UNIVERSITY OF WISCONSIN

DEPARTMENT OF BIOSTATISTICS

Technical Report #120

June 1997

A GENERAL CLASS OF FUNCTION-INDEXED NONPARAMETRIC TESTS FOR SURVIVAL ANALYSIS

Chin-Yu Lin
and
Michael R. Kosorok, PhD
Department of Biostatistics

MADISON, WISCONSIN

A GENERAL CLASS OF FUNCTION-INDEXED NONPARAMETRIC TESTS FOR SURVIVAL ANALYSIS

BY CHIN-YU LIN AND MICHAEL R. KOSOROK

University of Wisconsin–Madison

Many of the popular nonparametric test statistics for censored survival data used in two-sample, s-sample trend, and continuous covariate situations are special cases of a general statistic, differing only in the choice of the covariate-based label and weight function. A weight function determines the asymptotic efficiency of its corresponding statistic in this general class. Due to the mysterious nature of the possible alternatives, we are often unable to foresee which weight function is the best for the data. We show in this paper that certain large families of these statistics form stochastic processes doubly indexed by the weight function and the time scale which converge weakly to Gaussian processes also indexed by both the weight function and the time scale. These function-indexed processes are further generalized to accommodate settings with finite strata and are also shown to converge weakly to function-indexed Gaussian processes. These asymptotic properties allow us to develop versatile test procedures which are simultaneously sensitive to a reasonably large collection of alternatives. Due to the complexity of the Gaussian processes, a Monte Carlo approach is utilized to obtain the distributional characteristics of these statistics under the null hypothesis.

AMS 1991 subject classifications. Primary 62G10, 60F05; secondary 60G44, 62E25.

Key words and phrases. Censored data, contiguous alternatives, counting process, function-indexed stochastic process, martingale, Monte Carlo method, nonparametrics, weak convergence.

1. Introduction. Many popular nonparametric two-sample test statistics for censored survival data such as the log-rank, Peto-Peto (1972), and Gehan-Wilcoxon (Gehan, 1965) statistics have been shown to be special cases of a two-sample weighted log-rank statistic, differing only in the choice of weight function (Tarone and Ware, 1977; Gill, 1980). A poorly chosen weight function can result in less sensitivity to the actual observed treatment effects. For example, in Kosorok and Lin (1997), we observe that the effect of propranolol on patients having at least one myocardial infarction can be detected by the $G^{20,0}$-weighted log-rank statistics (from the $G^{\rho,\gamma}$ family of Harrington and Fleming, 1982) at a much earlier stage than the log-rank statistic after the β-Blocker Heart Attack Trial (BHAT) started in 1978 (β-Blocker Heart Attack Trial Research Group, 1982). To resolve problems like this, test procedures sensitive to a range of alternatives have been developed. For any given weight satisfying mild regularity conditions, there exists an alternative hypothesis for which the corresponding weighted log-rank statistic has maximum asymptotic relative efficiency (ARE) over all other weighted log-rank tests. Tarone (1981) and Fleming and Harrington (1991, Chapter 7) suggest selecting a finite number of relevant contiguous alternatives and then using the maximum of the corresponding collection of maximum ARE weighted log-rank statistics. Gastwirth (1985) proposed a similar idea; however, instead of taking the maximum, he used the linear combination of the same collection of statistics which maximized the minimum ARE over the set of alternatives. This Maximin Efficiency-Robust Test procedure can be applied to certain infinite collections of alternatives. For stochastic ordering alternatives, Fleming, Harrington, and O'Sullivan (1987) recommended a Renyi-type statistic which takes the supremum over time of the weighted log-rank statistics. Two-sample weighted log-rank statistics can be formulated as integrated, weighted differences of the estimated intensity processes from the two samples, with weights composed of nonnegative bounded predictable processes of bounded variation. Considering these stochastic processes to be doubly indexed by both the weight function and the time scale, Kosorok (1997) showed that they converge weakly on a useful large compact index set. This result offers us greater flexibility in selecting the alternative space over previous testing procedures and permits us to develop more efficient testing approaches.

In the BHAT study mentioned earlier, we may be interested in knowing the impact of weight, hypertention, and/or cigarette smoking on survival rates after adjusting for treatment effect in order to develop interventions for prolonging patients' lives. The previously described class of statistics does not allow us to do this, nor does it allow us to investigate the differences in survival rates for multiple treatment groups. Jones & Crowley (1989 & 1990) propose a class of single-covariate nonparametric tests for right-censored survival data that includes the Tarone-Ware (1977) two-sample class, the Cox (1972) score test, the Tarone (1975) s-sample trend statistics, the Brown et al (1974) modification of the Kendall rank statistic, the Prentice (1978) linear rank statistics, and the logit rank statistic of O'Brien (1978) as special cases. Certain large families of these statistics form stochastic processes doubly indexed by both the weight function and the time scale. In this paper, we generalize these single-covariate processes to allow for multiple covariates as well as for multiple covariates with finite strata, and establish their weak convergence over a useful function space and over the time scale. This class of statistics includes the class of Kosorok (1997) as a special case and can be applied to address the BHAT intervention question raised above. The formulation of multi-covariate processes in a single stratum as well as the hypothesis tests of interest are given in Section 2. Section 3.1 summarizes a general weak convergence theorem developed by Kosorok (1997) for stochastic processes which take on values in a complete metric space and which are indexed by elements in a compact set, and its application to the function-indexed processes on $(\mathbf{D}[0,\infty])^p$. Section 3.2 contains the asymptotic properties of the processes described in Section 2. In Section 4, we propose some test procedures which are simultaneously sensitive to ordered hazard and stochastic ordering alternatives and develop a Monte Carlo approach to obtain the null distributions. In Section 5, we generalize the statistics of Section 2 to accommodate the situation where strata are present.

2. Function-indexed Stochastic Processes. We will introduce the general class of nonparametric tests mentioned above and state the hypotheses of interest in this section.

For a sample of survival data of size n, let T_j and C_j represent the times to failure and censoring, respectively, and let $Z_j(t)$ be the covariates measured at time t for individual j. Define the observed failure counting process

$$N_j(t) = I_{\{T_j \wedge C_j \ \le t, \delta_j = 1\}},$$

and the at-risk process

$$Y_j(t) = I_{\{T_j \wedge C_j \ \ge t\}},$$

where I is an indicator function and $j = 1, ..., n$. Let

$$\overline{N}(t) = \sum_{j=1}^{n} N_j(t) \ \text{ and } \ \overline{Y}(t) = \sum_{j=1}^{n} Y_j(t),$$

and define the filtrations to be the histories of the study up to and including time t :

$$\mathcal{F}_t^n = \sigma\{N_j(s),\ Y_j(s+),\ Z_j(s+),\ j = 1 \ldots n,\ s \le t\}, \quad n = 1, 2, ...,$$

where $\sigma\{A\}$ is the smallest σ−field making all of A measurable. We will work only under the general random censorship model, i.e.,

$$P\{T \in [t, t+\Delta t), C \in [t, t+\Delta t) \mid \underline{Z}(t)\} = P\{T \in [t, t+\Delta t) \mid \underline{Z}(t)\} P\{C \in [t, t+\Delta t) \mid \underline{Z}(t)\},$$

where $\underline{Z}(t) = \{Z(s) : 0 \le s \le t\}$.

