ARTICLE

Radical Prostatectomy Versus Watchful Waiting in Localized Prostate Cancer: the Scandinavian Prostate Cancer Group-4 Randomized Trial

Anna Bill-Axelson, Lars Holmberg, Frej Filén, Mirja Ruutu, Hans Garmo, Christer Busch, Stig Nordling, Michael Häggman, Swen-Olof Andersson, Stefan Bratell, Anders Spångberg, Juni Palmgren, Hans-Olov Adami, Jan-Erik Johansson; for the Scandinavian Prostate Cancer Group Study Number 4

Background   The benefit of radical prostatectomy in patients with early prostate cancer has been assessed in only one randomized trial. In 2005, we reported that radical prostatectomy improved prostate cancer survival compared with watchful waiting after a median of 8.2 years of follow-up. We now report results after 3 more years of follow-up.

Methods   From October 1, 1989, through February 28, 1999, 695 men with clinically localized prostate cancer were randomly assigned to radical prostatectomy (n = 347) or watchful waiting (n = 348). Follow-up was complete through December 31, 2006, with histopathologic review and blinded evaluation of causes of death. Relative risks (RRs) were estimated using the Cox proportional hazards model. Statistical tests were two-sided.

Results   During a median of 10.8 years of follow-up (range = 3 weeks to 17.2 years), 137 men in the surgery group and 156 in the watchful waiting group died (P = .09). For 47 of the 347 men (13.5%) who were randomly assigned to surgery and 68 of the 348 men (19.5%) who were not, death was due to prostate cancer. The difference in cumulative incidence of death due to prostate cancer remained stable after about 10 years of follow-up. At 12 years, 12.5% of the surgery group and 17.9% of the watchful waiting group had died of prostate cancer (difference = 5.4%, 95% confidence interval [CI] = 0.2 to 11.1%), for a relative risk of 0.65 (95% CI = 0.45 to 0.94; P = .03). The difference in cumulative incidence of distant metastases did not increase beyond 10 years of follow-up. At 12 years, 19.3% of men in the surgery group and 26% of men in the watchful waiting group had been diagnosed with distant metastases (difference = 6.7%, 95% CI = 0.2 to 13.2%), for a relative risk of 0.65 (95% CI = 0.47 to 0.88; P = .006). Among men who underwent radical prostatectomy, those with extracapsular tumor growth had 14 times the risk of prostate cancer death as those without it (RR = 14.2, 95% CI = 3.3 to 61.8; P < .001).

Conclusion   Radical prostatectomy reduces prostate cancer mortality and risk of metastases with little or no further increase in benefit 10 or more years after surgery.

J Natl Cancer Inst 2008;100:1144–1154

Following widespread introduction of prostate-specific antigen (PSA) testing, prostate cancer incidence has increased dramatically in the Western world, including the Nordic countries (1). Many of these diagnoses represent clinically irrelevant tumors, and the risk of overdiagnosis and overtreatment poses a tremendous clinical and ethical dilemma (2,3). Central to this dilemma is the knowledge about the effects of localized radical treatment. The Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4)—which predominantly included men whose prostate cancer was not detected by PSA screening—was the first randomized trial to show that radical prostatectomy reduces the risk of prostate cancer mortality and the risk of developing metastases (4,5).

Key questions following our previous analyses of this trial, which were based on a median of 8.2 years of follow-up, include whether the absolute and relative benefits of surgical treatment would increase during longer follow-up, as we hypothesized;

Affiliations of authors: Department of Urology (AB-A, FF, MH), Regional Oncologic Center (LH, HG), Department of Pathology (CB), University Hospital, Uppsala, Sweden; Department of Urology (MR), Department of Pathology (SN), Helsinki University Central Hospital, Helsinki, Finland; King's College London, School of Medicine, Division of Cancer Studies, London, UK (LH); Department of Urology (SOA, JEJ), Center for Assessment of Medical Technology (JEJ), Örebro University Hospital, Örebro, Sweden; Department of Urology, Borås Hospital, Borås, Sweden (SB); Department of Urology, Linköping University Hospital, Linköping, Sweden (AS); Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (HOA, JP); Department of Epidemiology, Harvard School of Public Health, Boston, MA (HOA).

Correspondence to: Lars Holmberg, PhD, MD, Professor, King's College, Medical School, Division of Cancer Studies, Academic Oncology, 3rd Floor, Bermondsey Wing, Guy's Hospital, London SE1 9RT, UK (e-mail: lars.holmberg@kcl.ac.uk).

See "Funding" and "Notes" following "References."

DOI: 10.1093/jnci/djn255

whether overall mortality would remain reduced; whether the benefit is larger among younger than among older patients; and whether histopathologic parameters can predict lethal outcome following surgery (4).

We now present 12-year estimates of mortality and the risk of developing metastases. The results include 115 prostate cancer deaths and 163 instances of distant metastases. A histopathologic review of the prostatectomy specimens and its association with prognosis are presented.

Subjects and Methods

Study Design

All details concerning the study design have been published before (5), and the study protocol, as defined in 1988, is available at www.roc.se. The study was approved in all participating centers' regional ethics committees.

From October 1, 1989, to February 28, 1999, a total of 695 men were enrolled in the study by urologists at 14 centers in Sweden, Finland, and Iceland. After oral informed consent, participants were randomly assigned to radical prostatectomy or to watchful waiting, with stratification according to tumor grade and randomization center. The randomization list was computer generated, and the block size was unknown to the investigators. All random assignments were done by a telephone service from the trial office located outside the clinics. Initially, the cumulative 5-year disease-specific survival in the watchful waiting group was assumed to be 85%. We wished to detect a reduction of mortality from prostate cancer in the surgery group leading to a disease-specific 5-year survival of greater than 95%. The risk of type I error was accepted as 5% with a two-sided test and the risk of type II error as 20%, leading to a requirement of 233 patients for each group. The initial target was set at 520 patients. We planned two interim analyses of disease-specific survival using Kaplan–Meier curves and the log-rank test with group identification blinded to the analyst. One analysis was scheduled after inclusion of 300 patients, the other after 520 were randomly assigned. The results of both interim analyses were compatible with safety, but the overall event rate was lower than anticipated. Therefore, the Steering Committee decided to increase the target to 700 patients. With that sample size, and preserved levels of type one and two errors, we would be able to detect a 6% absolute survival difference between the study arms if the prostate cancer–specific survival was 95% in the best arm. The eligibility criteria were as follows: younger than 75 years of age, an estimated life expectancy of more than 10 years, and no other known malignancies. The tumor had to be newly diagnosed based on core biopsy or needle aspiration with high to moderately high differentiation, according to the definition of the World Health Organization. The tumor also had to be localized, including stage T0d (later changed to T1b), T1, or T2, according to the 1978 criteria of the International Union against Cancer; after 1994, men with T1c tumors were also accepted, according to the revised criteria of 1987 (6,7). Eligible patients also had to have a PSA level less than 50 ng/mL and a negative bone scan. In all but two centers, a modified version of Zelen's randomization model (8) was allowed in 1988–1990. The use of Zelen's model implied that only men in the experimental group received complete information

CONTEXT AND CAVEATS

Prior knowledge
The Scandanavian Prostate Cancer Group (SPCG)-4 randomized trial began in 1989 in Sweden to compare the outcomes of prostate cancer patients who were assigned to watchful waiting vs radical prostatectomy. In the previous report in 2005, at a median 8.2 years of follow-up, men who underwent radical prostatectomy had better disease-specific survival rates than men in the watchful waiting group.

