5-31-2005

Attributable Risk Function in the Proportional Hazards Model

Ying Qing Chen
Fred Hutchinson Cancer Research Center, yqchen@u.washington.edu

Chengcheng Hu
Harvard University, cchu@hsph.harvard.edu

Yan Wang
University of California, Berkeley, yanw@stat.berkeley.edu

UW Biostatistics Working Paper Series

Year 2005 *Paper* 254

Attributable Risk Function in the Proportional Hazards Model

Ying Qing Chen* Chengcheng Hu†
Yan Wang‡

*Fred Hutchinson Cancer Research Center
†Harvard University
‡University of California, Berkeley

Attributable Risk Function in the Proportional Hazards Model

Abstract

As an epidemiological parameter, the population attributable fraction is an important measure to quantify the public health attributable risk of an exposure to morbidity and mortality. In this article, we extend this parameter to the attributable fraction function in survival analysis of time-to-event outcomes, and further establish its estimation and inference procedures based on the widely used proportional hazards models. Numerical examples and simulations studies are presented to validate and demonstrate the proposed methods.

1 Introduction

Time-to-event outcomes have been collected in many epidemiological cohort studies as end-points, e.g., in risk assessment of hazardous exposures on disease incidences. The association between the time-to-event outcomes and the exposures is usually measured by relative risk. It is often quantified by the incidence ratio of events between those exposed and unexposed. For instance, in the widely used proportional hazards model (Cox, 1972),

$$\lambda(t \mid Z) = \lambda_0(t) \exp(\beta^{\mathrm{T}} Z), \tag{1}$$

the regression parameter $\beta \in \mathcal{B}$ is the hazards ratio in log-scale and thus quantifies the relative risk, when the covariate $Z \in \mathcal{Z}$ is the exposure indicator. Here, $\lambda(\cdot \mid Z)$ is the hazard function for Z and $\lambda_0(\cdot)$ is the baseline hazard function, respectively.

As pointed in Greenland (2001), however, there is also substantial public health interest in the disease risk attributable to the exposure for a given population. The quantity to measure such attribution is often referred as the population attributable fraction. Let D be the binary disease indicator. When Z takes value of 0 or 1, the attributable fraction (Levin, 1953) is usually defined as

$$\varphi = \frac{\mathrm{pr}\{D = 1\} - \mathrm{pr}\{D = 1 \mid Z = 0\}}{\mathrm{pr}\{D = 1\}}. \tag{2}$$

As noted in Greenland & Robins (1988), the attributable fraction φ may differ in the person-time analysis and the proportion analysis, unless the disease is rare. However, much of the previous methodology development involving the attributable faction has been focused on its estimation and inference when the disease is binary under the logistic regression models.

In this article, we consider an extended attributable fraction function of $\varphi(t)$ in the situation when the outcome of interest is continuous time-to-event and subject to potential censoring. A simple estimator is proposed and studied with the proportional hazards model. Simulation studies are conducted to evaluate its validity and performance. The proposed methods are applied to a dataset of multicenter AIDS cohort study for demonstration.

2 Attributable Fraction Function

Let T be the nonnegative random variable of the time-to-event. A straightforward extension of φ for T is, for some $t > 0$,

$$\widetilde{\varphi}(t) = \frac{\mathrm{pr}\{T \le t\} - \mathrm{pr}\{T \le t \mid Z = 0\}}{\mathrm{pr}\{T \le t\}} = \frac{F(t) - F(t \mid Z = 0)}{F(t)},$$

where $F(\cdot)$ is the cumulative distribution function of T. That is, the attributable fraction of disease risk due to an exposure can be time-dependent. When t is the end of follow-up period for a cohort study, τ, say, $\widetilde{\varphi}(\tau)$ is the attributable fraction in (2). For rare diseases, when $F(\cdot)$ are usually approximated by their respective cumulative hazard functions of $\Lambda(\cdot)$, $\widetilde{\varphi}(t)$ can be also expressed in $\{\Lambda(t) - \Lambda(t \mid Z = 0)\}/\Lambda(t)$. Within an infinitesimal neighbourhood of t, a reasonable measure of the attributable fraction function for T is thus

$$\varphi(t) = \frac{\lambda(t) - \lambda(t \mid Z = 0)}{\lambda(t)}. \tag{3}$$

An equivalent measure of $\varphi(t)$ is $\bar{\varphi}(t) = t^{-1} \int_0^t \varphi(u)du$. It is called the average attributable fraction function. In particular, $\bar{\varphi} = \bar{\varphi}(\tau)$ can be a useful summary measure of $\varphi(\cdot)$. When the reference population is a subset, $\mathcal{Z}_0 \subset \mathcal{Z}$, a general form of $\varphi(t)$ is

$$\varphi(t \mid \mathcal{Z}_0) = \frac{\lambda(t) - \lambda(t \mid \mathcal{Z}_0)}{\lambda(t)}.$$

As shown in the rest of this article, however, the estimation and inference procedures established for $\varphi(t)$ can be mostly generalised to $\varphi(t \mid \mathcal{Z}_0)$ as well. We would thus focus on $\varphi(t)$ for simpler demonstration.

The range of $\varphi(\cdot)$ is $(-\infty, 1]$. Under the proportional hazards model (1), $\varphi(t) \ge 0$ for all $t \ge 0$ if and only if $\beta \ge 0$. In addition, since $\lambda(t \mid Z = 0) = \lambda_0(t)$ in (1), $\varphi(t) = \{\lambda(t) - \lambda_0(t)\}/\lambda(t) = 1 - \lambda_0(t)/\lambda(t)$. Note that here $\lambda(t)$ is the hazard function of the marginal distribution $F(t)$, by ignoring the heterogeneity among the subjects in the given population. It usually does not equal $\int_z \lambda(t \mid z)dF_z(z)$, where $F_z(\cdot)$ is the distribution

function of Z. However, by the Bayes Theorem, we have

$$\begin{aligned}\lambda_0(t)\bar{F}(t) &= \int_{\mathcal{Z}} \bar{F}(t \mid z)\lambda_0(t)dF_Z(z) = \int_{\mathcal{Z}} \bar{F}(t \mid z)\lambda(t \mid z)\exp(-\beta^{\mathrm{T}}z)dF_Z(z) \\ &= \int_{\mathcal{Z}} f(t \mid z)\exp(-\beta^{\mathrm{T}}z)dF_Z(z) = \int_{\mathcal{Z}} \frac{f_{Z \mid T}(z \mid t)f(t)}{f_Z(z)}\exp(-\beta^{\mathrm{T}}z)dF_Z(z) \\ &= f(t)\int_{\mathcal{Z}} \exp(-\beta^{\mathrm{T}}z)dF_{Z \mid T}(z \mid t),\end{aligned}$$

where $\bar{F} = 1 - F$, $f = F'$, and $F_{Z \mid T}(z \mid t)$ is the conditional distribution function of Z given $T = t$, respectively. As a result,

$$\varphi(t) = 1 - \int_{\mathcal{Z}} \exp(-\beta^{\mathrm{T}}z)dF_{Z \mid T}(z \mid t). \tag{4}$$

