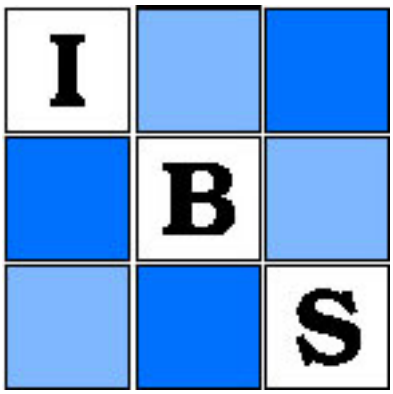

BIOMETRICS 46, 813–826
September 1990

Estimates of Absolute Cause-Specific Risk in Cohort Studies

J. Benichou

National Cancer Institute, Bethesda, Maryland 20892, U.S.A. and Département de Biostatistique et d'Informatique Médicale, Hôpital Saint-Louis, Paris, France

and

Mitchell H. Gail

National Cancer Institute, Bethesda, Maryland 20892, U.S.A.

SUMMARY

In this paper we study methods for estimating the absolute risk of an event c_1 in a time interval $[t_1, t_2)$, given that the individual is at risk at t_1 and given the presence of competing risks. We discuss some advantages of absolute risk for measuring the prognosis of an individual patient and some difficulties of interpretation for comparing two treatment groups. We also discuss the importance of the concept of absolute risk in evaluating public health measures to prevent disease. Variance calculations permit one to gauge the relative importance of random and systematic errors in estimating absolute risk. Efficiency calculations were also performed to determine how much precision is lost in estimating absolute risk with a nonparametric approach or with a flexible piecewise exponential model rather than a simple exponential model, and other calculations indicate the extent of bias that arises with the simple exponential model when that model is invalid. Such calculations suggest that the more flexible models will be useful in practice. Simulations confirm that asymptotic methods yield reliable variance estimates and confidence interval coverages in samples of practical size.

1. Introduction

A common outcome measurement in clinical trials of patients with surgically resected cancer is time to recurrence. Two treatments are often compared both in terms of their overall survival distributions and in terms of their time-to-recurrence distributions. Time-to-recurrence distributions are often computed by the procedure of Kaplan and Meier (1958), in which deaths from noncancer causes are treated as live withdrawals. Thus, time-to-recurrence distributions might be regarded as the "pure" distributions that would describe time to recurrence if no other risks were acting to kill the patient first. In deciding whether to give a toxic adjuvant therapy in the hope of preventing recurrence, it seems more appropriate to consider the absolute risk of recurrence, in the presence of competing risks, namely

$$\pi(t; \mathbf{x}) = \int_0^t h_1(u; \mathbf{x}) \exp\left[-\int_0^u \{h_1(v; \mathbf{x}) + h_2(v; \mathbf{x})\}\, dv\right] du, \tag{1.1}$$

where $h_1(u; \mathbf{x})$ is the cause-specific hazard of interest (e.g., cancer recurrence, c_1) for an individual with initial covariates $\mathbf{x}$, and $h_2(u; \mathbf{x})$ is the cause-specific hazard for other risks (e.g., all noncancer causes of death, c_2). One might be reluctant to try out a toxic cancer treatment in an elderly patient in whom $h_2 \gg h_1$ and for whom the absolute risk of recurrence in 5 years, say, $\pi(5; \mathbf{x})$, is small.

Key words: Absolute risk projection; Clinical trials; Cohort studies; Competing risks; Piecewise exponential model.

We shall consider the more general absolute risk of disease in $[t_1, t_2)$ given survival without recurrence to t_1, namely

$$\pi(t_1, t_2; \mathbf{x}) = \frac{\int_{t_1}^{t_2} h_1(u; \mathbf{x}) \exp[-\int_0^u \{h_1(v; \mathbf{x}) + h_2(v; \mathbf{x})\}\, dv]\, du}{\exp[-\int_0^{t_1} \{h_1(v; \mathbf{x}) + h_2(v; \mathbf{x})\}\, dv]}. \tag{1.2}$$

This quantity might be useful in managing a patient who has already survived a time t_1 by evaluating future absolute risk of recurrence in $[t_1, t_2)$.

The concept of absolute risk is not only useful in clinical settings but is also important in decisions affecting public health. For example, in order to estimate the absolute reduction in lung cancer incidence that might result from measures to reduce exposure to radon, one could categorize a general population into subgroups based on age, sex, smoking status, and current radon exposure levels and then estimate the absolute reduction in lung cancer incidence, in the presence of competing risks, that would result from lowering radon levels in each subgroup. This application would require the more general equation (1.2) because one would need to calculate future absolute risks for members of a subgroup who had already attained a given age.

The quantity $\pi(t_1, t_2; \mathbf{x})$ is estimable in principle without any competing risk assumptions, because, as Prentice et al. (1978) emphasize, all functions of the cause-specific hazards are estimable. Indeed, if all patients in a cohort of age t_1 are followed to death, $\pi(t_1, t_2; \mathbf{x})$ is just the expected proportion of them who are observed to have the event of interest before year t_2. Although we shall discuss these ideas with reference to cancer recurrence, they apply to any risk of interest, c_1. Chiang (1968) has used the term "crude" probability to describe $\pi(t; \mathbf{x})$, the probability of experiencing c_1 in the presence of competing risks, c_2.

To illustrate these ideas, we have computed estimates of the survival distribution, of the time-to-recurrence distribution, and of $1 - \pi(t)$ for lung cancer patients (see Gail et al., 1984) with resected T2N0 squamous cell disease (Figure 1). The two survival distributions are calculated as in Kaplan and Meier (1958), and the locus $1 - \hat{\pi}(t)$ is estimated nonparametrically as described in Section 2. This locus is slightly above the curve for time to recurrence and the two would tend to coincide only if there were no deaths from noncancer causes. If the hazard from noncancer causes predominates, $1 - \pi(t)$ will remain near 1.

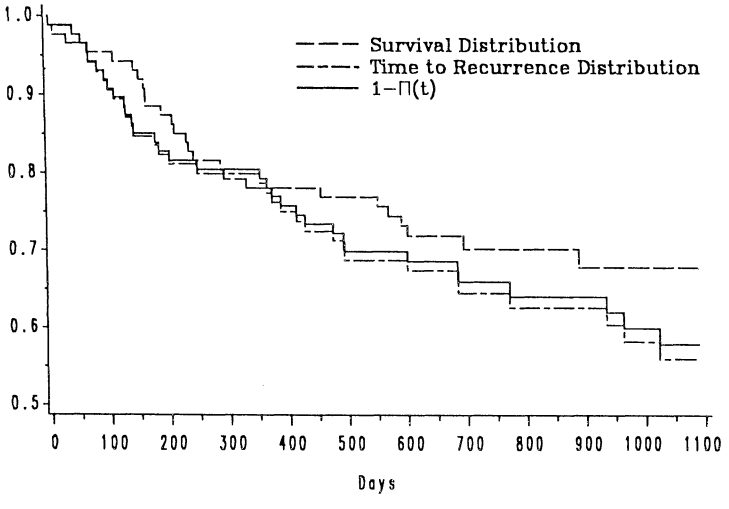

Figure 1. Overall survival, time to recurrence, and complement to 1 of the absolute risk of recurrence in patients with resected T2N0 squamous cell lung cancer (from data in Gail et al., 1984).

