Treatment of local–regional prostate cancer detected by PSA screening: benefits and harms according to prognostic factors

E M Wever*[,1], E A M Heijnsdijk[1], G Draisma[1], C H Bangma[2], M J Roobol[2], F H Schröder[2] and H J de Koning[1]

[1]Department of Public Health, Erasmus Medical Center, P.O. Box 2040, Rotterdam, 3000, CA, The Netherlands and [2]Department of Urology, Erasmus Medical Center, P.O. Box 2040, Rotterdam, 3000, CA, The Netherlands

Background: Men with screen-detected prostate cancer can choose to undergo immediate curative treatment or enter into an expectant management programme. We quantified how the benefits and harms of immediate treatment vary according to the prognostic factors of clinical T-stage, Gleason score, and patient age.

Methods: A microsimulation model based on European Randomized Study of Screening for Prostate Cancer data was used to predict the benefits and harms of immediate treatment versus delayed treatment of local–regional prostate cancer in men aged 55–74 years. Benefits included life-years gained and reduced probability of death from prostate cancer. Harms included lead time and probability of overdiagnosis.

Results: The ratio of mean lead time to mean life-years gained ranged from 1.8 to 31.2, and the additional number of treatments required per prostate cancer death prevented ranged from 0.3 to 11.6 across the different prognostic groups. Both harm–benefit ratios were lowest, most favourable, for men aged 55–59 years and diagnosed with moderate-risk prostate cancer. Ratios were high for men aged 70–74 years regardless of clinical T-stage and Gleason score.

Conclusion: Men aged 55–59 years with moderate-risk prostate cancer are predicted to derive greatest benefit from immediate curative treatment. Immediate treatment is least favourable for men aged 70–74 years with either low-risk or high-risk prostate cancer.

The European Randomized Study of Screening for Prostate Cancer (ERSPC) and the Gothenburg trial (part of the ERSPC) have shown that a reduction in prostate cancer mortality can be achieved using prostate-specific antigen (PSA) screening (Schroder et al, 2009; Hugosson et al, 2010; Schroder et al, 2012). It has been shown that after adjusting for quality of life there is still a substantial benefit (Heijnsdijk et al, 2012).

Men with screen-detected prostate cancer can choose between receiving immediate active treatment or entering into an expectant management programme. The potential benefit of actively treating the cancer immediately after diagnosis is an increase in life expectancy. The potential harm is the risk of living for many years with the side effects of treatment, years which might otherwise have been symptom free (Korfage et al, 2005). Alternatively, the potential benefit can be expressed as the reduction of prostate cancer-specific mortality, and the potential harm expressed as the percentage of overdiagnosis, that is, the proportion of men with screen-detected prostate cancer that, in the absence of screening, would die from other causes before the time of clinical diagnosis.

The aim of the present study was to quantify the potential benefits and harms of immediate versus delayed active treatment for local–regional prostate cancer according to the following factors at time of detection: clinical T-stage, Gleason score, and patient age. The results presented in this study are intended to help clinicians and patients decide whether or not treatment is favourable immediately after early detection.

MATERIALS AND METHODS

ERSPC trial: Rotterdam and Gothenburg sections. The ERSPC trial was initiated in the early 1990s in order to evaluate the effect of PSA screening on prostate cancer mortality. In the Rotterdam and Gothenburg sections, 42 376 men aged 55–74 years and 19 946 men aged 50–64 years were randomised, respectively. The time interval between the screening rounds was 4 years in Rotterdam and 2 years in Gothenburg. An overall reduction in prostate cancer mortality of 29% at a median follow-up of 11 years was observed in the ERSPC (Schroder *et al*, 2012).

MISCAN model. We used the MIcrosimulation SCreening Analysis (MISCAN) prostate cancer model (Draisma *et al*, 2003, 2006, 2009; Wever *et al*, 2010a). MIcrosimulation SCreening ANalysis is a microsimulation programme that simulates progression and screening of prostate cancer within a population. The model was validated using prostate cancer detection data from the Rotterdam (Draisma *et al*, 2003, 2006; Wever *et al*, 2010a) and Gothenburg (Hugosson *et al*, 2004) sections of the ERSPC, as well as the mortality reduction data from the overall ERSPC trial (Schroder *et al*, 2012). A summary of the assumptions in the model and the data used for calibration are presented in Table 1 and are outlined below. A more detailed description of the model can be found in earlier publications (Draisma *et al*, 2003, 2006) and in a standardised model profile (Wever *et al*, 2010b).