3 Estimation and Inferences

Suppose that there are n subjects assembled in an epidemiological cohort study. The observed outcomes consist of n iid copies of (X_i, Δ_i, Z_i), $i = 1, 2, \ldots, n$, where $X_i = \min(T_i, C_i)$ and $\Delta_i = I(T_i \leq C_i)$, respectively. Here C_i is the potential censoring time. Denote the at-risk indicator $Y_i(t) = I(X_i \geq t)$. Let $s(t) = \lim_n S(t) = \lim_n n^{-1}\sum_j Y_j(t)$ and $s_k(t;\beta) = \lim_n S_k(t;\beta) = \lim_n n^{-1}\sum_j Y_j(t)\exp(\beta^{\mathrm{T}}Z_j)Z_j^{\otimes k}$, $k = 0, 1, 2$, respectively. Assume that β_* is the true value of β in the semiparametric proportional hazards model (1), in which $\lambda_0(\cdot)$ is unspecified. The maximum partial likelihood estimator, $\widehat{\beta}$, can be then obtained by solving the partial score equations

$$\sum_{i=1}^{n}\int_0^{\tau}\left\{Z_i - \bar{Z}(t;\beta)\right\}dN_i(t) = 0,$$

where $\bar{Z}(t;\beta) = S_1(t;\beta)/S_0(t;\beta)$ and $N_i(t) = I(X_i \leq t, \Delta_i = 1)$, respectively. Let $\Sigma(\beta_*) = n^{-1}\sum_i \int_0^{\tau}\{Z_i - \bar{Z}(t;\beta_*)\}^{\otimes 2}Y_i(u)\exp(\beta_*^{\mathrm{T}}Z_i)\lambda_0(u)du$. Standard martingale theory of counting processes in Andersen & Gill (1982) shows that $\widehat{\beta}$ is consistent and $n^{1/2}(\widehat{\beta} - \beta_*)$ is asymptotically equivalent to

$$\Sigma^{-1}(\beta_*)\cdot n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\left\{Z_i - \bar{Z}(t;\beta_*)\right\}dM_i(t), \tag{5}$$

where $\{M_i(t) = N_i(t) - \int_0^t Y_i(u)\exp(\beta_*^{\mathrm{T}} Z_i)\lambda_0(u)du; i = 1, 2, \ldots, n\}$ are martingales with respect to the filtration of $\mathcal{F}_t = \sigma\{N_i(u), Y_i(u), Z_i, u \le t; i = 1, 2, \ldots, n\}$.

Denote $p_i(t;\beta) = Y_i(t)\exp(\beta^{\mathrm{T}} Z_i)/S_0(t;\beta)$. As derived in Xu & O'Quigley (2000), when C is independent of T and Z, p_i's are the conditional probabilities of the subjects observed to fail at t given that one of the at-risk subjects would fail at the same time. Therefore, the conditional distribution function of Z given $T = t$, $F_{Z\mid T}(z \mid t)$, can be consistently estimated by

$$\widehat{F}_{Z\mid T}(z \mid t) = \sum_{i=1}^{n} I(Z_i \le z)p_i(t;\widehat{\beta}).$$

A simple estimator of the attributable fraction function in (4) is thus

$$\widehat{\varphi}(t;\widehat{\beta}) = 1 - \int_{\mathcal{Z}} \exp(-\widehat{\beta}^{\mathrm{T}} z)d\widehat{F}_{Z\mid T}(z \mid t) = 1 - \frac{S(t)}{S_0(t;\widehat{\beta})}.$$

The following theorem gives the asymptotic properties of this estimator:

Theorem 1. Under the regularity conditions 1-4 specified in the appendix, $\widehat{\varphi}(t;\widehat{\beta})$ is uniformly consistent for $\varphi(t)$ for $t \in [0, \tau]$, i.e.,

$$\sup_{t\in[0,\tau]} \left|\widehat{\varphi}(t;\widehat{\beta}) - \varphi(t)\right| \to_p 0.$$

Furthermore, $n^{1/2}\{\widehat{\varphi}(t;\widehat{\beta}) - \varphi(t)\}$ converges weakly to a zero-mean Gaussian process. Its covariance function of $\sigma_\varphi(s,t)$, $s, t \in (0,\tau)$, can be consistently estimated by $\widehat{\sigma}_\varphi(s,t) = n^{-1}\sum_i \widehat{v}_i(s)\widehat{v}_i(t)$, where $\widehat{v}_i(t)$ is

$$\frac{S(t)\exp(\widehat{\beta}^{\mathrm{T}} Z_i)Y_i(t)}{S_0(t,\widehat{\beta})^2} - \frac{Y_i(t)}{S_0(t,\widehat{\beta})} + \frac{S(t)S_1(t,\widehat{\beta})^{\mathrm{T}}\widehat{\Sigma}^{-1}(\widehat{\beta})}{S_0(t,\widehat{\beta})^2}\int_0^\tau \left\{Z_i - \bar{Z}(u,\widehat{\beta})\right\} d\widehat{M}_i(u),$$

and $\widehat{M}_i(t) = N_i(t) - \int_0^t Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}} Z_i)d\widehat{\Lambda}(t)$, respectively.

As a result, the variance of $n^{1/2}\{\widehat{\varphi}(t,\widehat{\beta}) - \varphi(t)\}$ is approximately $\widehat{\sigma}^2_\varphi(t;\widehat{\beta}) = n^{-1}\sum_{i=1}^n \widehat{v}_i(t)^2$, and the pointwise $100(1-\alpha)\%$ confidence intervals for $\widehat{\varphi}(t)$ can be constructed as $\widehat{\varphi}(t) \mp z_{1-\alpha/2}n^{-1/2}\widehat{\sigma}_\varphi(t;\widehat{\beta})$, where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution.

