EUROPEAN JOURNAL OF CANCER 47 (2011) 2698–2705

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

False-positive screening results in the European randomized study of screening for prostate cancer

Tuomas P. Kilpeläinen [a,b,*], *Teuvo L.J. Tammela* [a], *Monique Roobol* [c], *Jonas Hugosson* [d], *Stefano Ciatto* [e], *Vera Nelen* [f], *Sue Moss* [g], *Liisa Määttänen* [h], *Anssi Auvinen* [b]

[a] *Dept. of Urology, Tampere University Hospital and University of Tampere, Tampere, Finland*
[b] *Tampere School of Public Health, University of Tampere, Tampere, Finland*
[c] *Dept. of Urology, Erasmus, MC, Rotterdam, The Netherlands*
[d] *Dept. of Urology, Sahlgrenska University Hospital, Göteborg, Sweden*
[e] *Dept. of Diagnostic Medical Imaging, ISPO, Firenze, Italy*
[f] *Provinciaal Instituut voor Hygiëne, Antwerp, Belgium*
[g] *Cancer Screening Evaluation Unit, Institute of Cancer Research, Surrey, UK*
[h] *Finnish Cancer Registry, Helsinki, Finland*

ARTICLE INFO

Article history:
Available online 23 July 2011

Keywords:
Mass screening
Prostatic neoplasms
PSA
Randomized controlled trials
Sensitivity and specificity

ABSTRACT

Background: Screening for prostate cancer (PC) with prostate-specific antigen (PSA) has been shown to decrease mortality, but has adverse effects, such as false-positive (FP) screening results. We describe the frequency of FP results and assess their relation to subsequent screening attendance, test results and prostate cancer risk in a large randomized trial.

Materials and methods: We included data from five centres of the European Randomized Study of Screening for Prostate Cancer, altogether over 61,000 screened men. Men were screened with PSA test at a 2–7 year interval depending on the centre; PSA cut-off was 3.0–4.0 ng/ml. A positive screen with no histologically confirmed PC in biopsy within 1 year was defined as an FP result.

Results: Of the 61,604 men who were screened at least once, 17.8% had one or more FP result(s). Almost 20% of men who participated at all screening rounds had one or more FP result(s). More than half of the men with an FP result had another FP if screened again. Men with FP results had a fourfold risk of PC at subsequent screen (depending on the round, 10.0% versus 2.6–2.7% of men with negative screen, risk ratio 3.8–3.9). The PCs following an FP result were in 92.8% of cases localised and low-grade versus 90.4% following a screen-negative result.

Conclusions: Our results show that FP results are common adverse effects in PC screening, as they affect at least one in six screened men. False-positive men are more prone to be diagnosed with PC but are also likely to have consistently high PSA levels.

1. Introduction

The benefits of screening for prostate cancer (PC) with prostate-specific antigen (PSA) test are becoming clearer with evidence from the European Randomized Study of Screening for Prostate Cancer (ERSPC), which has shown a 20% relative decrease in mortality in the screening arm.[1] However, the adverse effects and cost-effectiveness of screening need to be

* *Corresponding author:* Address: Tampere School of Public Health and University of Tampere, FIN-33014 Tampere, Finland. Fax: +358 335516057.
E-mail address: tuomas.kilpelainen@uta.fi (T.P. Kilpeläinen).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.06.055

thoroughly investigated before decisions regarding population-based screening can be made.

Serum PSA is an organ-specific marker that may be affected by any prostatic disease. Therefore, as all screening tests, it is not perfect in sensitivity or specificity. Undetected disease constitutes a false negative (FN) finding and a positive screening result in the absence of disease is a false positive (FP) result (Table 1). The challenge in PC screening is to define and predict the disease status based on the PSA test, as not all subjects can undergo the diagnostic test, the prostatic biopsy. Even the biopsy has uncertainties: first, the needle biopsy provides only a small sample of the prostate tissue potentially missing the cancer lesion and second, the presence of malignant tissue does not necessarily mean clinically significant PC (resulting in overdiagnosis). Overdiagnosis and overtreatment occur when PCs that would not have been diagnosed in the absence of screening are detected by screening and treated.[2,3]

In addition, screening for PC with PSA has relatively high FP rate, or conversely, low specificity.[4] Previously, the results from the Finnish component of the ERSPC trial have shown that 12.5% of the screened men (at 4-year interval) had an FP result at least once during three screening rounds.[5] Similarly, 10.4% of men in the Prostate, Lung, Colorectal and Ovarian cancer screening trial (PLCO) had an FP result during four PSA tests and 3 years of follow-up.[6]

We present the proportion of FP results during three screening rounds in five centres of the ERSPC trial: Belgium, Finland, Italy, the Netherlands and Sweden – with more than 61,000 screened men. We also investigated subsequent screening compliance, PC risk and repeated FP result(s).

