Quantifying the Amount of Variation in Survival Explained by Prostate-specific Antigen[1]

David A. Verbel,[2] Glenn Heller, William K. Kelly, and Howard I. Scher

Departments of Epidemiology and Biostatistics [D. A. V., G. H.] and Medicine [W. K. K., H. I. S.], Memorial Sloan-Kettering Cancer Center, New York, New York 10021

ABSTRACT

Purpose: **Post-therapy changes in prostate-specific antigen (PSA) have been proposed as surrogates for survival in clinical trials due to observed statistical associations. However, association alone does not satisfy the conditions of surrogacy. A measure that quantifies the amount of association is important. Using a population-based approach, we explore the relationship between PSA and survival and generate a measure to demonstrate the amount of the variation in survival explained by PSA.**

Experimental Design: **With serial PSA measurements from 254 patients with androgen-independent prostate cancer, we use a time-dependent Cox model with a nonlinear log relative risk function to quantify the strength of the association between time-dependent PSA and survival. The nonlinear log relative risk function provides a flexible Cox model and a more accurate measure of the strength of association between PSA and survival.**

Results: **Among these 254 patients, there were 247 deaths (median survival time = 13.0 months). Median follow-up time for those alive or censored was 59.3 months (range, 36.8–71.3 months). An association was observed between PSA and survival ($P < 0.01$, two-sided test). However, time-dependent PSA explains only 17% of the variation in survival.**

Conclusions: **Use of this methodology demonstrates that there remains sufficient variation in survival unaccounted for by PSA measurements in this patient cohort. Other factors, perhaps unknown, exist that determine survival outcome. Consideration of PSA alone as a surrogate can produce misleading information regarding the risk of death; its use as a surrogate for survival is not warranted when designing a clinical trial in this patient population.**

INTRODUCTION

PSA[3] is the most widely used marker in the management of prostate cancer today. It has been evaluated for screening, to define prognosis, and to assess therapeutic outcomes in several well-defined clinical states (1). The extensive use of PSA in these areas and observed statistical associations between various defined changes in post-therapy PSA and progression-free survival and overall survival has led to the hypothesis that a defined post-therapy change in PSA has value as a surrogate end point for survival in clinical trials (2, 3). This idea is appealing because PSA is easy to measure in an inexpensive and reproducible way, and PSA changes usually occur earlier than a progression or survival end point. Furthermore, PSA is always attainable, whereas traditional outcome measures may not be in cases of advanced metastatic prostate cancer (4).

These studies, however, have focused only on establishing a statistical association between PSA and survival without quantifying the strength of the association. This is troubling because the establishment of a statistical association through an appreciably small *P* value is not sufficient to justify its use as a surrogate for survival (5). Focusing on PSA, when survival is only marginally affected by biological mechanisms (including treatment), is misguided; instead the investigative effort should be expanded to include the identification and understanding of the effect of other biological mechanisms on survival (6). An example of this phenomenon can be seen in a recent study of hormone therapy in the noncastrate metastatic population conducted by Eisenberger *et al.* (7). Specifically, the trial, using a PSA end point, demonstrated a higher proportion of patients achieving a normalization of PSA using a combined androgen blockade *versus* a monotherapy approach. Despite this, no difference in survival was observed between the two cohorts (7).

In this study, our objective is to quantify the level of association between PSA and survival. To accomplish this, we built a time-dependent Cox model with a general, nonlinear log relative risk function. The patient population used for the analysis is a cohort of castrate metastatic (androgen-independent) prostate cancer patients, with progressive disease, enrolled on consecutive therapeutic trials. We chose the castrate metastatic state because it is reasonable to assume that the survival end point will be observed and that there will be a reduced confounding effect of testosterone modulation, which is known to affect serum PSA levels. For this analysis, a post-therapy change in PSA is not specifically defined; rather, we use every PSA measurement available from all patients post-therapy to estimate the log relative risk function. Our approach is population-based. PSA values are not identified by individual subjects but rather are aggregated to assess the overall relationship in the population between PSA and survival time. Using the nonlinear

Received 1/25/02; revised 4/12/02; accepted 5/1/02.
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.
[1] Supported in part by CaPCURE.
[2] To whom requests for reprints should be addressed, at Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, Box 44, New York, NY 10021. Phone: (212) 639-2145; Fax: (212) 717-3137; E-mail: verbeld@mskcc.org.