In addition to the pointwise confidence intervals, it is as well of practical interest to consider simultaneous $100(1-\alpha)$th percentile confidence bands, $\varphi_l(\cdot)$ and $\varphi_u(\cdot)$, say, such that

$$\mathrm{pr}\{\varphi_l(t) \le \varphi(t) \le \varphi_u(t), 0 \le t \le \tau\} = 1-\alpha.$$

Due to the fact that there is no independent increment structure in the limiting process of $n^{1/2}\{\widehat{\varphi}(\cdot;\widehat{\beta}) - \varphi(\cdot)\}$, it is not straightforward to be transformed into the standard Brownian bridge in direct confidence bands calculation. To find appropriate confidence bands, however, the simulation approach in Lin, Fleming & Wei (1994) can be adapted for relatively easy implementation. Specifically, consider n iid standard normal deviates $\{\varepsilon_i, i = 1, 2, \ldots, n\}$ in

$$\widehat{\psi}(t;\widehat{\beta}) = \frac{S(t)}{S_0(\beta_*,t)^2}\left\{\frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)\varepsilon_i\right\} - \frac{1}{S_0(\widehat{\beta},t)}\left\{\frac{1}{n}\sum_{i=1}^{n} Y_i(t)\varepsilon_i\right\} + \frac{S(t)S_1(\widehat{\beta},t)^{\mathrm{T}}\widehat{\Sigma}^{-1}(\widehat{\beta})}{S_0(\widehat{\beta},t)^2}\left[\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}\left\{Z_i - \bar{Z}(t;\widehat{\beta})\right\}\varepsilon_i dN_i(t)\right].$$

For any set of finite number of time points $(t_1, t_2, \ldots, t_m)$, $0 \le t_1, \ldots, t_m \le \tau$, the conditional limiting distribution of $(\widehat{\psi}(t_1;\widehat{\beta}), \widehat{\psi}(t_2;\widehat{\beta}), \ldots, \widehat{\psi}(t_m;\widehat{\beta}))^{\mathrm{T}}$ given the observed $\{(X_i, \Delta_i, Z_i)\}$ is the same as the unconditional distribution of $(\psi(t_1;\beta_*), \psi(t_2;\beta_*), \ldots, \psi(t_m;\beta_*))^{\mathrm{T}}$. As a result, $n^{1/2}\widehat{\psi}(\cdot;\widehat{\beta})$ and $n^{1/2}\{\widehat{\varphi}(\cdot;\widehat{\beta}) - \varphi(\cdot)\}$ have the same limiting distribution by the tightness of $\widehat{\psi}(t;\widehat{\beta})$. Therefore, $100(1-\alpha)$th percentile simultaneous confidence bands can be constructed as $\widehat{\varphi}(t;\widehat{\beta}) \mp \widetilde{z}_{1-\alpha/2} n^{-1/2}\widehat{\sigma}_{\varphi}(t;\widehat{\beta})$, where $\widetilde{z}_{1-\alpha/2}$ is computed such that

$$\mathrm{pr}\left\{\sup_{t\in[0,\tau]} \frac{n^{1/2}|\widehat{\psi}(t;\widehat{\beta})|}{\widehat{\sigma}_{\varphi}(t;\widehat{\beta})} \le \widetilde{z}_{1-\alpha/2}\right\} \approx 1-\alpha.$$

4 Numerical Studies and Examples

To gain some concrete sense of the proposed attributable fraction functions of $\widetilde{\varphi}(\cdot)$ and $\varphi(\cdot)$ in §2, we assume that the proportional hazards model (1) holds for the exponential baseline hazard functions of 1 and 0.01, respectively. Let $\beta = \log 2$ for the exposed $Z = 1$ against the unexposed $Z = 0$. Three proportions of exposure are considered: 25%, 50% and 75%,

respectively. As shown in Figure (1), the attributable fraction function defined by either $\widetilde{\varphi}(\cdot)$ or $\varphi(\cdot)$ is not necessarily constant over time, even when the baseline hazard function themselves are constant. When the baseline hazard function is larger, the attribution fraction functions change more rapidly over time. They change less when otherwise. That is, when the disease is less (more) frequent among the unexposed subjects, the disease risk attributable to the exposure tend to change less (more) rapidly over time. In addition, by comparing $\varphi(\cdot)$ with $\widetilde{\varphi}(\cdot)$, we find that $\varphi(\cdot)$ better approximates $\widetilde{\varphi}(\cdot)$ for the less frequent disease and the smaller proportion of exposure.

[Figure (1) about here]

Simulations are conducted to evaluate the validity and performance of the proposed estimator of $\varphi(\cdot)$ in §3. In addition to assuming that the baseline hazard functions are constant of 0.01 and 1.00, respectively, time-to-events are generated according to the proportional hazards model (1) with $\beta = 0$ and $\log 2$, respectively. Sample sizes are selected to be 200 and 500, respectively. Each subject's binary exposure indicator is generated according to the Bernoulli trial with the exposure probability of 25% and 50%, respectively. Censoring times are generated to yield about 30% and 10% of censored observations. The estimators and their associated variances are calculated at the 75%-tile and median of the marginal survival distribution, t_1 and t_2, respectively. Simulation results are listed in Table (1). For each entry in the table, 1000 simulated data sets are generated to calculate the bias and 95% nominal coverage probability. Here the bias is the difference between the average of the 1000 estimates and the true attributable fraction, and the 95% nominal coverage probability is the percentage of 1000 95% confidence intervals containing the true attributable fraction. As shown in the table, the proposed estimators are virtually unbiased and their confidence intervals maintain the desired coverage probabilities. In addition, the sample standard errors of each 1000 $\widehat{\varphi}(t)$ and the average of 1000 $\widehat{\sigma}_{\varphi}(t)$ are calculated, respectively. It is shown that they are close to each other, which suggest the accuracy of variance calculation in Theorem 1. Given the plethora of statistical literature on simulated confidence bands approaches, we

skip their simulation and demonstration in this article.