Denote the cumulative hazard by Λ and allow it to depend on n. Throughout, the covariate is assumed to be well-constructed so that the jth individual's hazard at time t is a function of $Z_j(t)$, i.e. $d\Lambda_j(t) = d\Lambda(t \mid \underline{Z}_j(t)) = d\Lambda(t \mid Z_j(t))$. Under the above assumptions and certain regularity conditions, Dolivo (1974) showed that

$$M_j(t) = N_j(t) - \int_0^t Y_j(s) d\Lambda(s \mid Z_j(s)) \tag{2.1}$$

are square integrable martingales over $[0, \infty)$ with predictable covariation

$$\langle M_i, M_j \rangle (t) = I_{\{i=j\}} \int_0^t Y_j(s) \left[1 - \Delta\Lambda(s \mid Z_j(s))\right] d\Lambda(s \mid Z_j(s)). \tag{2.2}$$

Jones and Crowley (1989) proposed a class of single-covariate nonparametric tests:

$$X^n(t) = n^{-1/2} \int_0^t w(s) \sum_{j=1}^{n} Y_j(s)[Z_j(s) - \overline{Z}(s)] dN_j(s), \tag{2.3}$$

where $w(s)$ is a locally bounded, predictable weight function, and $\overline{Z}(s) = \overline{Y}^{\,-1}(s)\sum Y_j(s)Z_j(s)$. They showed that X^n, specified by particular choices of covariates and weight functions, are equivalent to several well-known test statistics. For example, if the Z_j are either 0 or 1 (group indicator), then the X^n are two-sample weighted log-rank statistics. They also proposed a variance estimator for $X^n(t)$:

$$(2.4) \qquad V^n(t) = n^{-1}\int_0^t w^2(s)\sum_{j=1}^{n} Y_j(s)[Z_j(s) - \overline{Z}(s)]^2 \frac{\overline{Y}(s) - \Delta\overline{N}(s)}{\overline{Y}(s) - 1}\frac{d\overline{N}(s)}{\overline{Y}(s)},$$

where $\Delta\overline{N}(s) = \overline{N}(s) - \overline{N}(s-)$.

This paper focuses on the family with weight functions of the form $f\{\boldsymbol{b}^n(s)\}$, where $f : [0,1]^r \mapsto [0,1]$ is the "function-index" and is an element of a function space $\mathbf{H}$ which is either equal to or is a uniform-metric closed subset of $\mathbf{G}_r^+(K)$ (defined below), where $K < \infty$; $\boldsymbol{b}^n(t) = \{b_1^n(t), \ldots, b_r^n(t)\}^T$, where superscript T denotes transpose, each $b_l^n : [0,u] \mapsto [0,1]$ is an $\mathcal{F}_t^n$-predictable process, and $u \in (0,\infty]$; and where r is finite.

Before defining $\mathbf{G}_r^+(K)$, we need to introduce some additional notation. Let N_1^r be the set of all multi-indexes $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_r\}$, where α_l is either 0 or 1, $l = 1 \ldots r$; and define the linear operator

$$D_0^{\boldsymbol{\alpha}} \equiv \prod_{l=1}^{r}\left(\frac{\partial}{\partial x_l}\right)^{\alpha_l}$$

evaluated at $\boldsymbol{y} = \{y_1, \ldots, y_r\}^T$, where $y_l = \alpha_l x_l$, $l = 1 \ldots r$. For each $f : [0,1]^r \mapsto [0,1]$, also define

$$\|f\|_*^r \equiv \left(\sum_{\boldsymbol{\alpha}\in N_1^r}\int_{[\mathbf{0},\mathbf{1}]}\left[D_0^{\boldsymbol{\alpha}} f(\boldsymbol{s})\right]^2 d\boldsymbol{s}\right)^{1/2},$$

where $[\mathbf{0},\mathbf{1}] = [0,1]^r$ and $\boldsymbol{s} = \{s_1, \ldots, s_r\}^T$.

DEFINITION 1 *Let $\mathbf{G}_r^+(K)$ denote the space of absolutely continuous functions f mapping from $[0,1]^r$ to $[0,1]$ for which the total of the L_2-norms of all first cross-partial derivatives are bounded by K, in the sense that $\|f - f(\mathbf{0})\|_*^r \leq K$.*

We generalize Jones and Crowley's single-covariate statistics to accommodate the multiple-covariate situation. Suppose p covariates are obtained for each subject. Denote the covariate vector by $Z_j(t) = \{Z_{j1}(t), ..., Z_{jp}(t)\}^T$, and define $X^n(f,t) = \left\{X_1^n(f,t), ..., X_p^n(f,t)\right\}^T$ with the kth element

$$X_k^n(f,t) \quad = \quad n^{-1/2} \int_0^t f\{\boldsymbol{b}^n(s)\} \sum_{j=1}^{n} Y_j(s)[Z_{jk}(s) - \overline{Z}_k(s)] dN_j(s), \tag{2.5}$$

where $\overline{Z}_k(s) = \overline{Y}^{\,-1}(s) \sum Y_j(s) Z_{jk}(s)$, $k = 1, ..., p$, and $j = 1, ..., n$. Let $\boldsymbol{V}^n$ be the covariance estimate of $X^n(f, \cdot)$ with (k,l)th element defined as

$$V_{kl}^n(s,t) = n^{-1} \int_0^{s \wedge t} f^2\{\boldsymbol{b}^n(x)\} \sum_{j=1}^{n} Y_j(x)[Z_{jk}(x) - \overline{Z}_k(x)][Z_{jl}(x) - \overline{Z}_l(x)] \frac{\overline{Y}(s) - \Delta\overline{N}(s)}{\overline{Y}(s) - 1} \frac{d\overline{N}(x)}{\overline{Y}(x)},$$

k, $l = 1, ..., p$.

Let $\mathcal{Z}$ represent the space of potential covariate paths or some subspace of it. All the processes in $\mathcal{Z}$ are assumed to be adapted, bounded, and left-continuous with right-hand limits without loss of practical generality. Let Λ_z denote the cumulative hazard for the path $z \in \mathcal{Z}$. We are interested in testing the null hypothesis $H_0 : \Lambda_z^n = \Lambda$ over $[0, \infty)$ for all $z \in \mathcal{Z}$, against the proportional odds contiguous alternatives of the form

$$d\Lambda^n(t \mid z(t)) = \frac{\exp\left[\frac{\beta^T z(t)}{\sqrt{n}} g(t)\right] d\Lambda(t)}{1 + (\exp\left[\frac{\beta^T z(t)}{\sqrt{n}} g(t)\right] - 1)\Delta\Lambda(t)}, \tag{2.6}$$

for some finite $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^T$, some baseline hazard $\Lambda(t)$, and some real function g such that g is bounded by $G < \infty$. Note that (2.6) becomes the proportional hazards contiguous alternative model when the baseline hazard is continuous.

3. Weak Convergence of Function-Indexed Stochastic Processes.

A general weak convergence theorem developed by Kosorok (1997) is used to establish the asymptotic properties of X^n in (2.5). This theorem and its application to the function-indexed stochastic processes on $(\mathbf{D}[0,\infty])^p$ will be stated in Section 3.1, and the main results for X^n will be given in Section 3.2.

