

## Comparing two failure time distributions in the presence of dependent censoring

BY D. Y. LIN

*Department of Biostatistics, Box 357232, University of Washington, Seattle,  
Washington 98195, U.S.A.*

J. M. ROBINS

*Department of Epidemiology, Harvard University, Boston, Massachusetts 02115, U.S.A.*

AND L. J. WEI

*Department of Biostatistics, Harvard University, Boston, Massachusetts 02115, U.S.A.*

### SUMMARY

In a randomised clinical trial to compare two groups of patients, suppose that the time to disease occurrence, the major response variable, may be subject to dependent censoring by death or selective patient withdrawal, while the patients who have reached the disease endpoint are followed for their secondary endpoints or survival information. To adjust for the dependent censoring in assessing the group difference in disease occurrence, we assume that, on a logarithmic scale, the times to disease occurrence and dependent censoring for the two groups satisfy a bivariate location-shift model with a completely unspecified underlying distribution. Rank-based procedures are constructed for making inferences about the location-shift parameter. Model checking techniques are also developed. Numerical studies show that the proposed methods are appropriate for practical use. An illustration with data taken from a recent AIDS clinical trial is provided.

*Some key words:* Cause-specific hazard; Clinical trial; Competing risk; Informative censoring; Log-rank statistic; Semiparametric inference; Survival analysis.

### 1. INTRODUCTION

In a randomised clinical trial to assess which of two agents is more effective in delaying the onset of a particular disease, the time to disease occurrence may be subject to dependent censoring by death or selective patient withdrawal in addition to the usual independent censoring. On the other hand, the patients who have reached the disease endpoint may still be followed for their secondary endpoints or survival information. Such examples arise frequently in cancer, AIDS and cardiovascular diseases research. For instance, in a recent randomised study on the prophylaxis of pneumocystis carinii pneumonia, PCP, 154 AIDS patients who had recovered from an initial episode of PCP received trimethoprim-sulfamethoxazole, TS, and another 156 received aerosolised pentamidine, AP (Hardy et al., 1992). By the end of the trial, 14 patients in the TS group and 36 patients in the AP group had PCP recurrences. There were 43 and 47 deaths in the TS and AP groups, respectively. Most of those deaths, 36 in each group, occurred prior to recurrences of PCP. The patients were followed for their other opportunistic infections and mortality information even after

they had experienced the recurrences of PCP. A number of patients, however, were withdrawn from the study prematurely for various health-related reasons. The main challenge here is how to adjust for the dependent censoring due to death and selective withdrawal in comparing the distributions of PCP recurrence time between the two groups.

In the presence of dependent censoring, the conventional log-rank statistic and two-sample proportional hazards model are comparing two cause-specific hazard functions (Kalbfleisch & Prentice, 1980, p. 167). Unfortunately, the cause-specific hazard function does not have the same interpretation as the usual net hazard function. In particular, the exponential negative cumulative cause-specific hazard function is not the marginal survival function for disease occurrence. The clinical relevance of the cause-specific hazard function is especially questionable if there is appreciable selective patient withdrawal.

In a purely nonparametric setting, the distribution function for the failure time of interest is not identifiable if the censoring time is dependent on the failure time (Tsiatis, 1975). Most of the literature on dependent censoring has focused on modelling the dependence structure between the failure time and dependent censoring variable to make the problem identifiable. For example, Link (1989) proposed a model in which dependent censoring only occurs in a subpopulation defined by a frailty distribution and obtained a self-consistent estimator of the failure time distribution. Emoto & Matthews (1990) postulated a bivariate Weibull model and showed that the maximum likelihood estimators for the parameters of the joint distribution are consistent. In addition, Robins (1987a, Theorem AD.1, §§ 5, 6) showed that the group-specific marginal distribution of time to disease occurrence is estimable if data are available on some time-varying marker processes and the cause-specific hazard of censoring does not depend on the failure time conditional on the history of the markers. Semiparametric inferences of marginal survival curves can then be performed using the G-computation algorithm of Robins (1987b, Appendix 1; 1989, p. 129) or the inverse probability of censoring weighted estimators of Robins & Rotnitzky (1992) and Robins (1993a,b). When the data on such marker processes are not available, Robins (1989, p. 158; 1993a, Appendix 4), Robins & Tsiatis (1991) and Robins & Rotnitzky (1992, Appendix 4) demonstrated that it remains possible to estimate the group difference in the marginal failure time distribution if the failure times for the two groups follow a transformation model, the potential censoring times of all subjects including failures are observed and the dependent censoring distributions are identical between the two groups.

In this paper, we propose a new semiparametric approach to adjusting for dependent censoring when comparing the two distributions of disease occurrence time. Our approach requires weaker conditions than the aforementioned methods. Specifically, we assume that, on a logarithmic scale, the times to disease occurrence and dependent censoring for the two groups satisfy a bivariate location-shift model with a completely arbitrary underlying distribution. In the next section, we describe this model in detail and show how to draw valid statistical inference about the group difference. In § 3, we report the results of our simulation studies on the finite-sample properties of the proposed methods and provide an illustration with the PCP prophylaxis trial described in the opening paragraph.

## 2. INFERENCE PROCEDURES

### 2.1. *Data and model*

Let  $X^0$  and  $Y^0$  be the times to disease occurrence and dependent censoring, let  $C$  be the censoring time due to random loss to follow-up and study termination, and let  $Z$  be the group indicator. To ease our discussion, we shall at times refer to  $Y^0$  as the survival

time and  $C$  as the censoring time. In general,  $X^0$  and  $Y^0$  are correlated. We assume that  $C$  is independent of  $(X^0, Y^0)$  conditional on  $Z$ . The data consist of  $n$  independent replicates of the random vector  $(X, \delta, Y, \xi, Z)$ , where  $X = X^0 \wedge Y^0 \wedge C$ ,  $\delta = I(X^0 \leq Y^0 \wedge C)$ ,  $Y = Y^0 \wedge C$  and  $\xi = I(Y^0 \leq C)$ . Here and in the sequel,  $a \wedge b$  denotes the minimum of  $a$  and  $b$ , and  $I(\cdot)$  is the indicator function. Note that  $Y^0$  may censor  $X^0$  but not vice versa.

Suppose that all the time variables are measured on a logarithmic scale. The bivariate location-shift model assumes that there exist some unknown constants  $\theta_0$  and  $\eta_0$  such that the bivariate random vectors  $(X_i^0 - \theta_0 Z_i, Y_i^0 - \eta_0 Z_i)$  ( $i = 1, \dots, n$ ) have a common, but completely unspecified, joint distribution. Our objective is to use the above data to draw inferences about  $\beta_0 = (\eta_0, \theta_0)$ .