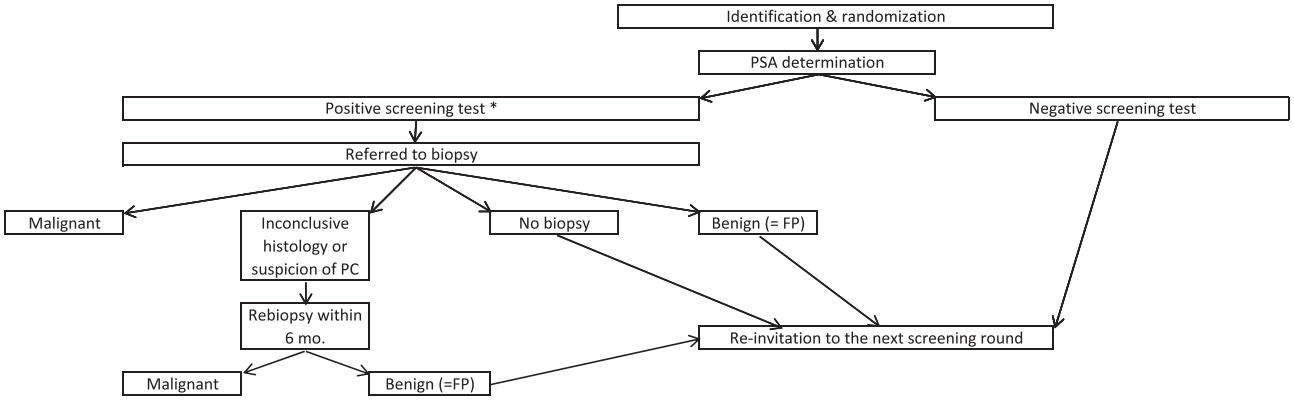
2. Materials and methods

The ERSPC trial is a multicentre study in eight European countries. In this study, we analysed data from five centres: Belgium, Finland, Italy, the Netherlands and Sweden. These five centres had data from at least three screening rounds and the largest numbers of men.

There was some variation between the centres in the screening protocol (Graph 1), mainly in the screening interval, PSA threshold, age of screened men and the mode of recruitment (Table 2). The screening protocols in the ERSPC centres have been described in detail elsewhere.[7–10] Men with a PC diagnosis and those who had emigrated from the study region were no longer invited. Men who chose not to participate were re-invited to the following rounds, except in the Netherlands. Due to lack of funding, the first screening interval in Belgium was delayed up to seven years. The Swedish centre used biennial screening and therefore had six screening rounds. In Sweden, the men with PSA <1.0 ng/ml in round 2

Table 1 – Relationship between terms regarding sensitivity and specificity.

		DISEASE		
		Positive	Negative	
TEST	Positive	True positive (TP)	False positive (FP)	Positive predictive value = TP / (TP + FP)
	Negative	False negative (FN)	True negative (TN)	Negative predictive value = TN / (FN+TN)
		Sensitivity TP / (TP+FN)	Specificity TN / (FP + TN)	



* Defined by criteria in each center, see text for details.
PC = prostate cancer; FP = false positive

Graph 1 – A general flow chart of the screening process.

Table 2 – Characteristics of the screening protocols in the ERSPC centres (screening interval, PSA cut-off, age range, recruitment mode and time of data collection).

	Interval, years	PSA cut-off, ng/ml	Age at entry (mean age)	Recruitment	Screening began	Followed up to
Belgium[a]	4–7	3.0	55–75 (64.3)	Volunteer	Jun. 1991–Dec. 2003	31.12.2007
Finland[b]	4	4.0	55–67 (60.1)	Population	Jan. 1996–Jan. 1999	31.5.2008
Italy[c]	4	4.0	55–71 (62.4)	Population	Oct. 1996–Oct. 2000	31.7.2008
Netherlands[d]	4	3.0/4.0	55–75 (63.6)	Volunteer	Nov. 1993–Mar. 2000	31.8.2008
Sweden[e]	2	3.0	50–64 (56.3)	Population	Dec. 1994	30.6.2008

PSA, prostate-specific antigen; DRE, digital rectal examination; TRUS, transrectal ultrasound.

[a] In the first round, PSA, DRE and TRUS to all. In the second round, PSA and DRE to all. PSA cut-off for biopsy 10 ng/ml in 1992–1994, 4 ng/ml in 1995–1998 and 3 ng/ml from 1999 onwards.

[b] For PSA 3.0–3.9 ng/ml: DRE in 1996–1998; DRE replaced by free/total PSA ratio with cut-off 16% from 1999 onwards.

[c] For PSA 2.5–4.0 ng/ml: DRE and TRUS (biopsy if suspicious). Biopsy for all with PSA >4.0 ng/ml.

[d] DRE and TRUS initially to all men 1993–1995; those with PSA >1.0, 1995–1997, abandoned 1997 onwards. PSA cut-off 4.0 ng/ml was lowered to 3.0 ng/ml in May 1997.

[e] No ancillary test.

were not invited to round 3, but were subsequently invited to rounds 4–6.

