

Pathologic Features of Prostate Cancer Found at Population-Based Screening With a Four-Year Interval

Robert F. Hoedemaeker, Theodorus H. van der Kwast, Rob Boer, Harry J. de Koning, Monique Roobol, André N. Vis, Fritz H. Schröder

Background: The currently recommended frequency for prostate-specific antigen (PSA) screening tests for prostate cancer is 1 year, but the optimal screening interval is not known. Our goal was to determine if a longer interval would compromise the detection of curable prostate cancer. **Methods:** A cohort of 4491 men aged 55–75 years, all participants in the Rotterdam section of the European Randomized Study of (population-based) Screening for Prostate Cancer, were invited to participate in an initial PSA screening. Men who received that screening were invited for a second screen 4 years later. Pathology findings from needle biopsy cores were compared for men in both rounds. Statistical tests were two-sided. **Results:** A total of 4133 men were screened in the first round (the prevalence screen), and 2385 were screened in the second round. The median amount of cancer in needle biopsy sets was 7.0 mm (95% confidence interval [CI] = 5.4 mm to 8.6 mm) in the first round and 4.1 mm (95% CI = 2.6 mm to 5.6 mm) in the second round ($P = .001$). Thirty-six percent of the adenocarcinomas detected in the first round but only 16% of those detected in the second round had a Gleason score of 7 or higher (mean difference = 20% [95% CI = 10% to 30%]; $P < .001$). Whereas 25% of the adenocarcinomas detected in the first round had adverse prognostic features, only 6% of those detected in the second round did (mean difference = 19% [95% CI = 11% to 26%]; $P < .001$). Baseline PSA values were predictive for the amount of tumor in biopsies in men with cancer in the first round but not for that in the second round. **Conclusion:** Most large prostate cancers with high serum PSA levels were effectively detected in a prevalence screen. In this population, a screening interval

of 4 years appears to be short enough to constrain the development of large tumors, although it is inconclusive whether this will result in a survival benefit. [J Natl Cancer Inst 2001;93:1153–8]

Prostate cancer is rapidly becoming a major health problem in Western countries. Epidemiologic studies show that it is now the second most commonly diagnosed malignancy in men after nonmelanotic skin cancer. In the United States, prostate cancer mortality is second only to mortality caused by lung cancer (1). The clinical incidence of prostate cancer has risen substantially during the last decade since the introduction of serum prostate-specific antigen (PSA) measurement as a tool for identifying men at risk for prostate cancer. The introduction and widespread use of this relatively cheap and simple test have resulted in mass screening of clinically asymptomatic men, especially in Western countries. Since 1993, the American Cancer Society has recommended annual serum PSA tests in asymptomatic men 50 years of age or older (2).

A potential and undesirable side effect caused by increased efforts for early detection of prostate cancer is an increased chance of detecting carcinomas that, had they remained undetected, would never have led to morbidity or mortality. At autopsy, the presence of clinically undiagnosed tumors in men 50 years of age or older is estimated to be at least five times higher than the lifetime risk for tumors that lead to morbidity or mortality (3,4). As of yet, the effects of PSA screening on overall morbidity, mortality, and quality of life in the screened population are not known. Several randomized trials on screening for prostate cancer are under way to investigate these questions (5). Variations in the screening protocol, including the interval between different rounds of screening for prostate cancer, are likely to have a considerable influence on the costs of prostate cancer screening and on the morbidity, mortality, and quality of life of the screened population.

The results of some studies (6–8) indicate that biannual screening would not compromise the detection of curable prostate cancers in some or possibly all of the men at risk for the disease and would lead to a substantial reduction in health-care costs. These studies, however, are based not on clinical experience but rather on models predicting the probability of pros-

tate cancer in the population at risk using retrospective data.

To determine whether a longer interval between screening rounds would compromise the detection of curable prostate cancer, we studied prostate cancer characteristics in a cohort of men during two rounds of population-based screening for prostate cancer that were performed in the Rotterdam University Hospital, The Netherlands. The two rounds were separated by an interval of 4 years. Histopathologically assessed tumor characteristics of screen-detected cancer on needle biopsy specimens were compared between the two rounds of screening.