Study design
Continued follow-up of the SPCG-4 randomized trial.

Contributions
After a median follow-up of 10.8 years, 13.5% of men in the radical prostatectomy group and 19.5% of men in the watchful waiting group died from prostate cancer; at 12 years, the percentages were 12.5% and 17.9%, respectively. However, at 12 years, the overall mortality in the two arms was not statistically significantly different (32.7% vs 39.8%).

Implications
No increase in benefit of radical prostatecomy was observed after 10 years.

Limitations
It is unclear whether these results would apply to today's Western male populations, who, unlike the men in the SPCG-4 trial, are diagnosed with prostate cancer mainly by prostate-specific antigen screening. Quality of life comparisons were not performed.

From the Editors

about the study after randomization, but all patients were informed that they were taking part in a clinical study. In 1990, when 68 men had been randomly assigned from these centers, it became clear that Zelen's model was not beneficial for randomization and thus all men were fully informed thereafter.

The surgery started with a pelvic lymphadenectomy of the obturator fossa (9). If intraoperative histolopathologic evaluation of sections of the removed lymph nodes showed no evidence of tumor spread at frozen section, a radical prostatectomy was performed (10). Men in the watchful waiting group received no immediate treatment. Until 2003, hormonal therapy was recommended to be initiated for men in the radical prostatectomy group if there were signs of tumor recurrence, whereas men in the watchful waiting group with symptoms of urethral obstruction underwent transurethral resection as first-line treatment. Both groups received hormonal therapy if metastases were confirmed. The hormonal therapies used were mainly orchiectomy or gonadotropin-releasing hormone analogs as lifelong therapy. In 2003, an amendment to the protocol allowed all men to have hormonal therapy if their physician advised it due to any sign of tumor progression, including elevation of PSA levels in asymptomatic patients.

Histopathologic Review

Using uniform criteria, all diagnostic core biopsies that were obtained before randomization were centrally reviewed in 1999

by study uropathologists and graded according to Gleason (11). In 2006, the radical prostatectomy specimens for all men in the radical prostatectomy group who had undergone radical prostatectomy within 1 year after randomization were retrieved and centrally reviewed by five uropathologists. The prostatectomy specimens had been fixed in 10% neutral-buffered formalin for 48 hours, surgical margins were marked with ink for the purpose of orientation, and fixed specimens were then sliced at 4-mm intervals. The basal and apical ends were cut at 4-mm intervals in the sagittal plane. The seminal vesicles were cut off and sectioned separately. In the central review, Gleason grade, Gleason score, the extent of extracapsular growth, and evidence of positive surgical margins were assessed. Seminal vesicle involvement was classified as extracapsular growth. The pathologists were blinded to patient outcome and assignment. In this report, only the results from the central review are used.

Follow-up and Definition of Clinical Events

All randomly assigned men were followed every 6 months during the first 2 years and annually thereafter for a clinical examination and blood tests, including PSA assay, at all participating centers. Initially, a bone scan was obtained annually; after 2003, it was done every second year in patients with no clinical signs or PSA level increases that would suggest disease progression. Metastases were considered to be present if there was a positive bone scan. Local recurrence was defined as a histologically verified local lesion or a palpable mass at the place of the prostate in the radical prostatectomy group and as a palpable transcapsular tumor growth or voiding disorder necessitating intervention in the watchful waiting group. Deaths were ascertained via the continuous monitoring of the trial and the official vital statistics in the participating countries. Follow-up was complete for all participants through December 31, 2006.

Two of the investigators (A. Bill-Axelson, F. Filén) and the principal study monitor collected all information according to a specific protocol concerning the men who had died and made standardized extracts of the patients' history without revealing allocated group. An independent endpoint committee with two urologists (J.-E. Damber, E. Varenhorst) and one pathologist (A. Lindgren) then individually determined death category in one of six categories: death from prostate cancer; death from another main cause but with distant metastases, regardless of local status; death from another main cause but with local progression without distant metastases; death from another cause but with local progression and with unknown status concerning distant metastases; death from another main cause without any evidence of tumor recurrence; and death from another main cause within the first month after random assignment.

Members of the endpoint committee were blinded to patients' group assignment and treatment received. The endpoint committee reached consensus for cases they initially classified differently.

Statistical Analysis

Analyses were based on a protocol decision from the steering committee before the first analysis of the trial, namely, to analyze data every third year. We analyzed three main endpoints: death from any cause, death from prostate cancer (with death from other causes treated as competing risk), and distant metastases (with death from other causes treated as competing risk). All analyses were carried out according to the intention-to-treat principle and included the 68 men who were randomly assigned according to Zelen's model. Relative risks (RRs) with 95% confidence intervals (CIs) and differences in cumulative incidence (with 95% CIs) were used as effect measures for each endpoint. We used Gray's test (12) to assess differences between the treatment groups, with a P value of less than .05 (two-sided) to indicate statistical significance. Relative risks were estimated from a Cox proportional hazards model. Proportionality was verified by visual inspection of the parallelism of the logarithms of the estimated cumulative hazards. Cumulative incidence was used to account for the fact that the endpoints constitute competing events (13).

The effect of radical prostatectomy on nonmortality outcomes was estimated with the difference in cumulative incidence (with 95% CIs) between the study arms regarding: 1) local recurrence and/or progression, 2) start of hormonal treatment, and 3) initiation of palliative treatment other than hormones, including palliative radiation, cytotoxic drugs, and laminectomy.

To assess a possible modification of the effect of radical prostatectomy by age, PSA level at diagnosis, or Gleason score, three subgroup analyses were carried out. The subgroups were not defined in the protocol but were specified before any of the data were seen: analyses according to age at diagnosis (<65 years of age vs ≥65 years); analyses according to PSA level at diagnosis (<10 ng/mL vs ≥10 ng/mL); and analyses according to the Gleason score of the prerandomization biopsy specimen (<7 vs ≥7, on a scale of 2–10, with 10 indicating the most poorly differentiated tumors). Any modification of the effect of radical prostatectomy according to subgroup was tested by a Cox proportional hazards model that included an interaction term between subgroup category and randomization group. In a second step, we further explored the interaction by including possible effect modifiers (age, PSA level at diagnosis, Gleason score) as continuous variables in the Cox proportional hazards model. When there was an indication of effect modification, we further controlled for age, PSA level, tumor stage (T1b, T1c, and T2), Gleason score, and year at inclusion (before 1994 or thereafter) by adding these as additional covariates in a Cox proportional hazards model.

