1. Introduction

Prostate cancer is the most commonly diagnosed non-cutaneous cancer and the third most common cause of death from cancer in males in the US, estimated at 27,350 deaths in 2006 [1]. The prostate-specific antigen (PSA) test for the early detection of prostate cancer, introduced in late 1980's, has rapidly gained popularity, not only among healthcare providers and their patients, but also among the “healthy” male population. This increasing public awareness and the subsequently rising use of the...
PSA test, as a means of early detection, demands an answer on the question whether early detection of prostate cancer reduces prostate cancer specific mortality.

Currently two randomized controlled trials (RCT) are under way, the European Randomized Study of Screening for Prostate Cancer (ERSPC) [2] and the Prostate Lung Colorectal and Ovary cancer screening trials (PLCO) [3], to study the effectiveness of screening for prostate cancer. The ERSPC was initiated in 1993 and has approximately 205,000 men randomized in eight different centers in Europe. The PLCO trial was also initiated in 1993 and has included approximately 74,000 men, randomized to the prostate arm of the trial. Conditions for a possible common analysis of ERSPC and PLCO trial have been discussed [4].

The ongoing randomized trials will not produce results for several years [5]. Furthermore, the benefit of the RCT design may not be realized if screening in men allocated to the control arm is widespread. This together with the knowledge that prostate cancer has a relatively long natural history, several case-control studies were already performed to try speeding up the process of answering the question on the value of prostate cancer screening. Here we review the results of published case-control studies on the efficacy of prostate cancer screening.

2. Materials and methods

2.1. Search strategy

The PubMed database was searched for papers published until April 2006, in the English language. Search query included the terms prostate cancer (prostate carcinoma) and case-control (case control) in title and/or abstract. For each retrieved paper, the reference lists were checked upon relevant (accessible) literature and the PubMed option Related Articles was consulted.

2.2. Basic methodological concept of case-control study

It is well known that the RCT is the study design that provides the most reliable unbiased estimates of the effect of the intervention under study. The RCT is the preferred approach for evaluation of cancer screening programs. In this study design, subjects consenting to participate are randomly allocated to either the screening group, offered screening, or to the control group, that is not offered any screening. The final outcome is assessed according to the intention-to-screen principle where the rate of the cancer-specific mortality is compared between the screening and the control group. In absence of an RCT, e.g. screening test is already widely disseminated [6], information on screening can also be achieved from observational (non-randomized) designs as cohort studies, where groups of individuals is followed up over a period of time, and from time trend analyses, which are based on examination of (aggregated) mortality data of populations across time. Another observational design used to evaluate cancer screening programs is the case-control study. Fig. 1 presents the basic concept of this study design.

In a case-control study of screening, the potential effect of the screening test for prevention of an adverse outcome (e.g. late stage of a disease or mortality) is examined by comparing the exposure to the screening test between those subjects with the adverse outcome, the cases, and those without this outcome, the controls [7]. The evaluation of the efficacy of a test in cancer screening focuses on reducing cancer-specific mortality, hence, the control subject should be alive at the date of death of the case subject and cancer-free at diagnosis of the case subject. As the starting-point of the case-control study is the presence of the adverse outcome in the cases, the exposure to the screening test has to be examined retrospectively, i.e. the screening history of each case and his matched control(s) has to be assessed in the period preceding the diagnosis of the case [7,8]. The frequency of screening (ever or never screened) is then compared between the cases and their matched controls which gives an estimate of the strength of the association (Odds ratio, OR), whether screening prevents the adverse outcome (Fig. 1). Importantly, as the screening test is often the same used for diagnostic purposes, true screenings should be distinguished from tests done in relation to symptoms; e.g. both digital rectal examination (DRE) and PSA are used for screening asymptomatic men and for examination of symptomatic subjects.

The cases and controls are selected from the same source population (Fig. 1) and age range within which screening is available. Any difference in screening history between cases and controls would be due to different rates of acceptance of the screening program (or opportunistic screening). Hence, cases and controls may not be appropriately “representative” as might be the case in screening for prostate cancer. It is known that African-Americans have a high-risk for prostate cancer, but the frequency of PSA testing in this population is low compared with the Whites [9].