SUBJECTS AND METHODS

The European Randomized Study of Screening for Prostate Cancer (ERSPC) is a multicenter, randomized, population-based trial that investigates the impact of systematic PSA screening for prostate cancer on prostate cancer mortality and quality of life. The study was approved by a government commission supervising the compliance with the Dutch law on population screening, and written informed consent was obtained from every participant before randomization. This report concerns a cohort of 4491 men aged 55–75 years, all of whom had been randomly assigned to the screening group in the Rotterdam section of ERSPC from June 1994 to March 1996. None of the participants had a previous diagnosis of prostate cancer. During the study period, screening was discontinued in all participants who reached the age of 75 years. Men who did not respond to the invitation for screening were excluded from further evaluation in this report. Information on various reasons for not visiting (e.g., because men died of other causes or moved out of the area) was obtained from the local government.

Screening Protocols

First round of screening (prevalence and interim screen). The first round of screening (the prevalence screen) took place from June 1994 to March 1996. The 4133 men who accepted the invitation to the first screen underwent serum PSA measurement, digital rectal examination (DRE), and transrectal ultrasound investigation (TRUS). Biopsies were recommended for men whose serum PSA level was 4 ng/mL or greater or whose DRE or

Affiliations of authors: R. F. Hoedemaeker, T. H. van der Kwast, A. N. Vis (Department of Pathology), R. Boer, H. J. de Koning (Department of Public Health), M. Roobol, F. H. Schröder (Department of Urology), Erasmus University Rotterdam, The Netherlands.

Correspondence to: Robert F. Hoedemaeker, M.D., Department of Pathology, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: Hoedemaeker@path.fgg.eur.nl).

See "Note" following "References."

© Oxford University Press

TRUS result was abnormal. Men who were recommended for biopsy but who either refused a biopsy or could not undergo a biopsy for medical reasons (e.g., because they were receiving anticoagulant therapy or had a comorbid condition) were excluded from further evaluation. For men with a benign biopsy outcome, an interim round of screening was conducted after 1 year. A small number of men (seven) with benign biopsy outcomes in the first interim round were reinvited for a second interim round, which took place 1 year later.

Second round of screening. The second round of screening took place from June 1998 to March 2000. All 2385 participants in the second round of screening had been screened 4 years earlier, and 645 (27%) of them had undergone an interim screen 3 years earlier after being recommended for biopsy in the first round. By the time the second round of screening began, the screening protocol had changed (9). In brief, because of the low positive predictive value and sensitivity of DRE and TRUS, biopsies were now recommended to all participants with a serum PSA level of 3 ng/mL or higher, regardless of the outcome of DRE or TRUS. In addition, the 1-year interval rescreen after a benign biopsy outcome was omitted.

Biopsy technique. In participants who complied with the recommendation for biopsy, systematic sextant biopsies were obtained during longitudinal ultrasonographic scanning of the prostate. All biopsies were performed with 18-gauge biopsy needles (Bard Inc., Murray Hill, NJ) driven by a pro-mag spring-loaded biopsy gun (Manan, Northbrook, IL). The needles were directed cranially at an angle of approximately 45 degrees from the transversal plane and outward at approximately 30 degrees from the sagittal plane. Each biopsy core was inked at its capsular end, numbered according to its site of origin, and sent in a separate container to the pathology department.

Histopathologic Examination

Processing and examination of biopsy cores. All biopsy cores were processed for routine histopathologic examination, as described previously (10). In brief, after fixation in a 4% saline-buffered formalin solution, every core was embedded sepa-

rately in paraffin, sectioned longitudinally into 5- μ m sections, and stained with hematoxylin-eosin. At least three histologic sections of different cutting levels of each biopsy core were examined.

In each case of prostate cancer, the reference pathologist for urologic pathology (T. H. van der Kwast) recorded the number of positive biopsy cores in each sextant set. The length (in millimeters) of cancer involvement was measured in each core and calculated for the sextant biopsy set as a whole. All adenocarcinomas were graded according to the Gleason scoring system (11). In addition, the length (in millimeters) and the percentage of high-grade tumor (i.e., Gleason growth pattern 4 or 5) were calculated for each sextant biopsy set.

Categorization of cancer involvement and grade in biopsy sets. Both the amount of tumor present in biopsy sets and the prostate cancer grade are prognostic features for biochemical relapse after treatment with curative intent (12). To account for the fact that both large well-differentiated tumors and small poorly differentiated tumors could have a poor prognosis, we constructed an arbitrary categorization model that combines these tumor features (Table 1). Category A contains biopsy sets with only a single small focus of well-differentiated adenocarcinoma. In addition to small but potentially dangerous tumors, this category is also likely to contain clinically insignificant tumors. Category B includes biopsy sets with larger amounts of adenocarcinoma, which sometimes contain high-grade cancer (Gleason growth pattern 4 or 5). This category is likely to contain prostate cancers that pose a threat to their host if left untreated. Biopsy sets in category C contain either large amounts of adenocarcinoma or carcinomas that consist largely of poorly differentiated tumor (Gleason growth pattern 4 or 5). Men whose tumors fall into this category would have a considerable risk of therapy failure.