MIcrosimulation SCreening ANalysis is a microsimulation programme, which simulates the progression of prostate cancer in individuals as a sequence of preclinical, clinical, and screen-detected tumour states. First, the age at death from other causes is simulated per individual using Dutch life tables (Statistics Netherlands, 2000–2007). Next, the progression of prostate cancer in the absence of screening is simulated. Prostate cancer may develop from no prostate cancer to a clinically diagnosed cancer through one or more screen-detectable preclinical stages. From each preclinical stage, a tumour may grow to the next clinical T-stage (T1, impalpable; T2, palpable, confined to the prostate; T3 +, palpable, with extensions beyond the prostatic capsule); it may dedifferentiate to a higher Gleason score (well differentiated, Gleason score 2–6; moderately differentiated, Gleason score 7;

poorly differentiated, Gleason score 8–10); or it may be clinically diagnosed. For these transitions, the time spent in the current stage is generated from a Weibull distribution, where the parameters depend on the current stage and the choice of next stage is determined by transition probabilities. In addition, there is a risk that a tumour in the local–regional stage (M0) will develop into distant disease (M1), which is modelled by using a stage and Gleason score-specific hazard function. Depending on the frequency and sensitivity of the screening test, preclinical cancers may be detected by screening. PSA test and subsequent biopsy were modelled as a single test, where the sensitivity parameter was assumed to be clinical T-stage-dependent. In the model, sensitivity is defined as the probability that a preclinical tumour is detected by a screening test at the time the test is taken.

Model parameters, including transition probabilities, mean dwelling times (the time from one preclinical state to another preclinical or clinical state), and stage-specific test sensitivities were estimated by constructing models for the ERSPC-Rotterdam and -Gothenburg, and by calibrating the model to the following data observed at these centres: baseline incidence (national cancer registry data for 1991 (Visser *et al*, 1994)) and stage distribution in the Netherlands (Rotterdam cancer registry data 1992–1993 (Spapen *et al*, 2000)); baseline incidence in Sweden (1988–1992 (Parkin *et al*, 1997)); incidence, Gleason and stage distributions in the control arms of ERSPC-Rotterdam and -Gothenburg; and detection rates, interval cancer rates, Gleason and stage distributions in the screen arms of ERSPC-Rotterdam (Draisma *et al*, 2003, 2006; Wever *et al*, 2010a) and Gothenburg (Hugosson *et al*, 2004). Number of cases diagnosed, and Gleason and stage distributions in the control arms versus those in the screen arms provide insight into disease progression through the various preclinical phases. Parameters were estimated by numerically minimising the deviance between the number of cases observed and the number of cases predicted by the models. Deviances were calculated by assuming Poisson likelihood for incidence data and by assuming multinomial likelihood for stage-distribution data.

Survival after clinical diagnosis of untreated prostate cancer was assumed to be according to the Gleason score-specific survival curves of (Albertsen *et al*, 2005), but we added clinical T-stage as an explanatory factor. For this, we used the Cox proportional hazard estimates from (Aus *et al*, 2005; T1 1, T2 1.51, T3 2.77), and

Table 1. Modelling assumptions and data used in the present study

Variable	Assumption	Calibration data
Other cause of death	Life tables	Statistics Netherlands, 2000–2007
Disease progression before diagnosis	Semi-Markov model	National Cancer Registry data, 1991 (Visser *et al*, 1994) Rotterdam Cancer Registry data, 1992–1993 (Spapen *et al*, 2000) Incidence in Sweden, 1988–1992 (Parkin *et al*, 1997) ERSPC Rotterdam trial (Draisma *et al*, 2003; Wever *et al*, 2010a) ERSPC Gothenburg trial (Hugosson *et al*, 2004)
Baseline prostate cancer-specific survival dependent on Gleason score	Poisson regression model	Connecticut Tumor Registry data (Albertsen *et al*, 2005)
Baseline prostate cancer-specific survival dependent on clinical T-stage	Relative risk compared with baseline survival dependent on Gleason score	Population-based National Prostate Cancer Registry of Sweden (Aus *et al*, 2005)
Treatment effect	Relative risk compared with baseline survival	SPCG-4 study (Bill-Axelson *et al*, 2011)
Screening effect	A proportion of men with screen-detected prostate cancer are cured when treated early	ERSPC trial (Schroder *et al*, 2012)

Abbreviations: ERSPC = European Randomized Study of Screening for Prostate Cancer; SPCG = Scandinavian Prostate Cancer Group.

changed the Albertsen model such that the weighted sum of hazards by clinical T-stage add up to the same overall level by age and Gleason score.