The key advantages of a case-control study are that it is less costly, less lengthy compared to RCT, does not require a large number of subjects, is appropriate for studying rare conditions
## Table 1 - Description of the case-control studies on prostate cancer screening

<table>
<thead>
<tr>
<th>Source population</th>
<th>Age at entry</th>
<th>Cases</th>
<th>Controls</th>
<th>Index date</th>
<th>Screening exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Years</td>
<td>N</td>
<td>Diagnosis PCa</td>
<td>Follow-up endpoint</td>
<td>N&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>A. Metastasis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friedman et al. [19]</td>
<td>health plan</td>
<td>39–95&lt;sup&gt;b&lt;/sup&gt;</td>
<td>139</td>
<td>1975–1985</td>
<td>at diagnosis or later</td>
</tr>
<tr>
<td>Kopec et al [16]</td>
<td>geographic area</td>
<td>40–84&lt;sup&gt;c&lt;/sup&gt;</td>
<td>236</td>
<td>≥1 Jan 1990</td>
<td>8/1999–5/2002</td>
</tr>
<tr>
<td>B. Death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richert-Boe et al. [18]</td>
<td>health plan</td>
<td>40–84&lt;sup&gt;b&lt;/sup&gt;</td>
<td>150</td>
<td>*</td>
<td>1981–1990</td>
</tr>
<tr>
<td>Weinmann et al. [13]</td>
<td>health plan</td>
<td>45–84&lt;sup&gt;c&lt;/sup&gt;</td>
<td>171</td>
<td>*</td>
<td>1992–1999</td>
</tr>
<tr>
<td>Weinmann et al. [14]&lt;sup&gt;h&lt;/sup&gt;</td>
<td>health plans</td>
<td>45–84&lt;sup&gt;e&lt;/sup&gt;</td>
<td>608/161</td>
<td>*</td>
<td>1997–6/2000</td>
</tr>
</tbody>
</table>

Abbreviation; PCa, prostate cancer.

<sup>a</sup> Endpoint free (no metastasis or alive and prostate cancer free at diagnosis of case).

<sup>b</sup> At prostate cancer diagnosis.

<sup>c</sup> At endpoint.

<sup>d</sup> Date of diagnosis case.

<sup>e</sup> Poorly or not described/defined.

<sup>f</sup> Well or moderately described/defined.

<sup>g</sup> Not selection criterion.

<sup>h</sup> White, Black men separately.

<sup>i</sup> Nested case control study.

<sup>j</sup> PCa free as of 1 Jan 1991 & ambulatory care 1989–1990.
and diseases with prolonged induction or latency, and it gives the possibility of examining multiple exposures, [8]. However, case-control studies are very susceptible to biases, and the major bias inherent to this study design are misclassification bias and (self-)selection bias. As mentioned above, the screening and diagnostic test are often the same. If during review of medical records, a test performed for diagnostic purposes in a control is erroneously assigned as screening test or a screening test in a case is incorrectly classified as diagnostic test, this ascertainment procedure would introduce misclassification bias which would lead to a bias in favor of screening or against screening, respectively [10]. Subjects attending a screening program or seeking opportunistic screening might differ in the risk of the disease under study, in social-economic status and in access to the screening test. High-risk subjects often decline participation in screening programs. These aspects confer (self-)selection bias, affecting the frequency of screening either in favor or against screening [7].

Although there are several methods described in order to correct for the above-mentioned biases [11,12], the generally regarded optimal design for the evaluation of a population-based screening program is an RCT. With the RCT design, the study population consists of subjects free of the disease at start of the study, randomly divided into two groups, but perhaps most importantly the investigators assigned the exposure under study. Therefore, an RCT provides the strongest evidence for causality and controls for “unknown” confounders.

3. Results

Only seven case-control studies of screening for prostate cancer were identified in the PubMed search, published from 1991 onwards, all conducted in North America. Of these, only three studies assessed the efficacy of either PSA alone or PSA and DRE on prostate cancer mortality [13–15]. One study concerned the effect of PSA screening on the development of metastatic disease [16] and the three remaining examined whether screening by DRE is associated with a reduction in prostate cancer mortality [17,18] or metastatic disease [19]. A description of these studies is presented in Table 1.

The source population from which both cases and controls were selected was composed of medical registries of health-care plans and centers (Kaiser Permanente, [13,14,18,19]; VA Outpatient Clinic System Database, [15,20]) or defined geographical areas (Metropolitan Toronto and surrounding counties, [16], and Olmsted County, Rochester Epidemiology Project, [17]). The subjects were aged 39 years and older at diagnosis of prostate cancer or at reaching study endpoint (metastasis or death from prostate cancer). An exception is the study by Jacobsen et al. [17], where no age range is given, as age was not an inclusion criterion, but prostate cancer death in a defined time period (1976–1991).