To validate our categorization model for sextant biopsy sets, the categorization in the model for sextant biopsy sets of 79 men in the first round of screening who underwent radical prostatectomy was compared with clinical follow-up data, i.e., biochemical failure during a 4-year follow-up period (defined as three consecutive postoperative PSA levels >0.1 ng/mL). Table 1 shows that biopsy catego-

ries of the 79 men were associated with the chance for extraprostatic tumor growth (Pearson chi-squared test, two-tailed $P = .01$). Biopsy categories were also associated with the chance for biochemical progression after surgery, although this association was not statistically significant (Pearson chi-squared test, two-tailed $P = .06$).

Comparison of Adenocarcinomas Found in Different Rounds of Screening and Statistical Analysis

The characteristics (i.e., amount and grade) of the adenocarcinomas that were detected in the first round of screening (the prevalence and interim screens combined) were compared with those of the adenocarcinomas that were detected in the second round of screening. To investigate whether characteristics of second-round adenocarcinomas were different in men who had undergone previous biopsies, we also compared the characteristics of the adenocarcinomas of participants with adenocarcinomas detected in the second round screen who had and who had not undergone biopsy during the first round of screening. Statistical analyses of comparisons between subsets of distributions based on ordinal variables, such as Gleason score or the arbitrary biopsy categories, were performed with a two-sided Pearson chi-squared test. The null hypothesis (statistical independence of the tested variables in subsets of participants) was rejected for P values under .05. Statistical analyses of comparisons of characteristics based on continuous variables with a skewed distribution, such as PSA values or the amount of tumor or high-grade cancer in biopsy sets, were performed with a two-tailed Mann-Whitney U test. The null hypothesis (a similar rank distribution of the tested variable in subsets of participants) was rejected for P values under .05.

RESULTS

Participation Rates

Of the 4491 men invited for PSA screening (after randomization) in the first round, 4133 (92%) accepted (Table

Table 1. Categories for amount and grade of adenocarcinoma on sextant prostate biopsy sets and comparison with progression after radical prostatectomy in 79 participants

Category	Criteria		Outcome after radical prostatectomy (n = 79)		
	Tumor extent	Grade	No. of cases (%)	Extraprostatic tumor growth, i.e., stage higher than pT2, at radical prostatectomy (% of cases)*	Biochemical progression within 4 y after surgery (%)†
A	Only one positive biopsy core, with $\leq 30\%$ of the core involved	Only Gleason growth patterns 1-3	22 (28)	3 (14)	1 (5)
B	Only one positive biopsy core, with >30% of the core involved, or more than one positive biopsy core, with a total percentage of cancer involvement of $\leq 30\%$	Gleason score ≤ 7	45 (57)	21 (47)	4 (9)
C	All others	Any	12 (15)	7 (58)	3 (25)
Total (%)			79 (100)	31 (39)	8 (10)

*The association between biopsy category and the frequency of extraprostatic tumor growth at radical prostatectomy is statistically significant (Pearson chi-squared test, two-tailed $P = .01$).

†Although an association between biopsy category and biochemical progression is also visible, it is not statistically significant (Pearson chi-squared test, two-tailed $P = .06$).

2). Of 1129 men with biopsy recommendations, 102 (9%) did not undergo biopsy because of refusal (49, or 48%) or medical reasons (53, or 52%). Of the 850 men with a benign biopsy outcome during the first round, 823 (97%) were invited for an interim rescreen (27 men were not invited back because they had reached the age of 76 years). Of these 823 men, 662 (80%) accepted (Table 2). Of the 161 men who did not visit, two (1%) had moved out of the area and 16 (10%) had died of causes other than prostate cancer. The reason for the failure to visit was unknown for 143 men (89%). Of the 301 men recommended for biopsy during the interim round, 17 (6%) did not undergo biopsy because of refusal (15, or 88%) or medical reasons (2, 12%).

