Empirical Estimates of the Lead Time Distribution for Prostate Cancer Based on Two Independent Representative Cohorts of Men Not Subject to Prostate-Specific Antigen Screening

Caroline J. Savage1, Hans Lilja2,3,4, Angel M. Cronin1, David Ulmert3,5, and Andrew J. Vickers1

Abstract

Background: Lead time, the estimated time by which screening advances the date of diagnosis, is used to calculate the risk of overdiagnosis. We sought to describe empirically the distribution of lead times between an elevated prostate-specific antigen (PSA) and subsequent prostate cancer diagnosis.

Methods: We linked the Swedish cancer registry to two independent cohorts: 60-year-olds sampled in 1981-1982 and 51- to 56-year-olds sampled in 1982-1985. We used univariate kernel density estimation to characterize the lead time distribution. Linear regression was used to model the lead time as a function of baseline PSA and logistic regression was used to test for an association between lead time and either stage or grade at diagnosis.

Results: Of 1,167 older men, 132 were diagnosed with prostate cancer, of which 57 had PSA ≥3 ng/mL at baseline; 495 of 4,260 younger men were diagnosed with prostate cancer, of which 116 had PSA ≥3 ng/mL at baseline. The median lead time was slightly longer in the younger men (12.8 versus 11.8 years). In both cohorts, wide variation in lead times followed an approximately normal distribution. Longer lead times were significantly associated with a lower risk of high-grade disease in older and younger men [odds ratio, 0.82 (P = 0.023) and 0.77 (P < 0.001)].

Conclusion: Our findings suggest that early changes in the natural history of the disease are associated with high-grade cancer at diagnosis.

Impact: The distinct differences between the observed distribution of lead times and those used in modeling studies illustrate the need to model overdiagnosis rates using empirical data. Cancer Epidemiol Biomarkers Prev; 19(5); 1201–7. ©2010 AACR.
date of diagnosis, and therefore screening is unlikely to be beneficial. In contrast, a very long lead time suggests that men are more likely to die from other causes before their disease would ever clinically manifest, resulting in overdiagnosis. The lead time of prostate cancer is therefore a critical part of estimating overdiagnosis rates and assessing the effect of screening.

To model overdiagnosis rates accurately, two items of information about lead time are needed. The first is the central estimate of the lead time, such as the mean or median. These estimates have been calculated in several studies, both in retrospective studies that used stored blood samples from men who were later clinically diagnosed with prostate cancer (11-14) and in modeling and simulation studies (7, 10, 15-17). Mean lead time estimates range from approximately 5 to 12 years, depending on the modeling assumptions and length of follow-up.

Calculation of overdiagnosis rates also requires the shape of the lead time distribution. To our knowledge, no study has described the empirical distribution of lead times. In the absence of an empirical shape of the lead time distribution, several studies have used theoretical parametric distributions in their calculations of overdiagnosis rates. The choice of distribution varies by study and seems to be driven by convenience rather than considerations of tumor biology (8, 18-20).

In this study, we characterize the distribution of lead times using the time between an elevated PSA and subsequent clinical diagnosis of prostate cancer in two independent cohorts of Swedish men. These estimates do not rely on the assumptions typical of simulation studies and are based on more than 25 years of follow-up, considerably longer than previous retrospective studies.

**Materials and Methods**

**Patient cohort.** Our study used two independent cohorts of Swedish men who participated in the Malmo Preventive Project (21, 22). The first cohort included 1,167 men who were born in 1921 and received a baseline evaluation and provided a blood sample in 1981. The second cohort included 4,260 men who were born between 1926 and 1931 and were sampled in 1982-1985. Participation rates in both cohorts were above 70%. EDTA-anticoagulated blood was collected at each occasion and archived at −20°C. The Cancer Registry at the National of Health and Welfare in Sweden was used to identify men diagnosed with prostate cancer before December 31, 2006. In both cohorts, prostate cancer cases were matched with three controls, matched on date of birth and date of baseline venipuncture (both ±2 months for all cases), who were selected at random from the group of men alive and without prostate cancer diagnosis at the follow-up time at which the case was diagnosed.

**Laboratory methods.** Stored frozen anti-coagulated blood plasma samples were used to determine men’s PSA values at baseline. After thawing, blood samples were measured using the dual-label DELFIA Prostate total/free PSA-Assay (Perkin-Elmer). All analyses were conducted at the Wallenberg Research Laboratories, Department of Laboratory Medicine, Lund University, University Hospital UMAS in Malmö, Sweden.