In the model, we assumed that all men diagnosed with prostate cancer receive radical prostatectomy. According to published results, (Bill-Axelson *et al*, 2011), we assumed that men receiving radical prostatectomy have a relative risk of 0.62 of dying from prostate cancer compared with men receiving no initial treatment. This analysis did not consider other treatments, such as radiation therapy and active surveillance, because of the limited published results about the effectiveness of these treatments. If the effectiveness of radiation therapy and active surveillance are different than that of radical prostatectomy, the results for these treatments will be different than those presented. For distant prostate cancer, we assumed that treatment has no effect. Some treatments might increase survival of distant disease slightly. However, in our view this is a minor limitation of the model.

The effect of early detection through screening on survival was included by assuming that a fraction of local–regional tumours detected by screening are cured because the tumours are treated earlier. The Gleason score-dependent cure rates were estimated by calibrating the ERSPC-Rotterdam model to the observed 29% prostate cancer mortality reduction in the overall ERSPC at a median follow-up of 11 years (Schroder *et al*, 2012).

Analysis. For the present study, we constructed a model in which individuals were initially screened between the ages of 50 and 74 years, and then subsequently every 4 years until the age of 75 years. Two situations were simulated: one in which all men with screen-detected prostate cancer received treatment immediately after diagnosis; and another in which they received delayed treatment, that is, they received treatment at the time they would have if clinically diagnosed in the absence of screening. Comparing these two simulated populations, life-years gained, reduced probability of death from prostate cancer, lead time and probability of over-diagnosis were calculated and stratified according to the prognostic factors clinical T-stage, Gleason score, and patient age. Note that lead time is the period by which diagnosis is advanced due to screening; therefore, mean lead time shows the average potential life-years with no side effects of treatment in case treatment was delayed to the time of clinical diagnosis.

Two alternative harm–benefit ratios were subsequently calculated from the results: the ratio between the mean lead time and mean life-years gained (M, which represents the average loss of life-years free from the potential side effects of curative treatment per life year gained), and the ratio between the percentage of overdiagnosis and the percentage of prostate cancer deaths prevented by early treatment (NNT, which represents the additional number of patients who must be treated in order to avoid one prostate cancer death). Considering that the decision to treat immediately or not depends on the variable M, we determined for which combinations of prognostic factors it would be less favourable ($M > 9$), more favourable ($3 \leqslant M < 9$), and most favourable ($M \leqslant 3$) to treat immediately. The ranges for these three groups were chosen arbitrarily.

Sensitivity analyses. To evaluate the statistical variation and uncertainty of the observed data, we conducted a number of sensitivity analyses. The sensitivity analyses compared models with various lead times, survival curves, and cure rates. Penalised optimisation was used to obtain a range of models with various lead times. Parameters for these models were estimated by minimising the sum of total deviance and lead time penalty (mean lead time × penalty). Different survivals were considered by assuming a relative risk of 0.8, 1, or 1.2 on the hazard of prostate cancer death. We also varied assumed mortality reductions due to screening: 21% (observed for the screen group in ERSPC), 29%

(observed for attendees in ERSPC), or 44% (observed for the screen group in ERSPC-Gothenburg).

RESULTS

Values describing the potential benefits and harms of active treatment varied considerably between the different prognostic groups: mean lead time ranged from 2.9 to 12.2 years; mean life-years gained ranged from 0.1 to 4.0 years; the percentage of screen-detected cases that were overdiagnosed ranged from 2.7 to 60.1%; and the percentage of men who avoided death from prostate cancer ranged from 1.5 to 32.1% (Table 2).

To illustrate Table 2, consider the example of men with the following prognostic factors: diagnosed with T1G7 prostate cancer at 62 years of age. These men each have a 28.4% risk of being an overdiagnosed case. The mean lead time for such overdiagnosed men is 11.3 years, that is, if these men decide to be treated at the time of screen detection they will live an average of 11.3 years with the potential side effects of curative treatment. The non-over-diagnosed men have to live an average of 8.5 years with the potential side effects if they decide to be treated immediately. The life expectancy is 16.5 years and there is an 18.7% risk of dying from prostate cancer if the cancer is treated immediately. If the cancer is treated at the expected time of clinical diagnosis then their life expectancy is 15.3 years (1.2 years less) and the risk of dying from prostate cancer is 32.3% (an increase of 13.6% in absolute terms). The negative effect of treating cancer at screen detection instead of time of clinical diagnosis is that the patient has to live on average 9.3 years ($((11.3 \times 28.4) + (8.5 \times 71.6))/100$) longer with the potential side effects of treatment. The positive effect is that the patient's life expectancy is 1.2 years longer. The ratio between the negative effect and the positive effect is 8.0 (the ratio is not exactly 9.3/1.2 because of rounding of decimals). The negative effect can also be expressed as the percentage of overdiagnosed cases (28.4%) and the positive effect as the percentage of prostate cancer death avoided by treating early (13.6%). The ratio between these is the NNT (2.1 in this example). Lower values for these ratios imply lower expected negative effects in relation to the expected positive effects, and therefore immediate treatment is more favoured.