Matching of controls to cases was primarily based on age and the date of enrollment in a health plan, a site of ambulatory care in a medical center or an area of residence for the studies using geographical area as source population. The two most recent studies, Weinmann et al. [14] and Concato et al. [15], included the largest number of cases and matched controls (Table 1). The index or reference date was properly defined in all studies, as the first date of suspicion of prostate cancer [8]. In the studies by Weinmann et al., an additional symptoms reference date was selected, defined as first date on which symptoms were noted that were subsequently found to be related to prostate cancer.

The results of the studies examining the potential role of screening on the development of advanced prostate cancer and those on mortality are discussed below separately.

3.1. Outcome: metastatic prostate cancer

The results of the two studies with metastasis of the prostate as outcome are presented in Table 2A. The study by Friedman et al. [19] was conducted in men with prostate cancer diagnosed in the pre-PSA era (1979–1985), hence the exposure to DRE screening was identified through review of medical charts. DRE in the three months prior to diagnosis were not included in the analyses to keep the case-control status undisclosed to the chart reviewers. Both cases and matched controls arose from the database of a regional health plan and were within the age group 39–95 years. The race-adjusted odds ratio (OR), up to 10 years before diagnosis, indicated no benefit of screening in prevention of the development of metastatic disease (OR = 0.9, 95% confidence interval, CI, 0.5–1.7). When DRE performed for evaluation of intestinal or rectal symptoms were included as “screening”, the OR became 0.7, but still statistically not significant (95% CI, 0.4–1.4).

In contrast, Kopec et al. [16] found that PSA screening reduced the risk of metastatic prostate cancer by 35% (OR = 0.65, 95%CI, 0.45–0.93). This beneficial effect was seen in men aged 45–59 years (OR = 0.52, CI, 0.28–0.98), and to some extent in those aged 60–84 years (OR = 0.67, 95%CI, 0.41–1.09). Some methodological points of this study should also be noted. The strength of the study is the effort in differentiating between true screening and diagnostic PSA tests, thorough review of medical charts for adjustment of baseline prognostic factors and a sensitivity analysis correcting for positive/suspicious DRE. However, there is a potential non-response and information bias in this study as the number of participants returning the mailed
informed consent was low and differed among cases and controls (69% and 51%, respectively). The same was true for the availability of medical charts (93% and 88%, respectively). This problem is also pointed out in an editorial by Barry [21]. In contrast to the other case-control studies where cases and controls are matched on age and the length of exposure window is kept equal for the case-control pair, the authors of this study matched only on age and considered the difference in exposure window as a confounding factor, which they adjusted for. It may be possible that the adjustment for exposure time has led to significant difference in outcome.

3.2. Outcome: prostate cancer mortality

Table 2B shows the corresponding odds ratios of the case-control studies assessing the potential benefit of prostate cancer screening with prostate cancer mortality as endpoint. The two DRE case-control studies, conducted in the pre-PSA era, showed contradictory results, varying from a 68% significant preventive effect (OR 0.32, 95% CI 0.20–0.52; [17]) to no effect of DRE screening on prostate cancer mortality (the odds ratio was not significantly different from 1, OR 0.84 95%CI 0.48–1.46, [18]). In the study by Jacobsen et al. [17], the analyses were restricted by eliminating the first year before the index date to exclude DREs possibly performed for diagnostic purposes. This restriction could have introduced misclassification bias in favor of screening, hence the reason of the significant result of this study.

Conflicting results were found in the PSA/DRE studies as well. In their first study, Weinmann et al. [13] found no statistical significant effect of the combination of PSA and DRE. However, in their subsequent study with a larger population size [14], a significant 27% reduction was observed, but this concerned the White male cohort only. This effect seemed to be attributable to DRE screening, as analyses including only DRE revealed statistically significant OR of 0.65 (95% CI 0.45–0.85) in Whites. In Black men, there was no evidence that cases were less likely to have had PSA/DRE screening than the controls, even when DRE alone was considered. Then again the population size was smaller than that for the White cohort (Table 1). If the White and

<table>
<thead>
<tr>
<th>Table 2 – Adjusted odds ratios of case-control studies on prostate cancer screening with either DRE alone or PSA and DRE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OR (95% CI)</strong></td>
</tr>
<tr>
<td><strong>A. Outcome: metastatic prostate cancer</strong></td>
</tr>
</tbody>
</table>
| **PSA era** | Kopec et al. [16] | 0.65 (0.45–0.93) | 0.52 (0.28–0.98)
| **DRE era** | 0.67 (0.41–1.09) |
| **B. Outcome: prostate cancer mortality** | **DRE era** | Richert-Boe et al. [18] | – | 0.84 (0.48–1.46) |
| **PSA era** | Jacobsen et al. [17] | – | 0.53 (0.32–0.90)
| **PSA era** | 0.32 (0.20–0.52) |
| **Concato et al. [15]** | 1.13 (0.63–2.06) |

Abbreviations: OR, odds ratio, CI, confidence interval.