3.1. *A General Weak Convergence Theorem and its Application.* In order to resolve the measurability problems in non-separable metric space, the following theorems utilize Hoffman-Jorgensen-Dudley (HJD) weak convergence theory as described in van der Vaart and Wellner (1996). Let $\boldsymbol{A}(\mathbf{H},\mathbf{D})$ be the space of continuous mappings from the metric space $\{\mathbf{H},\rho\}$ to the metric space $\{\mathbf{D},d\}$, where ρ and d are the respective metrics, and let $\mathcal{D}$ be the Borel σ-algebra of $\{\mathbf{D},d\}$. For $x(\cdot),y(\cdot)\in\boldsymbol{A}(\mathbf{H},\mathbf{D})$, define the metric

$$a(x,y)=\sup_{f\in\mathbf{H}} d(x(f),y(f)).$$

Let $\mathcal{A}$ be the Borel σ-algebra of $\{\boldsymbol{A}(\mathbf{H},\mathbf{D}),a\}$.

THEOREM 1 *Suppose a sequence of stochastic processes $\{X^n(\cdot)\}$ in $\boldsymbol{A}(\mathbf{H},\mathbf{D})$, with $\{\mathbf{H},\rho\}$ compact and $\{\mathbf{D},d\}$ complete, satisfies the following conditions :*

1. *$X^n(f)$ is a $\mathcal{D}^*$-measurable random variable for all $f\in\mathbf{H}$ and all $n\geq 1$, where $\mathcal{D}^*\subset\mathcal{D}$; and all finite dimensional distributions of the form $\{X^n(f_1),X^n(f_2),\ldots,X^n(f_m)\}^T$ converge HJD-weakly on $(\{\mathbf{D},d\})^m$ to some $\{X(f_1),X(f_2),\ldots,X(f_m)\}^T$ as $n\to\infty$.*

2. $\forall f,g\in\mathbf{H}$,

$$d(X^n(f),X^n(g))\leq q\,(\rho(f,g))\,U_n,$$

where

(i) d is a bounded, continuous, $\mathcal{D}^*\times\mathcal{D}^*$-measurable mapping from $\mathbf{D}\times\mathbf{D}$ to $[0,\infty)$,

(ii) q is continuous and non-decreasing with $q(0)=0$, and

(iii) $\{U_n\}$ is a stochastically bounded sequence of real random variables, i.e., for any $\epsilon>0$, there exists a $\tau<\infty$ such that $P\{U_n\leq\tau\}>1-\epsilon$, $\forall n\geq 1$.

3. *$h(\cdot)$ is a uniformly continuous mapping from $\{\mathbf{D},d\}$ to some complete and separable metric space $\{\boldsymbol{V},v\}$ and is measurable with respect to $\mathcal{D}^*$. Let $\tilde{\mathbf{A}}\equiv\boldsymbol{A}(\mathbf{H},\boldsymbol{V})$; for any $x,y\in\tilde{\mathbf{A}}$, define*

$$\tilde{a}(x,y)\equiv\sup_{f\in\mathbf{H}} v(x(f),y(f)),$$

where v is continuous and $h(\mathcal{D}^*) \times h(\mathcal{D}^*)-$ measurable; and let $\tilde{\mathcal{A}}$ be the Borel σ-algebra of $\{\tilde{\mathbf{A}}, \tilde{a}\}$.

Then

(a) $X^n(\cdot)$ converges HJD-weakly in $\{\mathbf{A}(\mathbf{H}, \mathbf{D}), a\}$ to $X(\cdot)$ as $n \to \infty$,

(b) $\tilde{\mathcal{A}}$ is a measurable σ-algebra, and

(c) $h(X^n(\cdot))$ converges weakly (in the usual sense) to $h(X(\cdot))$ on the $\tilde{\mathcal{A}}$-topology.

See Kosorok (1997) for the proof.

THEOREM 2 *For each $f \in \mathbf{H}$, where $\mathbf{H}$ is either equal to or a closed subset of $\mathbf{G}_r^+(K)$ for some $K < \infty$ and $r < \infty$, let $X^n(f)$ be a p-dimensional vector of stochastic processes with kth element*

$$X_k^n(f) \equiv X_k^n(f,.) = \int_0^{(\cdot)} f\{\boldsymbol{b}^n(s)\} dU_k^n(s), \tag{3.1}$$

where $\boldsymbol{b}^n = \{b_1^n, ..., b_r^n\}^T$. Suppose we have the following conditions for some sequence of histories $\mathcal{F}^n = \{\mathcal{F}_t^n, t \geq 0\}$:

1. *For each $k = 1, ..., p$, U_k^n is a cadlag $\mathcal{F}_t^n$-adapted process on $[0, u]$ such that*

(i) U_k^n is locally of bounded variation on $[0, u]$ and $\Delta U_k^n(0) = 0$, $\forall n \geq 1$, and

(ii) $\exists C_k < \infty$ such that for any $\mathcal{F}_t^n$-predictable h^n, with $|h^n| \leq c$, there exists a nonnegative $\mathcal{F}_t^n$-adapted right-continuous submartingale $G_k^n(h^n, t)$ such that

$$\left\{ \int_0^t h^n(s) dU_k^n(s) \right\}^2 \leq G_k^n(h^n, t)$$

$\forall t \in [0, u]$ and $E[G_k^n(h^n, u)] \leq c^2 C_k$, $\forall n \geq 1$.

2. *$\forall n \geq 1$, each b_j^n, for $1 \leq j \leq r$, is an $\mathcal{F}_t^n$-predictable process mapping from $[0, u]$ to $[0, 1]$.*

3. $X^n(f)$ *is measurable with respect to the Skorohod topology on* $(\mathbf{D}[0,u])^p$ *for all* $f \in \mathbf{H}$ *and all* $n \geq 1$*; and all finite dimensional distributions of the form* $\{X^n(f_1), \ldots, X^n(f_m)\}^T$*, for* $m < \infty$*, converge HJD-weakly in the uniform topology on* $(\mathbf{D}[0,u])^{pm}$ *to some* $\{X(f_1), \ldots, X(f_m)\}^T$ *as* $n \to \infty$.

Then

(a) $X^n(\cdot)$ *converges HJD-weakly in the uniform topology on* $\boldsymbol{A}(\mathbf{H}, (\mathbf{D}[0,u])^p)$ *to* $X(\cdot)$*; and*

(b) for any $t \in [0,u]$*, the collection*

$$\left\{X^n(f,t), \sup_{s\in[0,t]} X_k^n(f,s), \inf_{s\in[0,t]} X_k^n(f,s), k=1,...,p, \sup_{s\in[0,t]} X^n(f,s)^T X^n(f,s)\right\}$$

consists of measurable random variables which jointly converge weakly in the uniform topology over all $f \in \mathbf{H}$ *to the corresponding collection*

$$\left\{X(f,t), \sup_{s\in[0,t]} X_k(f,s), \inf_{s\in[0,t]} X_k(f,s), k=1,...,p, \sup_{s\in[0,t]} X(f,s)^T X(f,s)\right\}.$$

REMARK 1 *Part (b) of Theorem 2 enables us to construct Renyi-type extensions of the supremum-over-*$f \in \mathbf{H}$ *test procedures, and gives us confidence that we will not lose too much power while using these supreme-type statistics over a reasonable large function space* $\mathbf{H}$.

