Results From the Scandinavian Prostate Cancer Group Trial Number 4: A Randomized Controlled Trial of Radical Prostatectomy Versus Watchful Waiting

Lars Holmberg, Anna Bill-Axelson, Gunnar Steineck, Hans Garmo, Juni Palmgren, Eva Johansson, Hans-Olov Adami, Jan-Erik Johansson

Correspondence to: Lars Holmberg, MD, PhD, Division of Cancer Studies, Guy’s Hospital, King’s College London, St Thomas Street, London, SE1 9RT, UK (e-mail: lars.holmberg@kcl.ac.uk).

In the Scandinavian Prostate Cancer Group Trial Number 4 (SPCG-4), 347 men were randomly assigned to radical prostatectomy and 348 to watchful waiting. In the most recent analysis (median follow-up time = 12.8 years), the cumulative mortality curves had been stable over the follow-up. At 15 years, the absolute risk reduction of dying from prostate cancer was 6.1% following randomization to radical prostatectomy, compared with watchful waiting. Hence, 17 need to be randomized to operation to avert one death. Data on self-reported symptoms, stress from symptoms, and quality of life were collected at 4 and 12.2 years of median follow-up. These questionnaire studies show an intricate pattern of symptoms evolving after surgery, hormonal treatments, signs of tumor progression, and also from natural aging. This article discusses some of the main findings of the SPCG-4 study.


The Scandinavian Prostate Cancer Group Trial Number 4 (SPCG-4) started in 1989 when radical prostatectomy was newly introduced in Scandinavia and when there was essentially no prostate-specific antigen (PSA) testing in asymptomatic men; such testing only became common at the end of the inclusion of the trial a decade later. However, the trial data continue to be important for several reasons. In many parts of the world, the clinical panorama of prostate cancer still resembles that in Sweden in the early 1990s. The trial results point to many of the issues that modern diagnosis and treatment have to solve. SPCG-4 is to date the only trial to inform about both forces of mortality and self-reported symptoms and quality of life in men after radical prostatectomy or watchful waiting two decades and more out after a primary diagnosis of prostate cancer.

According to the protocol (http://www.roc.se/prostata/SPCG-4.pdf), the main trial data have been updated every 3 years since 2002 (1–6). In this presentation, we highlight some of the main findings with bearing on the topic of this conference and discuss some issues that have been raised when the trial results have been presented.

Patients and Methods

The study design has been described in detail earlier (1–6). Between October 1989 and December 1999, we randomized 695 men with a newly diagnosed, localized prostate cancer. Men were eligible for inclusion if they were younger than age 75 years, had an estimated life expectancy of more than 10 years, had no other cancers, and had a tumor stage T0d (later named T1b), T1 or T2, as assessed under the Union for International Cancer Control 1978 criteria. After 1994, men with T1c tumors were also included. The tumor had to be well to moderately-well differentiated according to the World Health Organization classification, and all men were required to have a serum PSA level of less than 50 ng/ml and a negative bone scan.

Among men assigned to radical prostatectomy, the procedure started with a lymphadenectomy of the obturator fossa. If no nodal metastases were found in frozen sections, a radical prostatectomy was performed. Tumor radicality was given priority over nerve preservation. Men assigned to watchful waiting received no immediate treatment but were re-assessed to receive treatment only for symptoms and signs of progressive disease. Thus, they were not under an active surveillance protocol with repeated biopsies or other means of determining a break point where curative treatment could be offered.

In 1999, four uropathologists who were blinded to the patients’ treatment allocation and outcome reviewed all core biopsy specimens and graded them with the use of Gleason scores. In 2006, all radical prostatectomy specimens were also reviewed, graded according to Gleason scores, evaluated for extracapsular tumor growth, and examined for surgical margins. Seminal vesical involvement was assessed as extracapsular tumor growth.

During follow-up, all medical records have been reviewed for information on events associated with prostate cancer, as well as on side effects of surgery, reported by the attendant urologists. An endpoint committee blinded to the treatment assignment determined the cause of death on the basis of information extracted from the patients’ medical records using a protocol that recorded disease progression in a standardized way. For the last follow-up, as presented in 2011 (5), all participants were reviewed through December 2009, and no single patient was lost to follow-up. The study was approved by the regional ethics committees for each participating center.
Follow-up and statistical analysis have been updated every third year according to the study protocol. We analyzed the main endpoints in terms of differences in cumulative incidence and relative risks using Cox proportional hazards models. The null hypothesis for equality between the cumulative incidence curves was tested with Gray test (a modification of the log-rank test to test cumulative incidence estimates). We also assessed the possible modification of treatment effects by age, PSA level at time of diagnosis, tumor stage, and Gleason score by including an interaction term in the Cox proportional hazards model between subgroup category and randomization group. If there was any evidence of effect modification, a further control for age, PSA level, tumor stage, Gleason score, and year of enrollment was introduced.

