The Dynamics of Death in Prostate Cancer

Robin T. Vollmer, MD, MS

Key Words: Hazard function; Histology; Prostate cancer; Survival

Abstract

The hazard function provides the instantaneous probability of death (or other key end point) at various times after diagnosis. Unlike the survival curve, the hazard function illustrates graphically or through calculations when deaths are common or uncommon. In this study, hazard functions were derived for prostate cancer by using survival data on large numbers of patients with prostate cancer with data in the Surveillance, Epidemiology and End Results (SEER) database. The results demonstrate a form of prostate cancer that rapidly evolves to cause death within 5 years, and this form of tumor is only partly identified by routine prognostic variables such as serum prostate-specific antigen (PSA) level, histologic grade, and quantity of tumor. The results also validate the presence of a reservoir of nonfatal prostate cancers that have increased rapidly during the PSA era, and they demonstrate that the incidence of fatal prostate cancers has declined.

Prostate-specific antigen (PSA) was discovered in the 1970s, purified and named as such by 1979, and tested as a serum marker in the 1980s. In 1992, the US Food and Drug Administration approved PSA as a screening test for prostate cancer, and it was hailed as “the most useful tumor marker” and lauded for its ability to enhance the detection of early prostate cancer. Subsequently, the number of PSA tests soared, as did the number of needle biopsies of prostate. For example, Figure 1 shows the percentage of men without prostate cancer who were older than 65 years and who had at least 1 PSA test during the calendar year on the abscissa (dots for white men and line for black men). Nevertheless, as pathologists fretted about overlooking small foci of tumor and as urologists opined that the PSA thresholds for performing biopsies should be lowered, epidemiologists and statisticians began to argue that in the PSA era, prostate cancer was being overdiagnosed, that too many men were being treated, and that PSA-based screening for prostate cancer offered an unproven benefit. Although their conclusions were based on mathematical models and computer simulations, 2 prospective trials (US Prostate, Lung, Colorectal, and Ovarian Cancer Screening trial and the European Randomized Study of Screening for Prostate Cancer trial) demonstrated that PSA-based screening produces overdiagnosis and overtreatment. Taken collectively, the trials showed at most a modest reduction in prostate cancer–specific mortality. Nevertheless, both trial results were considered preliminary because they were based on interim analyses that included just 634 patients observed to die of prostate cancer.

Recently, Welch and Black emphasized that overdiagnosis in tumors is closely linked to the dynamics of tumor...
growth and progression. Rapidly growing tumors cause clinical symptoms, can be diagnosed without screening, and often progress to cause death. By contrast, slowly growing tumors may not cause symptoms and progress so slowly that many patients die of other causes. The slowly progressive tumors comprise a reservoir of asymptomatic cases detectable by screening tests such as PSA. In this manner, screening and overdiagnosis of any tumor are directly linked to the dynamics of the tumor’s progression. A probabilistic function that in turn is closely linked to the dynamics of the tumor’s progression is the hazard function, which provides the instantaneous probability of death at any time after diagnosis. In what follows, I derive hazard functions for death in prostate cancer using large numbers of patients with known outcomes in the Surveillance, Epidemiology and End Results (SEER) database. The hazard-derived results help identify an aggressive form of prostate cancer that existed before and after the PSA era, and, with these mature historical data, they validate the phenomenon of overdiagnosis.

Materials and Methods

Study Data

The primary data for this study are values for incidence and survival, and these were obtained from the SEER database (http://www.seer.cancer.gov/csr/1975_2007; accessed August 1, 2010). Specifically, 11 survival curves were obtained for patients diagnosed during the following 11 periods: 1975-1979 (median at 1977), 1980-1984 (median at 1982), 1985-1989 (median at 1987), 1990, 1991, 1992, 1993, 1994, 1995, 1996, and 1997. The follow-up for the patients represented in these data was such that there were at least 10 years of usable survival results for each survival curve. In addition, I used survival data from a previously published study of 663 American Veterans Affairs patients diagnosed with prostate cancer and who were followed up to death or for a median of 5 years. I obtained the numbers of practicing urologists in the United States from http://www.census.gov/prod/3/98pubs/98statab/ (accessed August 4, 2011) and the prevalence of testing for serum PSA from Etzioni et al.

The Hazard Function

Kaplan-Meier plots are plots of the survival probability, S(t), at various times, t, after diagnosis, and when there are sufficient numbers of patients, S(t) can be considered a continuous function of t. Thus, in what follows, we will take advantage of this feature of S(t). The hazard function, h(t), is often called the force of mortality because higher levels of h(t) coincide with times of frequent deaths, and h(t) is related to S(t) as follows:

\[ h(t) = \frac{\partial \log (S(t))}{\partial t} \]

Here, log stands for the natural logarithm, and \( \partial \) symbolizes the partial differential with respect to time. Through integration, the relationship between S(t) and h(t) can also be written as:

\[ S(t) = \exp \left( - \int h(x) \, dx \right) \]

Here, exp stands for exponentiation, and the integration limits for x are from 0 to t.