**Statistical analyses.** Our primary aim was to determine the lead time associated with PSA screening for prostate cancer. We assumed that an elevated PSA represented cancer that was already present among the men who later developed clinically diagnosed prostate cancer. Lead time was thus calculated as the time interval between an elevated baseline PSA (≥3 ng/mL) and subsequent clinical diagnosis of prostate cancer among men who eventually received a prostate cancer diagnosis. We did not calculate a lead time for patients who developed cancer but whose baseline PSA was <3 ng/mL at the time of blood draw. Like all retrospective studies that use stored blood samples, our study makes the assumption that all men with an elevated PSA who were later clinically diagnosed with prostate cancer already had screen-detectable prostate cancer at the time of blood sampling, despite the fact that men were not actually diagnosed at a preclinical stage. It is also possible that some men may have had an elevated PSA at baseline for another reason, such as benign prostatic hyperplasia or prostatitis, and then subsequently developed prostate cancer.

We used univariate kernel density estimation to characterize the distribution of time between an elevated PSA measurement and subsequent clinical diagnosis of cancer (23). The resulting figure can be thought of as a “smoothed histogram” of the lead times in which the “bins” are formed by a moving kernel function. We chose to use the Epanechnikov kernel; alternative kernels (such as Gaussian and rectangular shapes) did not importantly change our findings. The bandwidth designates the region within which the density of lead times will be computed according to the kernel function. We used a bandwidth of 2.5 years; bandwidths of 2 and 3 years resulted in very similar results. Bootstrap methods were used to calculate the confidence intervals.

Our second aim was to characterize the relationship between an elevated PSA measurement at baseline and the length of time until clinical diagnosis of prostate cancer. We hypothesized that men with higher PSAs would have shorter lead times than men with lower PSAs. To test this hypothesis, we modeled the lead time as a function of PSA. To account for the potential non-linear relationship between PSA and lead time, we included restricted cubic splines with knots at the tertiles.

We also conducted an analysis to look at whether the lead time was associated with either stage or grade of the disease at clinical presentation. We classified patients as either high stage (T3 or higher) or high grade (Gleason score 7+ or WHO grade 3) and used logistic regression to test for an association, both with and without adjustment for baseline PSA level. We hypothesized that longer lead times would be associated with more favorable clinical features at presentation. All statistical analyses were conducted using Stata 10.0 (StataCorp LP).
Results

In total, 132 older men (age 60 at blood draw) and 495 younger men (ages 51-56 at blood draw) were diagnosed with prostate cancer. Characteristics of stage and grade are shown in Table 1. Of the patients who were diagnosed with cancer, 57 of 132 (43%) older men and 116 of 495 (23%) younger men had an elevated PSA (≥3 ng/mL) at baseline. Overall, men with PSA <3 ng/mL who were subsequently diagnosed with prostate cancer tended to have lower-stage and lower-grade tumors at the time of clinical presentation than those with an elevated PSA at baseline. The majority of controls in both cohorts [349 of 394 (89%) and 1,371 of 1,434 (96%) in older and younger men] did not have an elevated PSA at baseline. Table 2 shows the proportion of men with various PSA levels at baseline.

The median time between an elevated PSA and subsequent clinical diagnosis was slightly longer in the younger men than in the older men [median, 12.8 (interquartile range, 9.4-17.3) versus 11.8 (interquartile range, 8.8-17.6) years]. Figure 1A and B shows the density distribution of lead times following an elevated PSA in both cohorts. These figures can be interpreted as smoothed histograms of the lead time. Overall, we observed considerable, but relatively symmetric, variation in individuals' lead times. Among the older men, 10% of patients had lead times shorter than 3 years, whereas another 10% had lead times longer than 23 years. In the younger men, there was also wide variation: 10% of patients had lead times shorter than 7 years and 10% longer than 21 years, but fewer participants had very short lead times (only 2 men had lead times shorter than 3 years).

Statistical modeling and simulations rely on mathematical formulas to characterize the distribution of lead times. Therefore, we attempted to determine which, if any, commonly used parametric distributions would best approximate the observed variation in lead times. The lead times were not well approximated by the exponential or gamma distribution; however, the normal distribution seemed to be a much better fit of the data. The mean square error was more than 10 times higher for the exponential or gamma distribution than for the normal distribution in both cohorts.