After the various exclusions, 3616 men were left at the beginning of the second screening round. Of these, 593 (16%) were not invited for the second round because they had reached the age of 76 years. Of the 3023 men who were ultimately invited for the second screen, 2385 (79%) accepted (Table 2). Of the 638 men who did not visit, 128 (20%) had moved out of the area and 155 (24%) had died of causes other than prostate cancer. In 12 men (2%), prostate cancer had been detected outside the study screening. The reason for not attending the second screen was unknown for 343 men (54%). Of 481 biopsy recommendations, 41 men (9%) did not undergo biopsy because of refusal (31, or 76%) or medical reasons (10, or 24%).

Prostate Cancer Detection Rates

We compared the prostate cancer detection rates during the first round screen (prevalence screen), interim screens, and second round of screening to determine whether the rate of detecting prostate cancer would differ between subsequent

screening rounds. Table 2 shows that prostate cancer detection rates were similar in all rounds. During the prevalence screen, the detection rate was 4.3% (5.1% if the interim round is included). The detection rate of 3.9% in the second round did not differ statistically significantly from the rates in the first round (Pearson chi-squared test, two-tailed $P = .51$ for the first round versus the second round and $P = .36$ for the total first round [including the interim rounds] versus the second round).

To correct for any possible bias caused by age difference (ages in the first round ranged from 55 to 75 years, while ages in the second round ranged from 59 to 75 years), we examined the prostate cancer detection rate specifically in those men who were aged 59–75 years in the first round (Table 2). We found that the detection rate was slightly higher in this group of men during the first round. However, this rate still did not differ statistically significantly from that in the second round (Pearson chi-squared test, two-tailed $P = .09$).

Interval Cancers

The number of interval cancers (i.e., prostate cancers detected in a screened population outside regular screening) gives an indication of the efficacy of the screening protocol. In the cohort studied, prostate cancer was diagnosed outside the regular screening rounds in only 12 men. Prostate cancer in this group was mostly clinically inapparent and was found coincidentally at transurethral resection for prostatism or cystoprostatectomy for bladder cancer.

Age, Serum PSA Levels, and Tumor Characteristics

To evaluate differences in cancer characteristics (such as grade and size) be-

tween the subsequent rounds of screening and their relationship with patient's age and serum PSA level, we compared these parameters. In all rounds, the age of the participants with prostate cancer did not differ substantially from the age of the participants who had undergone a biopsy.

Although biopsies were not performed at serum PSA levels below 3 ng/mL in the second round, the median serum PSA levels of participants with prostate cancer was lower in the second round than in the first round (Table 3). In addition, PSA levels at biopsy after the second round screen did not predict the presence of adenocarcinoma. That is, at the prevalence (the first round) screen, serum PSA levels at biopsy were statistically significantly higher in participants with adenocarcinoma (median PSA level = 6.1 ng/mL) than in participants with a benign biopsy outcome (median PSA level = 2.7 ng/mL) (Mann-Whitney U test, two-tailed $P < .001$). In the second round, however, serum PSA levels at biopsy did not differ between men with benign biopsy outcomes (median, 4.5 ng/mL) and those with adenocarcinoma (median, 4.3 ng/mL) (Mann-Whitney U test, two-tailed $P = .57$).

The amount of adenocarcinoma in sextant biopsy sets was statistically significantly lower in tumors detected at the second round of screening than in tumors detected at the first round (Mann-Whitney U test, two-tailed $P = .001$; Table 3). The median amount of cancer in needle biopsy sets was 7.0 mm (95% confidence interval [CI] = 5.4 mm to 8.6 mm) in the first round and 4.1 mm (95% CI = 2.6 mm to 5.6 mm) in the second round ($P = .001$; Table 3). Nevertheless, the average number of positive biopsy cores in men with adenocarcinoma decreased only slightly, from 2.5 in the first round to 2.2 in the second round (Pearson

Table 2. Prostate cancer detection rates at the different rounds of screening

	No. invited	Screened, No. (% of invited)	Biopsy recommendation, No. (% of screened)	Performed biopsies, No. (% of screened)	Prostate cancer		
					No.	% of performed biopsies	% of screened*
Round 1	4491	4133 (92.0)	1129 (27.3)	1027 (24.8)	177	17.2	4.3*
Round 1: men aged 59–75 y†	—	3032	881 (29.1)	839 (27.7)	149	17.8	4.9*
Interim screens	823	662 (80.4)	301 (45.5)	284 (42.9)	33‡	11.6	5.0
Total round 1					210		5.1*
Round 2	3023	2385 (78.9)	481 (20.2)	440 (18.4)	94	21.4	3.9*

*Statistical tests on prostate cancer detection rates (Pearson chi-squared test, two-tailed test): round 1 versus round 2, $P = .51$; total round 1 versus round 2, $P = .36$; round 1 (ages 59–75 years) versus round 2, $P = .09$.