The two harm–benefit ratios (M and NNT) also showed considerable variation between the different prognostic groups: M ranged from 1.8 to 31.2 and NNT ranged from 0.3 to 11.6 (Table 2). If the same age groups were compared, immediate treatment was increasingly favourable with increasing Gleason score for patients at clinical stage T1 or T2. In contrast, favourability decreased with increasing Gleason score for patients in same age group at clinical stage T3. Favourability of immediate treatment increased with increasing T score in patients of the same age and with a Gleason score < 7. In patients of the same age, with a Gleason score of 7, immediate treatment was most favourable in those with clinical stage T2, followed by stage T3, and then stage T1. In patients of the same age with a Gleason score > 7, immediate treatment was also most favourable in patients with clinical stage T2 (although ratios were only slightly lower compared with T1) and least favourable in T3.

Considering that the decision to treat immediately depends on the lead time divided by the life-years gained (M), the relationship between these two variables was plotted in order to better illustrate how the various combinations of prognostic factors influence the favourability of immediate treatment after screen detection (Figure 1). For all age groups, the same pattern can be observed, where immediate treatment was most favourable in men diagnosed with T3G6, and least favourable in men diagnosed with T1G6 or T3G8. The figure illustrates clearly that both the benefits and

Table 2. Predicted percentage of overdiagnosis,[a] mean lead time,[b] life expectancy,[c] and percentage of prostate cancer (PC) death[d] according to prognostic factors of clinical T-stage, Gleason score, and patient age

| Clinical T-stage | Gleason score | Age at diagnosis (years) | Over-diagnosis (%) | Mean lead time (years) | | Treatment at screen-detection | | Treatment at symptoms presentation | | M[e] | NNT[e] |
				Over-diagnosed	Non-over-diagnosed	Life expectancy (years)	PC death (%)	Life expectancy (years)	PC death (%)	Mean lead time/mean life-years gained	Overdiagnosis/ Δ PC death
T1	<7	55–59	30.5	15.7	10.7	21.2	9.2	19.7	26.3	7.8	1.8
		60–64	39.9	13.4	9.5	17.6	6.5	16.7	18.7	12.1	3.3
		65–69	49.8	11.2	8.3	14.2	4.4	13.7	12.7	18.8	6.0
		70–74	60.1	8.9	6.9	10.9	2.8	10.7	8.0	30.4	11.6
	7	55–59	19.0	12.6	9.0	19.3	26.1	17.2	45.4	4.8	1.0
		60–64	28.4	11.3	8.5	16.5	18.7	15.3	32.3	8.0	2.1
		65–69	39.0	9.8	7.7	13.6	12.3	13.0	21.4	13.6	4.3
		70–74	50.1	7.9	6.7	10.7	7.4	10.4	12.9	24.3	9.2
	>7	55–59	5.4	6.3	4.8	16.1	47.3	14.1	62.7	2.5	0.4
		60–64	9.9	6.4	4.8	14.3	37.6	13.1	49.2	4.3	0.9
		65–69	17.4	6.4	4.8	12.4	26.0	11.7	34.3	7.7	2.1
		70–74	26.6	5.5	4.4	10.0	17.6	9.6	23.1	13.1	4.8
T2	<7	55–59	18.7	12.5	9.0	20.2	17.4	17.4	43.5	3.5	0.7
		60–64	27.9	11.2	8.5	17.0	13.0	15.4	32.2	5.7	1.5
		65–69	38.1	9.7	7.6	13.8	9.1	12.9	22.7	9.2	2.8
		70–74	49.3	7.9	6.6	10.8	6.0	10.3	14.9	14.9	5.5
	7	55–59	9.2	9.7	5.8	16.8	42.2	14.6	60.9	2.7	0.5
		60–64	15.0	8.6	5.8	14.9	32.7	13.4	47.8	4.2	1.0
		65–69	24.1	8.0	5.6	12.7	23.2	11.8	33.7	7.5	2.3
		70–74	34.5	6.8	4.9	10.2	15.1	9.8	22.0	13.5	5.0
	>7	55–59	4.0	5.1	3.6	12.5	67.4	11.1	77.4	2.6	0.4
		60–64	6.5	4.2	3.6	11.6	57.0	10.7	65.5	3.8	0.8
		65–69	11.6	4.6	3.9	10.5	44.7	10.0	51.2	7.0	1.8
		70–74	21.1	4.7	3.8	9.0	31.3	8.7	35.9	12.8	4.5
T3	<7	55–59	10.6	9.0	7.0	17.7	34.2	13.7	66.3	1.8	0.3
		60–64	17.7	8.5	6.9	15.3	27.5	12.8	53.4	2.8	0.7
		65–69	26.9	7.9	6.5	12.9	20.4	11.4	39.5	4.7	1.4
		70–74	38.3	6.7	5.8	10.3	13.6	9.5	26.5	8.1	3.0
	7	55–59	6.7	7.1	5.2	13.1	65.0	11.5	76.3	3.4	0.6
		60–64	11.3	6.7	5.2	12.1	55.1	11.1	64.7	5.2	1.2
		65–69	18.8	6.6	5.1	10.8	43.1	10.2	50.5	8.8	2.6
		70–74	28.8	5.8	4.7	9.1	30.3	8.8	35.5	15.0	5.5
	>7	55–59	2.7	4.1	2.9	9.0	84.0	8.5	86.6	6.5	1.1
		60–64	4.6	3.5	2.9	8.7	76.0	8.5	78.2	10.7	2.1
		65–69	9.0	3.7	3.1	8.4	63.8	8.2	65.5	19.0	5.2
		70–74	15.5	3.9	3.2	7.6	48.9	7.5	50.4	31.2	10.5