To evaluate how well PSA levels can predict development of metastatic prostate cancer under watchful waiting, we assessed serial measurements of PSA in 267 men in the watchful waiting arm of the study during the first 2 years of follow-up after randomization (7). We fitted individual regression lines to the PSA levels for each individual and evaluated PSA curve characteristics as determinants of diagnosis of metastatic prostate cancer.

During 1997 and 1998, Swedish participants in the study were asked to report symptoms and self-assessment of quality of life. Those who consented were sent a questionnaire where they were asked about quality, frequency, and intensity of symptoms and, for selected symptoms, about the corresponding distress. Likewise, psychological symptoms, well-being, and subjective quality of life were assessed.

In a health economy analysis (8), the long-term total cost per patient was estimated following radical prostatectomy and watchful waiting, including costs of all in-patient and out-patient hospital care. The study included data from 212 participants living in two counties in central Sweden, and all costs from randomization date until death or end of follow-up in July 2007 were included.

### Results and Comments

A total of 347 men were assigned to the radical prostatectomy group and 348 to the watchful waiting group. Reflecting the prescreening era, only 12% of the men had a nonpalpable T1c tumor at the time of enrollment. By the end of 2009, 284 men in the radical prostatectomy group had undergone a radical prostatectomy, and 302 men in the watchful waiting group had not undergone curative treatment. In the most recent analysis of the study (5), the median follow-up time was 12.8 years. Table 1 summarizes the main findings hitherto in the study.

### Mortality

The absolute risk reductions in overall and disease-specific mortality are of the same magnitude, indicating that the number needed to treat in order to avert one prostate cancer death is 17 at a 15-year horizon (Table 1). The unfolding cumulative mortality curves show a consistent pattern of a difference between the two groups over the follow-up (1,3–5) excluding that the findings in any of the reports represent just a random high.

### Distant Metastases

The absolute risk reduction in the cumulative incidence of distant metastases following radical prostatectomy compared with watchful waiting has remained substantially larger than the

---

**Table 1. Main endpoints in the Scandinavian Prostate Cancer Group Trial Number 4 (SPCG-4) trial analyzed according to intention to treat**

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Result by intention to treat</th>
<th>Comment and Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall mortality</td>
<td>RP 46.1%, WW 52.7%, ARR 6.6%, Gray’s test</td>
<td>Cumulative incidence at 15 years (5)</td>
</tr>
<tr>
<td>Disease-specific mortality</td>
<td>RP 14.6%, WW 20.7%, ARR 6.1%, Gray’s test</td>
<td>Cumulative incidence at 15 years (5)</td>
</tr>
<tr>
<td>Distant metastases</td>
<td>RP 21.7%, WW 33.4%, ARR 11.7%, Gray’s test</td>
<td>Cumulative incidence at 15 years (5)</td>
</tr>
<tr>
<td>Disease-specific mortality in men &lt;65 y</td>
<td>RP 16.4%, WW 25.8%, ARR 9.4%, Gray’s test</td>
<td>Cumulative incidence at 15 years (5)</td>
</tr>
<tr>
<td>Disease-specific mortality after RP by capsular extension</td>
<td>No capsular extension 3.2%, capsular extension 20.5%, ARR 173.3%, Gray’s test P &lt; .001</td>
<td>Cumulative incidence at 15 years [analyzed on data in (5) follow-up]</td>
</tr>
<tr>
<td>Prevalence of erectile dysfunction</td>
<td>RP 80%, WW 45%, RR for RP vs WW 1.8 (1.5–2.2)</td>
<td>Median follow-up at 4 years; range 1–73 years (2)</td>
</tr>
<tr>
<td>Prevalence of distress from compromised sexuality</td>
<td>RP 56%, WW 40%, RR for RP vs WW 1.4 (1.1–1.8)</td>
<td>Median follow-up at 4 years; range 1–73 years (2)</td>
</tr>
<tr>
<td>Prevalence of urinary leakage</td>
<td>RP 49%, WW 21%, RR for RP vs WW 2.3 (1.6–3.2)</td>
<td>Median follow-up at 4 years; range 1–73 years (2)</td>
</tr>
<tr>
<td>Prevalence of distress from urinary leakage</td>
<td>RP 29%, WW 9%, RR for RP vs WW 3.0 (1.8–5.2)</td>
<td>Median follow-up at 4 years; range 1–73 years (2)</td>
</tr>
<tr>
<td>Prevalence of moderate or high anxiety</td>
<td>RP 43%, WW 31%, RR for RP vs WW 0.7 (0.5–1.1)</td>
<td>Median follow-up at 4 years; range 1–73 years (2)</td>
</tr>
<tr>
<td>Low or moderate subjective quality of life</td>
<td>RP 40%, WW 45%, RR for RP vs WW 0.9 (0.7–1.2)</td>
<td>Median follow-up at 4 years; range 1–73 years (2)</td>
</tr>
</tbody>
</table>