†To correct for age differences between the two rounds of screening, a separate calculation was done for men aged 59–75 years in round 1.

‡One additional case of prostate cancer was detected at a second interim screen that was performed in seven participants 1 year after the first interim screen.

Table 3. Age, serum PSA levels, and tumor characteristics of participants at biopsy and participants with adenocarcinoma in the different rounds of screening

	Round 1 prevalence screen		Round 1 interim screen		
	Benign biopsies	Cancer at biopsy	Benign biopsies	Cancer at biopsy	
No. of participants	850	177	251	33	
Mean age, y (standard deviation)	65.4 (5.8)	65.7 (5.6)	66.9 (5.1)	66.3 (6.0)	
Serum PSA level, ng/mL, median (range)	2.7 (0.1–49.4)	6.1 (0.3–304.0)	5.1 (0.1–26.2)	5.4 (1.0–24.8)	
Tumor length at biopsy, mm, median (range)		8.4 (0.8–72.5)		3.2 (0.7–10.6)	
High-grade tumor length on biopsy, mm, mean (standard deviation)/median (range)		4.4 (10.3)/0.0 (0.0–72.5), available for 175 cases		0.7 (1.7)/0.0 (0.0–8.4), available for 32 cases	
Gleason score, No. (%)					
4		6 (3)		1 (3)	
5		12 (7)		2 (6)	
6		91 (51)		22 (67)	
7		42 (24)		7 (21)	
8		22 (12)		1 (3)	
9		3 (2)		0	
10		1 (1)		0	

	Total round 1, cancer at biopsy	Round 2 screen		Statistical tests on difference between total round 1 and round 2
		Benign biopsies	Cancer at biopsy	
No. of participants	210	346	94	
Mean age, y (standard deviation)	65.8 (5.7)	67.4 (4.5)	66.8 (4.6)	
Serum PSA level, ng/mL, median (range)	5.8 (0.3–304.0)	4.5 (3.0–36.0)	4.3 (3.0–15.1)	Mann-Whitney <i>U</i> test, two-tailed <i>P</i> = .001
Tumor length at biopsy, mm, median (range)	7.0 (0.7–72.5)		4.1 (1.1–56.1)	Mann-Whitney <i>U</i> test, two-tailed <i>P</i> < .001
High-grade tumor length on biopsy, mm, mean (standard deviation)/median (range)	3.8 (9.6)/0.0 (0.0–72.5), available for 208 cases		0.6 (1.7)/0.0 (0.0–10.6), available for 93 cases	Pearson chi-squared test, two-tailed <i>P</i> = .001
Gleason score, No. (%)				
4	7 (3)		1 (1)	
5	14 (7)		3 (3)	
6	113 (54)		75 (80)	
7	49 (23)		14 (15)	
8	23 (11)		1 (1)	
9	3 (1)		0	
10	1 (1)		0	

chi-squared test, *P* = .13; data not shown).

In addition to the smaller amount of cancer, prostate cancers detected at the second round were better differentiated than adenocarcinomas detected during the combined first rounds (i.e., the prevalence and interim screens). The median amount of high-grade cancer (expressed in millimeters of cancer with Gleason growth pattern 4 or 5) was statistically significantly lower in adenocarcinomas detected at the second round of screening than in adenocarcinomas detected at the first

round (Mann-Whitney *U* test, two-tailed *P* < .001; Table 3). Gleason score was statistically significantly lower in adenocarcinomas detected at the second round (Pearson chi-squared test, two-tailed *P* = .001; Table 3). Thirty-six percent of the adenocarcinomas detected in the first round but only 16% of those detected in the second round had a Gleason score of 7 or higher (mean difference = 20% [95% CI = 10% to 30%]; *P* < .001).