Table 1. Characteristics of the patients who were subsequently diagnosed with prostate cancer

<table>
<thead>
<tr>
<th>Clinical stage</th>
<th>Men age 60 at time of blood draw, n (%)</th>
<th>Men ages 51-56 at time of blood draw, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSA ≥3 ng/mL at baseline, n = 57</td>
<td>PSA &lt;3 ng/mL at baseline, n = 75</td>
</tr>
<tr>
<td></td>
<td>PSA ≥3 ng/mL at baseline, n = 116</td>
<td>PSA &lt;3 ng/mL at baseline, n = 379</td>
</tr>
<tr>
<td>T0/T1</td>
<td>5 (9)</td>
<td>26 (35)</td>
</tr>
<tr>
<td>T2</td>
<td>26 (46)</td>
<td>21 (28)</td>
</tr>
<tr>
<td>T3/T4</td>
<td>21 (37)</td>
<td>15 (20)</td>
</tr>
<tr>
<td>Unknown</td>
<td>5 (9)</td>
<td>13 (17)</td>
</tr>
<tr>
<td></td>
<td>25 (22)</td>
<td>135 (36)</td>
</tr>
<tr>
<td></td>
<td>34 (29)</td>
<td>129 (34)</td>
</tr>
<tr>
<td></td>
<td>49 (42)</td>
<td>97 (26)</td>
</tr>
<tr>
<td></td>
<td>8 (7)</td>
<td>18 (5)</td>
</tr>
<tr>
<td>Biopsy grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gleason ≤6 or WHO 1</td>
<td>21 (37)</td>
<td>31 (41)</td>
</tr>
<tr>
<td>Gleason 7 or WHO 2</td>
<td>20 (35)</td>
<td>16 (21)</td>
</tr>
<tr>
<td>Gleason ≥8 or WHO 3</td>
<td>11 (19)</td>
<td>15 (20)</td>
</tr>
<tr>
<td>Unknown</td>
<td>5 (9)</td>
<td>13 (17)</td>
</tr>
<tr>
<td>Died</td>
<td>49 (86)</td>
<td>44 (59)</td>
</tr>
<tr>
<td>Alive at last follow-up</td>
<td>8 (14)</td>
<td>31 (41)</td>
</tr>
<tr>
<td></td>
<td>116 (100)</td>
<td>379 (100)</td>
</tr>
</tbody>
</table>

Table 2. Cases and controls by PSA range in the older and younger men in the Malmo Preventive Medicine Cohort

<table>
<thead>
<tr>
<th>Baseline PSA level (ng/mL) in anti-coagulated blood plasma</th>
<th>Men age 60 at time of blood draw, n (%)</th>
<th>Men ages 51-56 at time of blood draw, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>0-2.9</td>
<td>75 (57)</td>
<td>349 (89)</td>
</tr>
<tr>
<td>3.0-9.9</td>
<td>40 (30)</td>
<td>39 (10)</td>
</tr>
<tr>
<td>≥10.0</td>
<td>17 (13)</td>
<td>6 (2)</td>
</tr>
<tr>
<td>Total</td>
<td>132</td>
<td>394</td>
</tr>
</tbody>
</table>
Figure 2 shows the relationship between PSA levels at baseline and time to clinical diagnosis of prostate cancer for those with an elevated PSA. Overall, higher PSA levels were significantly associated with shorter lead times ($P = 0.032$ and $P = 0.004$ for the older and younger men, respectively); however, this association was largely driven by a small number of patients with very high PSAs and short lead times. In the older men, the association between PSA level and lead time was no longer significant when restricted to patients with PSA <20 ng/mL (excluding 6 cases; $P = 0.5$). Similarly, the association was no longer significant among the younger men when restricted to those with PSA <20 ng/mL (excluding 10 cases; $P = 0.12$). We observed substantial variation in individual lead times across PSA levels. For example, in the older men, the lead times ranged from 1.8 to 25.1 years for PSA 3-10 ng/mL and from 1.4 to 20.6 years for PSA 10-15 ng/mL.

We also found that longer lead times were significantly associated with a lower risk of high-grade disease at clinical presentation [odds ratio (95% confidence interval) for 1-year increase in lead time, $0.82$ ($0.69-0.97$; $P = 0.023$) and $0.77$ ($0.67-0.89$; $P < 0.001$) for older and younger men, respectively]. Adjusting for baseline PSA levels did not importantly change these results [odds ratio, $0.81$ ($P = 0.024$) and $0.78$ ($P < 0.001$)]. We found no evidence...
of an association between the lead time and high-stage disease (defined as T3 or greater) at clinical diagnosis in either older or younger men [odds ratio (95% confidence interval), 1.02 (0.94-1.10; P = 0.7) and 0.96 (0.89-1.03; P = 0.2), respectively].