Results describe men diagnosed at the local–regional stage of prostate cancer.
[a]Overdiagnosis: proportion of men with screen-detected prostate cancer who, in the absence of screening, would die from other causes before the time of clinical diagnosis.
[b]Lead time: interval from time of screen detection to time of clinical diagnosis in the absence of screening (for non-overdiagnosed) or to time of death (for overdiagnosed).
[c]Life expectancy: mean interval from screen-detection to time of death.
[d]Percentage of prostate cancer death: proportion of men with screen-detected prostate cancer who die from prostate cancer.
[e]Ratios presented may differ slightly from values calculated using the estimates presented in the table due to the rounding of decimals within the model.
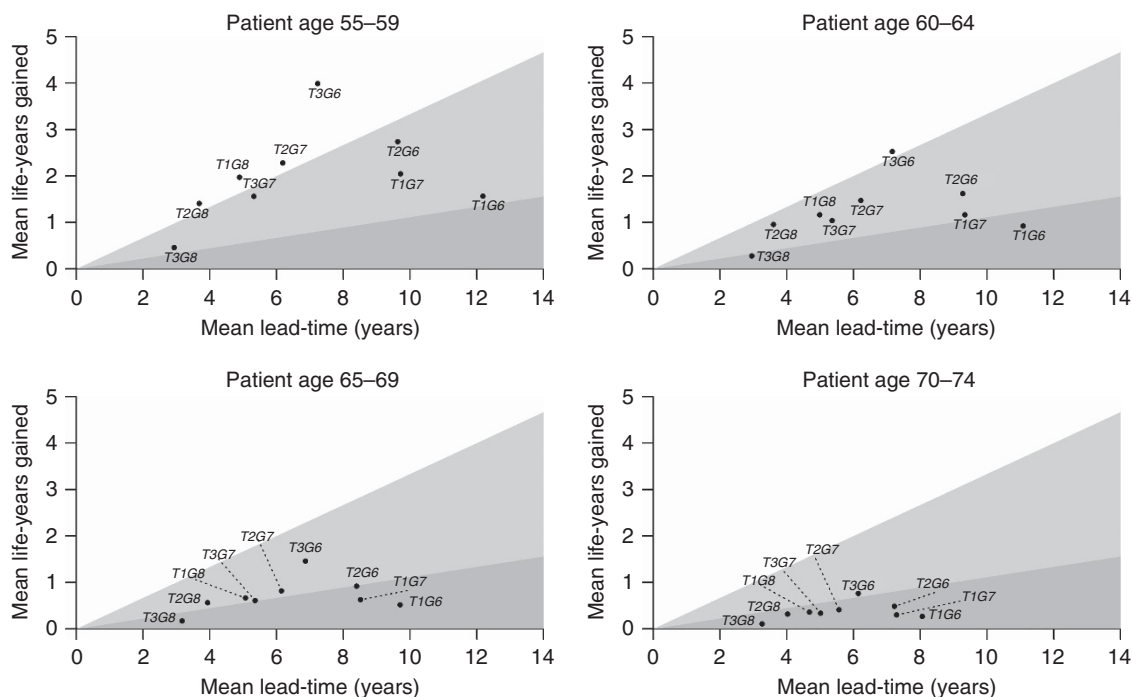
Figure 1. Comparison of mean life-years gained and mean lead time for different combinations of prognostic factors. The dark grey zone represents the area where treatment would be considered less favourable ($M > 9$), the light grey zone represents the area where treatment would be considered more favourable ($3 \leqslant M < 9$), and the white zone represents the area where treatment would be considered most favourable ($M \leqslant 3$). The variable M is the ratio of the mean lead time to the mean life-years gained, which represents the average loss of life-years free from the potential side effects of curative treatment per life-year gained.

harms decrease with increasing age, and that the benefits decrease relatively more than the harms. Therefore, immediate active treatment after screen detection was less favourable with increasing age. It also shows that with increasing age there is less variation in the harms and benefits between the different prognostic groups.