[Table (1) about here]

We demonstrate the developed methods with a publically released dataset of the Multicenter AIDS Cohort Study. This is an ongoing prospective study of the natural and treatment histories of HIV-1 infection in homosexual and bisexual men in four U.S. cities of Baltimore, Chicago, Pittsburgh and Los Angeles since 1984 (Kaslow, et al., 1987). For the demonstration purpose, we use a subset of 3341 subjects of the original cohort who were HIV-1 infection free at the initial enrollment. Among these subjects, a total of 508 seroconversions are reported in the dataset through 1999. Two risk factors are considered for the attributable risk calculation for this cohort: ever having sex with AIDS partner ($Z_1 = 1$) and ever having anal receptive sex ($Z_2 = 1$). Some summary statistics and the estimates of the hazards ratio are listed in Table (2). For the risk factor Z_1, the HIV infection incidence rate ratio is 17.8%/11.5%=1.54, which is consistent with the proportional hazards model estimate of $\widehat{\beta}$ of $\exp(0.470) = 1.599$. That means, the relative risk for HIV infection is about 1.5-1.6 times for those ever having sex with AIDS partners of those without. The finding is similar for the risk factor Z_2. As far as the attributable risks are concerned, the risk factor Z_1 attributes about 24.1% and Z_2 attributes about 18.6% to the overall risk, respectively. If either activity is involved, then it attributes about 37.2% to the overall risk. Their estimated attributable fraction functions are also plotted in Figure 2, respectively. For this particular dataset, the model-based $\widehat{\varphi}(\cdot)$ for the times-to-HIV infection tend to be uniformly larger than that of the crude $\widehat{\varphi}$ for the binary HIV infection outcomes. It is also interesting to see that the attributable fraction functions of $\varphi(\cdot)$ are not much varying over time, although they appear monotonically decreasing. In the future work, rigorous statistical procedures need to be developed for testing the constancy.

[Table (2) and Figure (2) about here]

5 Discussion

When the outcomes of interest are binary, i.e., the actual times of their occurrence are ignored, the logistic regression model, such as

$$\mathrm{pr}\{D = 1 \mid Z\} = \{1 + \exp(-\alpha - \beta^{\mathrm{T}} Z)\}^{-1},$$

is often used, where α and β are the regression parameters. It is then discovered by Drescher & Becher (1997) that the attributable fraction can be expressed as

$$\varphi = 1 - \int_{\mathcal{Z}} \exp(-\beta^{\mathrm{T}} z) dF_{Z \mid D}(z \mid D = 1)$$

for the rare diseases. In contrast with $\varphi(t)$ in (4), this would be exactly $\varphi(\tau)$ if the actual occurrence times of the events were scaled up to the same time of τ. Thus, $\varphi(\cdot)$ can be considered as a natural extension of φ to the time-to-event outcomes following the proportional hazards model.

In general, however, $\widetilde{\varphi}(\cdot)$ may be of more straightforward interpretation for being directly expressed in the cumulative risk without the assumption of rare disease. In presence of potential confounding variables, $W \in \mathcal{W}$, say, $\widetilde{\varphi}(\cdot)$ can be further extended to adjust for W. A W-specific attributable fraction function can be defined as

$$\widetilde{\varphi}(t \mid W) = \frac{F(t \mid W) - F(t \mid W, Z = 0)}{F(t \mid W)}.$$

Thus, the adjusted attributable fraction function, $\widetilde{\varphi}_{\mathrm{adj}}(t)$, can be defined by either $\int_{\mathcal{W}} \widetilde{\varphi}(t \mid W) dF_W(w)$ or

$$\widetilde{\varphi}_{\mathrm{adj}}(t) = \frac{\int_{\mathcal{W}} F(t \mid w) dF_W(w) - \int_{\mathcal{W}} F(t \mid W, Z = 0) dF_W(w)}{\int_{\mathcal{W}} F(t \mid w) dF_W(w)}.$$

The latter is considered as an extension of that in Whittemore (1982) to the time-to-event outcomes.

Similarly, the hazard-based $\varphi(\cdot)$ can be extended to adjust for W as well, i.e., the W-specific

$$\varphi(t \mid W) = \frac{\lambda(t \mid W) - \lambda(t \mid W, Z = 0)}{\lambda(t \mid W)}.$$

Under the proportional hazards model of $\lambda(t \mid Z, W) = \lambda_0(t)\exp(\beta^{\mathrm{T}}Z + \gamma^{\mathrm{T}}W)$, $\varphi(t \mid W)$ can be further expressed by $\varphi(t \mid W) = 1 - \int_{\mathcal{Z}} \exp(-\beta^{\mathrm{T}}Z) dF_{Z \mid T, W}(z \mid t, W)$. Thus, the adjusted attributable fraction function, $\varphi_{\mathrm{adj}}(t)$ can be simply defined as $\int_{\mathcal{W}} \varphi(t \mid W) dF_W(w)$. Compared with that of $\widetilde{\varphi}_{\mathrm{adj}}(t)$, $\varphi_{\mathrm{adj}}(t)$ may have more advantage in practical application with its straightforward adaptability with the proportional hazards model. It is yet necessary to point out that such applicability of $\varphi(\cdot)$ depends on the assumption of the underlying proportional hazards model. One natural way is use the flexible hazard-based models in $\varphi(\cdot)$ to relax the assumption, for instance, by the general relative risk model of $\lambda(t \mid Z, W) = \lambda_0(t) r(t; Z, W)$ in Prentice & Self (1983), where $r(\cdot)$ is a general form of hazards ratio.

In the estimation and inferences of the proposed $\widehat{\varphi}(\cdot)$, we adopted the stronger version of independence assumption for (T, C, Z) as in Xu & O'Quigley (2000), for ease of the presentation and asymptotic property derivation. The weaker version of the usual conditional independence assumption can be extended as well in the development, when the amount of exposure can be quantified and grouped into finite number of categories as seen in most of the epidemiological studies. The techniques and guidelines in Murray & Tsiatis (1996) for continuous exposure would be useful in constructing and extending $\widehat{\varphi}(\cdot)$, as cited in Xu & O'Quigley (2000).

Appendix

Proof of Theorem 1

To establish the asymptotic properties in Theorem 1, we assume the necessary regularity conditions specified in Theorem 4.1 of Andersen & Gill (1982):

1. $\Lambda_0(t)$ is continuous, nondecreasing, and $\Lambda_0(\tau) < \infty$.
2. There exists a compact neighborhood $\mathcal{B}$ of β_* such that
$$E\left\{\sup_{\beta \in \mathcal{B}} \|Z\|^2 \exp\left(\beta^t Z\right)\right\} < \infty;$$

3. $\mathrm{pr}\{Y(\tau)=1\}>0$;

4. $\Sigma=\int_0^\tau v(t,\beta_*)s_0(t,\beta_*)\lambda_0(t)dt$ is positive definite, where $v(t,\beta)=s_2(t,\beta)/s_0(t,\beta)-\{s_1(t,\beta)/s_0(t,\beta)\}^{\otimes 2}$.