The integral \( \int h(x) \, dx \) is often called the total hazard. Equation 2 implies that total fatality from 0 to any time, t, can be calculated as \( 1 - \exp ( - \int h(x) \, dx) \) with the integration limits being from 0 to t. Finally, in this study I used the sum of 2 \( \gamma \) functions for the mathematical form of h(t), because this form allowed for a near perfect fit of the survival data. (See Appendix 1 and the “Results” section for details.)

There are at least 2 advantages for examining h(t). First, graphs of h(t) show times of peak hazard and times of low hazard. Second, knowing the mathematical form for the hazard allows accurate estimation of the cumulative probability of death for any period during follow-up. Thus, in what follows, the derived mathematical hazard function will allow calculation of observed probabilities of death by 5 years after diagnosis and for all times of follow-up.

Results

The Fit of the Derived Hazard Function to SEER Data

Figure 2 shows a plot of observed values of S(t) (on the x-axis) vs that predicted by the derived hazard function h(t) and Equation 2 (on the y-axis) for all 11 SEER survival curves. The line on the plot shows where perfect agreement would occur, and the fact that the points lie close to the line indicates that the \( \gamma \) function model is an accurate model of the hazard
function for the observed SEER data. The median difference between observed \( S(t) \) and that predicted from the derived hazards was 0 (range, –0.015 to 0.008). Values of \( \alpha \) averaged 0.0758 (range 0.02 to 0.235), Values of \( \beta \) averaged 1.31 (range, 0.627 to 4.93). Values of \( \lambda \) averaged 0.00228 (range, 0.000217-0.00732), and values of \( \delta \) averaged 0.318 (range, 0.126-0.991).

Typical Hazard Functions

Figure 3 shows plots of 2 hazard functions: the first derived from SEER cases diagnosed in 1982 (upper curve) and the second derived from SEER cases diagnosed in 1993 (lower curve). The plots show that both hazard functions peak in the first 5 years after diagnosis and then gradually fall to near zero. All of the hazard functions in the SEER data manifest these early peaks, with an average peak at 1.2 years after diagnosis (range, 0.2-1.8 years). Furthermore, for all diagnostic years except 1992, the early peak was also the time for the highest hazard. The early peak in hazard implied that many deaths in prostate cancer occur during the first 5 years after diagnosis. For example, the percentage of total deaths in the SEER data occurring in the first 5 years averaged 44% (range, 20%-90%). Thus, the early peak in hazard function seem due to an aggressive form of prostate cancer that rapidly evolves to cause death. After 5 years, the fall in hazard function implies a lower death rate at scattered subsequent times.

Variables Related to Aggressive Prostate Cancer

The foregoing results suggest that death within 5 years of diagnosis is an end point that is closely related to an aggressive form of prostate cancer that rapidly evolves to cause death. Table I shows how several prognostic variables relate to this end point. For example, median values of PSA, Gleason score, tumor length, fraction of positive cores, and percentage of tumor were all significantly higher in patients observed to die within 5 years (\( P \approx 0 \); Wilcoxon tests). Nevertheless, the overlapping ranges for these variables suggest that none of them reliably discriminates aggressive tumors from the less aggressive tumors. Logistic regression analysis of these data showed that 2 variables, log(PSA) and primary Gleason grade, provided additive information about the probability of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fatal</th>
<th>Nonfatal</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSA level, ng/mL (μg/L)</td>
<td>13.6</td>
<td>7.2</td>
<td>~0</td>
</tr>
<tr>
<td>Median</td>
<td>0.2-2.691</td>
<td>0.4-201</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>5-10</td>
<td>4-10</td>
<td></td>
</tr>
<tr>
<td>Gleason score</td>
<td>7</td>
<td>6</td>
<td>~0</td>
</tr>
<tr>
<td>Median</td>
<td>0.02-167</td>
<td>0.02-136</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0.50</td>
<td>0.33</td>
<td>~0</td>
</tr>
<tr>
<td>Tumor length (mm)</td>
<td>10.6</td>
<td>5.0</td>
<td>~0</td>
</tr>
<tr>
<td>Median</td>
<td>0.02-167</td>
<td>0.02-136</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0.1-1</td>
<td>0.06-1</td>
<td></td>
</tr>
<tr>
<td>Fraction of positive cores</td>
<td>0.50</td>
<td>0.33</td>
<td>~0</td>
</tr>
<tr>
<td>Median</td>
<td>0.1-1</td>
<td>0.1-1</td>
<td>~0</td>
</tr>
<tr>
<td>Range</td>
<td>0.4-90</td>
<td>0.1-15</td>
<td>~0</td>
</tr>
<tr>
<td>Percentage of tumor</td>
<td>15</td>
<td>6</td>
<td>~0</td>
</tr>
<tr>
<td>Median</td>
<td>0.4-90</td>
<td>0.1-95</td>
<td>~0</td>
</tr>
<tr>
<td>Predictive probability</td>
<td>0.17</td>
<td>0.09</td>
<td>See text</td>
</tr>
<tr>
<td>Median</td>
<td>0.02-0.93</td>
<td>0.01-0.68</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PSA level and Gleason grade refer to the primary tumor. \( P \) values were calculated by using the nonparametric Wilcoxon test. The predictive probability comes from a logistic regression analysis using a 2-variable model with log(PSA) and the primary Gleason grade (see text).

