Apr 21;281(15):1395-400.
323. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. Jama. 2005 May 4;293(17):2095-101.
324. 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 Int. 2008 Jan;101(2):165-9.

PhD Portfolio

Erasmus MC
Universitair Medisch Centrum Rotterdam

Name PhD student:	Pim J. van Leeuwen
Erasmus MC Department:	Urology
PhD period:	2008 - 2011
Promotor(s):	Prof. Dr. C.H. Bangma
Supervisor:	Dr. M.J. Roobol

PhD training	**Year**	**Workload (ECTS)**
Erasmus Medical Centre Rotterdam		
Molmed Courses	2009 – 2010	2
Queen's university Belfast		
- English Speaking & Academic Writing	2008	2
Netherlands Institute for Health Sciences Rotterdam		
- Biostatistics for Clinicians	2009	1.9
- Regression Analysis for Clinicians	2009	1.9
- Survival Analysis for Clinicians	2009	1.9
- Advanced Analysis of Prognosis Studies	2010	1.9
- Planning and Evaluation of Screening	2010	1.9
The London School of Hygiene & Tropical Medicine		
- Cancer Survival: principles, methods and analysis	2009	2
Erasmus University Rotterdam		
- Biomedical English Writing and Communication	2010	1
Seminars and workshops		
Erasmus Medical Centre Rotterdam		
- Department Urology journal club	2008 – 2010	1
- Department Urology internal course	2008 – 2010	1
- Department Urology PhD meeting	2008 – 2010	1
Presentations		
Annual Meeting EAU, Spain	2009	0.5
Annual Meeting AUA, United States	2009	0.5
Najaarsvergadering NVU	2009	0.5
Academische Jaarprijs	2009	1
Voorjaarsvergadering NVU	2010	0.5

Annual Meeting EAU, Sweden	2010	0.5
Annual Meeting AUA, United States	2010	0.5
National Prostate Cancer Symposium Australia	2010	1
Najaarsvergadering NVU	2010	0.5
International Cancer Conference Tallin, Estonia	2011	1

(Inter)national conferences

Oral or poster presentations at AICC 2008, EAU 2009, EAU 2010, EAU 2011, ASCO 2009, AUA 2009, AUA 2010, ERSPC meeting 2009, ERSPC meeting 2010.	2008 – 2010	9

Other

Congress organization Symposium for Experimental Research in Surgical Specialties	2010	1

Total ECTS	**36.0**

Abbreviations

EAU European Association of Urology
AUA American Urological Association
ASCO American Society of Clinical Oncology
NVU Nederlandse Vereniging voor Urologie
ERSPC European Randomized Study of Screening for Prostate Cancer