To investigate whether the observed differences in tumor characteristics were reflected by serum PSA levels, we strati-

fied the frequency of biopsies, the frequency of prostate cancer at biopsy, and the median amount of tumor in biopsy specimens to range of serum PSA levels (Table 4). Both prostate cancer detection frequencies and the amount of prostate cancer in biopsy specimens at the prevalence screen were clearly associated with serum PSA levels at screening (Table 4). Despite overall higher serum PSA levels at the interim screen, both detection frequency and the amount of tumor in biopsy specimens were lower than during the prevalence screen, especially at high PSA

Table 4. Cancer detection and tumor length at biopsy of participants in different PSA ranges during screening

Serum PSA level range, ng/mL	Round 1 prevalence screen			Round 1 interim screen			Round 2 screen		
	No. of biopsies (%)	Cancer at biopsy (%/ % of biopsies)	Cancer length at biopsy, mm, median (range)	No. of biopsies (%)	Cancer at biopsy (%/ % of biopsies)	Cancer length at biopsy, mm, median (range)	No. of biopsies (%)	Cancer at biopsy (%/ % of biopsies)	Cancer length at biopsy mm, median (range)
0–2.9	469 (46)	27 (15/6)	4.4 (1.1–31.8)	66 (23)	6 (18/9)	1.9 (1.0–6.0)	0 (0)	0 (0)	—
3.0–3.9	42 (4)	10 (6/24)	5.4 (2.0–15.8)	13 (5)	3 (9/23)	4.4 (3.4–5.2)	157 (36)	35 (37/23)	3.3 (1.2–27.2)
4.0–9.9	437 (42)	95 (54/22)	9.2 (0.8–52.3)	178 (63)	20 (61/11)	3.0 (1.5–8.0)	252 (57)	52 (56/21)	5.2 (1.2–56.1)
≥10.0	79 (8)	45 (25/58)	14.6 (0.8–72.5)	27 (9)	4 (12/15)	2.6 (0.7–10.6)	31 (7)	7 (7/23)	4.2 (1.1–21.0)
Total	1027 (100)	177 (100/17.3)	8.4 (0.8–72.5)	284 (100)	33 (100/11.4)	3.2 (0.7–10.6)	440 (100)	94 (100/21.6)	4.1 (1.1–56.1)

levels (≥ 10 ng/mL, Table 4). The same trend was seen in the second round. That is, cancer detection rates in the second-round screen and the prevalence screen were similar at serum PSA levels of 3 ng/mL to 10 ng/mL, but the frequency of prostate cancer in biopsies performed at serum PSA levels of 10 ng/mL or above in the second-round screen was only half of the frequency that was observed in the prevalence screen.

In addition, the association between the range of serum PSA levels and the median length of tumor at biopsy that was clearly present in the first round was lost in the second round. Whereas in the first round a serum PSA level of 10 ng/mL or above was associated with a high amount of cancer in biopsies, the amount of adenocarcinoma present in biopsies performed at serum PSA levels of 10 ng/mL or above was comparatively low in the second round (Table 4).

Categorization Model

To examine whether the 4-year interval between PSA screenings in our study resulted in an increase in advanced-stage cancers at the second round, we compared the biopsy-determined categories of screen-detected prostate cancers in the first and second rounds. We found a substantial difference in the distribution of cancers in the three categories that we defined. There was a moderate increase in the number and the frequency of tumors of both categories A and B, from 39 (19%) and 119 (56%), respectively, in the first round to 28 (30%) and 60 (64%), respectively, in the second round. The number and the frequency of category C tumors at biopsy, however, dropped dramatically, from 52 (25%) in the first round to six (6%) in the second round (mean difference = 19% [95% CI = 11% to 26%]; $P < .001$). The category distribution of adenocarcinomas detected in the second round was statistically significantly lower than the distribution of adenocarcinomas detected in the first round (Pearson chi-squared test, two-tailed $P < .001$). These results suggest that the frequency of prostate cancer with adverse prognostic features did decrease after the interval of 4 years between the two screening rounds in this study.

Baseline PSA Level in Men Who Underwent a Biopsy in the Second Round

Of the 440 participants who underwent a biopsy in the second round, the baseline

serum PSA level (i.e., the PSA level found during the first round of the study) did not differ statistically significantly between men who turned out to have prostate cancer and those who had a benign biopsy outcome (Mann-Whitney U test, two-tailed $P = .13$). Baseline serum PSA levels were even slightly lower in men with prostate cancer detected in the second round (median, 3.2 ng/mL; range, 0.7–9.3 ng/mL) than in men with a benign biopsy outcome in the second round (median, 3.7 ng/mL; range, 0.5–35.7 ng/mL). Thus, PSA values obtained during the first round of the study (baseline PSA) did not predict prostate cancer in subsequent screening rounds.