The methods we use may be applied in the presence of loss to follow-up for which we make the usual competing risk assumption that those lost to follow-up are selected randomly from those at risk at the time of loss. Thus, competing risk assumptions are needed to account for loss to follow-up but not for deaths from competing causes, c_2.

In Section 2 we treat the case of a single stratum x. Parametric and nonparametric procedures are compared with respect to precision, and with respect to bias when the assumed parametric model fails to hold. The adequacy of large-sample variance approximations for $\hat{\pi}(t)$ is evaluated for the exponential and piecewise exponential models and for the nonparametric approach. In Section 3 we extend these ideas to multiple covariates under the proportional hazards model and emphasize exponential and piecewise exponential distributions. Analogous methods are given for the semiparametric approach in Appendix A.3. An example in Section 4 leads to a discussion of the relative importance of model misspecification and stochastic error in estimating $\pi(t_1, t_2; \mathbf{x})$. Coverage probabilities for $\pi(t_1, t_2; \mathbf{x})$ based on asymptotic theory are shown to have near nominal levels in simulations.

2. Notation and Results for a Single Stratum $X = x$

In this section we assume a homogeneous population with cause-specific hazard (cf. Prentice et al., 1978)

$$h_1(t) = \lim_{\Delta \downarrow 0} \Delta^{-1} \Pr(c_1 \text{ in } [t, t + \Delta) \mid \text{neither } c_1 \text{ nor } c_2 \text{ in } [0, t)).$$

We define $h_2(t)$ similarly. We seek to estimate $\pi(t)$ in (1.1) or, more generally, $\pi(t_1, t_2)$ in (1.2), and we suppress the redundant notation $\mathbf{x}$. For convenience we define $\bar{G}(t) = \exp[-\int_0^t \{h_1(u) + h_2(u)\}\, du]$. Then

$$\pi(t_1, t_2) = \{\bar{G}(t_1)\}^{-1} \int_{t_1}^{t_2} h_1(u)\bar{G}(u)\, du. \tag{2.1}$$

We shall consider three models for h_1 and h_2 as follows:

(1) *Exponential model:* $h_1(t) = h_1$ and $h_2(t) = h_2$ are constant.

(2) *Piecewise exponential model:* $h_1(t) = h_{1i}$ and $h_2(t) = h_{2i}$ are constant on time interval $\mathcal{I}_i = [\tau_{i-1}, \tau_i)$ for $i = 1, 2, \ldots, I$.

(3) *Nonparametric model:* $h_1(t)$ and $h_2(t)$ are functions to be estimated nonparametrically.

For the exponential model, $\hat{h}_1 \equiv d_1/T$ and $\hat{h}_2 \equiv d_2/T$ are independent and have consistent variance estimates d_1/T^2 and d_2/T^2, where d_1 and d_2 are the numbers of events observed and T is the corresponding total time on test (person-years exposure). The estimate of π under the exponential model is

$$\hat{\pi}_e(t_1, t_2) = \hat{h}_1(\hat{h}_1 + \hat{h}_2)^{-1}[1 - \exp\{-(\hat{h}_1 + \hat{h}_2)(t_2 - t_1)\}], \tag{2.2}$$

with variance, obtained by the delta method (Rao, 1965, pp. 319–322), as explained in connection with Section A.1 of the Appendix.

For the piecewise exponential model, $\hat{h}_{1i}$ and $\hat{h}_{2i}$ are mutually independent and independent of $\hat{h}_{1j}, \hat{h}_{2j}$ for all $j \neq i$. Also, $\operatorname{var}(\hat{h}_{1i})$ is consistently estimated by d_{1i}/T_i^2, in an extension of the previous notation. For simplicity, we only allow t_1 and t_2 to take on values $\tau_0 = 0$, $\tau_1, \tau_2, \ldots, \tau_I$. The estimate of π obtained under the piecewise exponential model is then

$$\hat{\pi}_p(t_1, t_2) = \sum_{i=i_1}^{i_2} \hat{h}_{1i}(\hat{h}_{1i} + \hat{h}_{2i})^{-1}[1 - \exp\{-(\hat{h}_{1i} + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}]A(i), \tag{2.3}$$

where i_1 indexes the interval $[t_1, \tau_{i_1})$, i_2 indexes $[\tau_{i_2-1}, t_2)$, $A(i_1) = 1$, and

$$A(i) = \prod_{j=i_1}^{i-1} \exp\{-(\hat{h}_{1j} + \hat{h}_{2j})(\tau_j - \tau_{j-1})\} \quad \text{for } i > i_1.$$

The variance of $\hat{\pi}_p$ is obtained by the delta method as in Section A.2 of the Appendix.

A nonparametric estimate $\hat{\pi}_{np}$ of π under model 3 may be obtained as in Aalen (1978) by substituting $\hat{\bar{G}}(t_1-)$, the right-continuous Kaplan–Meier estimate of surviving both c_1 (recurrence) and c_2 (death from noncancer causes) to time t_1, into the denominator of (2.1) and by replacing the numerator by $\Sigma\hat{\bar{G}}(t-)R^{-1}(t)$, where $R(t)$ is a left-continuous process defining the number at risk just before t. The summation is over distinct times in $[t_1, t_2)$ at which events c_1 occur. We are assuming continuous survival data without ties. Exactly the same estimator is discussed by Aalen and Johansen (1978), Gray (1988), Matthews (1988), and Keiding and Andersen (1989). This estimate is used to calculate $1 - \hat{\pi}(0, t)$ in Figure 1 for one stratum of lung cancer patients, who are described in Section 4. Kay and Schumacher (1983) give a similar estimate, but they were more interested in determining cause-specific hazard rates, $h_1(t)$, than in estimating absolute risks, $\pi(t_1, t_2)$. We derived $\text{var}(\hat{\pi}_{np})$ and an estimator $\widehat{\text{var}}(\hat{\pi}_{np})$, which were needed for computations in Tables 1 and 2, from Theorem 2 of Aalen (1978). These derivations, which are available from the authors, are not presented because they may be obtained more directly from Theorem 4.3 of Aalen and Johansen (1978), as discussed by Keiding and Andersen (1989).