Proof. For any f, $g \in \mathbf{H}\cup\{0\}$, define the metric

$$\rho(f,g) = \sup_{\boldsymbol{x}\in[0,1]} |f(\boldsymbol{x}) - g(\boldsymbol{x})|.$$

Let d_u be the uniform metric on $\mathbf{D}[0,u]$. Kosorok (1997) showed that the metric space $\{\mathbf{H}, \rho\}$ is compact, and for each $k = 1, ..., p$, there exists a sequence (in n) of stochastically bounded random variables $\{U_{nk}\}$ such that $\forall f, g \in \mathbf{H} \cup \{0\}$,

$$d_u(X_k^n(f), X_k^n(g)) \leq [\rho(f,g)]^{\frac{1}{4}} U_{nk}, \quad \forall n \geq 1.$$

Define $d(X^n(f), X^n(g)) = \max_{1\le k\le p} d_u(X_k^n(f), X_k^n(g))$, and U_n=$\max_{1\le k\le p} U_{nk}$. Then,

$$d(X^n(f), X^n(g)) \le [\rho(f,g)]^{\frac{1}{4}} U_n, \quad \forall n \ge 1,$$

and $\{U_n\}$ is also a sequence of stochastically bounded random variables. By Theorem 1, $X^n(\cdot)$ converges HJD-weakly in the uniform topology on $\boldsymbol{A}(\mathbf{H}, (\mathbf{D}[0,u])^p)$ to $X(\cdot)$.

Result (b) follows from Theorem 1 (c) by the Skorohod-measurability and uniform continuity of the mapping from $\mathbf{A}(\mathbf{H}, (\mathbf{D}[0,u])^p)$ to the finite-dimensional Euclidean space used to generate the given collection. □

3.2. *Asymptotic Properties.* We will show in this section that the function-indexed process X^n with elements as given in (2.5) converges HJD-weakly in the uniform topology on $\boldsymbol{A}(\mathbf{H}, (\mathbf{D}[0,\infty])^p)$ to a multivariate Gaussian process X.

THEOREM 3 *Let* $\boldsymbol{b}^n = (b_1^n, ..., b_r^n)^T$, $b_i^n \in \mathbf{B}(\mathcal{F}^n)$ *(defined in Definition 2 below),* $\forall i = 1, ..., r$. *Under the general random censorship model and the "contiguous alternative" sequence (2.6), suppose the following conditions hold :*

1. *There exists a function* $\pi : [0,\infty) \mapsto [0,1]$ *such that*

$$\sup_{t\in[0,\infty)} \left| \frac{\overline{Y}(t)}{n} - \pi(t) \right| \xrightarrow{P} 0 \quad as\ n \longrightarrow \infty.$$

2. *The covariate processes are adapted, bounded, and left-continuous with right-hand limits.*

Set $\mathcal{I}$ =*sup*$\{t : \pi(t) > 0\}$ *and* u =*sup* $\mathcal{I}$.

3. *For any* $t \in \mathcal{I}$,

   (i) $\Lambda(t) < \infty$.

(ii) *There exist left-continuous functions v_{kl}, k, $l = 1, \ldots, p$, with right-hand limits such that for all $t \in \mathcal{I}$,*

$$\sup_{s\in[0,t]} \left| \overline{Y}(s)^{-1} \sum_{j=1}^{n} Y_j(s) \left[Z_{jk}(s) - \overline{Z}_k(s)\right] \left[Z_{jl}(s) - \overline{Z}_l(s)\right] - v_{kl}(s) \right| \xrightarrow{P} 0$$

and $v_{kl}(s)$ are zero outside of $\mathcal{I}$.

Then,

(a) *$X^n(\cdot,\cdot)$ converges HJD-weakly in the uniform topology on $\mathbf{A}(\mathbf{H}, (\mathbf{D}[0,\infty])^p)$ to a multivariate Gaussian process $X(\cdot,\cdot)$ with mean function*

$$\boldsymbol{\mu}(f,t) = \int_0^t f\{\boldsymbol{b}(s)\} \left[\boldsymbol{v}(s)\boldsymbol{\beta}\right] g(s) d\Lambda(s), \tag{3.2}$$

and covariance function

$$\boldsymbol{V}_{fg}(s,t) = \int_0^{s\wedge t} f\{\boldsymbol{b}(x)\} g\{\boldsymbol{b}(x)\} \boldsymbol{v}(s)\pi(s) \left[1 - \Delta\Lambda(s)\right] d\Lambda(s), \tag{3.3}$$

for all $f, g \in \mathbf{H}$ and all $s,t \in [0,\infty]$, where $\boldsymbol{v}$ is a $p \times p$ matrix with (k,l)th element v_{kl};

(b) *For any $t \in [0,\infty]$, the collection*

$$\left\{ X^n(f,t), \sup_{s\in[0,t]} X_k^n(f,s), \inf_{s\in[0,t]} X_k^n(f,s), k = 1,\ldots,p, \sup_{s\in[0,t]} \left[X^n(f,s)\right]^T \left[X^n(f,s)\right] : t \in T \right\}$$

consists of measurable random variables which jointly converge weakly in the uniform topology over all $f \in \mathbf{H}$ to the corresponding collection

$$\left\{ X(f,t), \sup_{s\in[0,t]} X_k(f,s), \inf_{s\in[0,t]} X_k(f,s), k = 1,\ldots,p, \sup_{s\in[0,t]} X^T(f,s)X(f,s) : t \in T \right\};$$

(c) *The covariance estimator*

$$\begin{aligned} \boldsymbol{V}_{fg}^n(s,t) \;=\; & n^{-1} \int_0^{s\wedge t} f\{\boldsymbol{b}^n(x)\} g\{\boldsymbol{b}^n(x)\} \sum_{j=1}^{n} Y_j(x)[Z_j(x) - \overline{Z}(x)] \times \\ & [Z_j(x) - \overline{Z}(x)]^T \frac{\overline{Y}(x) - \Delta\overline{N}(x)}{\overline{Y}(x) - 1} \frac{d\overline{N}(x)}{\overline{Y}(x)} \end{aligned} \tag{3.4}$$

is uniformly consistent for $\boldsymbol{V}_{fg}$ over all $f, g \in \mathbf{H}$ and all $s,t \in [0,\infty]$.

DEFINITION 2 *Let $\mathbf{B}(\mathcal{F}^n)$ be the class of all random sequences of functions $b^n : [0,\infty) \mapsto [0,1]$ such that for each $n \geq 1$ and $t \in [0,\infty)$, $b^n(t)$ is $\mathcal{F}_t^n$-predictable and, for each closed subinterval of $\mathcal{I}$, $\mathcal{I}^* \subset \mathcal{I}$, the following holds:*

$$\sup_{s \in \mathcal{I}^*} |b^n(s) - b(s)| \to 0$$

in probability, as $n \to \infty$, for some $b : [0,\infty) \mapsto [0,1]$; where b is left-continuous with right hand limits and $db^+(t)$, where b^+ is the right-continuous version of b, changes sign only a finite number of times over $[0,\infty)$.