Varying lead time, survival, and mortality reduction caused the minimum of the harm–benefit ratio M to vary from 1.2 to 2.4 and the maximum to vary from 20.3 to 44.0 (Table 3). Considering the different assumptions in the sensitivity analysis, the same pattern was preserved where immediate treatment was most favourable in men diagnosed at age 55–59 with T3G6, and least favourable in men diagnosed at age 70–74 with T1G6 or T3G8.

DISCUSSION

The results from our model demonstrate that, in addition to potential life-years gained and probability of avoiding death from prostate cancer, the decision of whether a patient should receive immediate treatment or not should also depend on the potential lead time and the probability of overdiagnosis. Our model predicts that these factors vary considerably depending on the prognostic factors of clinical T-stage, Gleason score, and patient age.

The two harm–benefit ratios were lowest for men aged 55–59 years, and so immediate treatment would seem most favourable for this group. Both ratios were lowest for men with moderate-risk cancer, specifically those diagnosed with T3G6 (representing 3% of screen-detected cases at 55–59 years of age), T2G7 (7% of screen-detected cases at 55–59 years of age), T1G8 (1% of screen-detected cases at 55–59 years of age), and T2G8 (2% of screen-detected cases at 55–59 years of age). Our model suggests that patients with these prognostic factors would be likely to derive greatest benefit from immediate treatment. The majority of men aged 60–69 years at the time of detection belong to a group for which the harm–

benefit ratios could be considered 'intermediate'. Men aged 70–74 years with low-risk cancer (T1G6) or high-risk cancer (T3G8) had the highest ratios and, therefore, immediate treatment is least favourable for these patients. Indeed, the harm–benefit ratios are relatively high for all combinations of clinical T-stage and Gleason score in patients aged 70–74 years, which is due to the low probability of these patients living long enough to experience the benefit of curative treatment. The high harm–benefit ratios for immediate treatment of men aged 70–74 years also imply that the harm–benefit ratio for the screening of these men will also be high relative to younger men. Therefore, the negative impact of the screening process may be reduced by only screening younger men. Our results suggest that it is crucial to precisely determine which age groups would derive greatest benefit from screening and at what age it would be most favourable to cease.

In our models, the harm–benefit ratio is unfavourable in patients with either favourable prognostic factors or with unfavourable characteristics. In the first group, this is due to the extended lead times, in the second group due to the very modest benefits of immediate treatment. For the first group, patients with T1G6, T2G6, or T1G7 tumours, active surveillance might allow treatment to be delayed without impairing the chances of cure too much. For the second group, patients with T2G8, T3G7, or T3G8 tumours, watchful waiting, that is, waiting for symptoms to appear and starting palliative treatment when necessary, might be the most appropriate treatment.

In our analysis, we assumed that the favourability of immediate treatment for each combination of prognostic factors will depend on the ratio between mean lead time and mean life-years gained (M). It is important to note that the three ranges describing the favourability of the ratio M were chosen arbitrarily and do not imply that all patients in the least favourable group should not receive immediate treatment, and that all patients in the most favourable group should receive immediate treatment. The division between the groups only illustrates for which combination of

Table 3. Sensitivity analysis for uncertainty in the model and the data				
			M Mean lead time Life-years gained	
Mean lead time[a]	Relative risk on PC death[b]	Mortality reduction[c]	Minimum[d]	Maximum[d]
Base model				
8.0	1	29%	1.8 (age 55–59, T3G6)	31.2 (age 70–74, T3G8)
Varying lead time				
8.5	1	29%	2.4 (age 55–59, T3G6)	37.5 (age 70–74, T1G6)
7.1	1	29%	1.4 (age 55–59, T3G6)	25.2 (age 70–74, T3G8)
Varying survival				
8.0	0.8	29%	2.0 (age 55–59, T3G6)	36.5 (age 70–74, T3G8)
8.0	1.2	29%	1.7 (age 55–59, T3G6)	29.0 (age 70–74, T3G8)
Varying mortality reduction				
8.0	1	21%	2.6 (age 55–59, T3G6)	44.0 (age 70–74, T1G6)
8.0	1	44%	1.2 (age 55–59, T3G6)	20.3 (age 70–74, T3G8)

Abbreviation: PC = prostate cancer.
[a] We used penalised optimisation to obtain a range of models with different lead times (Draisma et al, 2003).
[b] A relative risk of 0.8 on the hazard of prostate cancer death increases PC-specific survival and a relative risk of 1.2 decreases PC-specific survival.
[c] We assumed various mortality reductions: 21% (observed for the screen group in European Randomized Study of Screening for Prostate Cancer (ERSPC)), 29% (observed for attendees in ERSPC), or 44% (observed for the screen group in ERSPC-Gothenburg).
[d] Between braces is indicated for which age category, clinical T-stage, and Gleason score the minimum (most favourable) or maximum (least favourable) value belongs to.

prognostic factors it is relatively less favourable to treat immediately and for which prognostic factors it is relatively more favourable to treat immediately.