Our results were based on the assumption that all men with an elevated PSA who were subsequently diagnosed had screen-detectable prostate cancer at the time of blood sampling. Men may have had an elevated PSA at baseline for another reason and then subsequently developed prostate cancer. A recent study by Robool et al. reported that in a sample of 3,056 men with elevated PSA, 1,077 had an initial positive biopsy, and 287 with an initially negative biopsy were subsequently diagnosed with prostate cancer over 11 years of follow-up (24). If we conservatively assume that all subsequently diagnosed cancers represent de novo cancers, then 20% of lead times are overestimated. Using the distribution of time from initial negative biopsy until diagnosis from the above study, we can adjust our estimates of lead time accordingly (i.e., 10% of lead times would be overestimated by 2 years; 8% by 6 years, and 2% by 10 years). This adjustment would only slightly lower the mean lead times: 12.1 versus 12.8 years and 12.5 versus 13.4 years in the older and younger men, respectively. Furthermore, the distribution of lead times was not importantly changed and the normal distribution remained a substantially better fit than the exponential or gamma distribution.

Lastly, we conducted a sensitivity analysis restricted to men with a PSA between 3 and 5 ng/mL at baseline. The rationale is that this range is characteristic of regularly screened men; those with higher PSA values would likely have been identified from an earlier screening. In total, 63 younger men had a PSA 3-5 ng/mL and were subsequently diagnosed with prostate cancer. Overall, our main findings were not importantly changed when restricted to this subgroup. The median estimate was 13.9 years (interquartile range, 11.1-17.7) and there remained considerable variation in the lead times that was best approximated by the normal distribution (the mean square error was nearly 20 times higher for the exponential or gamma distribution than for the normal distribution). The small number of older men with PSA 3-5 ng/mL (n = 22) precluded a similar analysis in this cohort.

Discussion

We have shown that there is substantial variation in the time from an elevated PSA measurement to a subsequent clinical diagnosis of prostate cancer in two unscreened cohorts of Swedish men. To our knowledge, this is the first study to empirically characterize the distribution of lead times associated with PSA screening. We also found that longer lead times were associated with a lower risk of high-grade disease at clinical presentation. Higher baseline PSA measurements were associated with shorter lead times; however, this association was almost entirely driven by men with high PSA values.

Previous studies have reported mean or median lead time estimates ranging from approximately 5 to 12 years (7, 10-13, 15, 16). The variation in estimates from statistical modeling and simulation studies reflects the various different assumptions, models, and populations used to calculate the lead time (7, 10, 15, 16). Our median estimates of the lead time, 11.8 and 12.8 years for older and younger men, were slightly higher than those of earlier retrospective studies (11-13), most likely because our study included longer follow-up. Lead time estimates are heavily influenced by the length of follow-up: an individual with a 20-year lead time would not be captured by a study with only 15 years of follow-up.

In the absence of an empirical distribution, several authors have relied on theoretical distributions in calculations of overdiagnosis calculations and simulation studies. Etzioni reported that overdiagnosis rates of 29% for whites and 44% for blacks were most consistent with observed incidence trends of prostate cancer (8). These estimates used three mean lead times derived from the literature and modeled the lead time with a gamma distribution. In a subsequent article, the mean lead time was estimated to be 4.6 years for whites and 6.8 years for blacks, with corresponding overdiagnosis rates of 22.7% and 34.4%; these analyses assumed that the lead time distribution was exponential.

Using theoretical parametric distributions simplifies calculations; however, the choice of distribution can have a profound effect on the resulting estimates of overdiagnosis. For example, if we assume that the lead time followed an exponential distribution, with the mean from older men in this study (12.8 years), approximately 55% of men would be expected to have lead times less than 10 years, and 30% of men with lead times less than 5 years. However, in our empirical study, less than 35% of 60-year-old men had lead times less than 10 years and only 20% had lead times less than 5 years. Assuming that the lead time followed a gamma distribution was similarly a poor model for our data. Although there are clear differences between the shape of the lead time distribution and the normal distribution—most notably that the lead times were somewhat overdispersed—the normal distribution represents a substantial improvement over the exponential or gamma distribution.