[3] The abbreviation used is: PSA, prostate-specific antigen.

log relative risk Cox model, we demonstrate that there remains sufficient variation in survival unaccounted for by PSA. This suggests that focusing on PSA alone, while important, does not provide sufficient information to predict survival. As a result, we conclude that the various post-therapy change definitions of PSA such as percentage decrease from baseline (at varying landmark time points), PSA slope, PSA doubling time, and PSA velocity (3, 7–13) should not be used as surrogates for survival time when designing and evaluating clinical trials.

MATERIALS AND METHODS

The conventional time-dependent Cox model is written as follows:

$$h(t/Z_t) = h_0(t)\exp[Z_t\beta]$$

where Z_t represents the PSA value recorded at time t, and $Z_t\beta$ is termed a linear log relative risk function. This specific form of the relative risk function, which has been used in the majority of analyses of PSA and its effect on survival (3, 9, 11, 12, 14, 15), can be restrictive because by defining the log relative risk as a linear function, the log relative risk is always estimated by multiplying the PSA value recorded at time t by a constant slope parameter β. This relationship is often not supported by the data.

To quantify the level of association between survival and PSA, we propose a model that builds upon the above time-dependent Cox model but uses a nonlinear specification of the log relative risk function (16–19). This general, more flexible log relative risk function does not assume a linear relationship between the log relative risk of death and PSA but rather can take many forms, including that of a linear relationship if in fact that is appropriate. This has the practical advantage of allowing the Cox model to be applicable in a wider range of data sets. The model is written as follows:

$$h(t/Z_t) = h_0(t)\exp[S(Z_t)]$$

where $S(Z_t)$ is the general, nonlinear log relative risk function.

The reason for using the time-dependent Cox model with a general, nonlinear log relative risk function is to add flexibility when we use a measure of association. Model-based measures of association are influenced by (*a*) the strength of the relationship between the covariate and response and (*b*) proper model specification. In order for the measure to primarily reflect the strength of the association between time-dependent PSA values and survival, we replaced the linear log relative risk function $Z_t\beta$ found in the conventional Cox model with the less restrictive general log relative risk function $S(Z_t)$. Thus, in this flexible Cox model, we do not constrain the function to be linear, and furthermore, we allow the data to determine its functional form. The method of estimating the nonlinear function $S(Z_t)$ is described in the "Appendix."

The measure that quantifies the strength of association is the R^2 statistic adapted to the Cox model (20). This statistic, similar to the R^2 statistic used to summarize the goodness of fit in linear regression models, measures the amount of variability in survival explained by time-dependent PSA values (see "Appendix"). In the Cox model, R^2 is a function of the relative risk function $\exp[S(t)]$ over time and is formed by the ratio of the information provided by having the time-dependent PSA values

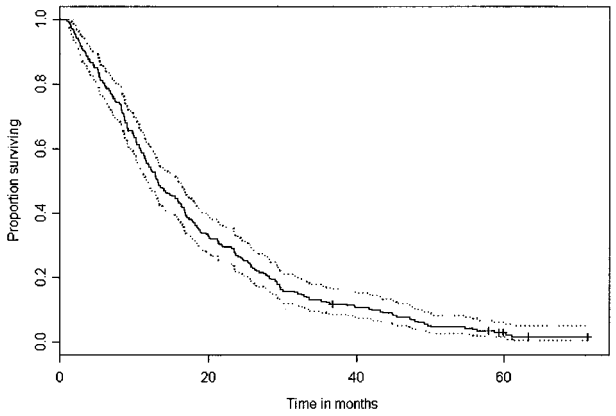

Fig. 1 Overall survival curve for 254 androgen-independent prostate cancer patients from Memorial Sloan-Kettering Cancer Center treated from 1987 through 1994. The *solid line* is the Kaplan-Meier estimate of the survival function. *Dashed lines* above and below the *solid line* are the 95% confidence bands around the Kaplan-Meier estimate.

(Z_t) in the model to the information provided by the model without PSA. This measure varies between 0 and 1, with high R^2 values indicating that PSA is discriminant with respect to the risk of death. Heuristically, when patients whose death is imminent have high PSA values (and conversely when patients with low and moderate PSA levels do not die in the short term), a high R^2 would result. All analyses were performed with S-Plus 2000 Professional software. Code is available from the authors upon request.