Our modelling study has some limitations. The predicted values are based on data from a specific population and on the assumptions underlying the model. However, using the best available data and most reasonable assumptions, we obtained the best possible estimates for the measures of interest. The results were calculated using data from the ERSPC Rotterdam and Gothenburg. Results may be different for other populations, as different incidences of and mortalities for prostate cancer have been observed in different countries (Ferlay et al, 2010). However, in the appendix of Heijnsdijk et al (2012), it is shown that the incidence of prostate cancer in the screen arm of all centres of the ERSPC is predicted well by the model. Also, in the United States, the estimated lifetime risk of prostate cancer diagnosis is 16.22% and of prostate cancer death is 2.79% (SEER, 2010), which are close to lifetime risks of this model (Wever et al, 2012). Therefore, the results can also be applied in the US population. Another limitation is that it was not possible to consider all available prognostic and predictive factors that would be available in clinical practice. For example, PSA measurements, MRI results, or the number of positive cores were not considered in our analysis. We

also made no distinction between Gleason scores $(3 + 4)$ and $(4 + 3)$. These could be important prognostic factors. However, this is the first study to quantify the benefits and harms of immediate treatment vs delayed treatment according to the prognostic factors of clinical T-stage, Gleason score, and patient age. Judging from our results, it would seem crucial that we understand the impact of at least these three factors. Another important prognostic factor that was not included in our model is the presence of co-morbidity. We suggest that the results obtained for older men in our study could be considered likely to be similar to individuals with high levels of co-morbidity. A final limitation is that we used survival curves based on non-contemporary data observed in the United States for the ERSPC population that was modelled (Albertsen et al, 2005). These survival data were used because they are one of the few data sets presenting survival of untreated prostate cancer as a function of age at diagnosis and Gleason score progression. The percentage of prostate cancer death that we report may seem higher than previous reports (Albertsen et al, 2005; Aus et al, 2005; Albertsen et al, 2011). However, the percentages of prostate cancer deaths are generally presented after a follow-up time of 10 or 20 years, whereas our results represent lifetime data. For example, Albertsen et al (2011) reported that, after 10 years of follow-up, the percentage of prostate cancer death in men aged 66–75 years and with T1cG8 is 13.7–25.7%, depending on co-morbidity. In the present study, the predicted lifetime percentage of prostate cancer death for men with T1G8 and aged 65–74 is 27.6%; however, the value is 14.0% after 10-year follow-up, which is on the low side compared with the numbers reported by Albertsen et al (2011).

In conclusion, the potential benefits and harms of immediate treatment of local–regional prostate cancer detected by PSA screening depend on the prognostic factors clinical T-stage, Gleason score, and patient age. The range of these values by the different prognostic groups is wide. Therefore, it would seem important to understand and consider these factors when making decisions regarding whether or not to treat a patient. Men aged 55–59 years and with moderate-risk cancer (T3G6, T2G7, T1G8, or T2G8) display the most favourable harm–benefit ratios, and would therefore seem most likely to derive greatest benefit from immediate curative treatment. Immediate curative treatment is least favourable for men aged 70–74 years with low-risk (T1G6) or high-risk cancer (T3G8).

CONFLICT OF INTEREST

Heijnsdijk and de Koning receive a grant from Beckman Coulter for a cost-effectiveness analysis of phi. The remaining authors declare no conflict of interest.

DISCLAIMER

The authors are solely responsible for the content. The views presented do not necessarily represent the official views of the National Cancer Institute, the National Institute of Health, or the Centers for Disease Control and Prevention. The funding sources had no role in the design or conduct of the study, no role in the

collection, management, analysis, or interpretation of the data, and no role in the writing of the report.

REFERENCES

Albertsen PC, Hanley JA, Fine J (2005) 20-year outcomes following conservative management of clinically localized prostate cancer. *JAMA* **293**: 2095–2101.

Albertsen PC, Moore DF, Shih W, Lin Y, Li H, Lu-Yao GL (2011) Impact of comorbidity on survival among men with localized prostate cancer. *J Clin Oncol* **29**: 1335–1341.