We assessed the association between prognosis and three histopathologic parameters—margins (positive, negative), extracapsular extension (not present, any extension), and Gleason score (2–6, 7, 8–10) in the radical prostatectomy specimens—expressed as relative risks obtained from multivariable Cox proportional hazards models. One model included the histopathologic parameters one by one and was adjusted for age group (<65 years vs ≥65 years). The other model included the histopathologic parameter of interest but was adjusted for the two other histopathologic parameters and for age group.

Results

The men who were randomly assigned to radical prostatectomy (n = 347) and to watchful waiting (n = 348) had similar baseline characteristics, with a mean age of 65 years. The vast majority had palpable tumors, with only 12% having T1c (nonpalpable)

Table 1. Baseline clinical characteristics of men in the SPCG-4 study by randomization group and age at diagnosis*

Characteristic	All ages		Age <65 y		Age ≥65 y	
	Radical prostatectomy (n = 347)	Watchful waiting (n = 348)	Radical prostatectomy (n = 157)	Watchful waiting (n = 166)	Radical prostatectomy (n = 190)	Watchful waiting (n = 182)
Mean age (SD), y	64.6 (5.1)	64.5 (5.0)	60 (3.5)	60.2 (3.4)	68.4 (2.5)	68.4 (2.4)
Mean PSA, ng/mL	13.5	12.3	12.7	12.4	14.2	12.2
Tumor stage, No. (%)						
T1b	33 (9.5)	50 (14.4)	14 (8.9)	22 (13.3)	19 (10.0)	28 (15.4)
T1c	43 (12.4)	38 (10.9)	24 (15.3)	21 (12.7)	19 (10.0)	17 (9.3)
T2	270 (77.8)	259 (74.4)	119 (75.8)	123 (74.1)	151 (79.5)	136 (74.7)
Unknown	1 (0.3)	1 (0.3)	0 (0.0)	0 (0.0)	1 (0.5)	1 (0.5)
WHO grade, No. (%)						
1	168 (48.4)	166 (47.7)	76 (48.4)	83 (50.)	92 (48.4)	83 (45.6)
2	178 (51.3)	182 (52.3)	81 (51.6)	83 (50.)	97 (51.1)	99 (54.4)
Unknown	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)	0 (0.0)
Gleason score at biopsy, No. (%)						
2–4	45 (13.0)	46 (13.2)	25 (15.9)	27 (16.3)	20 (10.5)	19 (10.4)
5–6	165 (47.6)	166 (47.7)	69 (43.9)	81 (48.8)	96 (50.5)	85 (46.7)
7	77 (22.2)	82 (23.6)	33 (21.0)	32 (19.3)	44 (23.2)	50 (27.5)
8–10	14 (4.0)	21 (6.0)	5 (3.2)	10 (6.0)	9 (4.7)	11 (6.0)
Unknown†	46 (13.3)	33 (9.5)	25 (15.9)	16 (9.6)	21 (11.1)	17 (9.3)
Method of detection, No. (%)						
Opportunistic screening‡	18 (5.2)	18 (5.2)	9 (5.7)	7 (4.2)	9 (4.7)	11 (6.0)
Coincidental§	87 (25.1)	91 (26.1)	41 (26.1)	46 (27.7)	46 (24.2)	45 (24.7)
TURP	40 (11.5)	56 (16.1)	19 (12.1)	29 (17.5)	21 (11.1)	27 (14.8)
Symptoms	152 (43.8)	138 (39.7)	68 (43.3)	66 (39.8)	84 (44.2)	72 (39.6)
Other	49 (14.1)	44 (12.6)	20 (12.7)	17 (10.2)	29 (15.3)	27 (14.8)
Unknown	1 (0.3)	1 (0.3)	0 (0.0)	1 (0.6)	1 (0.5)	0 (0.0)
PSA level, No. (%)						
<4 ng/mL	43 (12.4)	63 (18.1)	27 (17.2)	32 (19.3)	16 (8.4)	31 (17.0)
4–6.9 ng/mL	60 (17.3)	60 (17.2)	33 (21.0)	24 (14.5)	27 (14.2)	36 (19.8)
7–10 ng/mL	68 (19.6)	67 (19.3)	23 (14.6)	28 (16.9)	45 (23.7)	39 (21.4)
10.1–20 ng/mL	100 (28.8)	95 (27.3)	43 (27.4)	51 (30.7)	57 (30.0)	44 (24.2)
>20 ng/mL	69 (19.9)	60 (17.2)	30 (19.1)	28 (16.9)	39 (20.5)	32 (17.6)
Unknown	7 (2.0)	3 (0.9)	1 (0.6)	3 (1.8)	6 (3.2)	0 (0.0)
Positive margins in RPS, No. (%)						
0 mm	184 (64.8)		100 (73.0)		84 (57.1)	
1–9 mm	50 (17.6)		18 (13.1)		32 (21.8)	
10–19 mm	25 (8.8)		11 (8.0)		14 (9.5)	
≥20 mm	24 (8.5)		8 (5.8)		16 (10.9)	
Missing data	1 (0.4)		0 (0.0)		1 (0.7)	
Extracapsular extension in RPS, No. (%)						
0 mm	151 (53.2)		84 (61.3)		67 (45.6)	
1–9 mm	46 (16.2)		19 (13.9)		27 (18.4)	
10–19 mm	38 (13.4)		14 (10.2)		24 (16.3)	
≥20 mm	48 (16.9)		20 (14.6)		28 (19.0)	
Missing data	1 (0.4)		0 (0.0)		1 (0.7)	
Gleason score of RPS, No. (%)						
2–6	88 (31.0)		46 (33.6)		42 (28.6)	
7: 3+4	87 (30.6)		40 (29.2)		47 (32.0)	
7: 4+3	70 (24.6)		35 (25.5)		35 (23.8)	
8–10	38 (13.4)		15 (10.9)		23 (15.6)	
Missing data	1 (0.4)		1 (0.7)		0 (0.0)	

* SPCG = The Scandinavian Prostate Cancer Group Study Number 4; PSA = prostate-specific antigen; TURP = transurethral resection of the prostate; RPS = radical prostatectomy specimens; WHO =World Health Organization. Percentages may not add to 100% due to rounding error.

† Diagnosis was made by cytologic examination only in 55 patients; a biopsy specimen could not be retrieved in 24 patients.

‡ PSA levels assayed due to opportunistic screening in asymptomatic men.