2.2. Point estimation

The estimation of  $\eta_0$  has been studied extensively (Louis, 1981; Wei & Gail, 1983). Let  $\tilde{Y}_i(\eta) = Y_i - \eta Z_i$  ( $i = 1, \dots, n$ ). Then a reasonable estimating function for  $\eta_0$  is the log-rank statistic based on the transformed data  $\{\tilde{Y}_i(\eta), \xi_i, Z_i\}$  ( $i = 1, \dots, n$ ), namely

$$U_1(\eta) = n^{-1} \sum_{i=1}^n \xi_i \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{Y}_j(\eta) \geq \tilde{Y}_i(\eta)\} Z_j}{\sum_{j=1}^n I\{\tilde{Y}_j(\eta) \geq \tilde{Y}_i(\eta)\}} \right]. \tag{2.1}$$

Because the random vectors  $(Y_i^0 - \eta_0 Z_i)$  ( $i = 1, \dots, n$ ) have the same distribution and because the censoring times  $C_i$  are independent of the 'survival times'  $Y_i^0$  in each group, the statistic  $U_1(\eta_0)$  is asymptotically zero-mean normal (Fleming & Harrington, 1991, § 7.2). Let  $\hat{\eta}$  be the zero-crossing of  $U_1(\eta)$ . Then  $\hat{\eta}$  is consistent and asymptotically normal (Louis, 1981; Wei & Gail, 1983).

For estimating  $\theta_0$ , it seems natural to replace  $\{\tilde{Y}_i(\eta), \xi_i\}$  ( $i = 1, \dots, n$ ) in (2.1) with  $(X_i - \theta Z_i, \delta_i)$  ( $i = 1, \dots, n$ ) and use the resulting estimating function, denoted by  $S(\theta)$ . It can be shown that the mean of  $S(\theta_0)$  is the integral of a weighted difference between two cause-specific hazard functions  $g_0(\cdot)$  and  $g_1(\cdot)$ , where

$$g_Z(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{pr}(t \leq X^0 - \theta_0 Z < t + \Delta t | X^0 - \theta_0 Z \geq t, Y^0 - \theta_0 Z \geq t)}{\Delta t}$$

(Kalbfleisch & Prentice, 1989, p. 167; Fleming & Harrington, 1991, p. 265). Because the distribution of  $(Y^0 - \theta_0 Z)$  is not free of  $Z$  unless  $\theta_0 = \eta_0$ , the functions  $g_0(\cdot)$  and  $g_1(\cdot)$  are generally not equal to each other. Consequently, the use of  $S(\theta)$  would not yield a consistent estimator for  $\theta_0$ .

In order to construct valid estimating functions for  $\theta_0$ , we transform  $X_i$  and  $\delta_i$  to  $\tilde{X}_i(\beta)$  and  $\tilde{\delta}_i(\beta)$  as given in Table 1. Note that uncensored observations may become censored ones after the transformations. This artificial censoring occurs in only one of the two groups, and which group it occurs depends on the ordering of  $\theta$  and  $\eta$ . It is easy to verify that the transformations given in Table 1 may be written as

$$\tilde{X}_i(\beta) = (X_i^0 - \theta Z_i) \wedge (Y_i^0 - \eta Z_i - d) \wedge (C_i - \eta Z_i - d), \tag{2.2}$$

$$\tilde{\delta}_i(\beta) = I\{(X_i^0 - \theta Z_i) \leq (Y_i^0 - \eta Z_i - d) \wedge (C_i - \eta Z_i - d)\}, \tag{2.3}$$

where  $d = 0$  if  $\theta \leq \eta$  and  $d = \theta - \eta$  if  $\theta > \eta$ . Expressions (2.2) and (2.3) show that  $\tilde{X}_i(\beta)$  is the observation time for the failure time  $(X_i^0 - \theta Z_i)$  in the presence of dependent censoring by  $(Y_i^0 - \eta Z_i - d)$  and independent censoring by  $(C_i - \eta Z_i - d)$ , and  $\tilde{\delta}_i(\beta)$  is the associated failure indicator.

Table 1. Definitions of  $\tilde{X}_i(\beta)$  and  $\tilde{\delta}_i(\beta)$

		$\{\theta \leq \eta\}$		$\{\theta > \eta\}$	
	$\{Z_i = 0\}$	$\{Z_i = 1\}$		$\{Z_i = 0\}$	$\{Z_i = 1\}$
$\tilde{X}_i(\beta) =$	$X_i$	$(X_i - \theta) \wedge (Y_i - \eta)$		$X_i \wedge (Y_i - \theta + \eta)$	$X_i - \theta$
$\tilde{\delta}_i(\beta) =$	$\delta_i$	$\delta_i I(X_i - \theta \leq Y_i - \eta)$		$\delta_i I(X_i \leq Y_i - \theta + \eta)$	$\delta_i$

The log-rank statistic for the transformed data  $\{\tilde{X}_i(\beta), \tilde{\delta}_i(\beta), Z_i\}$  ( $i = 1, \dots, n$ ) is

$$U_2(\beta) = n^{-1/2} \sum_{i=1}^n \tilde{\delta}_i(\beta) \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\beta) \geq \tilde{X}_i(\beta)\} Z_j}{\sum_{j=1}^n I\{\tilde{X}_j(\beta) \geq \tilde{X}_i(\beta)\}} \right]. \tag{2.4}$$

Similar to  $S(\theta_0)$ , the mean of  $U_2(\beta_0)$  is the integral of a weighted difference of cause-specific hazard functions  $h_0(\cdot)$  and  $h_1(\cdot)$ , where

$$h_Z(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{pr}\{t \leq X^0 - \theta_0 Z < t + \Delta t \mid X^0 - \theta_0 Z \geq t, Y^0 - \eta_0 Z - d_0 \geq t\}}{\Delta t} \tag{2.5}$$

with  $d_0 = 0$  if  $\theta_0 \leq \eta_0$ , and  $d_0 = \theta_0 - \eta_0$  otherwise. Clearly,  $h_1(\cdot) = h_0(\cdot)$  under the assumed bivariate location-shift model, which implies that  $U_2(\beta_0)$  is centred around 0. Thus,  $U_2(\beta)$  is a reasonable estimating function. Given the estimator  $\hat{\eta}$  from  $U_1(\eta)$ , we then define the estimator  $\hat{\theta}$  for  $\theta_0$  as the zero-crossing of  $U_2(\hat{\eta}, \theta)$ , denoting  $(\hat{\eta}, \hat{\theta})'$  by  $\hat{\beta}$ . It is shown in Appendix 1 that  $\hat{\theta}$  is consistent and asymptotically normal.