DISCUSSION

Although annual PSA screening for prostate cancer has been recommended by the American Cancer Society since 1993 (2), to our knowledge, no study has yet demonstrated that this is the optimum interval for prostate cancer screening programs. Nevertheless, performing PSA tests at an annual interval is now the most commonly used method to screen for prostate cancer in the United States. A few reports (6–8), however, have postulated that biannual screening would be more cost-efficient because, although it is not likely to miss curable prostate cancers, it does lead to substantial savings in health-care costs. These reports are based on models constructed by use of historical data on prostate cancer incidence rates and associated PSA levels; to our knowledge, they have not yet been confirmed in clinical practice.

Our study shows a substantial decrease in both the amount and the grade of screen-detected prostate cancers 4 years after an initial prevalence screen (Tables 3 and 4). Relatively few advanced (category C) tumors were found after the 4-year interval. In addition, the frequency of prostate cancer in needle biopsies dropped at high PSA ranges (Table 4). These observations suggest that large prostate cancers with high PSA values are effectively detected during a prevalence screen and that even an interval of 4 years is not long enough for most large tumors to develop.

The findings in our study are in striking contrast to those in breast cancer screening: A pooled analysis of breast cancer screening programs with a 1.5- to 3-year screening interval shows that, despite a clear reduction in breast cancer-

related mortality associated with screening, the stage of the detected tumors does not change statistically significantly during subsequent rounds of screening (13).

It is important to note that the changes in prostate cancer characteristics that we observed in the current study took place after a prevalence screen. In a previous report (14), a comparison of characteristics of prostate cancer detected at the prevalence screen with those of a series of non-screen-detected prostate cancers showed a statistically significant drop in stage and grade in the series detected during the prevalence screen. The changes between cancer detected at the first round and cancer detected at the second round of screening that were observed in the current study suggest that periodic screens at intervals of up to 4 years will have an even greater beneficial effect on the grade and stage at which prostate cancer is detected.

Only very few reports have addressed the question of whether different screening intervals would be applicable for men with different clinical profiles (e.g., age, comorbidity, or baseline serum PSA level at the beginning of the screening program). Carter et al. (6) used PSA conversion rates to conclude that biannual screening for prostate cancer can be safely recommended in men with a baseline serum PSA level below 2.0 ng/mL. Annual testing would be required only for men with PSA levels of 2.0 ng/mL or above. This approach contends that baseline PSA levels can be used to determine the length of the interval by which screening should occur. Our findings do not confirm this assumption because the baseline PSA levels (observed at the prevalence screen) did not predict the occurrence of prostate cancer in the second round.

Despite the substantial reduction in both the amount and the grade of screen-detected prostate cancer over the 4-year interval between the first and second rounds, our findings do not prove that screening for prostate cancer has a beneficial effect on prostate cancer mortality. Moreover, screening for prostate cancer potentially has the undesirable effect of leading to overtreatment for clinically unimportant tumors. However, we saw little indication of this in our study because the fraction of category A tumors showed only a moderate increase in the second round. Category A tumors largely represent small well-differentiated adenocarci-

nomas with an uncertain clinical significance. Apart from potentially dangerous carcinomas at an early stage, this category is likely to harbor cancers that will not pose a threat to their hosts during their lifetime. Therefore, a substantial increase in category A tumors would have been more alarming than reassuring.

In screening for malignant diseases, the interval between subsequent screens is very important. When the interval between screens is too long, the chance increases for the development of large incurable tumors between rounds of screening. In the studied cohort of this report, the number of interval cancers was, however, limited. A total of 304 cancers were detected during the regular screening rounds, and only 12 (3.8% of all 316 cancers) were detected between screening rounds.

Our study has several limitations. One limitation, which applies to all studies that rely on prostate sextant biopsy outcome as an endpoint, is that tumor features observed on sextant prostate needle biopsies are not necessarily representative of adenocarcinoma in the prostate gland. However, Table 1 shows that the frequency of extraprostatic tumor growth is statistically significantly higher for higher biopsy categories. The model that we used is also associated with biochemical failure after surgery (although this association is not statistically significant). The observed drop in category C tumors between the two rounds, therefore, seems to indicate a more favorable outcome for men in whom cancer is detected at periodical rounds of screening.

Another limitation is the decreased attendance rate during the second round of screening. Although an attendance rate of 79% after a 4-year interval could be considered reasonable, it might have led to a bias in our results if men who forgo further regular screens have clinical characteristics that might favor the presence or absence of prostate cancer.