§ Coincidental detection due to other symptoms for which rectal examination was carried out.

tumors (Table 1). At the end of 2006, when the database was closed, 294 of the 347 men in the radical prostatectomy group had undergone radical prostatectomy and 298 of the 348 who were randomly assigned to watchful waiting had not undergone any curative treatment (Figure 1). The median follow-up time was 10.8 years (range = 3 weeks to 17.2 years).
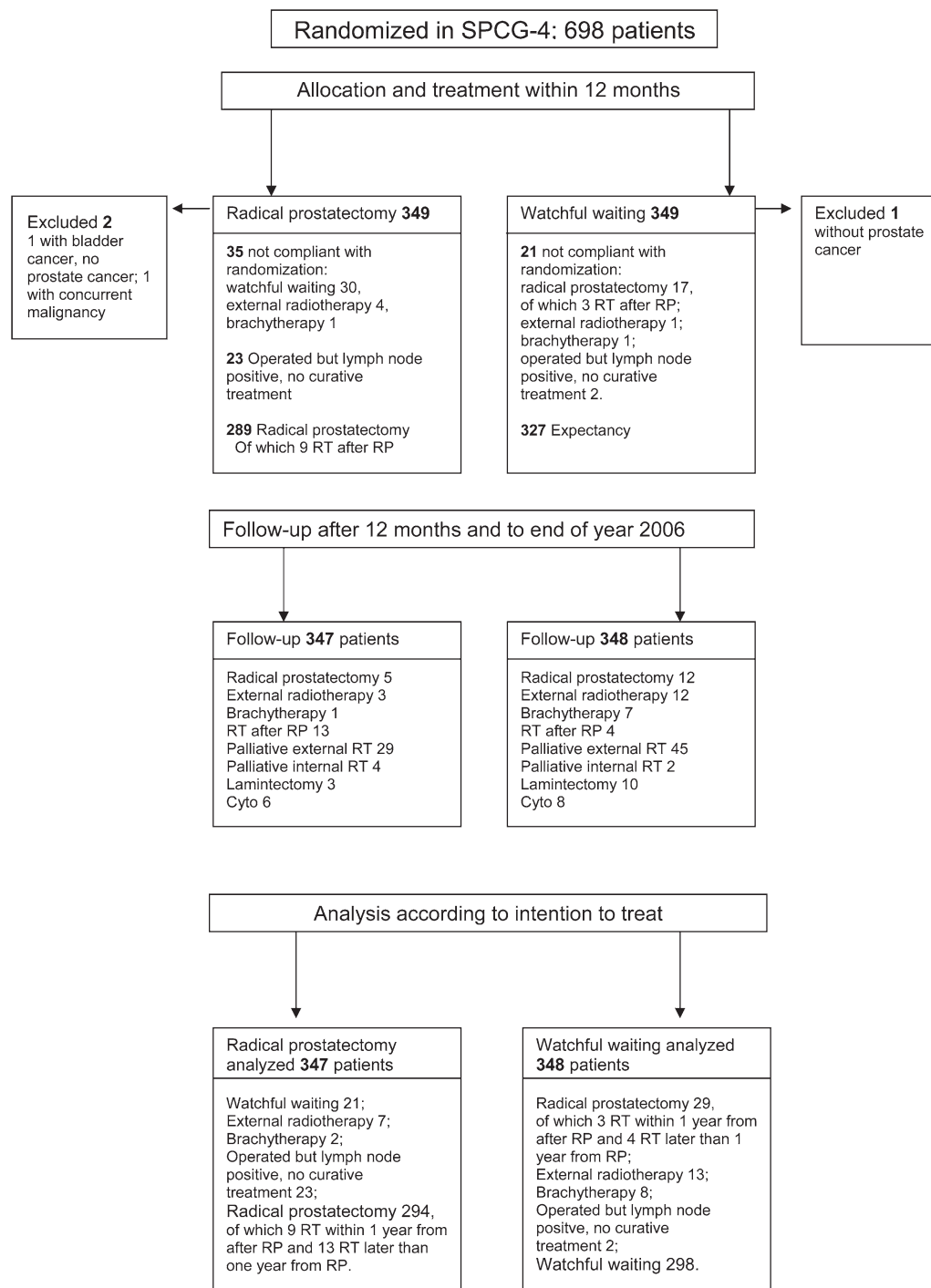
Figure 1. Trial flow diagram of the 695 men randomly assigned in the Scandinavian Prostate Cancer Group-4 study. RT = radiation therapy; RP = radical prostatectomy.

Mortality

By the end of 2006, 137 men in the radical prostatectomy group and 156 men in the watchful waiting group had died. The cumulative incidence at 12 years for overall mortality was 32.7% in the radical prostatectomy group and 39.8% in the watchful waiting group (difference = 7.1%, 95% CI = –0.5 to 14.7%). This nonstatistically significant difference corresponds to an RR of 0.82 (95% CI = 0.65 to 1.03; P = .09; Figure 2, A; Tables 2 and 3). The number of deaths from causes other than prostate cancer was 90 in the radical prostatectomy group and 88 in the watchful waiting group (P = .95).

Forty-seven of the 347 (13.5%) men in the radical prostatectomy group and 68 of the 348 (19.5%) men in the watchful waiting group died of prostate cancer by the end of 2006. In the radical prostatectomy group, the cumulative incidence at 12 years was 12.5%, and in the watchful waiting group, it was 17.9% (difference = 5.4%, 95% CI = 0.2 to 11.1%), corresponding to an RR of 0.65 (95% CI = 0.45 to 0.94; P = .03; Figure 2, B; Table 2). The difference between the two groups in the cumulative incidence of death from prostate cancer remained constant beyond 9 years of follow-up (Figure 2, B).
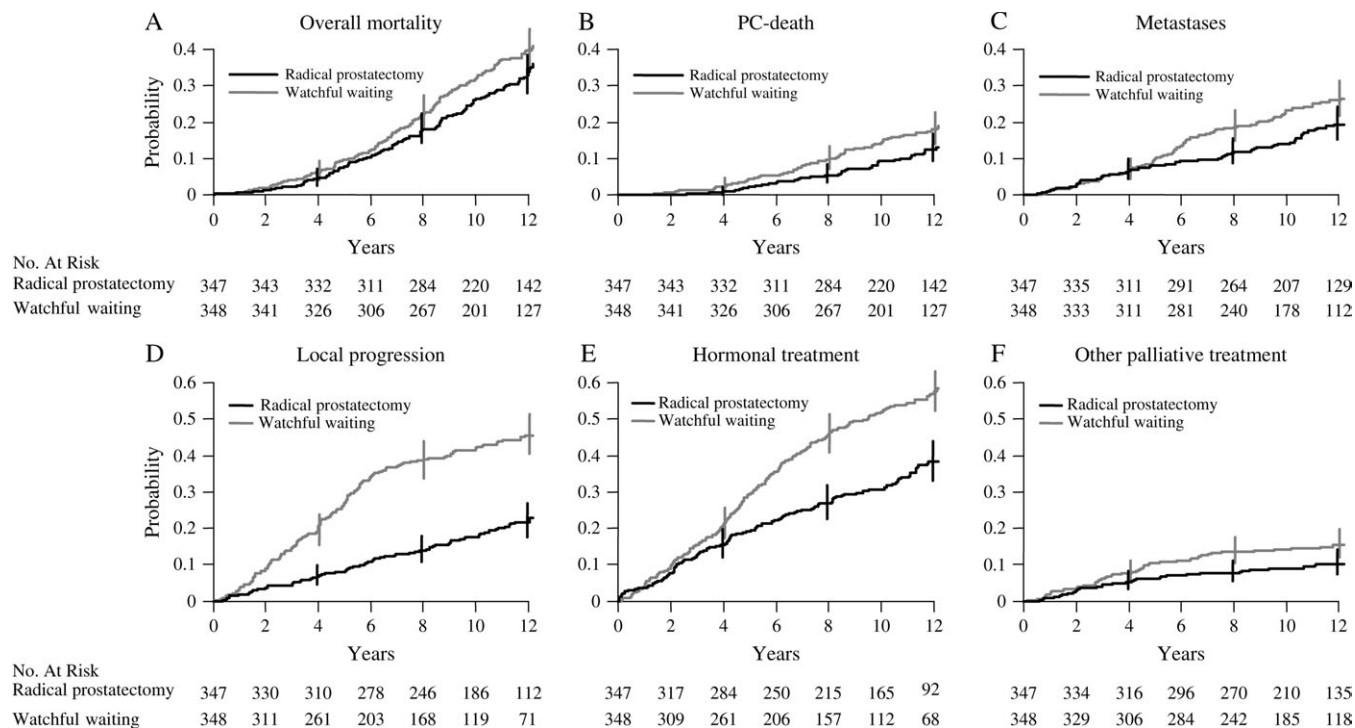
Figure 2. Cumulative incidence with 95% confidence intervals (CIs) at 4, 8, and 12 years of endpoints for all patients. **A**) Overall mortality: relative risk (RR) = 0.82, 95% CI = 0.65 to 1.03; *P* = .09. **B**) Prostate cancer (PC) death: RR = 0.65, 95% CI = 0.45 to 0.94; *P* = .03. **C**) Metastases: RR = 0.65, 95% CI = 0.47 to 0.88; *P* = .006. **D**) Local progression: RR = 0.36, 95% CI = 0.27 to 0.47; *P* < .001. **E**) Hormonal treatment: RR = 0.54, 95% CI = 0.44 to 0.68; *P* < .001. **F**) Other palliative treatment: RR = 0.63, 95% CI = 0.41 to 0.97; *P* = .04. *P* values (two-sided) were calculated using Gray's test.