Our device of artificial censoring to construct unbiased estimating functions in the presence of dependent censoring was previously used by Robins (1989, pp. 146, 158; 1993a, Appendix 4), Robins & Tsiatis (1991) and Robins & Rotnitzky (1992, Appendix 4) for the special case in which  $\eta_0$  is known to be zero. A different kind of recensoring was used by Efron (1967) to develop some Wilcoxon-type tests for the two-sample problem with independent censoring.

### 2.3. Interval estimation

We show in Appendix 1 that the bivariate random vector  $U(\beta_0) = \{U_1(\eta_0), U_2(\beta_0)\}'$  is asymptotically zero-mean normal with a covariance matrix that can be consistently estimated by

$$V = n^{-1} \begin{pmatrix} \sum_{i=1}^n W_{1i}^2 & \sum_{i=1}^n W_{1i} W_{2i} \\ \sum_{i=1}^n W_{1i} W_{2i} & \sum_{i=1}^n W_{2i}^2 \end{pmatrix}, \tag{2.6}$$

where

$$\begin{aligned} W_{1i} &= \xi_i \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq \tilde{Y}_i(\hat{\eta})\} Z_j}{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq \tilde{Y}_i(\hat{\eta})\}} \right] \\ &\quad - \sum_{i=1}^n \frac{\xi_i I\{\tilde{Y}_i(\hat{\eta}) \geq \tilde{Y}_i(\hat{\eta})\}}{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq \tilde{Y}_i(\hat{\eta})\}} \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq \tilde{Y}_i(\hat{\eta})\} Z_j}{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq \tilde{Y}_i(\hat{\eta})\}} \right], \\ W_{2i} &= \tilde{\delta}_i(\hat{\beta}) \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq \tilde{X}_i(\hat{\beta})\} Z_j}{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq \tilde{X}_i(\hat{\beta})\}} \right] \\ &\quad - \sum_{i=1}^n \frac{\tilde{\delta}_i(\hat{\beta}) I\{\tilde{X}_i(\hat{\beta}) \geq \tilde{X}_i(\hat{\beta})\}}{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq \tilde{X}_i(\hat{\beta})\}} \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq \tilde{X}_i(\hat{\beta})\} Z_j}{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq \tilde{X}_i(\hat{\beta})\}} \right]. \end{aligned}$$

In addition, the random vector  $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$  is asymptotically zero-mean normal. The limiting covariance matrix, however, is difficult to estimate directly.

One may perform hypothesis testing and construct confidence intervals for  $\eta_0$  using the asymptotic properties of  $U_1(\eta_0)$  (Wei & Gail, 1983). It is somewhat more complicated to draw inferences about  $\theta_0$  because  $U_2(\beta)$  involves not only  $\theta$  but also  $\eta$ . Let us define the minimum-dispersion statistic  $Q(\theta) = \min_{\eta} U'(\eta, \theta)V^{-1}U(\eta, \theta)$ . By the arguments of Wei, Ying & Lin (1990, Appendix 2), the statistic  $Q(\theta_0)$  is asymptotically  $\chi^2_1$ . Therefore, a  $(1 - \alpha)$  confidence interval for  $\theta_0$  is  $\{\theta: Q(\theta) \leq \chi^2_1(\alpha)\}$ , where  $\chi^2_1(\alpha)$  is the upper  $\alpha$  point of the  $\chi^2_1$  distribution.

We may also obtain confidence intervals for  $\eta_0$  and  $\theta_0$  by applying a resampling idea developed by Parzen, Wei & Ying (1994). Given the realisations of the data  $(X_i, \delta_i, Y_i, \xi_i, Z_i)$  ( $i = 1, \dots, n$ ), we calculate  $U_1(\eta)$ ,  $U_2(\beta)$  and  $(W_{1i}, W_{2i})$  ( $i = 1, \dots, n$ ), and construct the pair of equations

$$U_1(\eta) = -n^{-\frac{1}{2}} \sum_{i=1}^n W_{1i} G_i, \tag{2.7}$$

$$U_2(\beta) = -n^{-\frac{1}{2}} \sum_{i=1}^n W_{2i} G_i, \tag{2.8}$$

where  $(G_1, \dots, G_n)$  are independent standard normal variables. It is important to note that, for (2.7) and (2.8), we regard the data  $(X_i, \delta_i, Y_i, \xi_i, Z_i)$  ( $i = 1, \dots, n$ ) involved in  $U_1, U_2$  and  $(W_{1i}, W_{2i})$  ( $i = 1, \dots, n$ ) as fixed and the  $G_i$  as random. Let  $\beta^* = (\eta^*, \theta^*)'$  be the stochastic solution to (2.7) and (2.8). We show in Appendix 1 that the conditional distribution of  $n^{\frac{1}{2}}(\beta^* - \hat{\beta})$  given the observed data is asymptotically the same as the unconditional distribution of  $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$ . To approximate the distribution of  $\hat{\beta}$ , we obtain a large number of realisations of  $\beta^*$  by repeatedly generating the normal random samples  $(G_1, \dots, G_n)$  while fixing the data at their observed values. Confidence intervals for  $\eta_0$  and  $\theta_0$  or the joint confidence region for  $\beta_0 = (\eta_0, \theta_0)'$  can then be obtained from the percentiles of the empirical distribution of  $\beta^*$ . In the absence of selective patient withdrawals, the joint confidence region for  $\beta_0$  enables one to make simultaneous inference about the group differences in both disease occurrence and death.