Incident PC cases were identified from the trial database and local/national cancer registries. Only screen-detected PCs were analysed in this study (i.e. no interval cancers). An FP result was defined as a positive screening result without a cancer diagnosis in subsequent histological examination within 1 year from the screen. Men who did not undergo biopsy were not regarded FP. Men with PC diagnosis after 1 year in e.g. a re-biopsy were classified as with interval cancers. T3-4N0M0 or T1-4N0-2M1 or Gleason score $\geqslant 8$ cancer was defined as aggressive cases.

The study protocols were reviewed and approved by appropriate ethical committees in each participating country. Written informed consent was obtained from the screened men. In Belgium and the Netherlands, men were randomised after the informed consent was obtained due to legislative reasons. In Finland, Italy and Sweden, men were randomised to screening arm and control arm prior to informed consent and only men in the screening arm were contacted.

The 95% confidence intervals (CIs) for risks and proportions were calculated on the basis of basic standard error formulae. A generalised linear model for binomial distribution with a logarithmic link function was used to calculate age-adjusted risk ratios (RRs) and their 95% CIs. Age-standardisation for the prevalences of results was done using the entire study population as reference group, based on mean age group proportions. Spearman's correlation coefficient for proportion of PC and FP results was calculated summing up all rounds and centres. Statistical analyses were performed using Stata 8.2 (StataCorp, College Station, TX, USA).

3. Results

Overall, 61,604 men screened in the five centres of ERSPC trial were analysed in this study. Of them, 22,068 (35.8%) men participated in all rounds (three rounds, except in Sweden six rounds). Altogether 4733 PCs were detected by screening, yielding a detection probability of 3.4% (95% CI 3.2–3.5) in the first round, 3.4% (3.2–3.5) in the second round and 3.6% (3.4–3.8) in the third round (Table 3).

The proportion of FP results was 10.2% (95% CI 10.0–10.5; specificity 89.8%) in the first round, 11.0% (10.7–11.3; specificity 89.0%) in the second round and 11.1% (10.7–11.5; specificity 88.9%) in the third round. There was substantial variation between the centres (Table 3). The proportion of FP results per round varied from 3.5% to 20.6% depending on age, with a higher proportion in older ages (Table 4). The proportion of PC also increased with age. Spearman's correlation coefficient was 0.78 ($p < 0.001$) between the proportion of PC and FP results per round.

Altogether 10,972 men (17.8% of those who were screened at least once) had one or more FP result (Table 5). The proportion of men with FP result(s) varied by the centre from 11.1% in Italy to 26.4% in the Netherlands. The majority (74.7%) of the men with FP result(s) had only one FP result. Of the 22,068 men who participated in all (three to six) rounds, 19.0% had one or more FP result(s).

An FP result increased the risk for a next-round FP, with approximately 50% of the men having another FP result if they participated in the following round (Table 6). A negative screen carried a risk of 6.2–7.7% for an FP result in the next round (RR 6.5–8.6 by round); the RR relative to initially screen-negative men varied by the centre 2.5–15.0. A quarter of the men with FP results dropped out of the subsequent screening round, which was 1.6-fold following the first round and 1.5-fold after the second screen compared to the screen-negative men. RRs varied by the centre from 0.7 (Sweden) to 2.1 (Finland) (Table 6).

The absolute risk of a screen-detected PC in the next round following a previous FP (at first or second round) was 10%, which was roughly fourfold (ranging from 1.7 to 14.4 depending on the centre and screening round) compared to men with a negative screen in the previous round (Table 6). The positive predictive value (PPV) was highest (22.0% in the first, 22.7% in the second and 25.0% in the third round) among those screen-positive men who had not undergone a biopsy before. The PPV for the previous-round of FP men was 14.3% and 13.6% in the second and third rounds, respectively.

A total of 681 PCs were detected at screening following an FP result. Of those, 92.8% ($N = 632$) were non-aggressive and

Table 3 – The number and proportion of results in the five centres of the trial during 3–6 rounds.