To see how much precision is lost by using more general models when a simple exponential model, e, is correct, we considered: (a) a piecewise exponential model, p2, with hazards constant on the two intervals $[0, t_1)$ and $[t_1, \infty)$, (b) a piecewise exponential model, p, with hazards constant on intervals of unit width, and (c) a nonparametric model, np. We studied the ratios $\text{var}(\hat{\pi}_{np})/\text{var}(\hat{\pi}_e)$, $\text{var}(\hat{\pi}_p)/\text{var}(\hat{\pi}_e)$, and $\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e)$ for various values of h_1, h_2, and t_2 with t_1 fixed at 1 (Table 1). For simplicity we assume no loss to follow-up. The exponential model is most efficient because it utilizes events outside the interval $[t_1, t_2)$, as well as inside the interval. The piecewise model p2 does not use events occurring in $[0, t_1)$ directly. The piecewise model, p, which approximates the usual actuarial

Table 1
$\text{var}(\hat{\pi}_{np})/\text{var}(\hat{\pi}_e)$, $\text{var}(\hat{\pi}_p)/\text{var}(\hat{\pi}_e)$, $\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e)$, *and* $\text{var}(\hat{\pi}_e)$ *for various values of* h_1, h_2, *and* t_2 *with* t_1 *fixed at* 1[a]

		$t_2 = 2$	$t_2 = 3$	$t_2 = 5$	$t_2 = 10$
$h_1 = .2 \log 2$	$\text{var}(\hat{\pi}_{np})/\text{var}(\hat{\pi}_e)$	4.086	2.855	2.393	2.299
$h_2 = \log 2$	$\text{var}(\hat{\pi}_p)/\text{var}(\hat{\pi}_e)$	4.068	2.850	2.392	2.299
	$\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e)$	2.297	2.297	2.297	2.297
	$\text{var}(\hat{\pi}_e) \times 10^3$	.4794	.9401	1.295	1.387
$h_1 = .2 \log 2$	$\text{var}(\hat{\pi}_{np})/\text{var}(\hat{\pi}_e)$	5.463	3.129	2.018	1.482
$h_2 = .2 \log 2$	$\text{var}(\hat{\pi}_p)/\text{var}(\hat{\pi}_e)$	5.447	3.121	2.014	1.481
	$\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e)$	1.320	1.320	1.320	1.320
	$\text{var}(\hat{\pi}_e) \times 10^3$	.2570	.7065	1.457	2.211
$h_1 = \log 2$	$\text{var}(\hat{\pi}_{np})/\text{var}(\hat{\pi}_e)$	4.230	3.144	2.610	2.305
$h_2 = .2 \log 2$	$\text{var}(\hat{\pi}_p)/\text{var}(\hat{\pi}_e)$	4.103	3.059	2.588	2.305
	$\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e)$	2.297	2.297	2.297	2.297
	$\text{var}(\hat{\pi}_e) \times 10^3$	1.353	1.602	1.390	1.387

[a] Computations are based on two competing exponentials with hazards h_1 and h_2 and no loss to follow-up. Expressions for $\text{var}(\hat{\pi}_e)$, $\text{var}(\hat{\pi}_{p2})$, $\text{var}(\hat{\pi}_p)$, and $\text{var}(\hat{\pi}_{np})$ are given in the Appendix.

Table 2
Coverages of the 95% confidence interval for $\pi(1, t_2)$ estimated from 1,000 simulations

		$t_2 = 2$	$t_2 = 3$	$t_2 = 5$	$t_2 = 10$
Exponential case[a]					
$h_1 = .2 \log 2$	e	.946	.951	.951	.953
$h_2 = \log 2$	p2	.950	.955	.956	.956
	p	.933*	.958	.954	.931*
	np	.942	.969*	.959	.960
$h_1 = .2 \log 2$	e	.963	.962	.961	.960
$h_2 = .4 \log 2$	p2	.955	.952	.954	.956
	p	.957	.958	.953	.958
	np	.959	.966*	.962	.969*
$h_1 = .2 \log 2$	e	.960	.960	.954	.957
$h_2 = .2 \log 2$	p2	.958	.959	.960	.963
	p	.948	.951	.952	.952
	np	.953	.956	.961	.955
$h_1 = .4 \log 2$	e	.961	.961	.960	.952
$h_2 = .2 \log 2$	p2	.953	.952	.959	.948
	p	.960	.951	.946	.946
	np	.957	.957	.957	.956
$h_1 = \log 2$	e	.951	.949	.944	.936*
$h_2 = .2 \log 2$	p2	.945	.939	.915*	.918*
	p	.944	.934*	.930*	.924*
	np	.951	.944	.926*	.924*
Weibull case[a]					
$\lambda = .2 \log 2$	e	.972*	.088*	.019*	.566*
$h_2 = .2 \log 2$	p2	.211*	.928*	.958	.937
	p	.956	.955	.939	.941
	np	.954	.962	.947	.945
$\lambda = .1 \log 2$	e	.718*	.979*	.210*	.556*
$h_2 = .2 \log 2$	p2	.010*	.431*	.974*	.955
	p	.958	.959	.953	.962
	np	.958	.960	.954	.964*

* Asterisks indicate the coverage falls outside the interval [.937, .963], which should happen only 5% of the time by chance.

[a] There was no loss to follow-up. Both hazards were constant as shown for the exponential case. For the Weibull case, the hazard for c_1 is $2\lambda t$, corresponding to shape parameter 2.

analysis with unit intervals, uses only events within $[t_1, t_2)$ directly, as does the nonparametric model. The ratio $\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e) = \exp\{(h_1 + h_2)t_1\}$, because the expected number of patients available for estimating hazards on $[t_1, t_2)$ is depleted by the survival fraction $\exp\{-(h_1 + h_2)t_1\}$ in the model p2. This explains the constancy of the ratios $\text{var}(\hat{\pi}_{p2})/\text{var}(\hat{\pi}_e)$ in Table 1. It is perhaps surprising that $\hat{\pi}_p$ and $\hat{\pi}_{np}$ are nearly equally inefficient, even for small numbers of intervals. As the numbers of intervals increase, one would expect these estimators to have equivalent efficiency, and for ten intervals, the variances of $\hat{\pi}_p$ and $\hat{\pi}_{np}$ agree to three significant figures. Relative to the exponential model, the models p and np are least efficient for small t_2, because the interval of useable events $[t_1, t_2)$ is short.

Miller (1983) noted that Kaplan–Meier estimates had poor precision, compared to parametric models, and Gail and Byar (1986) found similar results for averages of survival curves obtained by direct standardization. However, those authors studied point estimates of the survival distribution and found that the efficiency of the Kaplan–Meier procedure

diminished for large values of time, whereas our results (Table 1) show that the relative performances of Kaplan–Meier and the piecewise exponential model improve with increasing t_2.

Simulations were carried out to determine whether reliable confidence intervals for π could be obtained from the variance calculations in Sections A.1 and A.2 of the Appendix for the parametric models and from the nonparametric estimate $\widehat{\text{var}}(\hat{\pi}_{\text{np}})$ mentioned above. These confidence intervals were computed from $\log \hat{\pi} \pm 1.96\{\widehat{\text{var}}(\log \hat{\pi})\}^{1/2}$, which was then exponentiated to obtain a confidence interval for π. However, the upper limit 1.0 was used whenever the previous upper limit exceeded 1.0. The log transform led to more nearly nominal coverage than the intervals $\hat{\pi} \pm 1.96\{\widehat{\text{var}}(\hat{\pi})\}^{1/2}$, and only the former intervals are presented.