#### 2.4. Model checking

We shall use some martingale-based residuals to examine the adequacy of the assumed location-shift model. To describe these residuals, we introduce the counting processes

$$N_{1i}(t; \eta) = \xi_i I\{\tilde{Y}_i(\eta) \leq t\}, \quad N_{2i}(t; \beta) = \delta_i(\beta) I\{\tilde{X}_i(\beta) \leq t\} \quad (i = 1, \dots, n).$$

Let  $\lambda_0(t)$  be the common hazard function for the transformed 'survival times'  $(Y_i^0 - \eta_0 Z_i)$  ( $i = 1, \dots, n$ ), and let  $h_0(t)$  be the common cause-specific hazard function for disease occurrence given in (2.5). Then

$$M_{1i}(t) = N_{1i}(t; \eta_0) - \int_{-\infty}^t I\{\tilde{Y}_i(\eta_0) \geq u\} \lambda_0(u) du,$$

$$M_{2i}(t) = N_{2i}(t; \beta_0) - \int_{-\infty}^t I\{\tilde{X}_i(\beta_0) \geq u\} h_0(u) du$$

are martingales marginally (Fleming & Harrington, 1991, § 1.3). The martingale residuals

are defined as

$$\begin{aligned}\hat{M}_{1i}(t; \hat{\eta}) &= N_{1i}(t; \hat{\eta}) - \int_{-\infty}^t I\{\tilde{Y}_i(\hat{\eta}) \geq u\} d\hat{\Lambda}_0(u; \hat{\eta}), \\ \hat{M}_{2i}(t; \hat{\beta}) &= N_{2i}(t; \hat{\beta}) - \int_{-\infty}^t I\{\tilde{X}_i(\hat{\beta}) \geq u\} d\hat{H}_0(u; \hat{\beta}),\end{aligned}$$

where

$$\hat{\Lambda}_0(t; \eta) = \int_{-\infty}^t \frac{\sum_{i=1}^n dN_{1i}(u; \eta)}{\sum_{j=1}^n I\{\tilde{Y}_j(\eta) \geq u\}}, \quad \hat{H}_0(t; \beta) = \int_{-\infty}^t \frac{\sum_{i=1}^n dN_{2i}(u; \beta)}{\sum_{j=1}^n I\{\tilde{X}_j(\beta) \geq u\}}.$$

Note that  $\hat{\Lambda}_0$  and  $\hat{H}_0$  are the Nelson–Aalen estimators for the cumulative functions of  $\lambda_0$  and  $h_0$ . The martingale residuals  $\hat{M}_{1i}(t; \hat{\eta})$  and  $\hat{M}_{2i}(t; \hat{\beta})$  can be interpreted as the differences at time  $t$  between the observed and predicted numbers of events for the  $i$ th subject with respect to death and disease occurrence, respectively. Because these residuals are approximately centred around zero under the assumed model, they provide the natural building blocks for model checking.

Let us consider the following two functions of the martingale residuals

$$U_1(t; \eta) = n^{-\frac{1}{2}} \sum_{i=1}^n Z_i \hat{M}_{1i}(t; \eta), \quad U_2(t; \beta) = n^{-\frac{1}{2}} \sum_{i=1}^n Z_i \hat{M}_{2i}(t; \beta).$$

Note that  $U_1(\eta) = U_1(\infty; \eta)$  and  $U_2(\beta) = U_2(\infty; \beta)$ . We shall call  $U_1(t; \hat{\eta})$  and  $U_2(t; \hat{\beta})$  the score processes for death and disease occurrence, respectively, as they are reminiscent of the partial likelihood score process for the Cox model (Wei, 1984). If the location-shift model is correct, then the score processes will fluctuate randomly around zero. In Appendix 2, we show that, under the assumed model, the bivariate process  $U(s, t; \hat{\beta}) = \{U_1(s; \hat{\eta}), U_2(t; \hat{\beta})\}'$  converges weakly to a zero-mean Gaussian process whose distribution can be approximated by that of  $\hat{U}(s, t) = \{\hat{U}_1(s), \hat{U}_2(t)\}'$ , where

$$\begin{aligned}\hat{U}_1(s) &= n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^s \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq u\} Z_j}{\sum_{j=1}^n I\{\tilde{Y}_j(\hat{\eta}) \geq u\}} \right] d\hat{M}_{1i}(u; \hat{\eta}) G_i + U_1(s; \eta^*) - U_1(s; \hat{\eta}), \\ \hat{U}_2(t) &= n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^t \left[ Z_i - \frac{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq u\} Z_j}{\sum_{j=1}^n I\{\tilde{X}_j(\hat{\beta}) \geq u\}} \right] d\hat{M}_{2i}(u; \hat{\beta}) G_i + U_2(t; \beta^*) - U_2(t; \hat{\beta}),\end{aligned}$$

and  $\{G_i\}$  and  $\beta^*$  are as defined near the end of § 2.3. As in (2.7) and (2.8), we regard  $\{G_i\}$  as random and  $\{X_i, \delta_i, Y_i, \xi_i, Z_i\}$  as fixed in  $\hat{U}(\cdot, \cdot)$ . To approximate the null distribution of  $U(\cdot, \cdot; \hat{\beta})$ , we simulate a number of realisations from  $\hat{U}(\cdot, \cdot)$  by repeatedly generating the normal random samples  $(G_1, \dots, G_n)$ . To assess visually how unusual the observed patterns of  $U_1(\cdot; \hat{\eta})$  and  $U_2(\cdot; \hat{\beta})$  are, we may plot them along with a few, say 20, realisations of  $\hat{U}_1(\cdot)$  and  $\hat{U}_2(\cdot)$ . We may also perform formal lack-of-fit tests using the supremum statistics  $\sup_t |U_1(t; \hat{\eta})|$  and  $\sup_t |U_2(t; \hat{\beta})|$ , the  $p$ -values being approximated by simulating the process  $\hat{U}(\cdot, \cdot)$ .

The bivariate location-shift model described in § 2.1 implies that:

- (i) the random variables  $(Y_i^0 - \eta_0 Z_i)$  ( $i = 1, \dots, n$ ) have a common marginal distribution, and
- (ii) subject to the dependent censoring by  $(Y_i^0 - \eta_0 Z_i - d_0)$  ( $i = 1, \dots, n$ ), the random variables  $(X_i^0 - \theta_0 Z_i)$  ( $i = 1, \dots, n$ ) have a common cause-specific hazard function.

It is evident from § 2.2 and Appendix 1 that conditions (i) and (ii) are sufficient for the validity of the proposed inference procedures. Using the techniques of Lin, Wei & Ying (1993, Appendix 3), we can show that the  $\sup_t |U_1(t; \hat{\eta})|$  test is consistent against the general alternative that condition (i) fails and that, given condition (i), the  $\sup_t |U_2(t; \hat{\beta})|$  test is consistent against the general alternative that condition (ii) fails. Note that conditions (i) and (ii) do not always imply that  $(X_i^0 - \theta_0 Z_i)$  ( $i = 1, \dots, n$ ) have a common marginal distribution for the same  $\theta_0$  satisfying condition (ii). This latter assumption is untestable, but is essential for interpreting  $\theta_0$  as a net causal effect. In an unpublished technical report, J. M. Robins provided an example in which (i) and (ii) hold for a unique nonzero  $\theta_0$  and yet  $X^0$  is marginally independent of  $Z$  and the sharp null hypothesis of no causal effect of treatment on disease occurrence is true.