Finally, our study may have been compromised by the use of different screening protocols during the first and second rounds of screening. In the second round, biopsy recommendations no longer relied on the outcome of DRE and TRUS, and the threshold of serum PSA levels as a sole tool for biopsy indication was low-

ered from 4 ng/mL to 3 ng/mL. Previous investigations of the participants in ERSPC (15,16) have shown that the number of aggressive tumors detected at low PSA levels is small. In addition, Table 4 clearly shows that differences in characteristics between tumors detected in the first round and tumors detected in the second round were most pronounced in the high PSA ranges (≥ 10 ng/mL), where the protocols of both rounds were the same. It is likely that the fraction of men with a high PSA level caused by diseases other than prostate cancer (e.g., chronic prostatitis or benign prostate hyperplasia) increased during the second round of screening, thereby accounting for the lower detection rate at high serum PSA levels and for the loss of the predictive value of serum PSA levels for both the presence and the amount of prostate cancer during the second round (Table 4).

Aside from these limitations, our results strongly suggest that, even over an interval of 4 years between screening rounds, there was no evidence of unfavorable changes in the characteristics of detected carcinomas in the subsequent rounds of prostate cancer screening. It appears that, during the prevalence screen, large prostate cancers manifested by high PSA levels are effectively detected. A screening interval of 4 years seems short enough to constrain the development of most large tumors. Moreover, baseline PSA levels found during a prevalence screen do not predict the chance of prostate cancer detection in subsequent screening rounds.

REFERENCES

- (1) Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1999. *CA Cancer J Clin* 1999;49:8-31, 1.
- (2) Mettlin C, Jones G, Averette H, Gusberg SB, Murphy GP. Defining and updating the American Cancer Society guidelines for the cancer-related checkup: prostate and endometrial cancers. *CA Cancer J Clin* 1993;43:42-6.
- (3) Franks LM. Latent carcinoma of the prostate. *J Pathol Bacteriol* 1954;68:603-16.
- (4) Sakr WA, Grignon DJ, Crissman JD, Heilbrun LK, Cassin BJ, Pontes JJ, et al. High grade prostatic intraepithelial neoplasia (HGPIN) and prostatic adenocarcinoma between the ages of 20-69: an autopsy study of 249 cases. *In Vivo* 1994;8:439-43.
- (5) Auvinen A, Rietbergen JB, Denis LJ, Schroder FH, Prorok PC. Prospective evaluation plan for randomised trials of prostate cancer screening. The International Prostate Cancer Screening Trial Evaluation Group. *J Med Screen* 1996;3: 97-104.
- (6) 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;277: 1456-60.
- (7) Etzioni R, Cha R, Cowen ME. Serial prostate specific antigen screening for prostate cancer: a computer model evaluates competing strategies. *J Urol* 1999;162:741-8.
- (8) Ross KS, Carter HB, Pearson JD, Guess HA. Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. *JAMA* 2000;284:1399-405.
- (9) Beemsterboer PM, Kranse R, de Koning HJ, Habbema JD, Schroder FH. Changing role of 3 screening modalities in the European randomized study of screening for prostate cancer (Rotterdam). *Int J Cancer* 1999;84:437-41.
- (10) Hoedemaeker RF, Kranse R, Rietbergen JB, Kruger AE, Schroder FH, van der Kwast TH. Evaluation of prostate needle biopsies in a population-based screening study: the impact of borderline lesions. *Cancer* 1999;85:145-52.
- (11) Gleason DF. Histologic grading of prostate cancer: a perspective. *Hum Pathol* 1992;23: 273-9.
- (12) Ravery V, Chastang C, Toubanc M, Boccon-Gibod L, Delmas V, Boccon-Gibod L. Percentage of cancer on biopsy cores accurately predicts extracapsular extension and biochemical relapse after radical prostatectomy for T1-T2 prostate cancer. *Eur Urol* 2000;37:449-55.
- (13) Boer R, de Koning H, van Oortmarssen G, Warmerdam P, van der Maas P. Stage distribution at first and repeat examinations in breast cancer screening. *J Med Screen* 1999;6:132-8.
- (14) Hoedemaeker RF, Rietbergen JB, Kranse R, Schroder FH, van der Kwast TH. Histopathological prostate cancer characteristics at radical prostatectomy after population based screening. *J Urol* 2000;164:411-5.
- (15) 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;90:1817-23.
- (16) Hoedemaeker RF, Rietbergen JB, Kranse R, van der Kwast TH, Schroder FH. Comparison of pathologic characteristics of T1c and non-T1c cancers detected in a population-based screening study, the European Randomized Study of Screening for Prostate Cancer. *World J Urol* 1997;15:339-45.

NOTE

Manuscript received January 5, 2001; revised May 23, 2001; accepted May 30, 2001.