Simulations are reported both for the case of two independent exponential competing risks, for which all parametric models are valid, and for the case of competing Weibull (c_1) and exponential (c_2) risks (Table 2). Each simulation experiment represented 100 subjects and was repeated 1,000 times, and there was no loss to follow-up. For the exponential case we report 20 independent simulation studies corresponding to four values of t_2 and five sets of values of h_1 and h_2 (Table 2). For each experiment the four analyses from models e, p2, p, and np are correlated, as they are based on the same data. The piecewise exponential model p used intervals of width 1. The estimates of π (not shown) were in good agreement with the theoretical values for all four models. The absolute bias was always less than 6×10^{-3}. Furthermore, the estimates of $\text{var}(\hat{\pi})$ (not shown) were in good agreement with the empirical sample variances for all four models. The confidence intervals yielded reliable coverage (Table 2). The coverage was slightly less than nominal in only one case with the exponential model and two cases with the piecewise exponential model p2. With the piecewise exponential, p, the coverage was lower than nominal levels in five cases, although the absolute difference was always less than or equal to 2.6%. These results from asymptotic theory are good in view of the fact that the trials had only 100 individuals. Moreover, for the piecewise exponential model, p, some time intervals had very few expected events. The most extreme example for cause c_1 is the case $t_2 = 10$, $h_1 = .2 \log 2$, and $h_2 = \log 2$, for which the tenth interval has only .01 expected event from cause c_1 and .03 expected event from c_2. With the nonparametric approach, the coverage was greater than nominal in three cases and lower than nominal in two cases, and the largest absolute deviation from nominal coverage was 2.6%.

The Weibull case in Table 2 allows us to assess the effect of model misspecification on the coverage probabilities. We have chosen a rather extreme Weibull shape parameter $p = 2$ to test the robustness of parametric methods that assume h_1 is constant or piecewise constant. Both models e and p2 exhibit seriously biased estimates of π (data not shown) and consequent failure of coverage (Table 2). In contrast, both the piecewise exponential model, p, and nonparametric methods perform well.

Taken together, the results of Tables 1 and 2 indicate that substantial efficiencies are achieved by using parametric models when sufficient data are available to validate these models, but that actuarial (p) or nonparametric methods should be used when little is known about the underlying distributions.

3. Covariates Under the Proportional Hazards Model

Under the proportional hazards model of Cox (1972), $h_1(t; \mathbf{x}) = h_{01}(t)\exp(\boldsymbol{\beta}^T\mathbf{x})$. We assume $h_2(t)$ does not depend on $\mathbf{x}$. However, there is no essential difficulty in allowing $h_2(t; \mathbf{x}) = h_{02}(t)\exp(\boldsymbol{\gamma}^T\mathbf{x})$, provided $\boldsymbol{\gamma}$ is functionally independent of $\boldsymbol{\beta}$. In this formulation, $\mathbf{x}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are vectors.

If $h_{01}(t; \boldsymbol{\theta})$ and $h_{02}(t; \boldsymbol{\varepsilon})$ are specified up to a finite set of parameters, as for example in exponential, piecewise exponential, or Weibull models, then $\hat{\pi}$ is a function of the estimates $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\varepsilon}}$, and $\hat{\boldsymbol{\beta}}$ and its variance may be obtained by Taylor series expansion from estimates of the covariance of $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\beta}})$. Variance estimates for $\hat{\pi}(t_1, t_2; \mathbf{x})$ are given for the exponential and piecewise exponential models in Sections A.1 and A.2 of the Appendix. These results and confidence intervals computed from them will be applied to data from a lung cancer clinical trial in Section 4.

Semiparametric estimates of $\pi(t_1, t_2; \mathbf{x})$ are also available. Indeed, one such estimate is

$$\hat{\pi}(t_1, t_2; \mathbf{x}) = \{\hat{S}_{01}(t_1)^{\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}\hat{S}_2(t_1)\}^{-1} \int_{t_1}^{t_2} \hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}\hat{S}_2(t)\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})\, d\hat{\Lambda}_{01}(t), \tag{3.1}$$

where $\hat{S}_2$ is the Kaplan–Meier estimate for cause c_2, and where

$$d\hat{\Lambda}_{01}(t) = 1/\Sigma \exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) \quad \text{if an event } c_1 \text{ occurs at } t, \tag{3.2}$$

and $d\hat{\Lambda}_{01}(t) = 0$ elsewhere. The summation in (3.2) is over individuals at risk at t. Finally, $\hat{S}_{01}(t)$ is the product of terms $\{1 - d\hat{\Lambda}_{01}(t)\}$ over distinct c_1 event times less than or equal to t, and $\hat{\boldsymbol{\beta}}$ is the usual partial likelihood estimator for $\boldsymbol{\beta}$.

From the results in Tsiatis (1981) or in Andersen and Gill (1982), variance estimates for $\hat{\pi}$ in (3.1) can be obtained as outlined in Appendix A.3.

4. An Example

4.1 *Data Description and Computation of* $\hat{\pi}$

Gail et al. (1984) report recurrence and survival data from 392 patients with resected stage I non–small-cell carcinoma of the lung. Although one ordinarily thinks of lung cancer as having a high recurrence rate, there were some subsets of lung cancer patients for whom the absolute risk of recurrence was small and for whom toxic adjuvant chemotherapy might not be indicated. Gail et al. noted that the recurrence data seemed consistent with an exponential survival model with separate hazards for each of six strata defined by tumor stage and histology. Strata 1, 2, 3, 4, 5, and 6 correspond respectively to the findings T1N0 squamous, T1N1 squamous, T1N0 nonsquamous, T2N0 squamous, T2N0 nonsquamous, and T1N1 nonsquamous. We refit the recurrence data, using an exponential model with hazard $\exp(\mu + \sum_{i=2}^{6} \beta_i x_i)$, and found estimates $\hat{\mu} = -9.1541$, $\hat{\boldsymbol{\beta}}_2 = -.0720$, $\hat{\boldsymbol{\beta}}_3 = 1.2539$, $\hat{\boldsymbol{\beta}}_4 = 1.5723$, $\hat{\boldsymbol{\beta}}_5 = 1.8970$, and $\hat{\boldsymbol{\beta}}_6 = 2.1332$. Here x_2, x_3, x_4, x_5, and x_6 are dummy variables such that $x_i = 1$ for stratum i and $x_i = 0$ otherwise. The hazard is expressed in days^{-1}. The risk of interest, c_1, is recurrence, and deaths in patients with no previous evidence of recurrence ("noncancer" deaths) comprise the competing risks, c_2.

We computed the values of $\hat{\pi}$ and $\widehat{\text{var}}(\hat{\pi})$ as in Section 3, based on the assumption that $h_2(t)$ does not depend on x (Table 3). Three models were used for the hazards $h_{01}(t)$ and $h_2(t)$, namely: (e) a simple exponential model; (p2) a piecewise exponential model with intervals $\mathcal{I}_1 = [0, 1 \text{ year})$ and $\mathcal{I}_2 = [1 \text{ year}, 5 \text{ years})$; and (p) a piecewise exponential model with five yearly intervals. Both p and p2 allow for the hazard to change following the first year. Semiparametric models based on equation (3.1) were not studied because the data were nearly exponential (Gail et al., 1984) and because models p and np perform in nearly identical fashion (Tables 1 and 2). Also shown in Table 3 is $\hat{\pi}^*$, the estimate obtained under an unsaturated covariate model with hazard $\exp(\mu^* + \sum_{i=1}^{3} \beta_i^* x_i^*)$, where now $x_1^* = 1$ if T1N1, $x_2^* = 1$ if T2N0, and $x_3^* = 1$ if nonsquamous. Otherwise, x_1^*, x_2^*, and x_3^* are zero. This model is additive in the effects of TN status and histology.