First we decompose the estimator $\widehat{\varphi}(t,\widehat{\beta})=1-\sum_i Y_i(t)/\sum_i Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)$ into

$$\left\{1-\frac{\sum_i Y_i(t)}{\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}}Z_i)}\right\}+\left\{\frac{\sum_i Y_i(t)}{\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}}Z_i)}-\frac{\sum_i Y_i(t)}{\sum_i Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)}\right\}=A(t)+B(t).$$

By Taylor's theorem, $B(t)$ equals

$$\frac{\sum_i Y_i(t)\{\sum_i Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)-\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}}Z_i)\}}{\sum_i Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}}Z_i)}=\frac{\sum_i Y_i(t)\sum_i Y_i(t)\exp(\widetilde{\beta}^{\mathrm{T}}Z_i)Z_i^{\mathrm{T}}(\widehat{\beta}-\beta_*)}{\sum_i Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}}Z_i)},$$

where $\widetilde{\beta}$ is on the line segment connecting $\widehat{\beta}$ and β_*. Following Theorem 4.1 and Corollary III.2 of Andersen & Gill (1982), we know that $s_0(t,\beta)$ and $s_1(t,\beta)$ are continuous in $\beta\in\mathcal{B}$. In addition, $s_0(t,\beta)$ is bounded away from zero on $[0,\tau]\times\mathcal{B}$. Since $\sup_{t\in[0,\tau]}|S(t)-s(t)|\to_p 0$ and $\sup_{t\in[0,\tau],\beta\in\mathcal{B}}|S_k(t,\beta)-s_k(t,\beta)|\to_p 0$, for $k=0,1,2$, respectively,

$$\frac{\sum_i Y_i(t)\sum_i Y_i(t)\exp(\widetilde{\beta}^{\mathrm{T}}Z_i)Z_i^{\mathrm{T}}}{\sum_i Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}}Z_i)\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}}Z_i)}\to\frac{s(t)s_1(t,\beta_*)^{\mathrm{T}}}{s_0(t,\beta_*)^2}$$

uniformly for $t\in[0,\tau]$. Due to the consistency of $\widehat{\beta}$ for β_* we know $B(t)\to_p 0$ uniformly on $[0,\tau]$. Also,

$$B(t)=\frac{s(t)s_1(t,\beta_*)^{\mathrm{T}}}{s_0(t,\beta_*)^2}(\widehat{\beta}-\beta_*)+O_p(n^{-1}). \qquad (\mathrm{A}\cdot 1)$$

It thus follows that $\widehat{\varphi}(t,\widehat{\beta})\to_p 1-s(t)/s_0(t,\beta_*)$ uniformly for $t\in[0,\tau]$.

To prove the uniform consistency of $\widehat{\varphi}(t,\widehat{\beta})$, we only need to show $\varphi(t)=1-s(t)/s_0(t,\beta_*)$. Under the proportional hazards model (1), $\lambda(t)$ equals

$$\lim_{\Delta t\to 0+}\mathrm{pr}\{T\in[t,t+\Delta t)\mid T\geq t\}/\Delta t=\lim_{\Delta t\to 0+}E\left[\mathrm{pr}\{T\in[t,t+\Delta t)\mid T\geq t,Z\}\mid T\geq t\right]/\Delta t$$
$$=E\left[\lim_{\Delta t\to 0+}\mathrm{pr}\{T\in[t,t+\Delta t)\mid T\geq t,Z\}/\Delta t\;\middle|\;T\geq t\right]=E\left\{\exp(\beta_*^{\mathrm{T}}Z)\mid T\geq t\right\}\lambda_0(t).$$

Therefore, $\varphi(t) = 1 - \lambda_0(t)/\lambda(t) = 1 - 1/E\{\exp(\beta_*^{\mathrm{T}} Z) \mid T \geq t\}$. On the other hand, $s_0(t, \beta_*)$ equals

$$E\left\{Y(t)\exp(\beta_*^{\mathrm{T}} Z)\right\} = E\left\{\exp(\beta_*^{\mathrm{T}} Z) \mid Y(t) = 1\right\} \mathrm{pr}\{Y(t) = 1\} = E\left\{\exp(\beta_*^{\mathrm{T}} Z) \mid Y(t) = 1\right\} s(t).$$

Hence, $\varphi(t) = 1 - s(t)/s_0(t, \beta_*) = 1 - 1/E\{\exp(\beta_*^{\mathrm{T}} Z) \mid Y(t) = 1\}$. Under the assumption that C is independent of (T, Z), $E\{\exp(\beta_*^{\mathrm{T}} Z) \mid Y(t) = 1\} = E\{\exp(\beta_*^{\mathrm{T}} Z) \mid T \geq t\}$. Therefore, the uniform consistency of $\widehat{\varphi}(t, \widehat{\beta})$ holds.

To prove the asymptotic normality, $\widehat{\varphi}(t, \widehat{\beta}) - \varphi(t)$ can be written as

$$\left\{1 - \frac{\sum_i Y_i(t)}{\sum_i Y_i(t)\exp(\beta_*^{\mathrm{T}} Z_i)}\right\} - \left\{1 - \frac{s(t)}{s_0(t, \beta_*)}\right\} + B(t).$$

By the expression of $B(t)$ in (A·1) and the martingale representation of $\widehat{\beta} - \beta_*$, it further equals

$$\frac{s(t)}{s_0(t, \beta_*)^2} n^{-1} \sum_{i=1}^{n} \left\{Y_i(t)\exp(\beta_*^{\mathrm{T}} Z_i) - s_0(t, \beta_*)\right\} - \frac{1}{s_0(t, \beta_*)} n^{-1} \sum_{i=1}^{n} \{Y_i(t) - s(t)\}$$
$$+ \frac{s(t) s_1(t, \beta_*)^{\mathrm{T}}}{s_0(t, \beta_*)^2} \Sigma^{-1} \cdot n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \left\{Z_i - \bar{z}(u, \beta_*)\right\} dM_i(u) + O_p(n^{-1}),$$

where $\bar{z}(t, \beta_*) = s_1(t, \beta_*)/s_0(t, \beta_*)$. Thus $n^{1/2}\{\widehat{\varphi}(t, \widehat{\beta}) - \varphi(t)\} = n^{-1/2} \sum_i v_i(t) + o_p(1)$. Here,

$$v_i(s) = \frac{s(t) Y_i(t) \exp\left(\beta_*^{\mathrm{T}} Z_i\right)}{s_0(t, \beta_*)^2} - \frac{Y_i(t)}{s_0(t, \beta_*)} + \frac{s(t) s_1(t, \beta_*)^{\mathrm{T}}}{s_0(t, \beta_*)^2} \Sigma^{-1} \int_0^{\tau} \left\{Z_i - \bar{z}(u, \beta_*)\right\} dM_i(u).$$