### 3. NUMERICAL RESULTS

#### 3.1. Simulation studies

Monte Carlo simulations were conducted to assess the performance of the proposed inference procedures. We considered randomised trials where  $n/2$  subjects receive each of the two groups. The times to disease occurrence and dependent censoring were given by

$$X_i^0 = \log T_{1i} + \theta_0 Z_i, \quad Y_i^0 = \log T_{2i} + \eta_0 Z_i,$$

where  $\eta_0 = -1$  and  $\theta_0 = 0$  or  $1$ . The baseline bivariate failure time vectors  $(T_{1i}, T_{2i})$  ( $i = 1, \dots, n$ ) were generated from two families of distributions: (i)  $T_{1i}$  and  $T_{2i}$  have Gumbel's bivariate exponential distribution with hazard rates of 1 and 0.5, respectively, and with correlation coefficient equal to  $r$  ( $0 \leq r \leq 0.25$ ); (ii)  $\log T_{1i}$  and  $\log T_{2i}$  have the bivariate normal distribution with means 0 and 1.2, respectively, with unit variances and with correlation coefficient equal to  $\rho$  ( $0 \leq \rho \leq 1$ ). The bivariate failure time vectors  $(X_i^0, Y_i^0)$  ( $i = 1, \dots, n$ ) were then subject to univariate censoring by the logarithm of an independent uniform  $(0, c)$  variable, where  $c = 10$  and  $20$  for the exponential and normal distributions, respectively.

Table 2 summarises the main results of the simulation studies. The entries were based on 10 000 and 2000 simulation samples for  $n = 100$  and  $500$ , respectively. The proposed estimator is virtually unbiased and the proposed confidence interval achieves accurate coverage probability. The coverage probability shown in the table pertains to the minimum-dispersion statistic. The results for the resampling method are similar.

For comparison, we also evaluated the naive approach which uses the log-rank estimating function  $S(\theta)$  mentioned in § 2.2. As expected, the naive approach tends to be more efficient than the proposed one when  $X^0$  and  $Y^0$  are independent because the latter approach censors some of the observed disease times in the analysis. When the two time variables are correlated, however, the naive approach yields a biased estimator and incorrect confidence interval. The bias of the estimator and the discrepancy of the coverage probability from the nominal level become worse as the sample size increases or as the correlation increases.

#### 3.2. Real example

We now return to the PCP prophylaxis study described in § 1. There is no significant difference between the two groups with respect to mortality, the  $p$ -value of the log-rank test being 0.32. In contrast, the observed log-rank  $\chi^2$ -statistic for the recurrence of PCP is 13.8, providing strong evidence for the superiority of TS over AP in reducing the cause-specific hazard function of PCP.

Table 2. Monte Carlo estimates for the biases and variances of the point estimators for  $\theta_0$  and for the coverage probabilities of the 0.95 confidence intervals for  $\theta_0$

Distribution	n	$\theta_0$	Bias		Variance		Cov. prob.	
			New	Naive	New	Naive	New	Naive
Exponential ( $r = 0$ )	100	0.0	0.019	0.008	0.110	0.090	0.96	0.95
		1.0	0.049	0.036	0.247	0.173	0.96	0.95
	500	0.0	0.003	0.002	0.021	0.018	0.95	0.95
		1.0	0.007	0.006	0.043	0.031	0.96	0.94
Exponential ( $r = 0.25$ )	100	0.0	0.018	0.114	0.100	0.097	0.96	0.93
		1.0	0.067	0.351	0.278	0.250	0.96	0.91
	500	0.0	0.003	0.106	0.018	0.018	0.96	0.87
		1.0	0.009	0.306	0.043	0.040	0.96	0.70
Normal ( $\rho = 0.25$ )	100	0.0	0.003	0.070	0.060	0.062	0.96	0.94
		1.0	0.017	0.191	0.095	0.087	0.97	0.91
	500	0.0	<0.001	0.073	0.012	0.013	0.96	0.89
		1.0	0.002	0.186	0.019	0.018	0.96	0.70
Normal ( $\rho = 0.5$ )	100	0.0	0.001	0.157	0.056	0.065	0.96	0.90
		1.0	0.018	0.456	0.094	0.104	0.97	0.72
	500	0.0	<0.001	0.161	0.012	0.014	0.95	0.70
		1.0	0.003	0.449	0.020	0.021	0.96	0.10

To assess the net PCP prophylaxis effects, we first consider death as the only dependent censoring factor. Assume that, on the natural logarithmic scale, the times to PCP recurrence and death satisfy the bivariate location-shift model. The group indicator  $Z$  takes the value 1 for TS, and the value 0 for AP. The main results of the analysis are given in Table 3. The point estimate of 0.72 suggests that, on the original scale, the average time to the recurrence of PCP in the TS group is about twice as long as that of the AP group. The confidence intervals based on the minimum-dispersion and resampling methods are very similar. Since the lower bound of the proposed 0.95 confidence interval excludes 0, we conclude that TS is more effective than AP in the PCP prophylaxis. The observed minimum-dispersion statistic for testing no group difference is 9.4, which is appreciably smaller than the aforementioned log-rank  $\chi^2$ -statistic. As shown in Table 3, the naive approach yields a slightly larger point estimate and a considerably higher lower bound for the confidence interval.

To check the assumed model, we plot in Fig. 1 the observed score processes along with

Table 3. Estimates of the location-shift parameters for the AIDS study

	Point estimate	0.95 confidence interval	
		Resampling†	Min. dispersion
Death	0.084	(-0.095, 0.26)	(-0.093, 0.27)
PCP			
proposed	0.720	(0.21, 2.05)	(0.20, 2.04)
naive	0.796	(0.41, 1.99)	(0.40, 1.99)

† Based on 10 000 simulation samples.