* ARR = absolute risk reduction following randomization to RP; RP = radical prostatectomy; RR = relative risk; WW = watchful waiting.
difference in mortality throughout the follow-up period (Table 1) (1,3–5). To date, this difference has not been carried through as a clearly increasing difference in mortality with longer follow-up. Thus, it seems as though some men with distant metastases respond well to therapy and live long enough to die from another cause. For instance, among older men, this group has become larger (3) and may indicate that despite an effect of radical prostatectomy on disease progression, we do not see a clear mortality reduction following allocation to surgery in men over 65.

Effect Modification and Subgroups
We have not found evidence that tumor stage, Gleason grade, or PSA level at diagnosis modifies the effect of radical prostatectomy in terms of differences in relative risk reduction (3–5,9). However, the absolute risk reduction differs between strata because the baseline risk of dying varies. Indeed, competing risk regression models show that at age 65, the absolute 10-year risk reduction in prostate cancer mortality following randomization to radical prostatectomy compared with watchful waiting ranges from 4.5% to 17.2% for low- versus high-risk patients at 10 years of follow-up (10). Thus, the estimate of benefit in the SPCG-4 study is an average of different levels of benefit, depending on both risk category of the prostate cancer and age.

We found evidence of a modification of the treatment effect by age at randomization (3–5,9). Table 1 shows the estimates for men younger than age 65, whereas the differences were small among men aged 65 years or over. Before the first subgroup analysis was undertaken, we had hypothesized that 65 years might be an age where the benefits of surgery would start to become less evident. Although there is evidence for this to be true, we have repeatedly cautioned that this is an exploratory subgroup analysis. Hence, our results should not guide clinical practice without careful considerations. First, it is biologically not plausible that there would be a very sudden change in treatment benefit at a specific age in all men. Second, more careful analysis suggests an effect modification by gradual age with little or no benefit among men who approach age group 70 rather than 65 (10).

In the radical prostatectomy group, the presence of capsular extension is a strong predictor for prostate cancer death (Table 1) and more so than positive margins (4). Extracapsular extension may be a relevant biomarker of a tumor that is difficult to remove radically and/or of already established micrometastatic disease; men with such features could possibly benefit from postoperative adjuvant therapy.

Early PSA Curves as Determinants of Lethal Prostate Cancer
During a mean follow-up of 8.5 years, out of the studied 267 men under watchful waiting, 34 died of prostate cancer and 18 developed metastases but were still alive at end of follow-up (7). In a log-linear model, both PSA value at baseline and the rate of PSA change were associated with the event of lethal prostate cancer. For example, PSA doubling time of less than 5 years was at 6 years of follow-up associated with a relative risk of 2.6 (95% confidence interval 1.5 to 4.5) to die from prostate cancer as compared to a PSA doubling time of more than 5 years. However, as measured by a receiver operating characteristic evaluation, the accuracy of the PSA curves during the first 2 years of follow-up in distinguishing lethal from good-prognosis prostate cancer was low regardless of cut-points chosen or of Gleason scores of the primary tumor.

Symptoms and Quality of Life
Table 1 also highlights the main findings concerning self-reported outcomes undertaken from 1997 to 1998 (2). Following a median follow-up of 4 years, the major differences are determined by the primary treatment. The prevalence of erectile dysfunction, urinary leakage, and the distress from these symptoms is higher in the radical prostatectomy group. Lower levels have been presented based on selected series from specialized institutions. However, evidence is emerging that in many population-based settings—even with modern surgical techniques—erectile dysfunction and urinary leakage are approaching our figures in men of a similar age and with similar disease status. Interestingly, the prevalence of anxiety or low or moderate subjective quality of life was at this time similar in the two groups.