Since $Y_i(t)\exp(\beta_*^{\mathrm{T}} Z_i)$ and $Y_i(t)$ are both monotonic processes, $n^{1/2}\{\widehat{\varphi}(t, \widehat{\beta}) - \varphi(t)\}$ converges weakly to a zero-mean Gaussian process with covariance function $\sigma_\varphi(s, t) = E\{v_1(s) v_1(t)\}$, as shown in the Example 2.11.16 of van der Vaart & Wellner (1996)

In order to prove that $\sigma_\varphi(s, t)$ is consistently estimated by $\widehat{\sigma}_\varphi(s, t)$, it suffices to show by Cauchy-Schwarz inequality that

$$n^{-1} \sum_{i=1}^{n} \left\{Y_i(t)\exp(\widehat{\beta}^{\mathrm{T}} Z_i) - Y_i(t)\exp\left(\beta_*^{\mathrm{T}} Z_i\right)\right\}^2 \to_p 0; \text{ and}$$
$$n^{-1} \sum_{i=1}^{n} \left[\int_0^{\tau} \{Z_i - \bar{Z}(u, \widehat{\beta})\} d\widehat{M}_i(u) - \int_0^{\tau} \{Z_i - \bar{z}(u, \beta_*)\} dM_i(u)\right]^2 \to_p 0.$$

These can be established by the consistencies of $\widehat{\beta}$, the uniform consistency of $\widehat{\Lambda}(\cdot)$, and $S_k(\cdot, \beta)$, $(k = 0, 1, 2)$, respectively, following Lemma 1 of Lin et al. (2000).

References

Andersen, P. K. & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.* **4**, 1100-1120.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *J. R. Statist. Soc.* B **34**, 187-220.

Drescher, K. & Schill, W. (1991). Attributable risk estimation from case-control data via logistic regression. *Biometrics* **47**, 1247-1256.

Greenland, S. (2001). Estimation of population attributable fractions from fitted incidence ratios and exposure survey data, with an application to electromagnetic fields and childhood leukemia. *Biometrics* **57**, 182-188.

Greenland, S. & Robins, J. M. (1988). Conceptual problems in the definition and interpretation of attributable fractions. *Amer. J. Epidemiol.* **128**, 1185-1197.

Kaslow, R. A., Ostrow, D. G., Deterls, R., et al. (1987). The multicenter AIDS cohort study: rationale, organization, and selected characteristics of the participants. *Am. J. Epidemiol.* **126**, 310-318.

Levin, M. L. (1953). The occurrence of lung cancer in man. *ACTA Unio Internationalis Contra Cancrum* **9**, 531-541.

Lin, D. Y., Fleming, T. R. & Wei, L. J. (1994). Confidence bands for survival curves under the proportional hazards model. *Biometrika* **81**, 73-81.

Lin, D. Y., Wei, L. J., Yang, I., & Ying, Z. (2000) Semiparametric regression for the mean and rate functions of recurrent events. *J.R. Statist. Soc.* **B 62**, 711-730.

Murray, S. & Tsiatis, A. A. (1996). Nonparametric survival estimation using prognostic longitudinal covariates. *Biometrics* **52**, 137-151.

Prentice, R. L. & Self, S. (1983). Asymptotic distribution theory for Cox-type regression models with general relative risk form. *Ann. of Statist.* **11**, 804-813.

Tsiatis, A. A. (1981). A large sample study of Cox's regression model. *Ann. Statist.* **9**, 93-108.

van der Vaart, A. W. & Wellner, J. A. (1996) *Weak convergence and empirical processes,* New York: Springer-Verlag.

Whittemore, A. S. (1982). Statistical methods for estimating attributable risk from retrospective data. *Statist. Med.* **1**, 229-243.

Xu, R. & O'Quigley (2000). Proportional hazards estimate of the conditional survival function. *J. R. Statist. Soc.* **B 62**, 667-680.

Figure 1: Attributable fraction functions in the proportional hazards model $\lambda(t \mid Z = 1) = 2 \cdot \lambda(t \mid Z = 0)$ with constant $\lambda(t \mid Z = 0)$. Solid lines are $\varphi(t) = 1 - \lambda(t \mid Z = 0)/\lambda(t)$. Dashed lines are $\widetilde{\varphi}(t) = 1 - F(t \mid Z = 0)/F(t)$.