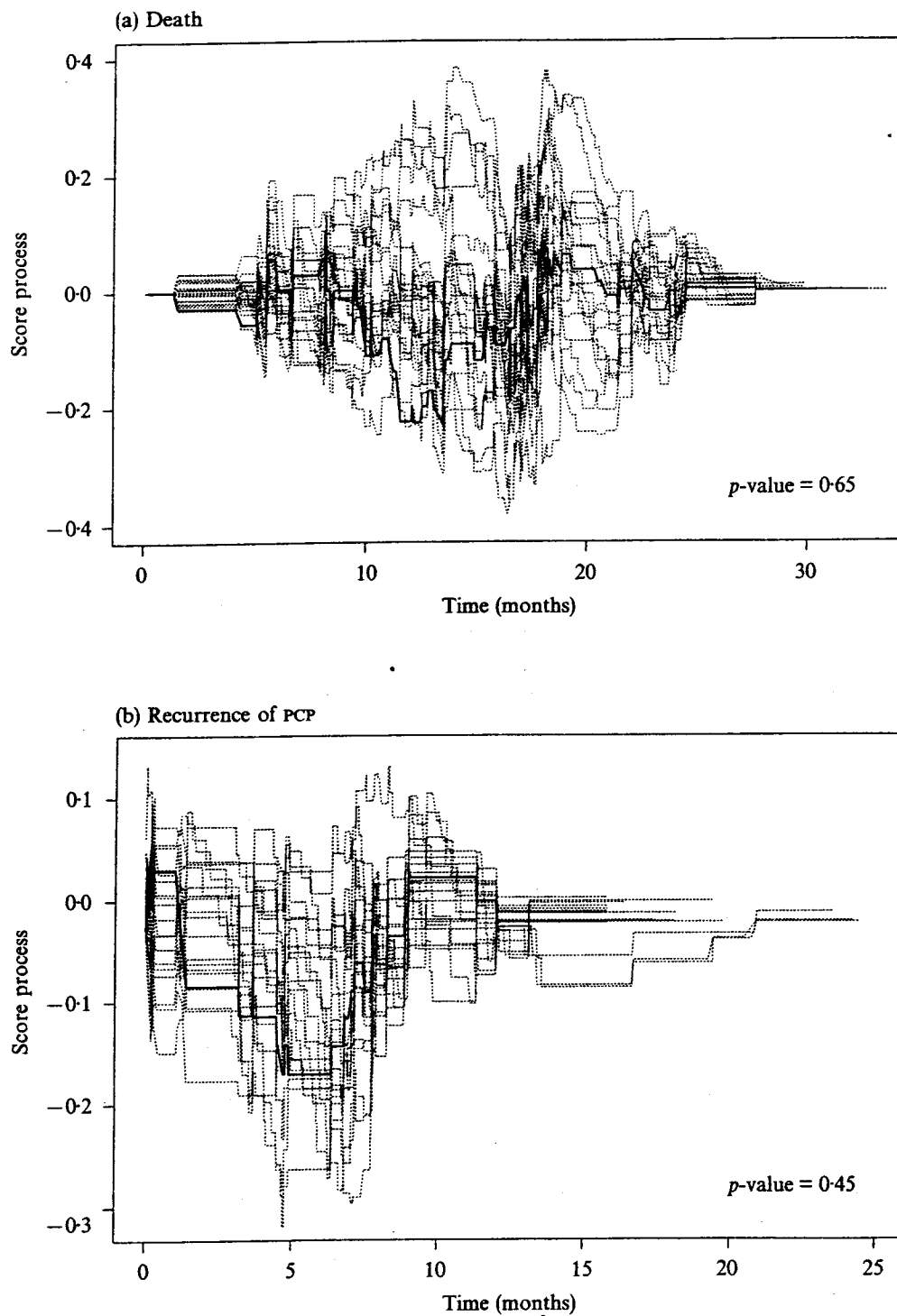


Fig. 1. Plots of score processes versus time for the AIDS study: (a) death, and (b) recurrence of PCP. The bold curve is the observed process. The dotted curves are 20 simulated realisations from the null distribution. The  $p$ -values are based on 1000 simulation samples.

20 realisations from the null distribution. These plots and the associated  $p$ -values for the supremum tests suggest that the model fits the data reasonably well.

As mentioned in § 1, quite a few patients were withdrawn from the prophylaxis study prematurely due to deteriorating health. Combining those early withdrawals with the deaths as the dependent censoring, we obtain 0.738 as the parameter estimate for PCP recurrence with a 0.95 confidence interval of (0.22, 2.00). The goodness-of-fit analysis indicates that the bivariate model is still adequate.

#### 4. REMARKS

The problem studied in this paper is somewhat different from the classical competing risks problem. In the latter case, one can only observe the minimum of the competing risk times, whereas in our setting disease occurrence does not preclude subsequent death. By taking advantage of this special data structure, we have been able to estimate the difference between the two marginal distributions of time to disease occurrence under the assumed model. The marginal distributions themselves, however, cannot be estimated without further assumptions. Like the existing methods for achieving identifiability in the competing risks problem, the proposed method depends on some assumptions that cannot be completely verified based on the available data. The minimal conditions necessary for our approach to lead to valid causal inferences are discussed by Robins (1995).

It is straightforward to extend the results of § 2 to sequential clinical trials. By expressing the log-rank estimating functions  $U_1$  and  $U_2$  in the forms of (A1.1) and (A1.2) at each interim look, we can obtain the asymptotic joint distribution over the looks by applying the multivariate central limit theorem. One may then use the resampling idea to obtain the repeated confidence intervals for  $\theta_0$  or the repeated confidence regions for  $\beta_0$ .

Following Wei & Gail (1983), we may also use weighted log-rank estimating functions to estimate  $\eta_0$  and  $\theta_0$ . The basic conclusions of § 2 hold for any weight function. In the presence of dependent censoring, however, the optimal weighted log-rank estimator for the independent censoring setting is no longer efficient. For the special case in which  $\eta_0$  is known to be zero and there is no independent censoring by  $C$ , Robins & Rotnitzky (1992, Appendix 4) and Robins (1993a, § 3.4, p. 281) proposed an  $M$ -estimator which is locally semiparametric efficient at a parametric submodel for the joint distribution of  $(X^0 - \theta_0 Z, Y^0)$ . In practice, however, it is difficult to obtain a truly optimal estimator because such a construction would inevitably require accurate estimation of  $\lambda_0(\cdot)$  and  $h_0(\cdot)$ .