Table 3

Estimates of $\hat{\pi}(1, t_2)$ *from the exponential model* (e), *the piecewise exponential model with unit intervals* (p), *and the piecewise exponential model with two intervals* (p2) *from data on lung cancer recurrence* (c_1) *and death from noncancer causes* (c_2) *in* 392 *patients*[a]

		$\hat{\pi}(1, 2)$			$\hat{\pi}(1, 3)$			$\hat{\pi}(1, 5)$		
		e	p2	p	e	p2	p	e	p2	p
Stratum 1	$\hat{\pi}$	.0373	.0341	.0389	.0721	.0664	.0602	.1348	.1257	.1682
n = 65	SD($\hat{\pi}$)	.0139	.0129	.0153	.0262	.0246	.0228	.0472	.0451	.0702
d_1 = 7	$\hat{\pi}$*	.0575	.0523	.0601	.1100	.1007	.0920	.2017	.1874	.2424
Stratum 2	$\hat{\pi}$	.0347	.0319	.0369	.0672	.0620	.0572	.1260	.1177	.1603
n = 20	SD($\hat{\pi}$)	.0241	.0223	.0260	.0459	.0427	.0396	.0830	.0784	.1105
d_1 = 2	$\hat{\pi}$*	.0823	.0744	.0867	.1554	.1418	.1317	.2781	.2581	.3332
Stratum 3	$\hat{\pi}$	.1246	.1131	.1283	.2302	.2111	.1934	.3952	.3700	.4685
n = 98	SD($\hat{\pi}$)	.0213	.0212	.0266	.0367	.0373	.0363	.0549	.0577	.1088
d_1 = 30	$\hat{\pi}$*	.1093	.0986	.1134	.2035	.1855	.1708	.3547	.3297	.4150
Stratum 4	$\hat{\pi}$	.1672	.1516	.1726	.3018	.2773	.2566	.4973	.4679	.5799
n = 87	SD($\hat{\pi}$)	.0265	.0269	.0347	.0435	.0451	.0445	.0592	.0641	.1167
d_1 = 33	$\hat{\pi}$*	.1354	.1221	.1412	.2486	.2269	.2108	.4224	.3940	.4912
Stratum 5	$\hat{\pi}$	.2236	.2022	.2306	.3911	.3598	.3363	.6108	.5782	.6950
n = 110	SD($\hat{\pi}$)	.0265	.0289	.0384	.0405	.0457	.0467	.0483	.0577	.1083
d_1 = 65	$\hat{\pi}$*	.2474	.2225	.2557	.4269	.3914	.3680	.6520	.6169	.7240
Stratum 6	$\hat{\pi}$	.2741	.2463	.2795	.4659	.4275	.4010	.6940	.6588	.7699
n = 12	SD($\hat{\pi}$)	.0948	.0892	.1022	.1355	.1332	.1291	.1413	.1498	.1595
d_1 = 6	$\hat{\pi}$*	.1545	.1389	.1615	.2808	.2558	.2396	.4684	.4370	.5416

[a] The estimate $\hat{\pi}$* is obtained under the unsaturated covariate model in Section 4, and n is the sample size and d_1 the number of recurrences (c_1) per stratum.

First we consider differences in estimates $\hat{\pi}$ for the three hazard models. In most cases, absolute differences in $\hat{\pi}$ among the exponential and piecewise exponential models are smaller than the standard deviations from these models; otherwise they are of the same magnitude. Thus if there are slight systematic deviations from exponentiality, their impact on estimates of π is no greater than the random uncertainty. In estimating $\pi(1, 2)$ and $\pi(1, 3)$ the differences among estimates are quite small. However, for $\pi(1, 5)$ absolute differences as large as .6950 − .5782 = .1168 were observed, but, again, these must be interpreted in view of the standard error .1083 for $\hat{\pi}_p(1, 5)$. The standard deviations indicate that one loses precision in exchange for the greater generality of the piecewise exponential model, especially for $\hat{\pi}(1, 5)$.

Misspecification of the form of the nuisance hazard is only one of several possible types of model misspecification. Another possible misspecification would occur if a saturated covariate model such as that used above for calculation of $\hat{\pi}$ was required, but instead a nonsaturated model like that used to calculate $\hat{\pi}$* was employed. In this case, $\hat{\pi}$* would usually be a biased estimate. Note that the quantities $\hat{\pi}$* in Table 3 often differ from corresponding values of $\hat{\pi}$ by more than the various estimates of $\hat{\pi}$ differ from each other. Nonetheless, the values of $\hat{\pi}$* are always within 2 standard deviations of the corresponding estimate $\hat{\pi}$. For this example, twice the log-likelihood ratio for comparing the saturated and unsaturated models was 8.4 or more, with 2 degrees of freedom ($P < .02$), which indicates lack of fit for the unsaturated model. In this example, random variation, measured by the standard deviation of $\hat{\pi}$, is comparable in magnitude to the systematic error that results from the use of an inappropriate unsaturated model for the effects of covariates. Other

types of misspecification, such as deviations from proportional hazards, should be investigated in particular cases to determine their possible importance compared to random error.

Regardless of the analysis model, it appears that those with T1N0 squamous disease (stratum 1) have modest absolute risk of recurrence beyond 1 year (Table 3). Unreported calculations show small absolute risk of recurrence from the date of resection and could be used to justify conservative management of such patients.

4.2 *Simulations to Study Properties of the Estimate $\hat{\pi}$ and of Related Confidence Intervals for the Lung Cancer Example*

To check the validity of the estimation procedures and the coverage of confidence intervals for π derived from the three nuisance hazard models used in Section 4.1, we carried out 1,000 simulations of the original clinical trial (Table 4) using the saturated exponential model in Section 4.1 to generate times to recurrence. Times to noncancer death, c_2, were generated from an exponential distribution with hazard $h_2 = \exp(-9.3705)$, independent of covariates. This hazard was estimated from the original data on noncancer deaths. In each simulated clinical trial, 392 recurrence times and times to noncancer deaths were generated according to these distributions, and with the same covariate distribution. Patients were assumed to enter the trial uniformly over 3 years, independently of the covariates, and to be followed for an additional 2-year period. This accrual pattern closely resembles that of the original trial. The estimates of π were in good agreement with the theoretical values for the three models, the absolute bias being always less than 10^{-2} (not shown). This result is not surprising because the simulated data truly follow exponential distributions. Furthermore, except for stratum 6 the estimates of $\text{var}(\hat{\pi})$ from the delta method were in good agreement with the sample variances (not shown), and the three models performed equally well in this respect. Likewise, the confidence intervals had nearly nominal coverage except in stratum 6 (Table 4). In the sixth stratum, which had only $n = 12$ patients, the confidence intervals for $\pi(1, 5)$ had coverages ranging from .881 and .888 for the three models. We repeated this simulation and found coverages ranging from .907 to .920.