RESULTS

The dataset used for this analysis contained 254 patients with castrate metastatic (androgen-independent) prostate cancer who were treated on 11 consecutive protocols from 1987 through 1994 at Memorial Sloan-Kettering Cancer Center (3). All patients had documented progressive disease despite castrate levels of testosterone. In general, therapeutic outcomes were evaluated at a minimum of every 8 weeks, whereas PSA was collected a minimum of every 4 weeks. All patients were followed until death or censored at last follow-up. There were 247 deaths during this time period, and the median survival time was 13.0 months. Median follow-up time (for those still alive or censored) was 59.3 months (range, 36.8–71.3 months). Fig. 1 displays the overall survival curve for these patients, calculated using the Kaplan-Meier method. Further detail on these patients can be found in Scher *et al.* (3).

Five patients were removed from this analysis because only baseline PSA measurements were available. All five of these patients died. The median PSA value for the remaining 249 patients was 138 ng/ml (range, 0–16,400 ng/ml). To create a more manageable scale in which to describe the PSA measurements and to eliminate right-skewed data, a natural log transformation was performed (all PSA measurements had a value of 1 added to avoid the situation of having left-skewed data and to have positive values after the transformation). The median $\log_e(1 + \text{PSA})$ value was 4.9 (range, 0.0–9.7).

Our objective was to quantify the strength of the association between the time-dependent PSA values and survival. In the

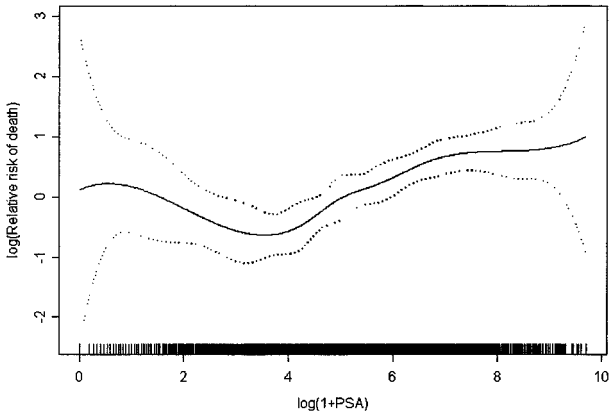

Fig. 2 Plot of $\log_e(1 + \text{PSA})$ and the log relative risk of death, estimated using a cubic spline with four knots. The *solid line* is the estimate of the log relative risk function. *Dashed lines* above and below the *solid line* are the 95% confidence bands around the estimate. *Hash marks* along the *X* axis indicate the density of $\log_e(1 + \text{PSA})$ values used in the estimation.

proposed Cox model, this association is described by the general, nonlinear log relative risk function *S*. The estimated relative risk function *S* is presented in Fig. 2. It is clear that the relationship between $\log_e(1 + \text{PSA})$ and the relative risk is not linear. The flexibility of this nonlinear function allows for this situation. We can see that the relative risk increases between 0 and 1, decreases for values of $\log_e(1 + \text{PSA})$ between 1 and 3, and then increases thereafter, more steeply between 3 and 6 and then moderately between 6 and 9, before rising again. This region of decreasing relative risk was unanticipated.

The log-likelihood ratio statistic from the proposed Cox model suggests that post-therapy PSA is associated with survival ($P < 0.01$, two-sided test). However, determination of a statistically significant association is not sufficient to justify its use as a surrogate for survival. It must be demonstrated, through some measure of association that quantifies the strength of association, that the PSA measurements explain a large amount of the variation in survival. Using the measure R^2 adapted to the Cox model, we found that 17% ($R^2 = 0.17$) of the variability in overall survival in this cohort is explained by the time-dependent PSA variable. This is a considerable amount, but sufficient variation in survival remains unaccounted for, even when using this flexible regression model.

DISCUSSION

Many authors have developed predictive models and/or examined time-dependent PSA values and their association with survival using the Cox model with conflicting results. Their conclusions have only been based upon evidence of a statistically significant association and not upon a measure of the strength of the covariate's association with a clinical outcome such as survival. Two main limitations of these approaches are as follows: (*a*) existence of a statistically significant association does not satisfy the conditions of surrogacy as some suggest; and (*b*) the Cox model used uses a linear log relative risk function that assumes a linear relationship between PSA and the log relative risk of death; this is a restrictive functional form, and one that is often not supported by the data. We derive a measure that quantifies the strength of the association between PSA and survival (R^2), and the general, nonlinear specification of the log relative risk function we propose allows for this measure to reflect the strength of any association noted while relying less upon the model being used.