We have confined our attention to the situation of a dichotomous covariate. It is straightforward to extend the proposed technique to an arbitrary bounded covariate. Suppose that  $-K_1 \leq Z \leq K_2$  with  $K_1, K_2 \geq 0$  and that the bivariate location-shift model of § 2.1 holds for this covariate. We shall replace  $d$  in (2.2) and (2.3) by  $-K_1(\theta - \eta)$  if  $\theta \leq \eta$ , and by  $K_2(\theta - \eta)$  otherwise. Then the subsequent results in § 2 still apply. A similar extension could be made for a multi-dimensional covariate vector with bounded components. It should be noted that this approach will be quite inefficient if  $K_1(\theta_0 - \eta_0)$  or  $K_2(\theta_0 - \eta_0)$  is large because then there will be excessive artificial censoring. In the special case of known zero  $\eta_0$ , the Robins–Rotnitzky  $M$ -estimator allows  $Z$  to be vector-valued with unbounded components. The generalisation of their estimator to the setting of unknown  $\eta_0$  is currently under investigation.

In many clinical trials, some measurements of biological activities are taken on the patients over the course of the study. These measurements may be of primary or secondary interest. For example, in AIDS clinical trials, CD4 counts are measured at certain time

intervals. The missing values on such measurements cannot be regarded as missing at random if the patients have died or withdrawn themselves from the study for health-related reasons. The idea presented in this paper may be used to adjust for such nonrandom missingness. The details will be communicated in separate reports.

APPENDIX 1

Asymptotic properties of  $U(\beta_0)$  and  $\hat{\beta}$

It is convenient to express the estimating function  $U(\beta)$  in the counting process notation introduced in § 2.4. Clearly, (2.1) and (2.4) can be written as

$$U_1(\eta) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^{\infty} \{Z_i - \bar{Z}^{(1)}(u; \eta)\} dN_{1i}(u; \eta),$$

$$U_2(\beta) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^{\infty} \{Z_i - \bar{Z}^{(2)}(u; \beta)\} dN_{2i}(u; \beta),$$

where

$$\bar{Z}^{(1)}(u; \eta) = \frac{\sum_{j=1}^n I\{\tilde{Y}_j(\eta) \geq u\} Z_j}{\sum_{j=1}^n I\{\tilde{Y}_j(\eta) \geq u\}}, \quad \bar{Z}^{(2)}(u; \beta) = \frac{\sum_{j=1}^n I\{\tilde{X}_j(\beta) \geq u\} Z_j}{\sum_{j=1}^n I\{\tilde{X}_j(\beta) \geq u\}}.$$

Simple algebraic manipulations yield

$$U_1(\eta_0) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^{\infty} \{Z_i - \bar{Z}^{(1)}(u; \eta_0)\} dM_{1i}(u),$$

$$U_2(\beta_0) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^{\infty} \{Z_i - \bar{Z}^{(2)}(u; \beta_0)\} dM_{2i}(u).$$

By the martingale central limit theorem (Fleming & Harrington, 1991, Theorem 5.3.5),

$$U_1(\eta_0) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^{\infty} \{Z_i - \bar{z}^{(1)}(u)\} dM_{1i}(u) + o_p(1), \tag{A1.1}$$

$$U_2(\beta_0) = n^{-\frac{1}{2}} \sum_{i=1}^n \int_{-\infty}^{\infty} \{Z_i - \bar{z}^{(2)}(u)\} dM_{2i}(u) + o_p(1), \tag{A1.2}$$

where  $\bar{z}^{(1)}(u)$  and  $\bar{z}^{(2)}(u)$  are the limits of  $\bar{Z}^{(1)}(u; \eta_0)$  and  $\bar{Z}^{(2)}(u; \beta_0)$ . Note that the right-hand sides of (A1.1) and (A1.2) are essentially sums of  $n$  independent and identically distributed random variables. It then follows from the multivariate central limit theorem and the Cramér-Wold device that the random vector  $U(\beta_0) = \{U_1(\eta_0), U_2(\beta_0)\}'$  is asymptotically zero-mean normal with covariance matrix

$$v = E \begin{pmatrix} w_{11}^2 & w_{11}w_{21} \\ w_{11}w_{21} & w_{21}^2 \end{pmatrix}, \tag{A1.3}$$

where

$$w_{1i} = \int_{-\infty}^{\infty} \{Z_i - \bar{z}^{(1)}(u)\} dM_{1i}(u), \quad w_{2i} = \int_{-\infty}^{\infty} \{Z_i - \bar{z}^{(2)}(u)\} dM_{2i}(u).$$

Note that the covariance matrix estimator  $V$  given in (2.6) is obtained from (A1.3) by replacing the unknown quantities in the  $w_{1i}$  and  $w_{2i}$  with their natural sample estimators.

Applying the techniques used in Theorem 1 of Ying (1993), we can show that, for  $\beta$  in a small

neighbourhood of  $\beta_0$ ,

$$U(\beta) = U(\beta_0) + \varphi n^{\frac{1}{2}}(\beta - \beta_0) + o_p(1), \quad (\text{A1.4})$$

where  $\varphi$  is a  $2 \times 2$  matrix of constants. Consequently,  $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$  is asymptotically zero-mean normal with covariance matrix  $\varphi^{-1}v\varphi^{-1}$ . The consistency of  $V$  follows from the arguments given in the proofs of Theorem 1 and Corollary 1 of Ying (1993).

We now justify the resampling approach. Recall that  $U(\beta^*) = -n^{-\frac{1}{2}} \sum W_i G_i$ , where  $W_i = (W_{1i}, W_{2i})'$ . It then follows from (A1.4) that

$$n^{\frac{1}{2}}(\beta^* - \hat{\beta}) = -\varphi^{-1}n^{-\frac{1}{2}} \sum_{i=1}^n W_i G_i + o_p(1). \quad (\text{A1.5})$$

The conditional distribution of the random vector  $n^{-\frac{1}{2}} \sum W_i G_i$  given the data  $(X_i, \delta_i, Y_i, \xi_i, Z_i)$  ( $i = 1, \dots, n$ ) is zero-mean normal with covariance matrix  $V$ . Hence, the conditional distribution of  $n^{\frac{1}{2}}(\beta^* - \hat{\beta})$  is asymptotically equivalent to the unconditional distribution of  $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$ .