* Including DREs for evaluation of intestinal or rectal symptoms.

† Age group 45–69 years.

‡ Age group 60–84 years.

§ Only DRE in presence of prostate cancer symptoms.

‖ Membership > = 10 years.

§§ Only screening test classified as definitely.

‖‖ Study reference date, W = white, B = black.
Black cohort data were analyzed jointly, most likely the study would not have yielded a statistically significant result.

Concato et al. [15], adjusting for race, showed no effect of PSA/DRE screening in their nested case-control study (OR 1.13, 95% CI 0.71–1.64). Additional analyses including tests performed in men with BPH symptoms or test for differential diagnosis did not alter the outcome.

Connor et al. [22] have shown that the criteria for selecting matched controls to cases, the choice of the time scale of screening and the method to measure exposure of screening in case-control studies can considerably affect the results due to biases introduced. The retrospective identification of screening exposure, particularly for prostate cancer, remains a difficult issue. In all studies reviewed, determination of the reason of the tests were done through review of medical records and in the most recent PSA/DRE studies [14,15], the authors made use of algorithms to classify the PSA and DRE findings as definitely screening, probably screening or diagnostic tests, accounting for BPH and its symptoms. The study of Jacobsen et al. [17], which assessed whether screening with DRE reduced prostate cancer mortality, did not take examinations that were performed for evaluation of BPH symptoms into account. Screening was defined as the absence of hematuria, bone pain and pelvic pain at the time of DRE. However, even when the analysis where performed with exposure being any DRE, an adjusted OR of 0.53 (95% CI 0.32–0.90) was observed. The importance of accurate classification of true screening and diagnostic tests in case-control studies of the efficacy of prostate cancer screening have been demonstrated by several authors [10,23–25]. If screening tests are not properly defined, the proportion of subjects classified as screened will be distorted leading to biased study results. Hence, the studies of Weinmann et al. [14] and Concato et al. [15] seem not to be weighed down by these aspects, but there is evidence of selection bias in Weinmann's study as a significant result was only seen in White men.

However, the significant outcome for DRE alone in White men in the large study of Weinmann et al. [14] might reflect (self-)selection bias as prostate cancer screening is opportunistic and, as mentioned before, the rate of screening in the US is higher among the White population.

4. Discussion

A large number of observational studies on the evidence of prostate cancer screening and mortality reduction have been published with contradictory outcome [26–29]. Coldman et al. [30] observed that high screening intensity is not associated with large reductions in prostate cancer mortality, when compared with geographic areas with low or medium screening intensity. Using incidence-based mortality rates, Chu et al. [31] demonstrated that the decrease in mortality rates among men with distant disease diagnosis accounted for the overall decline in prostate cancer mortality trends in the USA. In the non-randomized prospective Tyrol screening study initiated in 1993, men aged 45–74 years were offered the PSA test [28,32]. Over the period 1996–2000, the trends in prostate cancer mortality showed that the mortality in Tyrol was 35% lower than in the rest of Austria.

The results of the Quebec RCT, started in 1988, only further contributed to the controversy on prostate cancer screening. In the two publications of this RCT, decreases of 69% [33] and 62% [34] in the prostate cancer mortality were reported, when screened men were compared to unscreened men. The proper analyses according to the intention-to-screen principle yielded relative risks of 1.06 [33] and 1.08 [34], respectively, indicating no effect of screening, particularly due to lack of power.

From this current review of case-control studies we can conclude that the evidence for a beneficial effect of prostate cancer screening on mortality reduction is still not unambiguous. The major limitation of the case-control studies reviewed is the potential misclassification of the exposure, PSA tests and DREs, particularly in presence of other non-cancer symptoms of the prostate, and the accuracy of the index date. The studies by Concato et al. [15] and Weinmann et al. [14] seem not to be weighed down by these aspects, but there is evidence of selection bias in Weinmann’s study as a significant result was only seen in White men.