	Participation proportion, N (%)	Screen-negative, N (%)	Screen-positives, N (%)		
			False-positive	Not biopsied	Prostate cancer
All centres					
Round 1	56,064/72,210 (77.6)	47,461 (84.7)	5722 (10.2)	985 (1.8)	1896 (3.4)
Round 2	42,884/61,003 (70.3)	35,711 (83.3)	4732 (11.0)	992 (2.3)	1449 (3.4)
Round 3	27,835/42,248 (65.9)	22,929 (82.4)	3090 (11.1)	819 (2.9)	997 (3.6)
Belgium					
Round 1	4562/5178 (87.0*)	3916 (88.1*)	325 (6.1*)	214 (3.6*)	107 (2.2*)
Round 2	1987/3430 (56.8*)	1550 (80.9*)	237 (9.4*)	99 (5.8*)	101 (3.8*)
Round 3	718/1336 (51.3*)	593 (85.0*)	62 (8.5*)	49 (5.2*)	14 (1.4*)
Finland					
Round 1	20,789/30,197 (69.4*)	18,812 (90.0*)	1332 (7.0*)	102 (0.6*)	543 (2.8*)
Round 2	18,613/26,324 (71.2*)	16,309 (86.0*)	1467 (8.8*)	224 (1.4*)	613 (3.8*)
Round 3	12,739/18,376 (69.4*)	11,095 (86.3*)	978 (8.2*)	198 (1.7*)	468 (3.9*)
Italy					
Round 1	4908/5696 (85.4*)	4300 (88.0*)	377 (7.1*)	142 (3.3*)	89 (1.7*)
Round 2	4499/5607 (80.7*)	3942 (87.9*)	267 (5.9*)	217 (4.7*)	73 (1.5*)
Round 3	3292/5533 (62.2*)	2844 (86.1*)	145 (4.4*)	269 (8.5*)	34 (1.0*)
Netherlands					
Round 1	19,950/21,175 (94.3*)	15,240 (79.2*)	3225 (14.6*)	470 (1.9*)	1015 (4.3*)
Round 2	12,525/16,163 (77.4*)	9259 (74.1*)	2360 (18.7*)	355 (2.8*)	551 (4.4*)
Round 3	7711/9799 (72.5*)	5848 (75.8*)	1326 (17.2*)	217 (2.8*)	320 (4.1*)
Sweden					
Round 1	5855/9964 (57.0*)	5193 (85.1*)	463 (10.6*)	57 (1.1*)	142 (3.3*)
Round 2	5260/9479 (58.7*)	4651 (87.7*)	401 (8.0*)	97 (2.0*)	111 (2.3*)
Round 3	3375/7204 (63.0*)	2549 (76.8*)	579 (16.5*)	86 (2.4*)	161 (4.4*)
Round 4	4622/7851 (58.9)	3888 (84.1)	496 (10.7)	105 (2.3)	133 (2.9)
Round 5	4114/6674 (61.6)	3499 (85.1)	435 (10.6)	69 (1.7)	111 (2.7)
Round 6	3475/5688 (61.1)	2773 (79.8)	467 (13.4)	88 (2.5)	147 (4.2)

Proportions marked with an asterisk (*) are age-standardised to the mean age distribution of the round in all the centres.

Table 4 – Age-stratified aggregate results from the three rounds of the trial.

	Age at screen, years				
	<55	55–59	60–64	65–69	⩾70
Round 1					
Screen-negative	4579 (94.9)	18,299 (90.7)	12,614 (83.9)	9163 (77.7)	2806 (66.4)
False-positive	171 (3.5)	1299 (6.4)	1674 (11.1)	1706 (14.5)	872 (20.6)
Not biopsied	21 (0.4)	167 (0.8)	260 (1.7)	293 (2.5)	244 (5.8)
Screen-detected PC	55 (1.1)	414 (2.1)	491 (3.3)	634 (5.4)	302 (7.1)
Total	4826	20,179	15,039	11,796	4224
Round 2					
Screen-negative	792 (94.9)	8402 (91.4)	11,746 (84.9)	8741 (79.4)	6030 (75.3)
False-positive	34 (4.1)	552 (6.0)	1402 (10.1)	1470 (13.4)	1274 (15.9)
Not biopsied	7 (0.8)	82 (0.9)	278 (2.0)	326 (3.0)	299 (3.7)
Screen-detected PC	2 (0.2)	160 (1.7)	410 (3.0)	473 (4.3)	404 (5.0)
Total	835	9196	13,836	11,010	8007
Round 3					
Screen-negative	38 (84.4)	910 (82.2)	7450 (86.1)	8612 (81.9)	5919 (78.7)
False-positive	6 (13.3)	141 (12.7)	766 (8.9)	1162 (11.1)	1015 (13.5)
Not biopsied	0 (0.0)	16 (1.4)	166 (1.9)	334 (3.2)	303 (4.0)
Screen-detected PC	1 (2.2)	40 (3.6)	269 (3.1)	401 (3.8)	286 (3.8)
Total	45	1107	8651	10,509	7523

6.3% ($N = 43$) aggressive (missing information for 0.9% or six cases). Of the 1725 PCs following a screen-negative result, 90.4% ($N = 1560$) were non-aggressive and 7.8% ($N = 134$) aggressive (1.8%, $N = 31$ with missing information). The difference in proportion of aggressive cancers among those following an FP was statistically non-significantly lower compared with screen-detected cases subsequent to a screen-negative result (6.3% versus 7.8%, $p = 0.11$).

Table 5 – The prevalence of false-positive (FP) results in five centres of the trial.