## APPENDIX 2

### Asymptotic properties of $U(\cdot, \cdot; \hat{\beta})$

By the arguments leading to (A1.1), (A1.2), (A1.4) and (A1.5), we have

$$\begin{aligned} U(s, t; \hat{\beta}) &= U(s, t; \beta_0) - \varphi(s, t)\varphi^{-1}U(\beta_0) + o_p(1) \\ &= n^{-\frac{1}{2}} \sum_{i=1}^n \left( \int_{-\infty}^s \{Z_i - \bar{z}^{(1)}(u)\} dM_{1i}(u) \right. \\ &\quad \left. - \varphi(s, t)\varphi^{-1}n^{-\frac{1}{2}} \sum_{i=1}^n \left( \int_{-\infty}^{\infty} \{Z_i - \bar{z}^{(1)}(u)\} dM_{1i}(u) \right) \right. \\ &\quad \left. - \int_{-\infty}^t \{Z_i - \bar{z}^{(2)}(u)\} dM_{2i}(u) \right) + o_p(1), \end{aligned}$$

and

$$\begin{aligned} \hat{U}(s, t) &= n^{-\frac{1}{2}} \sum_{i=1}^n G_i \left( \int_{-\infty}^s \{Z_i - \bar{Z}^{(1)}(u; \hat{\eta})\} d\hat{M}_{1i}(u; \hat{\eta}) \right. \\ &\quad \left. - \int_{-\infty}^t \{Z_i - \bar{Z}^{(2)}(u; \hat{\beta})\} d\hat{M}_{2i}(u; \hat{\beta}) \right) + \varphi(s, t)n^{\frac{1}{2}}(\beta^* - \hat{\beta}) + o_p(1) \\ &= n^{-\frac{1}{2}} \sum_{i=1}^n G_i \left( \int_{-\infty}^s \{Z_i - \bar{Z}^{(1)}(u; \hat{\eta})\} d\hat{M}_{1i}(u; \hat{\eta}) \right. \\ &\quad \left. - \int_{-\infty}^t \{Z_i - \bar{Z}^{(2)}(u; \hat{\beta})\} d\hat{M}_{2i}(u; \hat{\beta}) \right) \\ &\quad - \varphi(s, t)\varphi^{-1}n^{-\frac{1}{2}} \sum_{i=1}^n G_i \left( \int_{-\infty}^{\infty} \{Z_i - \bar{Z}^{(1)}(u; \hat{\eta})\} d\hat{M}_{1i}(u; \hat{\eta}) \right. \\ &\quad \left. - \int_{-\infty}^{\infty} \{Z_i - \bar{Z}^{(2)}(u; \hat{\beta})\} d\hat{M}_{2i}(u; \hat{\beta}) \right) + o_p(1), \end{aligned}$$

where  $\varphi(s, t)$  is a deterministic matrix with  $\varphi(\infty, \infty) = \varphi$ . By the arguments given in Appendix 1 of Lin et al. (1993), the bivariate random processes  $U(\cdot, \cdot; \hat{\beta})$  and  $\hat{U}(\cdot, \cdot)$  are both asymptotically zero-mean Gaussian. Furthermore, analogous to the consistency of  $V$  for  $v$ , the conditional covariance function of  $\hat{U}(\cdot, \cdot)$  can be shown to converge to the limiting covariance function of  $U(\cdot, \cdot; \hat{\beta})$ . Therefore, the conditional distribution of  $\hat{U}(\cdot, \cdot)$  is the same in the limit as the unconditional distribution of  $U(\cdot, \cdot; \hat{\beta})$ .

## REFERENCES

- EFRON, B. (1967). The two sample problem with censored data. In *Proc. 5th Berkeley Symp. Math. Statist. Prob.*, 4, pp. 831–853. New York: Prentice-Hall.
- EMOTO, S. E. & MATTHEWS, P. C. (1990). A Weibull model for dependent censoring. *Ann. Statist.* 18, 1556–77.
- FLEMING, T. R. & HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. New York: Wiley.

- HARDY, W. D., FEINBERG, J., FINKELSTEIN, D. M. *et al.* (1992). A controlled trial of trimethoprim-sulfamethoxazole or aerosolized pentamidine for secondary prophylaxis of pneumocystis carinii pneumonia in patients with the acquired immunodeficiency syndrome. *New Engl. J. Med.* **327**, 1842–8.
- KALBFLEISCH, J. D. & PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- LIN, D. Y., WEI, L. J. & YING, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80**, 557–72.
- LINK, W. A. (1989). A model for informative censoring. *J. Am. Statist. Assoc.* **84**, 749–52.
- LOUIS, T. A. (1981). Non-parametric analysis of an accelerated failure time model. *Biometrika* **68**, 381–90.
- PARZEN, M. I., WEI, L. J. & YING, Z. (1994). A resampling method based on pivotal estimating functions. *Biometrika* **81**, 341–50.
- ROBINS, J. M. (1987a). Addendum to “A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect”. *Comp. Math. Applic.* **14**, 923–45.
- ROBINS, J. M. (1987b). A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J. Chron. Dis.* **40**, Suppl. 2: 139s–61s.
- ROBINS, J. M. (1989). The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology: A Focus on AIDS*, Ed. L. Sechrest, H. Freeman and A. Mulley, pp. 113–59. Washington, D.C.: NCHSR, U.S. Public Health Service.
- ROBINS, J. M. (1993a). Analytic methods for HIV treatment and cofactor effects. In *Methodological Issues of AIDS Behavioral Research*, Ed. D. G. Ostrow and R. Kessler, pp. 213–87. New York: Plenum.
- ROBINS, J. M. (1993b). Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In *Proc. Biopharm. Sect.* pp. 24–33. Washington, D.C.: Am. Statist. Assoc.
- ROBINS, J. M. (1995). An analytic method for randomized trials with informative censoring. *Life Time Data Anal.* To appear.
- ROBINS, J. M. & ROTNITZKY, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology—Methodological Issues*, Ed. N. Jewell, K. Dietz and V. Farewell, pp. 297–331. Boston, MA: Birkhäuser.
- ROBINS, J. M. & TSIATIS, A. A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Commun. Statist. A* **20**, 2609–31.
- TSIATIS, A. A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci.* **72**, 20–2.
- WEI, L. J. (1984). Testing goodness of fit for proportional hazards model with censored observations. *J. Am. Statist. Assoc.* **79**, 649–52.
- WEI, L. J. & GAIL, M. H. (1983). Nonparametric estimation for a scale-change with censored observations. *J. Am. Statist. Assoc.* **78**, 382–8.
- WEI, L. J., YING, Z. & LIN, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77**, 845–51.
- YING, Z. (1993). A large sample study of rank estimation for censored regression data. *Ann. Statist.* **21**, 76–99.

[Received May 1994. Revised July 1995]