Distant Metastases

Distant metastases were detected in 67 of the 347 men in the radical prostatectomy group and in 96 of the 348 men in the watchful wait-ing group by the end of 2006. The cumulative incidence of distant metastases at 12 years was 19.3% in the radical prostatectomy group and 26% in the watchful waiting group (difference = 6.7%,

Table 2. Cumulative incidence, absolute risk reductions, and relative risk, with 95% confidence intervals (CIs) for overall mortality, prostate cancer mortality, and distant metastases in men randomly assigned to radical prostatectomy vs watchful waiting*

	Cumulative incidence				Absolute risk reduction, % (95% CI)	Relative risk (95% CI)	*P*
	Radical prostatectomy		Watchful waiting				
Outcome	No.	% (95% CI)	No.	% (95% CI)			
Overall mortality	137		156				
All ages, 8 y		17.9 (14.3 to 22.4)		22.4 (18.4 to 27.3)	4.6 (−1.4 to 10.5)		
All ages, 12 y		32.7 (27.9 to 38.4)		39.8 (34.7 to 45.7)	7.1 (−0.5 to 14.7)	0.82 (0.65 to 1.03)	.09
Age < 65, 8 y		12.1 (7.9 to 18.5)		23.5 (17.8 to 30.9)	11.4 (3.1 to 19.6)		
Age < 65, 12 y		21.9 (16.1 to 29.9)		40.2 (33.0 to 49.0)	18.3 (7.8 to 28.8)	0.59 (0.41 to 0.85)	.004
Age ≥ 65, 8 y		22.6 (17.4 to 29.5)		21.4 (16.2 to 28.3)	−1.2 (−9.6 to 7.30)		
Age ≥ 65, 12 y		42 (35 to 50.5)		39.3 (32.5 to 47.7)	−2.7 (−13.5 to 8.0)	1.04 (0.77 to 1.40)	.81
Prostate cancer–specific mortality	47		68				
All ages, 8 y		5.5 (3.5 to 8.5)		9.8 (7.1 to 13.5)	4.3 (0.4 to 8.2)		
All ages, 12 y		12.5 (9.2 to 16.8)		17.9 (14.1 to 22.7)	5.4 (−0.2 to 11.1)	0.65 (0.45 to 0.94)	.03
Age < 65, 8 y		5.1 (2.6 to 10)		13.3 (9.0 to 19.6)	8.2 (1.9 to 14.4)		
Age < 65, 12 y		11.9 (7.5 to 18.7)		23.1 (17.2 to 30.9)	11.2 (2.6 to 19.8)	0.50 (0.30 to 0.84)	.014
Age ≥ 65, 8 y		5.8 (3.3 to 10.3)		6.6 (3.8 to 11.4)	0.8 (−4.1 to 5.7)		
Age ≥ 65, 12 y		13.1 (8.8 to 19.5)		13.2 (8.9 to 19.6)	0.1 (−7.3 to 7.5)	0.87 (0.51 to 1.49)	.55
Distant metastases	67		96				
All ages, 8 y		11.5 (8.6 to 15.4)		18.7 (15 to 23.3)	7.2 (1.8 to 12.5)		
All ages, 12 y		19.3 (15.3 to 24.2)		26 (21.6 to 31.2)	6.7 (0.2 to 13.2)	0.65 (0.47 to 0.88)	.006
Age < 65, 8 y		10.8 (6.9 to 17)		22.9 (17.3 to 30.3)	12.1 (4.0 to 20.1)		
Age < 65, 12 y		20.7 (15 to 28.6)		30.3 (23.8 to 38.5)	9.6 (−0.3 to 19.5)	0.52 (0.34 to 0.81)	.006
Age ≥ 65, 8 y		12.1 (8.2 to 17.8)		14.8 (10.5 to 21)	2.7 (−4.2 to 9.7)		
Age ≥ 65, 12 y		17.9 (13 to 24.6)		22 (16.5 to 29.3)	4.1 (−4.4 to 12.6)	0.80 (0.51 to 1.27)	.28

* Numbers of men with each outcome are as of the end of 2006. Results are shown by 8 and 12 years of follow-up and by age group. *P* values (two-sided) were calculated using Gray's test.

Table 3. Causes of death by randomization arm*

Reason for death	Radical prostatectomy	Watchful waiting
Prostate cancer, No.	47	68
Other causes, No.	90	88
Other main cause, with metastases	5	10
Other main cause, without metastases but with local progression or recurrence	10	20
Other main cause, with unknown status regarding metastases but with local progression	0	1
Other main cause, with no evidence of metastases or local progression or recurrence	74	56
Other main cause, within first month after randomization	1	1
Total deaths, No.	137	156

* All events were evaluated by the endpoint committee.

95% CI = 0.2 to 13.2%), corresponding to an RR of 0.65 (95% CI = 0.47 to 0.88; *P* = .006; Figure 2, C; Table 2). As with deaths from prostate cancer, the difference between the two groups in cumulative incidence of metastases remained constant beyond about 7 years of follow-up (Figure 2, C).

Nonmortality Outcomes

During follow-up, fewer men in the radical prostatectomy group than in the watchful waiting group developed local recurrence and/or progression (73 vs 161). The cumulative incidence at 12 years was 21.7% in the surgery group and 45.6% in the watchful waiting group (difference = 23.9, 95% CI = 16.8 to 30.9%) (Figure 2, D).