Table 4

The coverage and average width of the 95% confidence interval for $\pi(t_1, t_2)$ estimated from 1,000 simulations[a]

		$\pi(1, 2)$			$\pi(1, 3)$			$\pi(1, 5)$		
		e	p2	p	e	p2	p	e	p2	p
Coverage										
Stratum	$E(d_1)$									
1	7.7	.956	.959	.965*	.966*	.960	.961	.953	.957	.957
2	2.2	.956	.952	.956	.950	.946	.950	.948	.954	.956
3	34.7	.959	.961	.949	.945	.937	.942	.953	.950	.958
4	39.1	.948	.954	.959	.955	.950	.954	.927*	.937	.948
5	61.3	.943	.948	.945	.956	.949	.950	.954	.949	.939
6	7.7	.944	.945	.952	.931*	.929*	.928*	.888*	.888*	.881*
Average width ($\times 10^2$)										
Stratum	$E(d_1)$									
1	7.7	5.5	5.6	6.1	10.4	10.6	10.7	18.7	19.0	19.9
2	2.2	20.5	20.6	20.9	30.2	30.4	30.5	46.8	47.1	48.1
3	34.7	7.9	8.4	10.0	13.5	14.4	15.1	20.1	21.4	24.9
4	39.1	9.7	10.5	12.5	15.8	17.1	17.9	21.3	22.9	27.0
5	61.3	9.9	11.2	14.0	15.1	17.1	18.2	17.9	20.3	25.1
6	7.7	34.7	35.6	38.1	47.6	48.6	49.0	47.1	47.4	48.4

[a] The simulations are described in Section 4.2, and the asterisk is defined in the footnote to Table 2. The term $E(d_1)$ is the expected number of recurrences, c_1, in each stratum. These strata contained 65, 20, 98, 87, 110, and 12 patients, respectively.

The loss of precision inherent in the more flexible models may be measured by the mean width of the 95% confidence intervals. All models yield extremely wide confidence intervals in stratum 6, which has only 12 subjects, and for $\hat{\pi}(1, 5)$ in stratum 2, which has only 2.2 events, c_1, on average. The estimate $\hat{\pi}_p$ leads to appreciably wider confidence intervals than $\hat{\pi}_e$ for estimates of $\pi(1, 2)$ and $\pi(1, 5)$ in strata 3, 4, and 5. The theoretical relative efficiencies, measured by $\text{var}(\hat{\pi}_e)/\text{var}(\hat{\pi}_p)$ and calculated from known parameters in the simulation experiment, are .51, .68, and .50, respectively, for $\pi(1, 2)$, $\pi(1, 3)$, and $\pi(1, 5)$ in stratum 5, and this pattern is repeated in all strata, though the theoretical relative efficiency is never less than .50. The estimates $\hat{\pi}_{p2}$ were never less efficient than .77. These theoretical variance ratios are in good agreement with the squared ratios of average widths of confidence intervals in Table 4.

5. Discussion

The absolute risk, π, studied in this paper is the "crude" probability of experiencing the event of interest, c_1, in the presence of competing risks, c_2 (Chiang, 1968; Prentice et al., 1978). As described in the Introduction, absolute risk may be more meaningful than the cause-specific survival curve $S_1(t) = \exp(-\int_0^t h_1(u)\, du)$ for evaluating some issues in clinical management and public health. Moreover, unlike $S_1(t)$, the pure survival curve, π has a valid interpretation as a probability even when competing risks are not independent.

The methods in this paper allow one to estimate $\pi(t_1, t_2; \mathbf{x})$ for an individual characterized by covariates $\mathbf{x}$ and thus to make individualized absolute risk projections. One might wish to combine individualized estimates to obtain a direct adjusted value for the entire population,

$$\hat{\pi}_{\text{adj}}(t_1, t_2) = \sum_x f(\mathbf{x})\hat{\pi}(t_1, t_2; \mathbf{x}),$$

where $f(\mathbf{x})$ is the known proportion of individuals with covariate levels $\mathbf{x}$ in a reference population. This procedure would be analogous to the methods for direct adjustment of survival curves described by Murphy and Haywood (1981), Makuch (1982), and Chang, Gelman, and Pagano (1982). Taylor series estimates of $\text{cov}(\hat{\pi}(t_1, t_2; \mathbf{x}), \hat{\pi}(t_1, t_2; \mathbf{x}'))$ could be used to estimate the variance of $\hat{\pi}_{\text{adj}}$, as Gail and Byar (1986) did for direct adjusted survival curves.

We do not recommend the use of $\pi(0, t)$ alone as a means of testing for treatment effect. For examples, if a cancer treatment increases h_2 but leaves h_1 unaffected, $\pi(0, t)$ will diminish in the treated group, yet overall survival is reduced and c_1-specific survival, $S_1(t)$, is unchanged. Instead, one should compare overall survival and estimates of the cause-specific survival curves S_1 and S_2 in the treated and untreated groups, as is common practice. Here $S_2(t) = \exp(-\int_0^t h_2(u)\, du)$. If h_2 is not affected by treatment, the change in $\pi(0, t)$ is a more realistic gauge of treatment benefit than a comparison of S_1 curves. If both h_1 and h_2 are affected by treatment, $\pi(0, t)$ still gives useful descriptive information for summarizing the burden of recurrence in each of the treatment groups. Gray (1988) gives a formal test for comparing $\pi(0, t)$ in two treatment groups and discusses the correct interpretation of such a test.

The simulations we performed showed that asymptotic methods work well even for reasonably small sample sizes and for piecewise exponential models with as many as ten different hazard rates. The delta method leads to reliable estimates of the variance of $\hat{\pi}$. Confidence intervals based on the log transformation had nearly nominal coverage, and this procedure was noticeably better than the confidence interval, $\hat{\pi} \pm 1.96\{\text{var}(\hat{\pi})\}^{1/2}$. Asymptotic nonparametric theory also leads to reliable confidence intervals in samples of modest size.

The calculations of $\widehat{\text{var}}(\hat{\pi})$ allow us to gauge the relative importance of random and systematic error in estimating π. The expectation of the absolute value of a normally

distributed error is about .798 times the standard deviation. Our example shows that systematic differences in estimating π that arise from using unsaturated, rather than saturated, models for covariate effects tend to be larger than the expected random error, whereas differences that result from using exponential rather than piecewise exponential models tend to be smaller than expected random errors.

Although we present estimates and variance calculations (Appendix A.3) for semiparametric models with covariates, further numerical work is needed to investigate the small-sample properties of these procedures. There are certain advantages to the piecewise exponential model, which we have explored more thoroughly. Often only actuarial data on the numbers of events and numbers at risk in fixed time intervals are available, and one is quite willing to assume piecewise constant hazards on these intervals. The methods in this paper for model p may be used by estimating the total time on test T_i as the product of the width of interval $\mathcal{I}_i$ times the number who enter the interval less half the number who fail from any cause or are censored. For large cohorts, even when ungrouped time data are available, the piecewise exponential model is often preferred for ease of computation. Furthermore, nonparametric methods and actuarial methods (p) usually agree very well, as in Tables 1 and 2.

ACKNOWLEDGEMENTS

We thank Professor V. Susarla for helpful comments and Mrs. Jennifer Donaldson for typing the manuscript and assistance with graphics. We also appreciate the insightful comments of the reviewers and editors, which greatly improved our understanding and presentation of nonparametric and semiparametric methods. Dr Benichou was supported by a grant from the Fondation Louis of the Fondation de France and by a grant from the Association pour la Recherche sur le Cancer (ARC).