Anatomic Pathology / Original Article
early death, with \( P \) values, respectively, of approximately 0 for log(PSA) and .001 for the primary Gleason grade. After these 2, none of the remaining variables in Table 1, including full Gleason score, provided additional prognostic information \( (P > .2) \). Nevertheless, the predictive probability of early death formulated from the 2-variable logistic model once again demonstrated so much overlap of values between the rapidly evolving vs the more slowly evolving tumors that it also cannot be considered a reliable discriminating model.

The Hazard Functions Before and After Serum PSA

The dramatic differences in curve height and enclosed area for the 2 hazard functions illustrated in Figure 3 are a small sample of a downward trend in hazard during the PSA era. In Figure 3, the hazard function for cases diagnosed in 1982 (before PSA testing was used) is clearly higher than for cases diagnosed in 1993 (when PSA testing was commonly used). The areas under these 2 curves correspond to the integral in Equation 2, so that the larger area for the 1982 hazard implied a higher total death rate, which for cases diagnosed in 1982 was 57% compared with 14% for cases diagnosed in 1993. Figure 4 shows that there was a steady decline in the hazard peak during the years before and after PSA testing was used (significant, \( P \sim 0 \); linear regression). Total mortality also dropped from 61% in 1977 to 2% in 1997 \( (P \sim 0; \) linear regression). Nevertheless, the times of peak hazard did not change \( (P > .1; \) linear regression), implying that a component of rapidly progressing tumor remained present throughout the period.

Incidence of Nonfatal and Fatal Forms of Prostate Cancer

The incidence rates for nonfatal and fatal cases of prostate cancer were obtained as products of the incidence of prostate cancer for each diagnostic year and the probabilities of non-fatality and fatality, respectively, for those years as follows:

\[
\text{Nonfatal Cases} = \text{Incidence} \times \left[\exp \left( - \int h(x) \, dx \right) \right]
\]

\[
\text{Fatal Cases} = \text{Incidence} \times \left[1 - \exp \left( - \int h(x) \, dx \right) \right]
\]

Here, the hazard function used was specific to the year of diagnosis, and the integration limits were from 0 to \( \infty \). Figure 5 shows plots of the incidence of nonfatal and fatal prostate cancers vs the year of diagnosis. Whereas the incidence of nonfatal cases rose during the PSA era \( (P = .001; \) linear regression analysis of logarithm of incidence vs time), the incidence of fatal cases fell \( (P < .006 \) for linear regression analysis of incidence vs time and time squared). Figure 6 shows that during this time, the number of practicing urologists in the United States also rose steadily in nearly a linear manner \( (P \sim 0; \) linear regression analysis). Thus, the incidence of nonfatal cases was positively associated with the number of practicing urologists \( (P = .0002 \) for linear regression analysis of logarithm of incidence vs number of urologists), and the incidence of fatal cases dropped as the number of practicing urologists rose \( (P < .002 \) by linear regression for a parabolic relationship between the incidence of fatal cases and the number of urologists).

Although these results do not prove a causative relationship between the number of urologists and the incidence of nonfatal cases of prostate cancer, the results support the concept of Welch and Black\(^{23} \) of a reservoir of subclinical prostate cancers.

Discussion

This study demonstrates that the hazard function provides useful information about prostate cancer, and it demonstrates that the effects of PSA screening on the types of tumors detected could have been predicted from a careful analysis of mature SEER survival data, that is, without using the results of prospective randomized trials. Deriving
the hazard function requires neither complex modeling nor assumptions about stage progressions, and it requires no computer simulations. Its derivation comes straightforwardly from the calculus of survival, and its analysis requires nothing other than mature survival data. The results also documented a form of rapidly evolving fatal prostate cancer that persists into the PSA era, even though it has been diluted by many cases of nonfatal prostate cancer. This form of tumor relates to the observation that an average of 44% of deaths occur during the first 5 years after diagnosis. Because the hazard function peaks at less than 2 years after diagnosis, this form of fatal tumor may evolve so quickly that it may not be affected by screening or localized treatment. And it is not accurately recognized by routine prognostic factors. The results support conclusions about the rising incidence of a slowly evolving and nonfatal prostate cancer during the PSA era and document that the incidence of this reservoir of nonfatal tumors is directly related to the number of practicing urologists.