A total of 132 men in the radical prostatectomy group vs 205 in the watchful waiting group started hormonal therapy. The difference in cumulative incidence of use of hormonal therapy at 12 years was 19.3% (95% CI = 11.7 to 26.9%), with a lower probability of starting hormonal therapy in the radical prostatectomy group (Figure 2, E). There was no difference between the groups in time to starting hormonal therapy. All other palliative treatments were also less common in the radical prostatectomy group; palliative radiation (33 vs 47 men), treatment with cytotoxic drugs (6 vs 8 men), and laminectomy (3 vs 10 men). The difference in the cumulative incidence of palliative treatment at 12 years, at which the first of any of these events is included, was 5.2% (95% CI = 0.1 to 10.3%; Figure 2, E and F).

Subgroup Analyses

Neither PSA level at randomization nor Gleason score modified the effect of radical prostatectomy (*P* > .20, test of interaction). In contrast, age was an independent modifier of the effect of radical prostatectomy. For overall mortality, the interaction term between age (<65 years vs ≥65 years) and treatment was statistically significant (*P* = .015) and remained so also when age was considered as a continuous variable (*P* = .006). However, no statistically significant interaction was observed for age and prostate cancer death (*P* = .13, for interaction with age at cutoff of 65 years).

Among men who were younger than 65 years at diagnosis, statistically significant risk reductions in the radical prostatectomy group were observed for all three endpoints, varying between 40% and 50% in relative terms and 10% and 18% in absolute risk reductions at 12 years of follow-up (Figure 3, Table 2). For men 65 years and older, however, there was no discernible difference between the groups in any outcome (Figure 3, Table 2). The interaction between age and treatment did not change when the model was further adjusted for PSA level, Gleason score, and tumor stage.

Histopathologic Parameters

Positive surgical margins were present in 99 of the 283 (35%) evaluable prostatectomy specimens. Positive margins were associated with prognosis in a model that adjusted only for age.
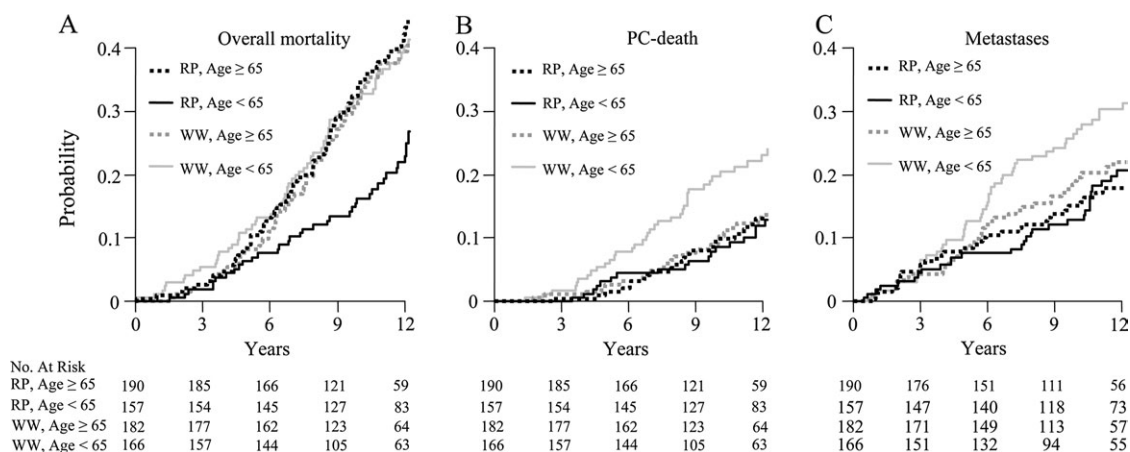
Figure 3. Cumulative incidence of endpoints by age group. **A**) Overall mortality: Age < 65: Relative risk (RR) 0.59, 95% confidence interval (CI) = 0.41 to 0.85; *P* = .004. Age ≥ 65: RR = 1.04, 95% CI = 0.77 to 1.4; *P* = .81. **B**) Prostate cancer (PC) death. Age < 65: RR = 0.5, 95% CI = 0.3 to 0.84; *P* = .01. Age ≥ 65: RR = 0.87, 95% CI = 0.51 to 1.49; *P* = .55. **C**) Metastases. Age < 65: RR = 0.52, 95% CI = 0.34 to 0.81; *P* = .006. Age ≥ 65: RR = 0.8, 95% CI = 0.51 to 1.27; *P* = .28. *P* values (two-sided) were calculated using Gray's test.

Table 4. The association between the histopathologic parameters from the review of the radical prostatectomy specimens and risk of prostate cancer death and distant metastases*

Outcome	No. of men	No. of events	HR (95% CI)†	HR (95% CI)‡
Disease-specific mortality				
Negative margins	184	14	Referent	Referent
Positive margins	99	15	2.28 (1.08 to 4.80)	1.09 (0.49 to 2.42)
Extracapsular extension 0 mm	151	2	Referent	Referent
Extracapsular extension > 0 mm	132	26	19.56 (4.62 to 82.8)	14.2 (3.26 to 61.8)
Gleason score, 2–7§	245	15	Referent	Referent
Gleason score, 8–10	38	14	6.92 (3.32 to 14.4)	5.05 (2.37 to 10.7)
Distant metastases				
Negative margins	184	22	Referent	Referent
Positive margins	99	21	2.02 (1.09 to 3.71)	0.88 (0.45 to 1.71)
Extracapsular extension 0 mm	151	7	Referent	Referent
Extracapsular extension >0 mm	132	35	7.59 (3.34 to 17.25)	5.32 (2.24 to 12.6)
Gleason score, 2–6	88	3	Referent	Referent
Gleason score 7	157	23	4.95 (1.48 to 16.5)	2.59 (0.72 to 9.36)
Gleason score, 8–10	38	17	18.8 (5.46 to 64.8)	9.90 (2.66 to 36.8)

* HR = hazard ratio; CI = confidence interval; PSA = prostate-specific antigen.

† Adjusted for age (≥65 years or <65 years).

‡ Adjusted for age (≥65 years or <65 years), PSA level (<10 ng/mL vs ≥10 ng/mL), and mutually for margins, extracapsular extension, and Gleason score at central review with categories for the latter three as specified in the table.

§ No events in men with Gleason score 2–6.

However, when extracapsular tumor growth, PSA level, and Gleason score from the preoperative biopsies were included in the model, the relative risk associated with positive margins was close to 1.0 (Table 4).

Extracapsular tumor growth was found in 132 of all 284 (46%) radical prostatectomy specimens. Compared with men without extracapsular tumor growth, those with extracapsular growth experienced a 14-fold (multivariable RR = 14.2, 95% CI = 3.26 to 61.8; $P < .001$) risk of prostate cancer death (Figure 4, A; Table 4). Only two men without evidence of extracapsular tumor growth died from prostate cancer, and the first death occurred after more than 11 years of follow-up. Furthermore, among men with Gleason score 2–6 in the radical prostatectomy specimens, no prostate cancer deaths occurred during follow-up (Figure 4, B; Table 4).