Examples of these function-indexed families include:

- $f\{\boldsymbol{b}^n(s)\} \equiv 1$;
- $r = 1$, $f(x) = x$, and $b_1^n(s) = \widehat{\pi}(s) = \overline{Y}(s)/n$; and
- $r = 2$, $f(x_1, x_2) = x_1^{\rho/\epsilon_1} x_2^{\gamma/\epsilon_2}$, and $b_1^n(s) = \{S_p^n(s-)\}^{\epsilon_1}$, $b_2^n(s) = \{1 - S_p^n(s-)\}^{\epsilon_2}$, where $S_p^n(s)$ is the pooled Kaplan-Meier estimator, $\rho \in \{0\} \cup [\epsilon_1, \tau_1]$, $\gamma \in \{0\} \cup [\epsilon_2, \tau_2]$, and $0 < \epsilon_i \leq \tau_i < \infty$, $i = 1, 2$. This is the $G^{\rho,\gamma}$-weight defined in Harrington and Fleming (1982).

Proof of Theorem 3. Theorem 2 is used to develop this theorem. Condition (2) of Theorem 2 follows by the definitions. Condition (3)—weak convergence of all finite dimensional distributions—is verified in the Appendix. Therefore, (a) and (b) are established if Condition (1) of Theorem 2 can be obtained. Let

$$U_k^n(\cdot) = n^{-1/2} \int_0^{(\cdot)} \sum_{j=1}^n Y_j(s)[Z_{jk}(s) - \overline{Z}_k(s)]dN_j(s),$$

then

$$X_k^n(f, \cdot) \;=\; \int_0^{(\cdot)} f\{\boldsymbol{b}^n(s)\}\, dU_k^n(s).$$

Because $Y_j(Z_{jk} - \overline{Z}_k)$, $j = 1, ..., n$, are bounded, left-continuous, predictable processes with right-hand limits, U_k^n is a cadlag adapted process and is of bounded variation on $[0,\infty)$, for $k = 1, ..., p$. In addition, $\Delta U_k^n(0) = 0$, $\forall n \geq 1$ and $\forall\, k$.

Since

$$U_k^n(t) = n^{-1/2} \int_0^t \sum_{j=1}^n Y_j(s)\,[Z_{jk}(s) - \overline{Z}_k(s)]\,dM_j(s) + n^{-1/2} \int_0^t \sum_{j=1}^n Y_j(s)\,[Z_{jk}(s) - \overline{Z}_k(s)]\,d\Lambda_j^n(s),$$

for any predictable h^n with $|h^n| \le c$, where $c < \infty$,

$$\left\{\int_0^t h^n dU_k^n\right\}^2 \le 2\left\{n^{-\frac{1}{2}} \int_0^t h^n \sum_{j=1}^n Y_j(Z_{jk} - \overline{Z}_k)dM_j\right\}^2 + 2\left\{n^{-\frac{1}{2}} \int_0^t h^n \sum_{j=1}^n Y_j(Z_{jk} - \overline{Z}_k)d\Lambda_j^n\right\}^2,$$
$$\equiv G_k^n(h^n, t), \qquad \forall t \in [0, \infty],$$

where $G_k^n(h^n, t)$ is a nonnegative right-continuous $\mathcal{F}_t^n$-submartingale, for $k = 1, ..., p$. In addition,

$$\mathrm{E}\left[n^{-1/2} \int_0^t h^n \sum_{j=1}^n Y_j\,(Z_{jk} - \overline{Z}_k)\,dM_j\right]^2$$
$$\le c^2\ \mathrm{E}\left[n^{-1} \int_0^t \sum_{j=1}^n Y_j\,\left(Z_{jk} - \overline{Z}_k\right)^2 d\Lambda_j^n\right]$$
$$= c^2\ \mathrm{E}\left[n^{-1} \int_0^t \sum_{j=1}^n \left(Z_{jk} - \overline{Z}_k\right)^2 dN_j\right] \le c^2\,(2M)^2,$$

for all $t \in [0, \infty]$, where $\sup_{j,k} \sup_{t\in[0,\infty)} |Z_{jk}(t)| \le M < \infty$.

Observe that if $\Lambda_j^n = \Lambda$,

$$n^{-1/2} \int_0^t h^n \sum_{j=1}^n Y_j\,(Z_{jk} - \overline{Z}_k)\,d\Lambda = 0, \quad \forall t \in [0, \infty].$$

Thus, condition (1) of Theorem 2 holds under the null hypothesis. Now,

$$n^{-1/2} \int_0^t h^n \sum_{j=1}^n Y_j\,\left[Z_{jk} - \overline{Z}_k\right]\,d\Lambda_j^n$$
$$= n^{-1} \int_0^t h^n \sum_{j=1}^n Y_j\,\left[Z_{jk} - \overline{Z}_k\right]\,\sqrt{n}\left[d\Lambda_j^n - d\Lambda\right].$$

Note that $x^{-1}\,|e^{ax} - 1| \le |a|e^{|a|}$, $\forall x \in (0, 1]$ and any real a. Therefore,

$$
\begin{aligned}
&\sqrt{n}\left|d\Lambda_j^n(s) - d\Lambda(s)\right| \\
&= \sqrt{n}\left|\exp\left[\frac{\beta^T z(s)g(s)}{\sqrt{n}}\right] - 1\right| \times \left|1 - \frac{\exp\left[\frac{\beta^T z(s)g(s)}{\sqrt{n}}\right]\Delta\Lambda(s)}{[1-\Delta\Lambda(s)] + \exp\left[\frac{\beta^T z(s)g(s)}{\sqrt{n}}\right]\Delta\Lambda(s)}\right| d\Lambda(s) \\
&\leq 2MG\sum_{l=1}^{p}|\beta_l| \exp\left(MG\sum_{l=1}^{p}|\beta_l|\right) d\Lambda(s).
\end{aligned}
$$

Thus

$$
\begin{aligned}
E\left[n^{-\frac{1}{2}}\int_0^t h^n \sum_{j=1}^{n} Y_j \left[Z_j - \overline{Z}\right] d\Lambda_j^n\right]^2 &\leq c^* E\left[n^{-1}\int_0^t \sum_{j=1}^{n} Y_j \, d\Lambda\right]^2 \\
&\leq c^{**} E\left[n^{-1}\int_0^\infty \sum_{j=1}^{n} Y_j \, d\Lambda_j^n\right]^2 \\
&= c^{**} E\left[\frac{\sum_{j=1}^n \Lambda_j^n(T_j)}{n}\right]^2 \\
&\leq c^{**} E\left[\frac{\sum_{j=1}^n \{\Lambda_j^n(T_j)\}^2}{n}\right], \qquad (3.5)
\end{aligned}
$$

for some finite c^* and c^{**}. The second inequality follows by (2.6) and the last by the Cauchy inequality. For any failure time distribution F with corresponding survival function S and cumulative hazard Λ, and any $t \in [0, \infty)$, we can show that $\int_0^t \Lambda^2(s)dF(s) \leq 2\int_0^t \Lambda(s)dF(s)$ and $\int_0^t \Lambda(s)dF(s) \leq 1$ by the integration by parts technique and the relationship between F, S, and Λ. Thus, (3.5) is bounded above by $2c^{**}$. (a) and (b) then follow by Theorem 2.