	Men participating at least once, *N*	Men with FP(s), % (*N*)[a]	1 FP, % (*N*)	2 FPs, % (*N*)	3 FPs, % (*N*)	Men participating every round, *N*	Men with FP(s), *N* (%)[a]
All centres	61,604	17.8 (10,972)	74.7 (7752)	20.1 (2089)	5.2 (538)	22,068	19.0 (4186)
Belgium	4677	11.0 (569)	90.7 (516)	9.0 (51)	0.4 (2)	584	15.6 (105)
Finland	23,771	13.0 (2934)	75.2 (2207)	20.8 (611)	4.0 (116)	10,326	11.9 (1184)
Italy	5696	10.5 (635)	78.7 (500)	18.3 (116)	3.0 (19)	2597	9.0 (286)
Netherlands	19,950	26.1 (5266)	74.3 (3912)	20.2 (1063)	5.5 (291)	7711	27.8 (2228)
Sweden[b]	7510	22.3 (1568)	55.9 (876)	22.1 (347)	12.4 (195)	850 2112[c]	44.9 (383) 20.8 (442)[c]

Sweden[b]	4 FPs %, (N)	5 FPs (%)	6 FPs (%)
	5.2 (81)	3.3 (52)	1.1 (17)

[a] Age-standardised proportion.
[b] Sweden has six screening rounds.
[c] Excluding the third round of the Swedish trial (the men with PSA <1.0 ng/ml in round 2 were not invited to round 3, but were subsequently invited to rounds 4–6).

If the PSA threshold would have been 4.0 ng/ml in all the centres, the proportion of FP results would have decreased from 17.8% to 11.7% (10,972 versus 7182). However, fewer PCs would have been detected: 3481 instead of 4733 (91.5% of these PCs would have been non-aggressive, 6.7% aggressive, 1.8% unknown). In Belgium, the proportion of FP results would have been 6.4% (instead of 11.0%); in Finland 12.0% (13.0%); in Italy 9.3% (10.5%); in the Netherlands 12.4% (26.1%) and in Sweden 14.0% (22.3%).

4. Discussion

The results from five centres of the ERSPC trial show that false-positive screening results affect one in six screening participants during the course of the screening programme. Almost 20% of the men who participate in every (three to six) screening round encounter an FP result at least once. Men with FP results are often diagnosed with PC in the next round and more than half have another FP result if re-screened. The men with FP results are also more likely to drop out of the subsequent screening rounds. Our results also show that there are marked differences in the prevalence of FP results between the ERSPC centres, most likely due to differences in PSA threshold, but this could also reflect the underlying PC risk.

The ERSPC trial was launched in the early 1990s to assess whether screening for PC with PSA decreases PC mortality. The early results have shown a relative mortality decrease of 20% in the screening versus control arm.[1] When adjusted for non-attendance and contamination, the relative decrease was approximately 31%.[11] However, despite these promising results, the adverse effects and cost-effectiveness of screening need to be evaluated thoroughly if screening is to be recommended in the future.

False-positive screening results represent one aspect of the adverse effects of PC screening, in addition to overdiagnosis and overtreatment. FP results can be problematic for several reasons, even though the prostate biopsy as such seldom results in complications.[12,13] Waiting for the biopsy and afterwards the result of the biopsy can be psychologically straining to the patient, even if the biopsy eventually turns out to be negative.[14] Men with FP results commonly undergo repeated follow-up biopsies, which increase the costs of screening and could reduce the compliance.[5,14,15]

We defined an FP result as a screen-positive result without a PC diagnosis in biopsy within a year from the PSA test. The one-year time limit was adopted to ensure comparability between the centres. If this time limit is extended, some missed PCs are detected in FP men (rendering them true positive) but also PCs arising de novo after screening become more common. The definition of an FP result is problematic, however, as elevated PSA resulting in FP may indicate a PC missed in biopsy (i.e. a true positive, which overestimates the FP prevalence) or, a biopsy may result in the diagnosis of an indolent PC (and the screening test could be interpreted as FP in the sense that no clinically significant disease was diagnosed, with underestimation of the FP frequency).

High prevalence of FP results is a well-known issue with screening for PC with PSA and active search for a better screening tool has been ongoing to increase specificity. There is evidence on the usefulness of the free/total PSA ratio,[16] especially when combined with PSA density and digital rectal examination in multivariate regression models (reduction in FP results was 22%).[17] PSA velocity has not been proven very effective in increasing specificity.[18,19] New biomarkers (e.g. kallikrein-related peptidase 2, proPSA, nicked PSA, p2PSA) have shown potential, but are not yet in widespread clinical use.[20–22]

Our aggregate results from the ERSPC trial confirm that FP results are common in PC screening – one in six men have at least one FP result during the screening protocol. Three quarters of them have only one FP, but 25% have two or three. There is, however, much variation between the centres in the risk of FP result(s).

Why is the prevalence of FP results so different between the centres? One explanation could be age, as the frequency of FP results increases with age. This is most likely due to other PSA-elevating prostatic diseases than PC, such as benign prostatic hyperplasia (BPH) and prostatitis, which become more common with age.[23–25] Most of our results are,

Table 6 – Risks for subsequent round prostate cancer (PC), false-positive result (FP) and non-participation after previous round FP result versus negative screening result.