Discussion

In our previous analyses, which were based on a median of 8.2 years of follow-up (4), we found relative reductions of 40% in risk of metastases, 44% in risk of prostate cancer mortality, and 26% in overall mortality in favor of radical prostatectomy. After added follow-up, now to a median of 10.8 years, we found a relative reduction of 35% in risk of metastases, 35% in prostate cancer death, and 18% in overall mortality in favor of radical prostatectomy. These relative risks correspond to absolute risk reductions at 12 years of 6.7%, 5.4%, and 7.1%, respectively. Contrary to our predictions based on shorter follow-up (5), the absolute difference in cumulative incidence of distant metastasis and prostate cancer death did not further increase after 7–9 years of follow-up. The relative risk of disease-specific mortality increased after 3 more years of follow-up
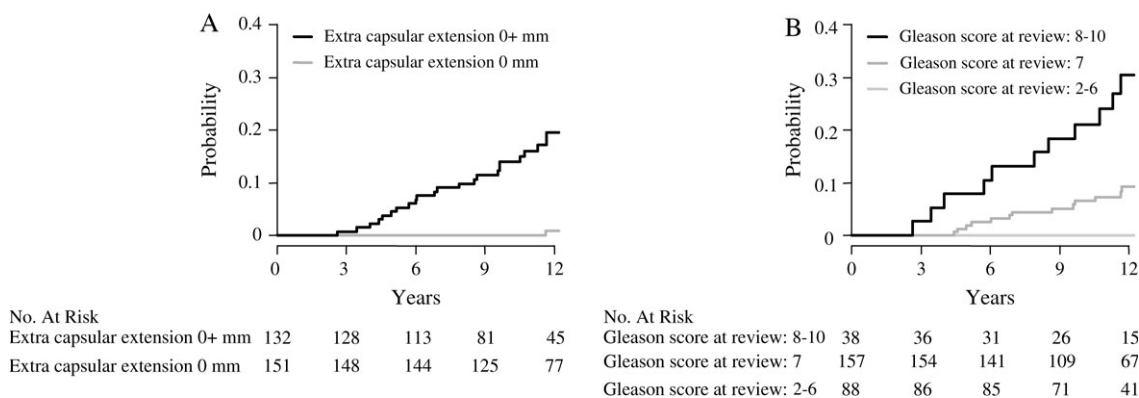
Figure 4. Cumulative incidence of prostate cancer mortality among men who were randomly assigned to radical prostatectomy and underwent radical prostatectomy within 1 year after randomization. **A**) Mortality by extracapsular tumor growth. Relative risk (RR) = 0.05, 95% confidence interval (CI) = 0.01 to 0.22; $P < .001$. **B**) Mortality by Gleason score of radical prostatectomy specimens. Gleason score 7 vs Gleason score 8–10, RR = 0.24, 95% CI = 0.11 to 0.49; $P < .001$. Gleason score 2–6, RR was not estimated (no events). *P* values (two-sided) were calculated using Gray's test.

from 0.50 (5), to 0.56 at a median follow-up of 8 years (4), to 0.65 in the present analysis. The relative reduction in all-cause mortality following radical prostatectomy also decreased over time and was no longer statistically significant after 12 years of follow-up. In the present analysis, we could add data from the radical prostatectomy specimens following a complete central review, and a key finding is that almost all men in the radical prostatectomy group who died from prostate cancer had tumor growth outside the prostate capsule.

Strengths of our study include its randomized design, complete follow-up, standardized histopathologic review, and use of clinically relevant outcomes, which were assessed blindly with regard to randomization group. Compliance with recommended treatment was high despite the drastic difference between the two interventions. Furthermore, a large proportion of patients in this trial belonged to a group with clinically significant disease who might benefit from radical local treatment.

A critical question is whether our results are generalizable to settings in which the majority of prostate cancers are detected by means of PSA testing among asymptomatic men, a situation that prevails in the United States and in a growing number of Western countries. Radical local treatment can by definition convey no survival benefit to those who are overdiagnosed because such men have a nonlethal disease. It is unlikely that the high risk of overdiagnosis and the long lead times estimated in screening (3) will be substantially offset by the proportion of men who are cured due to early treatment. In settings with a large proportion of PSA-detected tumors, the relative reduction in risk of death following radical prostatectomy might be somewhat larger or similar to that in our study, but the absolute reduction would be smaller. An absolute risk reduction of 5.4% implies a number needed to treat of 19 at 12 years in our patient population. However, estimates built on the SPCG-4 study and the current clinical situation (14) suggest that the number needed to treat might be up to five times higher. This perspective further underlines the importance of trials testing active surveillance in patient groups for whom the estimated cost utility of radical prostatectomy is low, or even negative.

As expected from the literature, Gleason score and PSA level at diagnosis were prognostic factors, as also shown in our previous analyses (15). However, we found that men with high Gleason score (≥7) or PSA level (≥10 ng/mL) had similar relative benefits from radical prostatectomy as men with low Gleason score or PSA less than 10 ng/mL, although the absolute benefits will differ by prognostic subgroup (15). These findings regarding outcome after radical prostatectomy by subgrouping the men by Gleason score and PSA level at diagnosis is potentially important for clinical decision making. However, our study was not powered to analyze the effect modification of different prognostic variables, and we could have missed moderate but clinically relevant interactions. The absence of strong interactions indicates, however, that our estimate of the effect of radical prostatectomy is generalizable to other settings without screening but with a different distribution of these prognostic factors. Further evidence regarding the generalizability of our findings might come from the Prostate Cancer Intervention Versus Observation Trial in the United States (16) and the Prostate Testing for Cancer and Treatment study (ProtecT) in the United Kingdom (17). Until the results of those studies are available, our study provides the only evidence from a randomized trial for the benefit of radical prostatectomy.

Local recurrence and/or progression was much lower in the radical prostatectomy group than in the watchful waiting group, as was the use of hormonal and other palliative treatments. Because local recurrence or progression could not be assessed uniformly in the study arms and because indications for palliative treatment could not be followed in detail, the quantitative estimates of the effect of radical prostatectomy on these endpoints are difficult to interpret. However, the more frequent use of hormonal therapy and palliative treatments in the watchful waiting arm reflects symptom burden and has consequences for the patients' quality of life. Androgen deprivation has been shown in two studies (18,19) to decrease general health and quality of life, with more fatigue, loss of energy, emotional distress, change in body image, and worries about cancer and dying. The quality of life of patients with hormone-refractory prostate cancer has also shown to be reduced, likely from both the palliative care and the knowledge of having an incurable disease (20). The higher cumulative incidence of all palliative interventions in the watchful waiting arm corresponds well with our previous finding (21) that the symptom burden in the watchful waiting group seems to outweigh the side effects of surgery, as measured by self-estimated global well-being.

The use of hormonal therapy and other palliative measures followed the pattern of risk of progression and metastatic disease regarding both quantitative levels of use and timing. Thus, these factors cannot explain the better prognosis in the group randomized to surgery. If anything, our data imply that the difference in time to progression may be underestimated because the occurrence of metastatic disease was in absolute terms 6.7% more common even though use of hormonal therapies was 19.3% more common in the watchful waiting arm at 12 years.