In our example, we examined the relationship between PSA, modeled as a time-dependent covariate, and survival in 254 castrate metastatic (androgen-independent) cancer patients. We observed a region of decreasing relative risk ($1 < \log_e(1 + \text{PSA}) < 3$), which was unanticipated; a possible biologically plausible explanation for this observation is that there is a heterogeneous group of patients within this region that contains patients with poorly differentiated metastatic tumors with low PSA values and patients with low PSA values but without metastatic disease. The former group of patients does much worse clinically and therefore causes the risk of death, relative to the other patients, to decrease. The flexibility of this nonlinear function allowed for this observation. Had we used the conventional Cox model with a linear log relative risk function, this group of patients would not have been identified.

We found that despite a statistically significant association between PSA and survival ($P < 0.01$, two-sided test), only 17% of the variability in survival could be explained by the time-dependent PSA values; other factors, many of which are unknown, account for the remaining variation. Until these factors are discovered and incorporated, the results suggest that PSA is not adequate as a surrogate end point for survival in this clinical state, and using it (or any of the various definitions of a post-therapy change in PSA) as a surrogate in clinical trials could lead to problems in the design and the final interpretation of outcome of these clinical trials. Despite this finding, PSA is important as a predictor for survival. This methodology could be extended by incorporating other predictors (*e.g.*, age, pretherapy serum lactate dehydrogenase, pretherapy hemoglobin, and other factors collected serially), in addition to PSA, into the nonparametric generalized additive Cox model to determine how much variability is unexplained by such a model.

This study is limited by a relatively small number of patients who were treated heterogeneously and by the fact that we looked only at castrate metastatic (androgen-independent) prostate cancer patients. In addition, the estimation of the nonlinear log relative risk function is sensitive to the number of knots chosen for the B-spline function (see "Appendix"). In general, an increase in the number of knots used to estimate the nonlinear log relative risk function will increase the R^2 value, but at the cost of increasing the complexity of the model. The goal is to balance the level of association against model parsimony, which we believe we have done.

Currently, PSA remains the best clinical marker we have, but the value of PSA as a surrogate for survival, as suggested by the results of this study, is limited. Efforts must continue to be made to find other predictors of survival.

ACKNOWLEDGMENTS

We thank Dr. E. S. Venkatraman and Dr. Michael Morris for their contributions.

APPENDIX

The function *S*, which is termed a cubic B-spline function, is made up of a linear combination of cubic polynomials (21). In general, spline functions provide greater flexibility in estimating the relative risk function (22) because no assumptions are made as to the relationship between the risk of the event and the covariate being explored. Rather, the data determine the functional form of the relative risk function.

Involved in the specification of this model is choosing the number of knots that will be used to estimate the cubic B-splines (23). These knots are break points in the data where each of the B-splines will be estimated. Our goal is to produce *S* as a simple, smooth function of the time-dependent PSA covariate. For this data, four knots are used to specify this relative risk covariate function, and we allow the data to determine where the knots are placed (24).

The measure of association chosen for this model is the R^2 statistic of Kent and O'Quigley shown below (20).

$$R^2 = 1 - exp[(-2PLR/n].$$

This statistic is analogous to the squared multiple correlation coefficient of the normal linear regression model. *PLR* denotes the log partial likelihood ratio statistic for testing the hypothesis that there is no time-dependent PSA effect on survival and is a function of the ratio of the Cox log partial likelihood with the time-dependent PSA information $[S(Z_t)]$ in the model relative to the log partial likelihood without PSA.

REFERENCES

1. Scher, H. I., and Heller, G. Clinical states in prostate cancer: toward a dynamic model of disease progression. Urology, *55:* 323–327, 2000.

2. Kelly, W. K., Scher, H. I., Mazumdar, M., Vlamis, V., Schwartz, M., and Fossa, S. D. Prostate specific antigen as a measure of disease outcome in hormone refractory prostatic cancer. J. Clin. Oncol., *11:* 607–615, 1993.

3. Scher, H. I., Kelly, W. K., Zhang, Z. F., Ouyang, P., Sun, M., Schwartz, M., Ding, C., Wang, W., Horak, I. D., and Kremer, A. B. Post-therapy serum prostate specific antigen level and survival in patients with androgen-independent prostate cancer. J. Natl. Cancer Inst. (Bethesda), *91:* 244–251, 1999.

4. Scher, H. I., Mazumdar, M., and Kelly, W. K. Clinical trials in relapsed prostate cancer: defining the target. J. Natl. Cancer Inst. (Bethesda), *88:* 1623–1634, 1996.