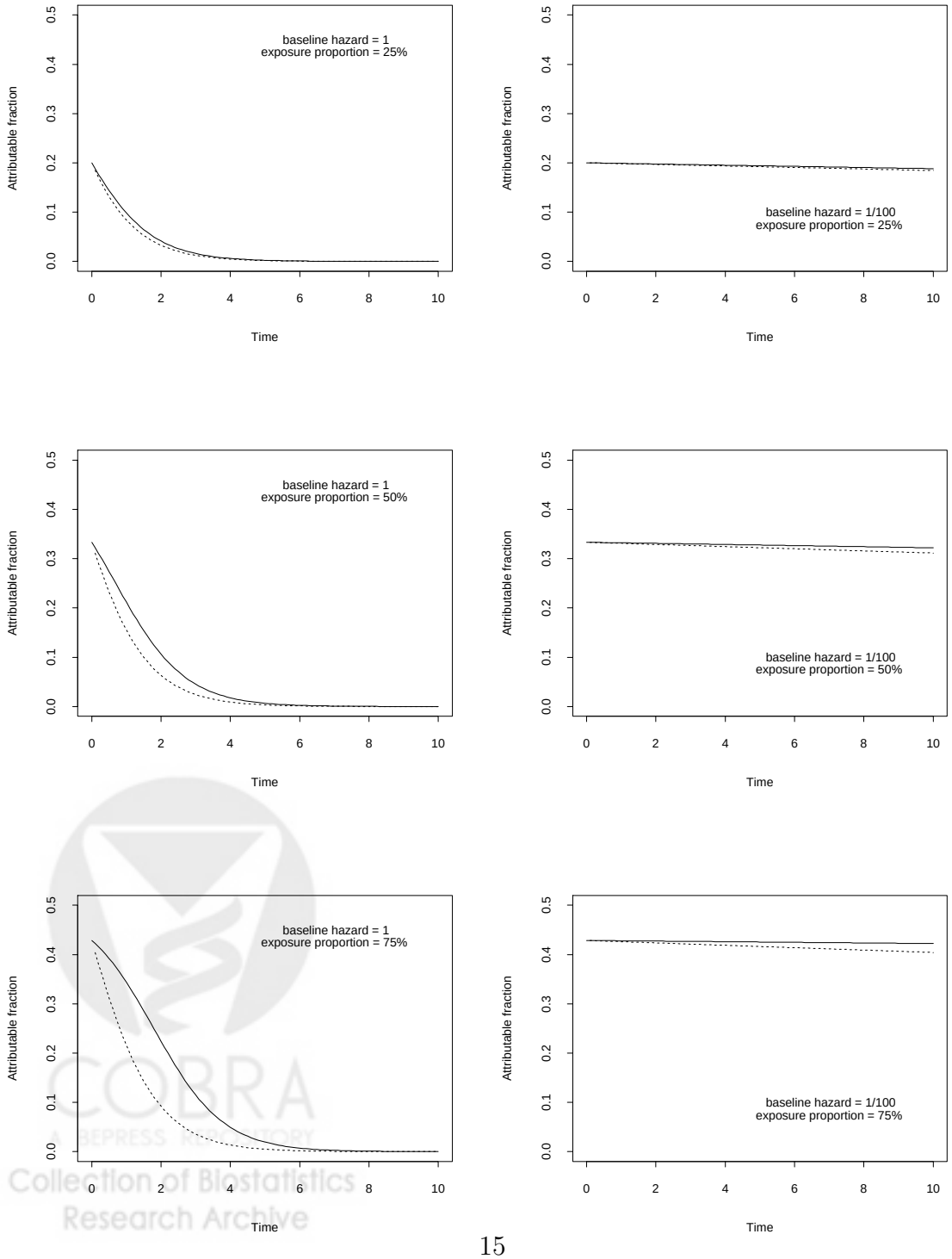

Figure 2: Estimated $\widehat{\varphi}(\cdot)$ and their confidence intervals in the proportional hazards model $\lambda(t \mid Z = 1) = 2 \cdot \lambda(t \mid Z = 0)$ for the Multicenter AIDS Cohort Study dataset: (a) risk factor Z_1; (b) risk factor Z_2; (c) combined risk factor of Z_1 or Z_2. The solid lines are $\widehat{\varphi}(\cdot)$. The dashed lines are the pointwise confidence intervals. The dotted lines are crude $\widehat{\varphi}$.

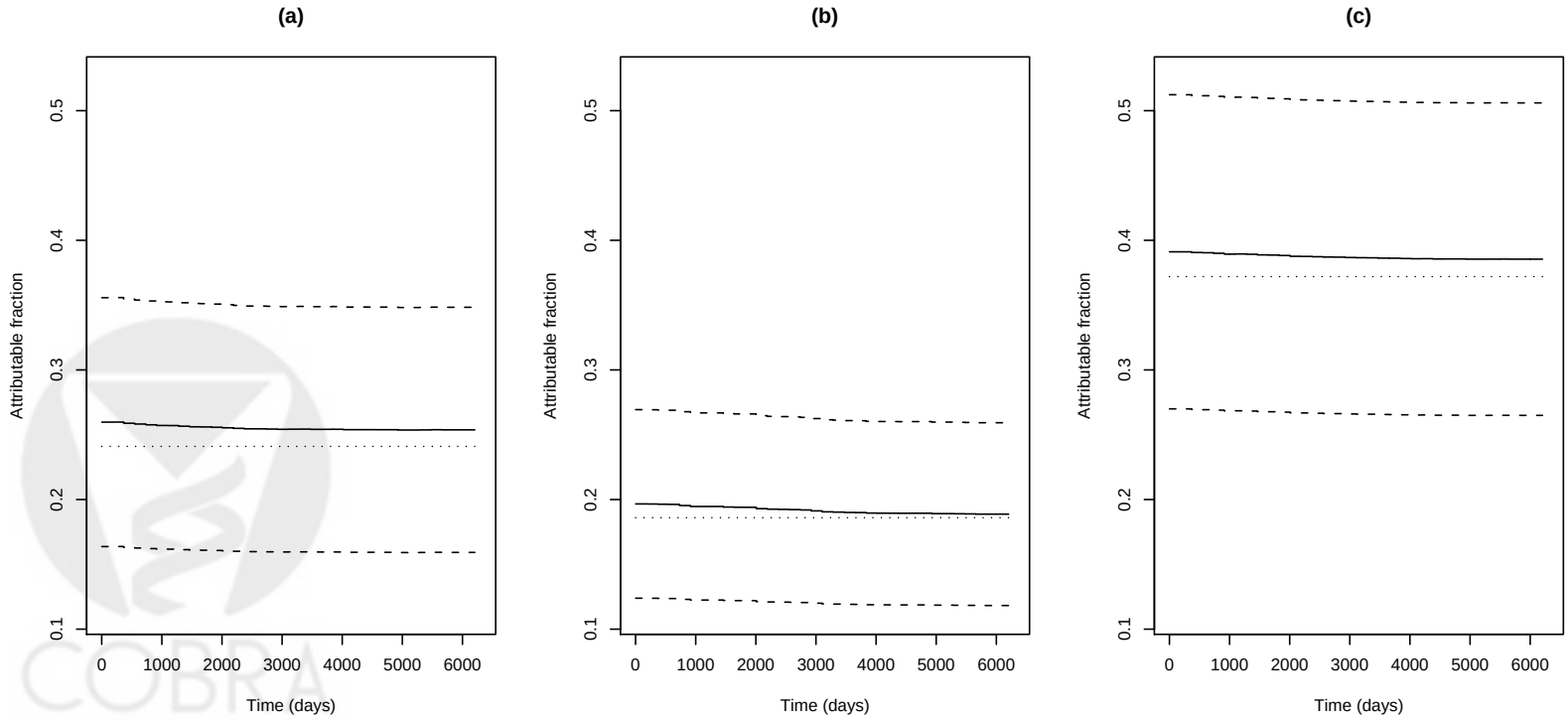

Table 1: Summary of Simulation Studies under the proportional hazards model $\lambda(t \mid Z) = \lambda_0(t)\exp(\beta_*^{\mathrm{T}} Z)$.