	Risk for PC after FP result (%)	Risk for PC after negative screen (%)	RR (95% CI)	Risk for FP after FP result (%)	Risk for FP after negative screen (%)	RR (95% CI)	Risk for non-participation after FP result (%)	Risk for non-participation after negative screen (%)	RR (95% CI)
All centres									
Round 1/2	10.0	2.7	3.7 (3.3–4.2)	50.1	7.7	6.5 (6.2–6.8)	26.8	17.2	1.6 (1.5–1.6)
Round 2/3	10.0	2.6	3.9 (3.3–4.5)	53.0	6.2	8.5 (8.0–9.0)	27.8	18.6	1.5 (1.4–1.6)
Belgium									
Round 1/2	7.9	4.6	1.7 (0.9–3.4)	26.7	10.6	2.5 (1.8–3.6)	49.0	36.5	1.3 (1.2–1.6)
Round 2/3	7.8	1.3	6.0 (2.0–18.3)	39.1	4.8	8.1 (5.0–13.1)	41.3	20.6	2.0 (1.5–2.6)
Finland									
Round 1/2	14.7	2.5	6.0 (4.9–7.2)	53.4	5.3	10.1 (9.2–11.0)	23.6	11.6	2.0 (1.8–2.3)
Round 2/3	11.8	2.8	4.2 (3.3–5.3)	49.8	4.6	10.7 (9.6–12.0)	27.3	13.3	2.1 (1.8–2.3)
Italy									
Round 1/2	7.0	1.0	6.7 (3.9–11.7)	37.4	3.3	11.2 (8.8–14.2)	31.8	17.2	1.8 (1.6–2.2)
Round 2/3	1.2	0.8	1.6 (0.4–6.6)	25.3	2.6	9.6 (6.8–13.7)	39.3	32.3	1.2 (1.0–1.4)
Netherlands									
Round 1/2	7.2	3.9	1.8 (1.5–2.2)	51.9	13.6	3.8 (3.6–4.1)	26.2	21.0	1.2 (1.2–1.3)
Round 2/3	9.9	2.9	3.4 (2.8–4.3)	55.2	9.2	6.0 (5.5–6.6)	28.2	18.7	1.5 (1.4–1.6)
Sweden									
Round 1/2	14.7	1.0	14.4 (9.8–21.3)	49.1	4.2	11.8 (9.9–14.1)	23.8	15.7	1.5 (1.3–1.8)
Round 2/3	11.7	3.3	3.5 (2.4–5.1)	66.9	10.8	6.2 (5.4–7.2)	15.1	22.5	0.7 (0.5–0.9)
Round 3/4	8.9	3.3	2.7 (1.8–4.1)	58.5	9.4	6.2 (5.2–7.4)	15.4	17.7	0.9 (0.7–1.1)
Round 4/5	10.4	1.5	7.0 (4.6–10.6)	57.8	3.9	15.0 (12.3–18.2)	14.5	10.2	1.4 (1.1–1.8)
Round 5/6	9.9	3.1	3.2 (2.1–4.7)	63.9	7.6	8.4 (7.2–9.9)	17.3	12.0	1.4 (1.1–1.8)

however, age-standardised, indicating that age cannot explain the differences between the centres.

Screening interval could also influence FP result prevalence, as it has been shown that a long screening interval (7 years in the Belgian centre) results in more interval cancers after 4 years from the screen.[26] No difference has been observed in the incidence of interval cancers between the biennial and 4-year intervals in the Swedish and Dutch centres,[10] and according to our results, no conclusive evidence is found to associate FP prevalence with shorter or longer screening intervals. Sweden (with shortest interval) and Belgium (with longest interval) were not the centres with most or least FP results.

Based on our results, the main reason for differences in FP prevalence appears to be PSA threshold which is related to sensitivity. Belgium, the Netherlands and Sweden all used the PSA cut-off of 3.0 ng/ml, while in Finland and Italy a higher PSA cutoff was used. Belgium, Finland and Italy had approximately 10–13% risk for an FP result, which is comparable to a previous estimate from the PLCO trial with a risk of 10% for at least one FP result in annual screening with the PSA cut-off point of 4.0 ng/ml during a 3-year screening period.[6]

This notion is further supported by the observation that if the PSA threshold would have been 4.0 ng/ml in all the centres, the FP prevalence would be 11.7% (instead of 17.8%) and would range from 6.4% (Belgium) to 14.0% (Sweden). The decrease is especially marked in the Netherlands and Sweden. FP prevalence of 11.7% would be relatively similar to the FP risk of 10.4% in the PLCO trial.[6]

The Swedish and Dutch centres had higher frequency of FP results, both exceeding 20%. A previous analysis from the ERSPC study showed that the test sensitivity was slightly higher in the Netherlands (0.93) and Sweden (0.90) compared to Finland (0.89).[27] Of the same five centres analysed in this study, the lead-time for PC has previously been estimated the longest in the Netherlands, possibly reflecting high sensitivity.[28] Based on our results, the enhanced sensitivity (relating to lower PSA threshold, shorter screening intervals) may in turn decrease specificity and cause higher prevalence of FP results. The observed correlation (coefficient 0.78) between PC detection proportion and FP proportion also reflects this.