Finally, the results demonstrated a decline in fatal forms of prostate cancer and showed that this decline is associated with the use of serum PSA testing and the number of practicing urologists. The reasons for the decline in fatal forms of prostate cancer are probably multiple. The decline may be due in part to earlier diagnosis and effective treatment of prostate cancers that evolve at a sufficiently slow rate that they can be detected by PSA screening and treated before they spread. The decline in fatality may also be due in part to how prostatectomy removes hyperplastic tissue, thereby preventing obstructive uropathy and renal failure. And the decline may relate to how serum PSA testing allows more accurate attribution of death as being due to prostate cancer. Before serum PSA testing, some men’s death may have been attributed to prostate cancer, when, in fact, their level of PSA, had it been measured, would have been found to be too low. Now with serum PSA testing, such deaths are attributed to other causes.

From the Department of Laboratory Medicine, VA Medical Center, and Department of Pathology, Duke University Medical Center, Durham, NC.

Address reprint requests to Dr Vollmer: Laboratory Medicine 113, VA Medical Center, 508 Fulton St, Durham, NC 27705.

References

To estimate the hazard function, \( h(t) \), I first obtained raw values for \( h(t) \) as follows. If the survival probabilities in a survival curve at times \( t_1 \) and \( t_2 \) are \( S(t_1) \) and \( S(t_2) \), respectively, then Equation 1 implies that \( h(t) \) can be approximated for a median time, \( tm \), between \( t_1 \) and \( t_2 \) as:

**Equation A1**

\[
h(tm) = -\log \left( \frac{S(t_2)}{S(t_1)} \right) / (t_2 - t_1)
\]

An example of a plot of such raw values for the hazard function of 1 Surveillance, Epidemiology and End Results (SEER) survival curve is shown as the points in Figure 7. Note that the raw estimates of hazard rise to a peak before 5 years and then gradually fall. This early peak in raw hazards occurred at an average of 1.8 years after diagnosis across the 11 SEER survival curves (range, 0.5-3.5 years), and the pattern of an early rise in hazard followed by a fall suggested that the sum of 2 \( \gamma \) functions would be a suitable mathematical form for \( h(t) \). For example, the smooth curve in the plot comes from the sum of 2 such \( \gamma \) functions, \( h_1(t) \) and \( h_2(t) \), given as follows:

**Equation A2**

\[
h_1(t) = 0.115 \times t \times \exp(-0.646 \times t)
\]

**Equation A3**

\[
h_2(t) = 0.027 \times t^2 \times \exp(-0.212 \times t)
\]

Consequently, the following general form for \( h(t) \) was adopted:

**Equation A4**

\[
h_1(t) = \alpha \times t \times \exp(-\beta \times t)
\]

**Equation A5**

\[
h_2(t) = \delta \times t^2 \times \exp(-\delta \times t)
\]

with \( \alpha, \beta, \lambda \), and \( \delta \) representing coefficients to be adjusted so that \( h(t) \) as used in Equation 2 would fit the observed values of \( S(t) \). The form for \( h_1 \) allows for an early peak in hazard at a time equaling \( 1/\beta \), while the form for \( h_2 \) allows for a later peak at a time equaling \( 2/\delta \). In fact, initial values for \( \beta \) and \( \delta \) were obtained for each survival curve using the locations of the first and secondary peaks in the raw hazards, and values of \( \alpha \) and \( \lambda \) were obtained from the heights of the raw hazard peaks. Final values for \( \alpha, \beta, \lambda \), and \( \delta \) were obtained by the iterative nonlinear least-squares fitting routine nls in SPLUS (MathSoft, Seattle, WA). Thus, for each of the SEER survival curves, \( h(t) \) was derived in a stepwise manner: first plotting raw estimates of \( h(t) \), then obtaining values of \( \alpha, \beta, \lambda \), and \( \delta \) to roughly match \( h(t) \), and then finally using these values and nonlinear least-squares fitting to obtain final values of \( \alpha, \beta, \lambda \), and \( \delta \). The goal in this process was to derive a mathematical form for \( h(t) \) so that the right side of Equation 2 would provide a close fit to observed values of \( S(t) \). Figure 8 shows an example of how well this process worked for one of the survival curves. In Figure 8, the points are the observed survival values from the SEER survival curve for cases diagnosed between 1980 and 1984, and the smooth curve comes from the fitted \( h(t) \) as used in Equation 2. See the “Results” section for how well the \( \gamma \) model for \( h(t) \) worked for all SEER data.