Aus G, Robinson D, Rosell J, Sandblom G, Varenhorst E (2005) Survival in prostate carcinoma–outcomes from a prospective, population-based cohort of 8887 men with up to 15 years of follow-up: results from three countries in the population-based National Prostate Cancer Registry of Sweden. *Cancer* **103**: 943–951.

Bill-Axelson A, Holmberg L, Ruutu M, Garmo H, Stark JR, Busch C, Nordling S, Haggman M, Andersson SO, Bratell S, Spangberg A, Palmgren J, Steineck G, Adami HO, Johansson JE (2011) Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med* **364**: 1708–1717.

Draisma G, Boer R, Otto SJ, van der Cruijsen IW, Damhuis RA, Schroder FH, de Koning HJ (2003) 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* **95**: 868–878.

Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Gulati R, Feuer E, de Koning H (2009) Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst* **101**: 374–383.

Draisma G, Postma R, Schroder FH, van der Kwast TH, de Koning HJ (2006) Gleason score, age and screening: modeling dedifferentiation in prostate cancer. *Int J Cancer* **119**: 2366–2371.

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin D (2010) *GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10*. International Agency for Research on Cancer: Lyon, France. http://www.globocan.iarc.fr (accessed 28 December 2011).

Heijnsdijk EA, Wever EM, Auvinen A, Hugosson J, Ciatto S, Nelen V, Kwiatkowski M, Villers A, Paez A, Moss SM, Zappa M, Tammela TL, Makinen T, Carlsson S, Korfage IJ, Essink-Bot ML, Otto SJ, Draisma G, Bangma CH, Roobol MJ, Schroder FH, de Koning HJ (2012) Quality-of-life effects of prostate-specific antigen screening. *N Engl J Med* **367**: 595–605.

Hugosson J, Aus G, Lilja H, Lodding P, Pihl CG (2004) Results of a randomized, population-based study of biennial screening using serum prostate-specific antigen measurement to detect prostate carcinoma. *Cancer* **100**: 1397–1405.

Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, Pihl CG, Stranne J, Holmberg E, Lilja H (2010) Mortality results from the Goteborg randomised population-based prostate-cancer screening trial. *Lancet Oncol* **11**: 725–732.

Korfage IJ, Essink-Bot ML, Borsboom GJ, Madalinska JB, Kirkels WJ, Habbema JD, Schroder FH, de Koning HJ (2005) Five-year follow-up of health-related quality of life after primary treatment of localized prostate cancer. *Int J Cancer* **116**: 291–296.

Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J (1997) Cancer Incidence in Five Continents, Vol. VII. In *IARC Scientific Publications No. 143* Vol. 5, IARC: Lyon.

Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Maattanen L, Bangma CH, Aus G, Villers A, Rebillard X, van der Kwast T, Blijenberg BG, Moss SM, de Koning HJ, Auvinen A (2009) Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* **360**: 1320–1328.

Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Paez A, Maattanen L, Bangma CH, Aus G, Carlsson S, Villers A, Rebillard X, van der Kwast T, Kujala PM, Blijenberg BG, Stenman UH, Huber A, Taari K, Hakama M, Moss SM, de Koning HJ, Auvinen A (2012) Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med* **366**: 981–990.

SEER (2010) *Cancer Statistics Review 1975–2007*. National Cancer Institute.

Spapen SJ, Damhuis RA, Kirkels WJ (2000) Trends in the curative treatment of localized prostate cancer after the introduction of prostate-specific antigen: data from the Rotterdam Cancer Registry. *BJU Int* **85**: 474–480.

Visser O, Coeberg JW, Schouten LJ (1994) *Incidence of Cancer in The Netherlands*, vol. 13, The Netherlands Cancer Registry: Utrecht, The Netherlands 45.

Wever EM, Draisma G, Heijnsdijk EA, Roobol MJ, Boer R, Otto SJ, de Koning HJ (2010a) Prostate-specific antigen screening in the United States vs in the European Randomized Study of Screening for Prostate Cancer-Rotterdam. *J Natl Cancer Inst* **102**: 352–355.

Wever EM, Heijnsdijk EA, Draisma G, de Koning HJ (2010b) *Model Profile of the MISCAN Prostate Cancer Model*. Department of Public Health, Erasmus MC: The Netherlands, accessed at https://cisnet.flexkb.net/mp/pub/CISNET_ModelProfile_PROSTATE_ERASMUS_001_12152009_69754.pdf.

Wever EM, Hugosson J, Heijnsdijk EA, Bangma CH, Draisma G, de Koning HJ (2012) To be screened or not to be screened? Modeling the consequences of PSA screening for the individual. *Br J Cancer* **107**: 778–784.