A further explanation for differences in the FP risk could be some underlying differences in the populations from which the men are selected. It could also be possible that the volunteer-based design of the Dutch centre is subject to selection bias, as e.g. men with BPH underlying lower urinary tract symptoms could be more inclined to participate in a screening study for prostate cancer, yielding the highest FP risk. However, such a bias should also be present in the Belgian results.

Our findings show that men with FP results are very likely to have another FP at re-screening. In fact, more than 50% of the men screened after an FP result still had elevated PSA but negative biopsy. This contradictory result – as 75% men only have one FP result during three rounds – is explained by the fact that many FP men choose not to participate at next screen. We do not know whether these men would be FP or have a PC if they were rescreened. Results from the PLCO trial indicate that men with FP results are almost twice more likely to decline subsequent screening compared to men with a negative screen.[29] Similar results have been shown from the Finnish trial.[5] In this combined analysis of the ERSPC trial, the effect of FP results on non-participation was also found, although weaker.

The PPV of a positive screening test for men without previous biopsy (i.e. new screen-positive men) was similar (22–25%) both at first screen and at later screens, showing that men who turn screen-positive from previous screen-negative results are relatively likely to be diagnosed with PC, consistent with previous studies on the PPV of elevated PSA.[30] This PPV was higher than in men with a previous round FP (14%), but one must bear in mind that from this population, the prevalent PCs have already been "harvested" and that a previous round PSA test cannot be as accurate a predictor as current round PSA test. Based on a single PSA test, the FP men still are at higher risk for next-round PC than screen-negative men. It remains unknown whether this 10% risk for PC in FP men is due to more aggressive follow-up or whether these men carry a genuinely increased risk for PC. The cancers diagnosed after an FP result were mostly localised and low-grade, but 6.3% of the cancers were of aggressive nature (similar to other screen-detected cases).

Our study has some limitations. Variations in the screening protocol in the centres make it difficult to interpret the results, especially if these protocols changed during the screening years (e.g. changes in the PSA threshold or ancillary tests). Variations in the screening protocol in different centres have been previously shown to influence PC detection rates.[31] The strengths of the study include large study size, prospective design and generalizable results.

Our results from a large randomised trial with over 61,000 screened men show that one in six men had at least one FP result during repeated screening protocol. Moreover, almost 20% of men who participated in all screening rounds had FP result(s). The men with FP results are more likely to drop out from subsequent screening but upon attending, they often have another FP result or are diagnosed with PC. The prevalence of FP results could be decreased with higher PSA threshold, but this would also mean missing some cancers. False positive results remain a challenge in screening for prostate cancer and novel approaches are needed to increase the specificity of repeated screening.

Conflict of interest statement

None declared.

Acknowledgements

The international coordination of the European Randomised Study of Screening for Prostate Cancer (ERSPC) has been supported since the study's initiation in 1991 by grants from Europe Against Cancer and the fifth and sixth framework programme of the European Union, by many grants from agencies in the individual participating countries and by unconditional grants from Beckman-Coulter-Hybritech Inc. The studies in each national centre were funded by numerous local grants.