Since conditions (1)—(3) of Theorem 2 imply conditions (1) and (2) of Theorem 1, the uniform consistency of $\boldsymbol{V}^n$ follows from Theorem 1.(c) by the Skorohod-measurability of the sup-norm metric on $(\mathbf{D}[0,\infty])^p$. □

Under some regularity conditions, the above results can be generalized to the contiguous alternative sequences discussed in Jones and Crowley (1990):

$$
\sup_{z} \sup_{t\in[0,\infty)} |d\Lambda_z^n(t) - d\Lambda(t)| \longrightarrow 0, \quad \text{as } n \longrightarrow \infty. \qquad (3.6)
$$

COROLLARY 1 *Under the contiguous alternative sub-model (3.6), suppose conditions (1)—(3) of Theorem 3 obtain and for any predictable process h^n with $|h^n| \leq c$, $c < \infty$,*

$$\sup_{1\leq k\leq p} E\left\{ n^{-1/2}\int_0^t h^n(s)\sum_{j=1}^n Y_j(s)[Z_{jk}(s)-\overline{Z}_k(s)]d\Lambda_j^n \right\}^2 \leq C,$$

for all $n \geq 1$, $t \in [0,\infty]$, and some $C < \infty$. Then, results (a)—(c) of Theorem 3 apply to $X^n(\cdot,\cdot)$ except that the mean function of the limiting process X is of a different from which depends on Λ_j^n.

The proof of this corollary follows along the lines of the proof of Theorem 3.

4. Test procedures and Monte Carlo estimation of P-values. Suppose that $\min_{1\leq k\leq p} \inf_{f\in \mathbf{H}} V_{ff}^{kk}(\infty,\infty) > 0$, where V_{ff}^{kk} is the (k,k)th element of $\boldsymbol{V}_{ff}$. We define $\mathcal{C}$ to be a collection of standardized statistics

$$\mathcal{C} = \left\{ W^n(f,t) \equiv \text{diag}[\boldsymbol{V}_{ff}^n(\infty,\infty)]^{-1/2} X^n(f,t), \;\; f \in S \right\}, \tag{4.1}$$

where $S \subset \mathbf{H}$ is a chosen index set and diag(A) is a diagonal matrix whose diagonal elements are the same as A's. Since $\boldsymbol{V}_{fg}^n(\infty,\infty)$ is a consistent estimator of $\boldsymbol{V}_{fg}(\infty,\infty)$, W^n converge HJD-weakly to multivariate Gaussian processes with mean $\text{diag}[\boldsymbol{V}_{f,f}(\infty,\infty)]^{-1/2}\,\boldsymbol{\mu}(f,t)$ and covariance $\text{diag}[\boldsymbol{V}_{fg}(\infty,\infty)]^{-1/2}\,\boldsymbol{V}_{fg}(s,t)\,\text{diag}[\boldsymbol{V}_{fg}(\infty,\infty)]^{-1/2}$.

The statistics we propose for applying to the data analysis settings are

$$\sup_{f\in S} [W^n(f,\infty)]^T[W^n(f,\infty)], \quad \sup_{f\in S}\sup_{t\in[0,\infty]} [W^n(f,t)]^T[W^n(f,t)],$$

$$\sup_{f\in S,\,k} |W_k^n(f,\infty)|, \text{ and } \sup_{f\in S,\,k}\sup_{t\in[0,\infty]} |W_k^n(f,t)|,$$

where W_k^n is the kth element of the vector W^n. These statistics should be sensitive to both the ordered hazard and the stochastic ordering alternatives since the supremum-over-function-space statistics give sensitivity to broad ordered hazards alternatives and the supremum-over-time statistics to stochastic ordering alternatives. See Kosorok and Lin (1997) for a further discussion of how to select these statistics. Note that for the single-covariate problem, the first and the second statistics are respectively equivalent to the third and the fourth statistics. Due to the complexity of the Gaussian processes, we generally are not able to obtain the P-values of these statistics through analytical means. Therefore, a Monte Carol approach is now proposed to simulate the null distribution of the processes W and estimate the P-values of these statistics.

Let $\widetilde{X}_q^n, q = 1, ..., Q$, be Q "artificial" realizations of X^n generated as follows. Obtain nQ independent standard normal random deviates, ω_{jq}, $j = 1 \dots n$, $q = 1 \dots Q$, and construct the corresponding artificial realization of X^n, $\widetilde{X}_q^n$, with the kth element

$$\widetilde{X}_{kq}^n(f, \cdot) \;=\; n^{-1/2} \int_0^{(\cdot)} f\{\boldsymbol{b}^n(s)\} \sum_{j=1}^{n} Y_j(s) \left[Z_{jk}(s) - \overline{Z}_k(s)\right] d\widetilde{M}_{jq}^n(s), \tag{4.2}$$

where

$$\widetilde{M}_{jq}^n(t) \;=\; \omega_{jq} \int_0^t \left\{ \frac{\overline{Y}(s) - \Delta \overline{N}(s)}{\overline{Y}(s) - 1} \right\}^{\frac{1}{2}} dN_j(s). \tag{4.3}$$

Define

$$\widetilde{\mathcal{F}}_t^n \;=\; \sigma\{\mathcal{F}_t^n,\; \widetilde{M}_{jq}^n(s),\; s \le t,\; j = 1 \dots n,\; q = 1, ..., Q\,\}.$$

REMARK 2 *$\widetilde{M}_{jq}$ are $\widetilde{\mathcal{F}}_t^n$-martingales since*

$$\begin{aligned} E\left[\, \widetilde{M}_{jq}^n(t) \mid \widetilde{\mathcal{F}}_t^n \,\right] \;&=\; \widetilde{M}_{jq}^n(t-) \;+\; E\left[\omega_{jq} \left\{ \frac{\overline{Y}(t) - \overline{N}(t)}{\overline{Y}(t) - 1} \right\}^{\frac{1}{2}} \Delta N_j(t) \;\Big|\; \widetilde{\mathcal{F}}_t^n \right], \\ &=\; \widetilde{M}_{jq}^n(t-). \end{aligned}$$

THEOREM 4 *Suppose that the model described in Theorem 3 obtains with the corresponding sequence of histories $\widetilde{\mathcal{F}}^n$ and that $\min_{1 \le k \le p} \inf_{f \in \mathbf{H}} V_{ff}^{kk}(\infty, \infty) > 0$, where V_{ff}^{kk} is the (k,k)th element of $\boldsymbol{V}_{ff}$. Then,*

(a) The collection $\{\widetilde{X}_q^n(\cdot,\cdot), q = 1, ..., Q\}$, where $\widetilde{X}_q^n(f, \cdot)$ are as given in (4.2), converges HJD-weakly in the uniform topology on $\{\boldsymbol{A}(\mathbf{H}, (\mathbf{D}[0,\infty])^p)\}^Q$ to a collection of Q independent multivariate Gaussian processes, $\{X_q(\cdot,\cdot), q = 1, ..., Q\}$, such that each X_q has mean $\mathbf{0}$ and covariance function $\boldsymbol{V}_{fg}$, for all $f, g \in \mathbf{H}$, where $\boldsymbol{V}_{fg}$ is as defined in (3.3); and