In the second round of the questionnaire study at a median follow-up of 12.2 years, it becomes more difficult to isolate the effects of surgery compared with watchful waiting because active treatment of disease progression, also in its early stages, influences symptom development in all ages and over a long-term follow-up (5,11). By the end of the most recent follow-up (5), a total of 139 men in the radical prostatectomy group and 223 men in the watchful waiting group had received hormonal therapy (corresponding to cumulative incidence of hormonal therapy at 15 years of 39.6% and 63.4%, respectively). Furthermore, the different patterns of disease progression in the two study groups and ensuing symptom development make the situation even more complex. The second questionnaire shows that both groups continue to develop symptoms but under different scenarios. In both study groups, men with prostate cancer have considerably higher prevalence of symptoms, distress, and low quality of life than men in the general population (6).

Health Economy
The overall cost in the radical prostatectomy group was 34% higher than in the watchful waiting group during a median follow-up of 12 years. The difference in costs between the two groups varied between €4062 and €9120 (corresponding to US $5361 and US $12,036 in 2007) in different subgroups defined by age and risk category, always more costly for the radical prostatectomy group. The total mean cost per patient was statistically significantly higher in the radical prostatectomy group than in the watchful waiting group with an average difference of €6123 (US $8081 in 2007). The difference was strongly driven by the initial cost of the radical prostatectomy procedure (8).

Conclusions
Assignment to radical prostatectomy rather than watchful waiting was associated with a substantial relative reduction in prostate cancer mortality, indicating that the operation indeed alters the natural history of the disease. The reduction in overall and prostate cancer mortality was modest in absolute terms, indicating that the number needed to treat is 17 to avoid one prostate cancer death at 15 years of follow-up. Of the 695 men randomized, 328 were still alive at the latest follow-up (5). Thus, events will continue to develop, and
SPCG-4 will be a source to analyze forces of mortality from prostate cancer and comorbid conditions over very long-term follow-up and when men grow old. These perspectives will be important for decision making also in current urological practice.

It is likely that the absolute mortality reduction following radical prostatectomy will be lower among men with lower risk tumors than among those in SPCG-4. On the other hand, the overall estimate of effect is an average of different benefits, and the data indicate that the gains are substantial in younger men with high-risk tumors. The lack of a benefit in men older than 65 years, which was a prespecified limit, is in a more advanced analysis depending on tumor characteristics, PSA level, and general health status, rather than approaching 70 years. It is not correct to assume a lack of benefit in men over 65 without considering tumor characteristics, PSA level, and general health status. Side effects occurred over the whole age-spectrum, so the diminishing return by higher age of life years saved by treatment was not accompanied by less side effects; the cost–benefit balance will evidently tend to be negative in any group with little effect from treatment on tumor progression.

Our analyses of self-reported symptoms and distress underline that an understanding of the pattern of side effects and time dimension of the occurrence for each treatment is important for patient counseling that should not be limited to the period around diagnosis. SPCG-4 is to date the only study to present a perspective of self-reported symptoms under radical prostatectomy and watchful waiting two decades out from the initial diagnosis of prostate cancer.

We had hypothesized that years of life could be won by radical prostatectomy at the cost of a lower quality of life, but the equation is more complex. If active surveillance is substituted for watchful waiting, dimensions also of curative interventions starting late after diagnosis and anxiety of active follow-up have to be considered. The long-term follow-up shows that men do not adjust over time but continue to be distressed from erectile dysfunction and other side effects. The results underline that especially in men with low-risk disease, decision making for treatment needs to include that the patient has had full information to be able to weigh anticipated benefits against side effects. And once treatment or surveillance of any kind is undertaken, side effects should be avoided by all means possible. These lessons have become increasingly important in the light of possible overdiagnosis following PSA screening. It is to be noted that our data reflecting side effects of tissue damage caused by surgical intervention apply also to patients today with on average much lower tumor burden and are likely to be applicable to patients treated in low-volume practices. Although a lead time following PSA screening could delay symptoms and bother under watchful waiting or active surveillance compared with the SPCG-4 results, there is no lead time for the tissue damage caused by surgery.

References


Funding

This work was supported by the Swedish Cancer Society and US National Institutes of Health.

Affiliations of authors: Regional Cancer Center Uppsala Örebro, Uppsala, Sweden (LH, HG); Division of Cancer Studies, Medical School, King’s College London, London, UK (LG, HG); Division of Clinical Cancer Epidemiology, Department of Oncology and Pathology (AB-A, GS, EJ) and Department of Medical Epidemiology and Biostatistics (H-OA, JP), Karolinska Institute, Stockholm, Sweden; Department of Surgical Sciences, Uppsala University, Uppsala, Sweden (LB-A, EJ); Division of Clinical Cancer Epidemiology, Sahlgrenska Academy, Gothenburg, Sweden (GS); Department of Epidemiology, Harvard School of Public Health, Boston, MA (H-OA); Center for Assessment of Medical Technology, Örebro, Sweden (J-EJ).