Evidence of extracapsular tumor growth in radical prostatectomy specimens was a strong predictor of metastases and prostate cancer death. Indeed, among patients with extracapsular growth, 26 of 132 (20%) died during follow-up. Thus, these men should be considered for postoperative radiotherapy (22,23). However, because most men with extracapsular extension did not die during this follow-up, a substantial proportion of them may have benefited from surgery. Positive surgical margins did not carry prognostic information independent of extracapsular growth in multivariable analyses. This finding is in contrast with those of other studies (23) that used PSA recurrence, rather than death, as an endpoint. In our study, no men who underwent radical prostatectomy and had specimen Gleason scores of 2–6 died from prostate cancer.

The subgroup analyses by age showed that the benefit of radical prostatectomy was limited to younger men. However, we caution that the study was not designed to look at age groups separately and, as for any of our subgroup analyses, all results should be interpreted cautiously. The age 65 cutoff was chosen because this was the mean age of the participants. However, we have no underlying empirical data to judge whether this cutpoint is biologically relevant. Other studies are needed to validate both the finding itself and whether the cutpoint is clinically relevant. In our data, the age-related difference in benefit could not be attributed to selection bias or confounding by any of the measured prognostic factors, and

age per se is not an important prognostic factor in the source population for this study (25). One possible interpretation of our findings is that in elderly patients all tumors with a potentially lethal phenotype have already metastasized at the time of diagnosis and cannot be cured by local treatment alone.

This is the third time an analysis of the main outcomes of this trial has been presented, and multiple statistical tests have been done with the risk of finding one or more randomly statistically significant results. However, our interpretation does rely mainly on stable long-term quantitative estimates as manifested in the cumulative incidence curves. The results of the subgroup analyses should be viewed as hypothesis generating rather than informing clinical action.

We conclude that radical prostatectomy results in a reduction in distant metastases and disease-specific death among patients with clinically localized prostate cancer not detected by PSA screening. Longer follow-up is needed to document whether this reduction is mainly restricted to the first decade of follow-up. The benefits of surgery radical prostatectomy have to be weighed against side effects, but in our study this was not a straightforward balance between years of life gained and side effects (21). Watchful waiting also has side effects, and men face two different scenarios of symptoms and distress. We caution that this balance may look quite different if radical prostatectomy were to be compared with active surveillance. The finding that almost all prostate cancer deaths following radical prostatectomy occurred among patients with extracapsular tumor growth appears clinically relevant when searching for those who can be cured only with the addition of adjuvant treatment. Finally, the finding that radical prostatectomy conveys benefit chiefly to younger patients needs to be applied cautiously in the management of clinically localized prostate cancer. Although the finding was statistically robust, it was from a subgroup analysis, and the possible biologic underpinning of this observation is still enigmatic.

References

1. Kvåle R, Auvinen A, Adami HO, et al. Interpreting trends in prostate cancer incidence and mortality in the five Nordic countries. *J Natl Cancer Inst.* 2007;99(24):1881–1887.
2. Adami HO, Baron JA, Rothman KJ. Ethics of a prostate cancer screening trial. *Lancet.* 1994;343(8903):958–960.
3. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst.* 2003; 95(12):868–878.
4. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med.* 2005;352(19): 1977–1984.
5. Holmberg L, Bill-Axelson A, Helgesen F, et al. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. *N Engl J Med.* 2002;347(11):781–799.
6. Harmer MH, ed. TNM Classification of Malignant Tumours. 3rd ed. Geneva: International Union Against Cancer; 1978.
7. Sobin LH, Hermanek P, eds. UICC TNM Classification of Malignant Tumours. 4th ed. Berlin, Germany: Springer-Verlag; 1987.
8. Zelen M. A new design for randomized clinical trials. *N Engl J Med.* 1979; 300(22):1242–1245.
9. Brendler CB, Cleeve LK, Anderson EE, et al. Staging pelvic lymphadenectomy for carcinoma of the prostate risk versus benefit. *J Urol.* 1980; 124(6):849–850.
10. Walsh PC, Lepor H. The role of radical prostatectomy in the management of prostatic cancer. *Cancer.* 1987;60(suppl 3):526–537.
11. Gleason DF. Histologic Grading and Clinical Staging of Prostatic Carcinoma. Philadelphia, PA: Lea & Febiger; 1977.
12. Gray R. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988;16(3):1141–1154.
13. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. 2nd ed. Hoboken, NJ: Wiley-Interscience; 2002.
14. Klotz LH, Nam RK. Active surveillance with selective delayed intervention for favorable risk prostate cancer: clinical experience and a number needed to treat analysis. *Can J Urol.* 2006;13(suppl 1):48–55.
15. Holmberg L, Bill-Axelson A, Garmo H, et al. Prognostic markers under watchful waiting and radical prostatectomy. *Hematol Oncol Clin North Am.* 2006;20(4):845–855.
16. Wilt TJ, Brawer MK. The prostate cancer intervention versus observation trial: a randomized trial comparing radical prostatectomy versus expectant management for the treatment of clinically localized prostate cancer. *J Urol.* 1994;152(5):1910–1914.
17. Donovan J, Hamdy F, Neal D, et al. Prostate testing for cancer and treatment (ProtecT) feasibility study. *Health Technol Assess.* 2003;7(14):1–88.
18. Herr HW, O'Sullivan M. Quality of life of asymptomatic men with non-metastatic prostate cancer on androgen deprivation therapy. *J Urol.* 2000; 163(6):1743–1746.
19. Fowler FJ, McNaughton M, Collins E, et al. The impact of androgen deprivation on quality of life after radical prostatectomy for prostate carcinoma. *Cancer.* 2002;95(2):287–295.
20. Sullivan PW, Mulani PM, Fishman M, et al. Quality of life findings from a multicenter, multinational, observational study of patients with metastatic hormone-refractory prostate cancer. *Qual Life Res.* 2007;16(4): 571–575.
21. Steineck G, Helgesen F, Adolfsson J, et al. Quality of life after radical prostatectomy or watchful waiting. *N Engl J Med.* 2002;347(11):790–796.
22. Bolla M, van Poppel H, Collette L, et al. Postoperative radiotherapy after radical prostatectomy: a randomised controlled trial (EORTC trial 22911). *Lancet.* 2005;366(9485):572–578.
23. Stephenson AJ, Scardino PT, Kattan MW, et al. Predicting the outcome of salvage radiation therapy for recurrent prostate cancer after radical prostatectomy. *J Clin Oncol.* 2007;25(15):2035–2041.
24. Chang SS, Cookson MS. Impact of positive surgical margins after radical prostatectomy. *Urology.* 2006;68(2):249–252.
25. Helgesen F, Holmberg L, Johansson J-E, et al. Trends in prostate cancer survival in Sweden 1960 through 1988: evidence of increasing diagnosis of non-lethal tumors. *J Natl Cancer Inst.* 1996;88(17):1216–1221.

Funding

Notes

Lindgren, MD, Department of Pathology, University Hospital, Uppsala; Eberhard Varenhorst, MD, PhD, Department of Urology, University Hospital, Linköping; and Former Principal Investigator Bo Johan Norlén, MD, PhD.