(b) Let

$$\widetilde{W}_q^n = diag[\mathbf{V}_{ff}^n(\infty,\infty)]^{-1/2}\widetilde{X}_q^n,$$
$$W_q = diag[\mathbf{V}_{ff}(\infty,\infty)]^{-1/2}\widetilde{X}_q.$$

Then, for any finite subset $T \subset [0,\infty]$, the collection

$$\Big\{\widetilde{W}_q^n(f,t),\ \sup_{s\in[0,t]}\widetilde{W}_{kq}^n(f,s),\ \inf_{s\in[0,t]}\widetilde{W}_{kq}^n(f,s),\ k=1,...,p,$$
$$\sup_{s\in[0,t]}\left[\widetilde{W}_q^n(f,s)\right]^T\left[\widetilde{W}_q^n(f,s)\right],\ q=1,...,Q:\ t\in T\Big\}$$

converge weakly in the uniform topology over all $f \in \mathbf{H}$ to the corresponding collection

$$\Big\{W_q(f,t),\ \sup_{s\in[0,t]}W_{kq}(f,s),\ \inf_{s\in[0,t]}W_{kq}(f,s),\ k=1,...,p,$$
$$\sup_{s\in[0,t]}W_q^T(f,s)W_q(f,s),\ q=1,...,Q:\ t\in T\Big\}.$$

Proof. For $q=1,...,Q$, define p-dimensional processes $\widetilde{X}_q^\diamond$ with kth element

$$\widetilde{X}_{kq}^\diamond(f,\cdot) = n^{-1/2}\int_0^{(\cdot)} f\{\boldsymbol{b}^n(s)\}\sum_{j=1}^n Y_j(s)\left[Z_{jk}(s)-\overline{Z}_k(s)\right]d\widetilde{M}_{jq}^\diamond(s),$$

where

$$\widetilde{M}_{jq}^\diamond(t) = \omega_{jq}\int_0^t\{1-\Delta\Lambda(s)\}^{\frac{1}{2}}\,dN_j(s).$$

It is not difficult to see that $\widetilde{M}_{jq}^\diamond(t), j=1...n, q=1...Q$, are uncorrelated square-integrable $\widetilde{\mathcal{F}}_t^n$-martingales. By Theorem 1.5.1 of Fleming and Harrington (1991), $\widetilde{X}_{kq}^n$, $\widetilde{X}_{kq}^\diamond$, and $\widetilde{X}_{kq}^n - \widetilde{X}_{kq}^\diamond$ are also martingales. For all $q=1,...,Q$,

$$\begin{aligned}
&\left[\widetilde{X}_{kq}^n(t)-\widetilde{X}_{kq}^\diamond(t)\right]^2 \\
&\le \left[\widetilde{X}_{kq}^n(\infty)-\widetilde{X}_{kq}^\diamond(\infty)\right]^2 \\
&= n^{-1}\Bigg[\sum_{j=1}^n f\{\boldsymbol{b}^n(X_j)\}\left[Z_{jk}(X_j)-\overline{Z}_k(X_j)\right]\times \\
&\qquad \left\{\left[\frac{\overline{Y}(X_j)-\Delta\overline{N}(X_j)}{\overline{Y}(X_j)-1}\right]^{\frac{1}{2}}-[1-\Delta\Lambda(X_j)]^{\frac{1}{2}}\right\}\omega_{jq}N_j(X_j)\Bigg]^2, \\
&\equiv n^{-1}A_{kq}^n,
\end{aligned}$$

where $X_j = T_j \wedge C_j$. For any $\epsilon > 0$,

$$\begin{aligned}
&\mathrm{P}\left\{ \max_{1\le q\le Q} \max_{1\le k\le p} \sup_{t\in[0,\infty]} \left|\widetilde{X}^n_{kq}(t) - \widetilde{X}^\diamond_{kq}(t)\right| > \epsilon \right\} \\
&= \mathrm{P}\left\{ \max_{1\le q\le Q} \max_{1\le k\le p} \sup_{t\in[0,\infty]} \left|\widetilde{X}^n_{kq}(t) - \widetilde{X}^\diamond_{kq}(t)\right|^2 > \epsilon^2 \right\} \\
&\le \frac{1}{\epsilon^2} \sum_{q=1}^{Q} \mathrm{E}\left[\max_{1\le k\le p} n^{-1} A^n_{kq} \right],
\end{aligned}$$

by the submartingale inequality [cf. Proposition 2.16 in Chapter 2 of Ethier and Kurtz (1986)]. Since w_{jq} are independent,

$$\begin{aligned}
\mathrm{E}\left(\max_{1\le k\le p} n^{-1} A^n_q\right) &\le \tau \mathrm{E}\left[n^{-1} \sum_{j=1}^{n} \left\{ \left[1 - \frac{\Delta\overline{N}(X_j) - 1}{\overline{Y}(X_j) - 1}\right]^{\frac{1}{2}} - [1 - \Delta\Lambda(X_j)]^{\frac{1}{2}} \right\}^2 N_j(X_j) \right], \\
&\le \tau \mathrm{E}\left[\sup_{s\in[0,\infty]} \left\{ \left[1 - \frac{\Delta\overline{N}(s) - 1}{\overline{Y}(s) - 1}\right]^{\frac{1}{2}} - [1 - \Delta\Lambda(s)]^{\frac{1}{2}} \right\}^2 I_{\{\overline{Y}(s)>1\}} \right] + \frac{\tau}{n}, \\
&\equiv \tau \mathrm{E}[B^n] + \frac{\tau}{n},
\end{aligned}$$

for some $\tau < \infty$ and $q = 1, ..., Q$. By (A.9) in the Appendix, $B^n \to 0$ in probability as $n \to \infty$. Along with the bounded convergence theorem, we have

$$\mathrm{P}\left\{ \max_{1\le q\le Q} \max_{1\le k\le p} \sup_{t\in[0,\infty]} \left|\widetilde{X}^n_{kq}(t) - \widetilde{X}^\diamond_{kq}(t)\right| > \epsilon \right\} \xrightarrow{P} 0, \qquad \text{as } n \longrightarrow \infty.$$

Therefore, the weak convergence of $\{\widetilde{X}^n_q, q = 1, ..., Q\}$ can be established by verifying that of $\{\widetilde{X}^\diamond_q, q = 1, ..., Q\}$.