5. Fleming, T. R., and DeMets, D. L. Surrogate end points in clinical trials: are we being misled? Ann. Intern. Med., *125:* 605–613, 1996.

6. Fazzari, M., Heller, G., and Scher, H. I. The Phase II/III transition: toward the proof of efficacy in cancer clinical trials. Control. Clin. Trials, *21:* 360–368, 2000.

7. Eisenberger, M. A., Blumenstein, B. A., Crawford, E. D., Miller, G., McLeod, D. G., Loehrer, P. J., Wilding, G., Sears, K., Culkin, D. J., Thompson, I. M., Bueschen, A. J., and Lowe, B. A. Bilateral orchiectomy with or without flutamide for metastatic prostate cancer. N. Engl. J. Med., *339:* 1036–1042, 1998.

8. Dowling, A., and Tannock, I. Systemic treatment for prostate cancer. Cancer Treat. Rev., *24:* 283–301, 1998.

9. Pruthi, R. S., Johnstone, I., Tu, I. P., and Stamey, T. A. Prostate-specific antigen doubling times in patients who have failed radical prostatectomy: correlation with histologic characteristics of the primary cancer. Urology, *49:* 737–742, 1997.

10. Leibman, B. D., Dillioglugil, O., Scardino, P. T., Abbas, F., Rogers, E., Wolfinger, R. D., and Kattan, M. W. Prostate-specific antigen doubling times are similar in patients with recurrence after radical prostatectomy or radiotherapy: a novel analysis. J. Clin. Oncol., *16:* 2267–2271, 1998.

11. Patel, A., Dorey, F., Franklin, J., and deKernion, J. B. Recurrence patterns after radical retropubic prostatectomy: clinical usefulness of prostate specific antigen doubling times and log slope prostate specific antigen. Urology, *158:* 1441–1445, 1997.

12. Vollmer, R. T., Dawson, N. A., and Vogelzang, N. J. The dynamics of prostate specific antigen in hormone refractory prostate carcinoma. Cancer (Phila.), *83:* 1989–1994, 1998.

13. Hudes, G., Einhorn, L., and Ross, E. Vinblastine *versus* vinblastine plus oral estramustine phosphate for patients with hormone-refractory prostate cancer: a Hoosier Oncology Group and Fox Chase Network Phase III trial. J. Clin. Oncol., *17:* 3160–3166, 1999.

14. Vollmer, R. T., Kantoff, P. W., Dawson, N. A., and Vogelzang, N. J. A prognostic score for hormone-refractory prostate cancer: analysis of two Cancer and Leukemia Group B studies. Clin. Cancer Res., *5:* 831–837, 1999.

15. Pound, C. R., Partin, A. W., Eisenberger, M. A., Chan, D. W., Pearson, J. D., and Walsh, P. C. Natural history of progression after PSA elevation following radical prostatectomy. J. Am. Med. Assoc., *281:* 1591–1597, 1999.

16. Korn, E. L., and Simon, R. Explained residual variation, explained risk, and goodness of fit. Am. Stat., *45:* 201–206, 1991.

17. Schemper, M. The explained variation in proportional hazards regression. Biometrika, *77:* 216–218, 1990.

18. Schemper, M. Further results on the explained variation in proportional hazards regression. Biometrika, *79:* 202–204, 1992.

19. Schemper, M., and Stare, J. Explained variation in survival analysis. Stat. Med., *15:* 1999–2012, 1996.

20. Kent, J. T., and O'Quigley, J. Measures of dependence for censored survival data. Biometrika, *75:* 525–534, 1988.

21. Heinzl, H., and Kaider, A. Gaining more flexibility in Cox proportional hazards regression models with cubic spline functions. Computer Methods Programs Biomed., *54:* 201–208, 1997.

22. Durrleman, S., and Simon, R. Flexible regression models with cubic splines. Stat. Med., *8:* 551–561, 1989.

23. Rosenberg, P. S. Hazard function estimation using B-splines. Biometrics, *51:* 874–887, 1995.

24. Sleeper, L. A., and Harrington, D. P. Regression splines in the Cox model with application to covariate effects in liver disease. J. Am. Stat. Assoc., *85:* 941–949, 1990.

Clinical Cancer Research

Quantifying the Amount of Variation in Survival Explained by Prostate-specific Antigen

David A. Verbel, Glenn Heller, William K. Kelly, et al.

Clin Cancer Res 2002;8:2576-2579.