Decisions on population-based screening require solid evidence of its benefits on the reduction of cancer-specific mortality. The costs and the effects of quality of life will be important additional outcomes. It is clear from the studies discussed above that observational study designs were not able to provide the evidence needed for this important decision. Level one evidence, i.e. from a prospective RCT is definitely still required. Waiting for the outcome of the two large-scaled randomized trials [3,35] seems the only rational approach. However, the question remains how a result of an intention-to-screen analysis from an RCT of screening (effectiveness estimate) can be translated to daily practice, since there is always a certain level of non-compliance in the screening arm and
contamination in the control arm. To answer the question of the value of early detection on an individual level (efficacy estimate) a secondary analysis with both correction for non-compliance and contamination is needed [36]. This type of analysis determines the magnitude of benefit in terms of prostate cancer risk/mortality reduction by attending a screening program.

When non-compliance and contamination are large in a RCT, investigators might make comparisons of those who actually received the intervention with those who did not, similar as in the publication of the Quebec trial [33,34]. This would lead to large effect estimates, but is not valid, as the randomization status is not complied with, introducing selection bias in the analyses. A method to adjust for non-compliance and contamination in an RCT has been developed by Cuzick et al. [36], which in the Rotterdam section of ERSPC, will be applied in addition to the main intention-to-screen analysis, to give an effect estimate after adjustment for non-compliance and contamination. A case-control study using the ERSPC data will also be conducted afterwards to give an impression of the magnitude of all known biases inherent to the case-control study on the outcome of a population-based trial.

5. Conclusion

In conclusion, as for other observational studies, the current available data from case-control studies on the value of early detection of prostate cancer in reducing disease specific mortality yielded inconclusive evidence. Hence, the presumably most debated topic in urologic practice remains uncertain at this time point. An answer has to come from well-conducted RCTs, the ERSPC in Europe and the PLCO trial in the US, of which the outcomes can be expected in 2007–2010.

With the growing awareness of prostate cancer screening not only among physicians, but also among the general population, a point of concern for these RTCs can be the generalizability to population level. As described above valid statistical techniques are available to correct for this issue to obtain a reliable estimate of the true effect of undergoing the screening test. Until then, it is advisable that men who wish to be screened should be well informed of both the benefits and possible harms (i.e. overdiagnosis, overtreatment, consequences of aggressive treatment as sexual, urinary and bowel dysfunctions) of prostate cancer screening, realizing that the major benefit on prostate cancer mortality reduction has not been shown yet.

References


CME questions

Please visit www.eu-acme.org/europeanurology to answer these EU-ACME questions on-line. The EU-ACME credits will then be attributed automatically.

1. The major drawback of the case control study design is that:
   A. It is not suitable for rare conditions.
   B. The retrospective design induces susceptibility for selection and misclassification bias.
   C. The intervention under study (cause) is not applied by the investigator.
   D. It is logistically inefficient.

2. The strongest evidence on the effect of an intervention on disease outcome is a RCT since
   A. It has a prospective study design.
   B. It is based on individual data.
   C. It has an observational approach.
   D. The exposure under study is applied by the investigator.

3. The term ‘index or reference date’ refers to the:
   A. date of diagnosis of the case
   B. date of death of the case
   C. date on which the case was screened
   D. date on which the control was screened
4. Which of the following statements is correct in a case-control study with prostate cancer death as endpoint?
   A. A case is a subject who was diagnosed with prostate cancer and died from lung cancer
   B. A control is a subject who was diagnosed with prostate cancer before his matched case and died from prostate cancer
   C. A case is a subject who was diagnosed with prostate cancer and still alive
   D. A control is a subject without prostate cancer before the date of diagnosis of his matched case and still alive at death of this case

5. Regarding the effect of screening on prostate cancer mortality following statement is correct:
   A. Case-control studies demonstrated the effectiveness of PSA screening
   B. The Quebec RCT showed conclusive and unbiased outcome
   C. Trend analyses are useful in analyzing the effect of screening
   D. The evidence of prostate cancer screening and mortality has yet to come from the RCTs

6. The outcome of a RCT on the effectiveness of a screening program can
   A. automatically be translated to daily practice.
   B. be interpreted as an efficacy estimate of the screening program applied.
   C. give an efficacy estimate after correction for both non-compliance and contamination.
   D. give a reliable estimate on the effectiveness of the program when using an intention-to-treat analysis.