REFERENCES

1. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality in a randomized European study. *N Engl J Med* 2009;**360**:1320–8.
2. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst* 2009;**101**:374–83.
3. Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from US prostate cancer incidence trends. *J Natl Cancer Inst* 2002;**94**:981–90.
4. Määttänen L, Hakama M, Tammela TLJ, et al. Specificity of serum prostate-specific antigen determination in the Finnish prostate cancer screening trial. *Br J Cancer* 2007;**96**:56–60.
5. Kilpeläinen TP, Tammela TL, Määttänen L, et al. False-positive screening results in the Finnish prostate cancer screening trial. *Br J Cancer* 2010;**102**:469–74.
6. Croswell JM, Kramer BS, Kreimer AR, et al. Cumulative incidence of false-positive results in repeated, multimodal cancer screening. *Ann Fam Med* 2009;**7**:212–22.
7. Nelen V, Thys G, Neels H, et al. ERSPC: features and preliminary results from the Antwerp study centre. *BJU Int* 2003;**92**(Suppl. 2):17–21.
8. Mäkinen T, Tammela TL, Stenman UH, et al. Second round results of the Finnish population-based prostate cancer screening trial. *Clin Cancer Res* 2004;**10**:2231–6.
9. Ciatto S, Gervasi G, Frullini P, et al. Specific features of the Italian section of the ERSPC. *BJU Int* 2003;**92**(Suppl. 2):30–2.
10. Roobol MJ, Grenabo A, Schröder FH, et al. Interval cancers in prostate cancer screening: comparing 2- and 4-year screening intervals in the European Randomized Study of Screening for Prostate Cancer, Gothenburg and Rotterdam. *J Natl Cancer Inst* 2007;**99**:1296–303.
11. Roobol MJ, Kerkhof M, Schröder FH, et al. Prostate cancer mortality reduction by prostate-specific antigen-based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol* 2009;**56**:584–91.
12. Mäkinen T, Auvinen A, Hakama M, et al. Acceptability and complications of prostate biopsy in population-based PSA screening versus routine clinical practice: a prospective, controlled study. *Urology* 2002;**60**:846–50.
13. Carlsson S, Holmberg E, Moss S, et al. No excess mortality after prostate biopsy – Results from the European Randomized Study of Screening for Prostate Cancer (ERSPC). *BJU Int* [Epub 2010, Oct 15].
14. Fowler Jr FJ, Barry MJ, Walker-Corkery B, et al. The impact of a suspicious prostate biopsy on patients' psychological, socio-behavioral and medical care outcomes. *J Gen Intern Med* 2006;**21**:715–21.
15. Lafata J, Simpkins J, Lamerato L, et al. The economic impact of false-positive cancer screens. *Cancer Epidemiol Biomarkers Prev* 2004;**13**:2126–32.
16. Finne P, Auvinen A, Määttänen L, et al. Diagnostic value of free prostate-specific antigen among men with a prostate-specific antigen level of <3.0 μg per liter. *Eur Urol* 2008;**54**:362–70.
17. Finne P, Finne R, Bangma C, et al. Algorithms based on prostate-specific antigen (PSA), free PSA, digital rectal examination and prostate volume reduce false-positive PSA results in prostate cancer screening. *Int J Cancer* 2004;**111**:310–5.
18. Schröder FH, Roobol MJ, van der Kwast TH, et al. Does PSA velocity predict prostate cancer in pre-screened populations? *Eur Urol* 2006;**49**:460–5.
19. Etzioni RD, Ankerst DP, Weiss NS, et al. Is prostate-specific antigen velocity useful in early detection of prostate cancer? A critical appraisal of the evidence. *J Natl Cancer Inst* 2007;**99**:1510–5.
20. Bangma CH, van Schaik RH, Blijenberg BG, et al. On the use of prostate-specific antigen for screening of prostate cancer in European Randomised Study for Screening of Prostate Cancer. *Eur J Cancer* 2010;**46**:3109–19.
21. Gupta A, Roobol MJ, Savage CJ, et al. A four-kallikrein panel for the prediction of repeat prostate biopsy: data from the European Randomized Study of Prostate Cancer screening in Rotterdam, Netherlands. *Br J Cancer* 2010;**103**:708–14.
22. Jansen FH, van Schaik RH, Kurstjens J, et al. Prostate-specific antigen (PSA) isoform p2PSA in combination with total PSA and free PSA improves diagnostic accuracy in prostate cancer detection. *Eur Urol* 2010;**57**:921–7.
23. Mehik A, Hellström P, Lukkarinen O, et al. Epidemiology of prostatitis in Finnish men: a population-based cross-sectional study. *BJU Int* 2000;**86**:443–8.
24. Rhodes T, Girman CJ, Jacobsen SJ, et al. Longitudinal prostate growth rates during 5 years in randomly selected community men 40–79 years old. *J Urol* 1999;**161**:1174–9.
25. Wright EJ, Fang J, Metter EJ, et al. Prostate specific antigen predicts the long-term risk of prostate enlargement: results from the Baltimore Longitudinal Study of Aging. *J Urol* 2002;**167**:2484–7.
26. Nelen V, Thys G, Hermans A, et al. Interval cancers in the Antwerp European randomised study of screening for prostate cancer study, using a 6 year screening interval. *Eur J Cancer* 2010;**46**:3090–4.
27. Auvinen A, Raitanen J, Moss S, et al. Test sensitivity in the European prostate cancer screening trial: results from Finland, Sweden, and the Netherlands. *Cancer Epidemiol Biomarkers Prev* 2009;**18**:2000–5.
28. Finne P, Fallah M, Hakama M, et al. Lead-time in the European randomised study of screening for prostate cancer. *Eur J Cancer* 2010;**46**:3102–8.
29. Ford ME, Havstad SL, Demers R, et al. Effects of false positive prostate cancer screening results on subsequent prostate cancer screening behavior. *Cancer Epidemiol Biomarkers Prev* 2005;**14**:190–4.
30. Roobol MJ, Steyerberg EW, Kranse R, et al. A risk-based strategy improves prostate-specific antigen-driven detection of prostate cancer. *Eur Urol* 2010;**57**:79–85.
31. Otto SJ, Moss SM, Määttänen L, et al. PSA levels and cancer detection rate by centre in the European randomized study of screening for prostate cancer. *Eur J Cancer* 2010;**46**:3053–60.