As hypothesized, higher baseline PSA levels were significantly associated with shorter lead times; however, in both cohorts, this association was driven by a few men with very high PSAs and short lead times. One prior study attempted to describe the relationship between PSA levels at baseline and median lead time (11). With 20 years of follow-up, Tornblom et al. found that median lead times were 3.6 years for men with PSA ≥10 ng/mL and 11.2 years for men with PSA 3-10 ng/mL. Our results showed a similar pattern, but the median lead times were longer overall. Furthermore, although the mean lead times seemed to decrease with higher PSA, there remained substantial variation in individuals’ lead times.
We also found evidence of an association between longer lead times and lower risk of high-grade, but not high-stage, disease at clinical presentation. These findings support the theory that early changes in the prostate are responsible for the tumor grade at presentation and the time from screen-detectable disease to clinical diagnosis. If high-grade disease was the product of a long series of genetic mutations, one would expect longer lead times to be associated with higher-grade prostate cancer, as these cancers would have had the longest time to acquire mutations. Instead, it seems that changes that occur early in the natural history of the disease are associated with the aggressiveness of the cancer, in terms of both how quickly the tumor progresses and how poorly differentiated the disease is at presentation.

There are several limitations of this study. First, it might be argued that 57 and 116 patients are small numbers; however, it is unlikely that more patients would importantly change the distribution of lead times or our main findings. Second, we assumed that all men with an elevated PSA who were later clinically diagnosed with prostate cancer already had screen-detectable prostate cancer at the time of blood sampling. It is plausible that some men may have had an elevated PSA at baseline for some other reason, such as benign prostatic hyperplasia or prostatitis, and then independently developed prostate cancer at a later time. Including these men would have overestimated the lead time; however, we showed that this is unlikely to be a large number of participants, and did not importantly change the central estimate or distribution of the lead time. Third, this study has characterized the distribution of lead times following a single PSA measurement at either age 60 or age 51 to 56. Some patients may have had an elevated PSA for several years and thus had cancer that was screen-detectable several years before the date of blood draw. However, when we only included subjects with PSA between 3 and 5 ng/mL—characteristic of men undergoing repeated screening—the distribution of lead times was nearly identical to that of all men with an elevated PSA. Lastly, follow-up was censored at 25 years. Therefore, some men with an elevated PSA at baseline who were cancer-free and alive at last follow-up could have subsequently developed prostate cancer. This is unlikely to be a large number of men or to importantly change our main findings; however, it could result in a slight underestimate of the lead time.

In sum, we showed that in two separate cohorts with 25 years of follow-up and very little, if any, screening contamination, the lead time following an elevated PSA varied considerably for individuals. These large differences seem to reflect the inherent variation in the aggressiveness of prostate cancer, as shorter lead times were associated with a higher grade of disease at clinical presentation. More research is needed to better understand other factors driving the large variation in lead times. The distinct differences between the observed distribution of lead times and those used in modeling studies illustrate the need to model overdiagnosis rates using empirical data.

**Disclosure of Potential Conflicts of Interest**

Dr. Hans Lilja holds patents for free PSA and hK2 assays. No other potential conflicts of interest were disclosed.

**Acknowledgments**

We thank Gun-Britt Eriksson, Kerstin Håkansson, and Mona Hassan Al-Battat for expert assistance with immunoassays.

**Grant Support**

Supported in part by a R21-CA127768-01A1 phased innovation research in cancer prognosis and prediction grant from the National Cancer Institute; Swedish Cancer Society [Project No. 3455]; Swedish Research Council (Medicine) [Project No. 20095]; Fundación Federico SA; David H. Koch through the Prostate Cancer Foundation, the Sidney Kimmel Center for Prostate and Urologic Cancers, and Specialized Program of Research Excellence grant P50-CA92629 from the National Cancer Institute to Dr. P.T. Scardino.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 12/11/2009; revised 02/10/2010; accepted 02/18/2010; published OnlineFirst 04/20/2010.

**References**

Empirical Estimates of the Lead Time Distribution for Prostate Cancer Based on Two Independent Representative Cohorts of Men Not Subject to Prostate-Specific Antigen Screening

Caroline J. Savage, Hans Lilja, Angel M. Cronin, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-09-1251

Cited articles
This article cites 23 articles, 2 of which you can access for free at:
http://cebp.aacrjournals.org/content/19/5/1201.full#ref-list-1

Citing articles
This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/19/5/1201.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.