				$\beta_* = 0$							
				$t_1 : S(t_1) = 0.75$				$t_2 : S(t_2) = 0.50$			
$\lambda_0(t) \equiv \lambda_0$	Expo. Prob.	Cens.%	n	|Bias|	Cov. Prob.	SE	Mean SE	|Bias|	Cov. Prob.	SE	Mean SE
0.01	0.25	10%	200	0.0002	0.933	0.0448	0.0422	0.0026	0.934	0.0448	0.0426
0.01	0.25	10%	500	0.0011	0.953	0.0265	0.0271	0.0003	0.956	0.0264	0.0271
0.01	0.25	30%	200	0.0020	0.941	0.0489	0.0482	0.0048	0.945	0.0497	0.0486
0.01	0.25	30%	500	0.0007	0.942	0.0315	0.0308	0.0019	0.944	0.0317	0.0308
0.01	0.50	10%	200	0.0015	0.944	0.0746	0.0742	0.0035	0.946	0.0752	0.0746
0.01	0.50	10%	500	0.0026	0.951	0.0474	0.0471	0.0036	0.952	0.0477	0.0472
0.01	0.50	30%	200	0.0027	0.948	0.0869	0.0862	0.0058	0.950	0.0878	0.0846
0.01	0.50	30%	500	0.0006	0.940	0.0552	0.0535	0.0018	0.940	0.0556	0.0536
1.00	0.25	10%	200	0.0010	0.955	0.0426	0.0425	0.0011	0.956	0.0429	0.0425
1.00	0.25	10%	500	0.0008	0.954	0.0275	0.0272	0.0000	0.954	0.0274	0.0272
1.00	0.25	30%	200	0.0006	0.944	0.0478	0.0483	0.0021	0.946	0.0483	0.0485
1.00	0.25	30%	500	0.0000	0.945	0.0315	0.0307	0.0012	0.945	0.0316	0.0307
1.00	0.50	10%	200	0.0006	0.948	0.0750	0.0747	0.0026	0.948	0.0756	0.0751
1.00	0.50	10%	500	0.0008	0.956	0.0470	0.0470	0.0017	0.957	0.0471	0.0471
1.00	0.50	30%	200	0.0028	0.947	0.0860	0.0840	0.0060	0.950	0.0873	0.0846
1.00	0.50	30%	500	0.0019	0.944	0.0553	0.0533	0.0007	0.945	0.0556	0.0553

(Continued in next page)

(Table 1 Continued)

				$\beta_* = \log 2$							
				$t_1 : S(t_1) = 0.75$				$t_2 : S(t_2) = 0.50$			
$\lambda_0(t) \equiv \lambda_0$	Expo. Prob.	Cens.%	n	Bias	Cov. Prob.	SE	Mean SE	Bias	Cov. Prob.	SE	Mean SE
0.01	0.25	10%	200	0.0017	0.945	0.0486	0.0481	0.0018	0.952	0.0329	0.0328
0.01	0.25	10%	500	0.0005	0.954	0.0301	0.0306	0.0004	0.951	0.0207	0.0251
0.01	0.25	30%	200	0.0004	0.960	0.0488	0.0522	0.0034	0.947	0.0330	0.0327
0.01	0.25	30%	500	0.0005	0.965	0.0306	0.0330	0.0011	0.949	0.272	0.0270
0.01	0.50	10%	200	0.0021	0.950	0.0643	0.0644	0.0030	0.950	0.0643	0.0644
0.01	0.50	10%	500	0.0005	0.945	0.0414	0.0414	0.0013	0.949	0.0346	0.0346
0.01	0.50	30%	200	0.0009	0.955	0.0688	0.0686	0.0019	0.957	0.0577	0.0589
0.01	0.50	30%	500	0.0022	0.955	0.0461	0.0461	0.0021	0.956	0.0381	0.0386
1.00	0.25	10%	200	0.0023	0.948	0.0490	0.0491	0.0014	0.951	0.0333	0.0332
1.00	0.25	10%	500	0.0000	0.953	0.0291	0.0307	0.0010	0.954	0.0203	0.0202
1.00	0.25	30%	200	0.0016	0.947	0.0512	0.0519	0.0030	0.956	0.0343	0.0343
1.00	0.25	30%	500	0.0006	0.955	0.0318	0.0311	0.0004	0.945	0.0217	0.0223
1.00	0.50	10%	200	0.0007	0.945	0.0651	0.0654	0.0020	0.950	0.0542	0.0541
1.00	0.50	10%	500	0.0021	0.954	0.0399	0.0414	0.0023	0.953	0.0336	0.0346
1.00	0.50	30%	200	0.0034	0.943	0.0737	0.0737	0.0040	0.954	0.0613	0.0620
1.00	0.50	30%	500	0.0008	0.952	0.0462	0.0461	0.0014	0.956	0.0379	0.0385

Expo. Prob., probability of $Z = 1$; Cens.%, censoring probability; t_1, 75%-tile of marginal survival function; t_2, median of marginal survival function; |Bias|, absolute difference between 1000 $\widehat{\varphi}(t)$ and the true value; Cov. Prob., percentage of 1000 95% nominal confidence intervals containing the true value; SE, sample standard error of 1000 $\widehat{\varphi}(t)$; Mean SE, average of 1000 $\widehat{\sigma}_{\varphi}(t)$

Table 2: Summary statistics and estimates of the proportional hazards model $\lambda(t \mid Z) = \lambda_0(t)\exp(\beta^{\mathrm{T}} Z)$ for the Multicenter AIDS Cohort Study dataset

Risk factor	Z	Size (%)	HIV incidence (%)	$\widehat{\beta}$ (SE)	$\widehat{\varphi}$
sex with AIDS partner	$Z_1 = 0$	1387 (58.5%)	160 (11.5%)		
	$Z_1 = 1$	1954 (41.5%)	348 (17.8%)	0.470 (0.095)	0.241
anal sex with partner	$Z_2 = 0$	1995 (59.7%)	247 (12.3%)		
	$Z_2 = 1$	1346 (40.3%)	261 (19.4%)	0.475 (0.089)	0.186
anal sex with AIDS partner	$Z_1 = Z_2 = 0$	871 (26.1%)	83 (9.5%)		
	otherwise	2470 (73.9%)	425 (15.5%)	0.625 (0.120)	0.372

Size % are the percentages of $Z = 0/1$ in the cohort, respectively; HIV incidence % are the percentages of HIV incidences in $Z = 0/1$, respectively; SE are the standard errors of $\widehat{\beta}$ in the proportional hazards model; $\widehat{\varphi}$ is the sample estimates of φ