RÉSUMÉ

Dans cet article, nous présentons des méthodes d'estimation du risque absolu de survenue d'un événement c_1 dans un intervalle de temps $[t_1, t_2)$, pour un sujet à risque au temps t_1 et en présence de risques compétitifs. Nous discutons les avantages du risque absolu pour déterminer le pronostic d'un sujet donné et ses difficultés d'interprétation dans la comparaison de deux groupes thérapeutiques. Nous discutons également de l'importance du concept de risque absolu dans l'évaluation de mesures de prévention en santé publique. Des calculs de variance permettent d'évaluer l'importance relative des erreurs aléatoire et systématique dans l'estimation du risque absolu. Des calculs d'efficacité sont aussi présentés afin de déterminer la perte de précision dans l'estimation du risque absolu avec une approche non paramétrique ou avec le modèle exponentiel par intervalle par rapport au modèle exponentiel simple. D'autres calculs indiquent le biais qui résulte de l'utilisation du modèle exponentiel simple quand celui-ci n'est pas valide. De tels calculs suggèrent que les modèles plus généraux sont utiles en pratique. Des simulations confirment que les méthodes asymptotiques conduisent à des estimateurs valides de la variance et à des intervalles de confiance ayant le recouvrement souhaité, ceci pour des échantillons de taille conforme à la pratique.

REFERENCES

Aalen, O. (1978). Nonparametric estimation of partial transition probabilities in multiple decrement models. *Annals of Statistics* **6,** 534–545.

Aalen, O. and Johansen, S. (1978). An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics* **5,** 141–150.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large-sample study. *Annals of Statistics* **4,** 1100–1120.

Breslow, N. and Crowley, J. (1974). A large-sample study of the life table and product limit estimates under random censorship. *Annals of Statistics* **2,** 437–453.

Chang, I., Gelman, R., and Pagano, M. (1982). Corrected group prognostic curves and summary statistics. *Journal of Chronic Diseases* **35,** 669–674.

Chiang, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics.* New York: Wiley.
Cox, D. R. (1972). Regression models and lifetables (with Discussion). *Journal of the Royal Statistical Society, Series B* **34,** 187–220.
Gail, M. H. and Byar, D. P. (1986). Variance calculations for direct adjusted survival curves, with applications to testing for no treatment effect. *Biometrical Journal* **28,** 587–599.
Gail, M. H., Eagan, R. T., Feld, R., Ginsberg, R., Goodell, B., Hill, L., Holmes, E. C., Lubeman, J. M., Mountain, C. F., Oldham, R. K., Pearson, F. G., Wright, P. W., Lake, W. H., and the Lung Cancer Study Group (1984). Prognostic factors in patients with resected stage I non-small-cell lung cancer. A report from the Lung Cancer Study Group. *Cancer* **54,** 1802–1813.
Gill, R. D. (1980). Censoring and stochastic integrals. Mathematical Centre Tracts 124. Amsterdam: Mathematisch Centrum.
Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics* **16,** 1141–1154.
Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53,** 457–481.
Kay, R. and Schumacher, M. (1983). Unbiased assessment of treatment effects on disease recurrence and survival in clinical trials. *Statistics in Medicine* **2,** 41–58.
Keiding, N. and Andersen, P. K. (1989). Nonparametric estimation of transition intensities and transition probabilities: A case study of a two-state Markov process. *Applied Statistics* **38,** 319–329.
Makuch, R. W. (1982). Adjusted survival curve estimation using covariates. *Journal of Chronic Diseases* **35,** 437–443.
Matthews, D. E. (1988). Likelihood-based confidence intervals for functions of many parameters. *Biometrika* **75,** 139–144.
Miller, R. G. (1983). What price Kaplan–Meier? *Biometrics* **39,** 1077–1081.
Murphy, V. K. and Haywood, L. J. (1981). Survival analysis by sex, age group and hemotype in sickle cell disease. *Journal of Chronic Diseases* **34,** 313–319.
Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics* **34,** 541–554.
Rao, C. R. (1965). *Linear Statistical Inference and Its Applications.* New York: Wiley.
Tsiatis, A. A. (1981). A large-sample study of Cox's regression model. *Annals of Statistics* **9,** 93–108.

Received June 1987; *revised September* 1988 *and July* 1989; *accepted September* 1989.

Appendix

Variance Calculations

A.1 *Variance of* $\hat{\pi}_e(t_1, t_2; \mathbf{x})$ *Under the Proportional Hazards Model*

First note that $\hat{\pi}_e(t_1, t_2; \mathbf{x})$ is given by (2.2) with $\hat{h}_1$ replaced by $\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})$. From the delta method,

$$\widehat{\text{var}}(\hat{\pi}_e(t_1, t_2; \mathbf{x})) = \left(\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{h}_{01}}\right)^2 \widehat{\text{var}}(\hat{h}_{01}) + 2\sum_{j=1}^{p}\left(\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{h}_{01}}\right)\left(\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{\beta}_j}\right)\widehat{\text{cov}}(\hat{h}_{01}, \hat{\beta}_j)$$
$$+ \sum_{j=1}^{p}\sum_{k=1}^{p}\left(\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{\beta}_j}\right)\left(\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{\beta}_k}\right)\widehat{\text{cov}}(\hat{\beta}_j, \hat{\beta}_k) + \left(\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{h}_2}\right)^2\widehat{\text{var}}(\hat{h}_2)$$

where

$$\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{h}_{01}} = \frac{\hat{h}_{01}(\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}))^2}{\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2}(t_2 - t_1)\exp\{-(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)(t_2 - t_1)\}$$
$$+ \frac{\hat{h}_2\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)^2}[1 - \exp\{-(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)(t_2 - t_1)\}],$$

$$\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{h}_2} = \frac{\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2}(t_2 - t_1)\exp\{-(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)(t_2 - t_1)\}$$
$$- \frac{\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)^2}[1 - \exp\{-(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)(t_2 - t_1)\}],$$

and

$$\frac{\partial\hat{\pi}_e(\mathbf{x})}{\partial\hat{\beta}_j} = \frac{\hat{h}_{01}^2(\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}))^2}{\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2}(t_2 - t_1)x_j\exp\{-(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)(t_2 - t_1)\},$$

$$+ \frac{\hat{h}_{01}\hat{h}_2\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)^2}x_j[1 - \exp\{-(\hat{h}_{01}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_2)(t_2 - t_1)\}].$$

Here $\mathbf{x}$ is the $p \times 1$ vector of covariates with jth component x_j, $\boldsymbol{\beta}$ is the $p \times 1$ vector of coefficients, variance and covariance terms for cause c_1 are obtained from the inverse of the observed information matrix, and $\widehat{\text{var}}(\hat{h}_2) = d_2/T^2$. For a single stratum (Section 2), the terms involving derivatives with respect to $\hat{\boldsymbol{\beta}}$ vanish and the remaining terms are the same except $\hat{h}_{01} = \hat{h}_1$ and $\exp(\hat{\boldsymbol{\beta}}\mathbf{x}) = 1$.

A.2 *Variance of* $\hat{\pi}_p(t_1, t_2; \mathbf{x})$ *Under the Proportional Hazards Model*

Similarly, $\hat{\pi}_p(t_1, t_2; \mathbf{x})$ is given by (2.3), with $\hat{h}_{1i}$ replaced by $\hat{h}_{01i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})$, which we denote by $\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})$. Then

$$\widehat{\text{var}}\{\hat{\pi}_p(t_1, t_2; \mathbf{x})\} = \sum_{i=i_1}^{i_2}\sum_{j=i_1}^{i_2}\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{h}_{1i}}\right)\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{h}_{1j}}\right)\widehat{\text{cov}}(\hat{h}_{1i}, \hat{h}_{1j}) + 2\sum_{i=i_1}^{i_2}\sum_{j=1}^{p}\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{h}_{1i}}\right)\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{\beta}_j}\right)\widehat{\text{cov}}(\hat{h}_{1i}, \hat{\beta}_j)$$

$$+ \sum_{j=1}^{p}\sum_{k=1}^{p}\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{\beta}_j}\right)\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{\beta}_k}\right)\widehat{\text{cov}}(\hat{\beta}_j, \hat{\beta}_k) + \sum_{i=i_1}^{i_2}\left(\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{h}_{2i}}\right)^2\widehat{\text{var}}(\hat{h}_{2i}),$$

where

$$\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{h}_{1i}} = \left[\frac{\hat{h}_{1i}(\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}))^2}{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i}}(\tau_i - \tau_{i-1})\exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}\right.$$

$$\left. + \frac{\hat{h}_{2i}(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})^2}[1 - \exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}]\right]A(i) - B(i)\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}),$$

$$\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{h}_{2i}} = \left[\frac{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i}}(\tau_i - \tau_{i-1})\exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}\right.$$

$$\left. - \frac{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})^2}[1 - \exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}]\right]A(i) - B(i),$$

and

$$\frac{\partial\hat{\pi}_p(\mathbf{x})}{\partial\hat{\beta}_j} = \sum_{i=i_1}^{i_2}\left[\frac{\hat{h}_{1i}^2(\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}))^2}{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i}}x_j(\tau_i - \tau_{i-1})\exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}\right.$$

$$\left. + \frac{\hat{h}_{1i}\hat{h}_{2i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})^2}x_j[1 - \exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}]\right]A(i) - \sum_{i=i_1}^{i_2}C(i).$$

In these expressions

$A(i_1) = 1,$

$$A(i) = \prod_{j=i_1}^{i-1}\exp\{-(\hat{h}_{1j}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2j})(\tau_j - \tau_{j-1})\} \quad \text{for } i > i_1,$$

$B(i_2) = 0,$

$$B(i) = (\tau_i - \tau_{i-1})\sum_{j=i+1}^{i_2}\left[\frac{\hat{h}_{1j}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x})}{\hat{h}_{1j}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2j}}[1 - \exp\{-(\hat{h}_{1j}\exp(\hat{\boldsymbol{\beta}}^T\mathbf{x}) + \hat{h}_{2j})(\tau_j - \tau_{j-1})\}]A(j)\right]$$

for $i < i_2$,

$C(i_1) = 0,$

and

$$C(i) = \frac{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}))^2}{\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}) + \hat{h}_{2i}} x_j[1 - \exp\{-(\hat{h}_{1i}\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}) + \hat{h}_{2i})(\tau_i - \tau_{i-1})\}]A(i) \sum_{j=i_1}^{i-1} \hat{h}_{1j}(\tau_j - \tau_{j-1})$$

for $i > i_1$.

Other terms are defined in Section A.1 and Section 2. For a single stratum (Section 2), the terms involving derivatives with respect to $\hat{\boldsymbol{\beta}}$ vanish and the remaining terms are the same except $\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x}) = 1$.

A.3 *Semiparametric Estimate of* $\mathrm{var}\{\hat{\pi}(t_1, t_2; \mathbf{x})\}$

First consider the numerator W of $\hat{\pi}(t_1, t_2; \mathbf{x})$ *in* (3.1), and write

$$\begin{aligned} W &\equiv \int_{t_1}^{t_2} \hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\hat{S}_2(t)\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})\, d\hat{\Lambda}_{01}(t) \\ &= -\int_{t_1}^{t_2} \hat{S}_2(t)d\{\hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\} \\ &= -\left[\int_{t_1}^{t_2} \{\hat{S}_2(t) - S_2(t)\}d\{\hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})} - S_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\}\right. \\ &\qquad \left. + \int_{t_1}^{t_2} S_2(t)d\{\hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})} - S_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\} + \int_{t_1}^{t_2} \hat{S}_2(t)d\{S_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\}\right] \\ &\equiv -(\mathrm{I} + \mathrm{II} + \mathrm{III}). \end{aligned}$$

Because the product of the first term and the square root of the cohort sample size converges in probability to zero, one obtains

$$\mathrm{var}(W) = \mathrm{var}(\mathrm{II}) + \mathrm{var}(\mathrm{III}) + 2\,\mathrm{cov}(\mathrm{II}, \mathrm{III}) = \mathrm{var}(\mathrm{II}) + \mathrm{var}(\mathrm{III}),$$

since the independence of the processes $\hat{S}_2(t)$ and $\hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}$ implies $\mathrm{cov}(\mathrm{II}, \mathrm{III}) = 0$. Further,

$$\mathrm{var}(\mathrm{III}) = \int\int \mathrm{cov}\{\hat{S}_2(u), \hat{S}_2(v)\}d\{S_{10}(u)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\}d\{S_{10}(v)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\}$$

can be evaluated from the known covariance of the process $\hat{S}_2(u)$, which is given in Breslow and Crowley (1974) and in Gill (1980, §4.2). The range of integration for this integral is the set $\{(u, v): t_1 \leq u \leq t_2, t_1 \leq v \leq t_2\}$. Furthermore, one obtains

$$\begin{aligned} \mathrm{var}(\mathrm{II}) &= \mathrm{var}\left[\int_{t_1}^{t_2} S_2(t)d\{\hat{S}_{01}(t)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}\}\right] \\ &= \mathrm{var}\left[S_2(t_2)\exp\{-\hat{\Lambda}_{01}(t_2)\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})\} - S_2(t_1)\exp\{-\hat{\Lambda}_{01}(t_1)\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})\}\right. \\ &\qquad \left. - \int_{t_1}^{t_2} \exp\{-\hat{\Lambda}_{01}(t)\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})\}dS_2(t)\right] \end{aligned}$$

by integrating by parts and by replacing $\hat{S}_{01}(t)$ by $\exp\{-\hat{\Lambda}_{01}(t)\}$. These variance terms can be computed using the joint distribution of $\{\hat{\Lambda}_{01}(t), \hat{\boldsymbol{\beta}}\}$ given in Tsiatis (1981) or in Andersen and Gill (1982) and the delta method.

Similarly, one can compute the variance of the denominator $U \equiv \hat{S}_{01}(t_1)^{\exp(\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{x})}S_{02}(t_1)$ and the covariance between U and W. The variance of $\hat{\pi}(t_1, t_2; \mathbf{x}) = W/U$ follows by another application of the delta method.