An adaption of the time-transformed method used by Gill (1980) in the proof of his Theorem 4.2.1 will be applied to establish the finite dimensional weak convergence of $\{\widetilde{X}^\diamond_q, q = 1, ..., Q\}$ to $\{X_q, q = 1, ..., Q\}$. For any collection of mQ bounded left-continuous step functions on $[0, \infty]$, $\{c_{hq}, h = 1...m, q = 1...Q\}$, $m < \infty$, let

$$C^n_q(t) = \sum_{h=1}^{m} c_{hq}(t) f_h\{\boldsymbol{b}^n(t)\}, \quad \text{and} \quad C_q(t) = \sum_{h=1}^{m} c_{hq}(t) f_h\{\boldsymbol{b}(t)\}.$$

By (A.2) in the Appendix, for all $t \in \mathcal{I}$,

$$\max_{1\le q\le Q} \sup_{s\in[0,t]} \left| C^n_q(s) - C_q(s) \right| \xrightarrow{P} 0. \tag{4.4}$$

Define a p-dimensional process $\boldsymbol{U}^n$ with the kth elements

$$\widetilde{U}_k^n(t) \;=\; n^{-1/2}\int_0^t \sum_{q=1}^{Q} C_q^n(s)\sum_{j=1}^{n} Y_j(s)\left[Z_{jk}(s)-\overline{Z}_k(s)\right]d\widetilde{M}_{jq}^{\circ}(s),$$

and $H_{jkq}(s)=n^{-1/2}C_q^n(s)Y_j(s)[Z_{jk}(s)-\overline{Z}_k(s)]$, for $k=1...p$, $j=1...n$, and $q=1...Q$.

We can enumerate all the discontinuities of Λ_j^n and Λ, for all $n\geq 1$, in a single sequence $t_1,t_2,...$, say. Choose $\delta_h>0$, $h=1,2,...$, such that $\sum_{h=1}^{\infty}\delta_h<\infty$. Define the time transformation $\phi^*:[0,\infty]\mapsto[0,\infty]$ by

$$\phi^*(t)=t+\sum_{h:t_h\leq t}\delta_h.$$

Let $\mathcal{I}^*=[0,\phi^*(u-))$ if $u\notin\mathcal{I}$ and $\mathcal{I}^*=[0,\phi^*(u-)]$ if $u\in\mathcal{I}$.

The processes N_j^*, Y_j^*, H_{jkq}^*, and $\widetilde{M}_{jq}^*$ are defined as follows. Firstly, if $t^*=\phi(t)$ for some t, we let $N_j^*(t^*)=N_j(t)$, $Y_j^*(t^*)=Y_j(t)$, $H_{jkq}^*(t^*)=H_{jkq}(t)$, and $\widetilde{M}_{jq}^*(t^*)=\widetilde{M}_{jq}^{\circ}(t)$. Next, we define N_j^* on the intervals $[\phi^*(t_h-),\phi^*(t_h))$ by letting N_j^*, conditional on $Y_j(t_h)$, make a single jump at the point R_{jh} with probability $Y_j(t_h)\Delta\Lambda_j^n(t_h)$, where R_{jh} is an independent random variable uniformly distributed on $(\phi^*(t_h-),\phi^*(t_h))$. Also, for $t^*\in[\phi^*(t_h-),\phi^*(t_h))$, $h=1,2,...$, we define $Y_j^*(t^*)=Y_j(t_h)$, $H_{jkq}^*(t^*)=H_{jkq}(t_h)$, and

$$\widetilde{M}_{jq}^*(t^*)=\widetilde{M}_{jq}^{\circ}(t_h-)+\omega_{jq}\;[1-\Delta\Lambda(t_h)]^{\frac{1}{2}}\;\left[N_j^*(t^*)-N_j(t_h-)\right].$$

Define

$$\mathcal{E}_{t^*}^{*n}\equiv\sigma\left\{R_{jh},\text{for all }h:\phi^*(t_h-)\leq t^*;N_j^*(s^*),\omega_{jq}N_j^*(s^*),s^*\leq t^*;j=1\ldots n,q=1\ldots Q\right\}$$

and

$$\mathcal{F}_{t^*}^{*n}\equiv\begin{cases}\sigma\left\{\mathcal{F}_t^n,\mathcal{E}_{t^*}^{*n}\right\}, \text{ if } t^*=\phi^*(t),\\[2ex] \sigma\left\{\mathcal{F}_{t-}^n,\mathcal{E}_{t^*}^{*n}\right\}, \text{ if } \phi^*(t-)\leq t^*<\phi^*(t).\end{cases}$$

We can see that $\widetilde{M}_{jq}^*(t^*)$, $j=1\ldots n$, are square integrable $\mathcal{F}_{t^*}^{*n}$-martingales, with $H_{jkq}^*(t^*)$ and $Y_j^*(t^*)$ being $\mathcal{F}_{t^*}^{*n}$-predictable.

Let

$$\widetilde{U}_k^{*n}(t^*)=\int_0^{t^*}\sum_{q=1}^{Q}\sum_{j=1}^{n}H_{jkq}^*(s)d\widetilde{M}_{jq}^*(s)$$

and obtain the predictable covariations $\left\langle \widetilde{M}^*_{jq}, \widetilde{M}^*_{j'q'} \right\rangle (\cdot) = 0$, for $j \neq j$ or $q \neq q'$;

$$\left\langle \widetilde{M}^*_{jq}, \widetilde{M}^*_{jq} \right\rangle (t^*) = \left\langle \widetilde{M}^\circ_{jq}, \widetilde{M}^\circ_{jq} \right\rangle (t) = \int_0^t \left[1 - \Delta\Lambda(s)\right] Y_j(s) d\Lambda^n_j(s),$$

for $t^* = \phi^*(t)$; and

$$\left\langle \widetilde{M}^*_{jq}, \widetilde{M}^*_{jq} \right\rangle (t^*) = \left\langle \widetilde{M}^\circ_{jq}, \widetilde{M}^\circ_{jq} \right\rangle (t_h-) + Y_j(t_h) I_{\{R_{jh} \leq t^*\}} \left[1 - \Delta\Lambda(t_h)\right] \Delta\Lambda^n_j(t_h),$$

for $\phi^*(t_h-) \leq t^* < \phi^*(t_h)$. The processes $\widetilde{U}^*_k$, $k = 1 \ldots p$, are square integrable $\mathcal{F}^{*n}$-martingales with the predictable variations

$$\left\langle \widetilde{U}^{*n}_k, \widetilde{U}^{*n}_{k'} \right\rangle (t^*) = \begin{cases} \int_0^t \sum_{q=1}^Q \sum_{j=1}^n H_{jkq}(s) H_{jk'q}(s) d \left\langle \widetilde{M}^\circ_{jq}, \widetilde{M}^\circ_{jq} \right\rangle (s), & \text{if } t^* = \phi^*(t), \\ \int_0^{t_h-} \sum_{q=1}^Q \sum_{j=1}^n H_{jkq}(s) H_{jk'q}(s) d \left\langle \widetilde{M}^\circ_{jq}, \widetilde{M}^\circ_{jq} \right\rangle (s) + \\ \quad \sum_{q=1}^Q \sum_{j=1}^n H_{jkq}(t_h) H_{jk'q}(t_h) Y_j(t_h) I_{\{t^* \leq R_{jh}\}} \left[1 - \Delta\Lambda(t_h)\right] \Delta\Lambda^n_j(t_h), \\ \qquad \text{if } \phi^*(t_h-) \leq t^* < \phi^*(t_h). \end{cases}$$

The rest of the proof follows along the lines of the proof of Theorem 3.

5. Stratified Stochastic Processes. For the β-Blocker Heart Attack Trial mentioned in Section 1, researchers are mainly interested in the ability of propranolol to reduce mortality rates of patients with at least one myocardial infarction. Since survival rates are usually expected to depend on age and clinic location, the investigation of the propranolol effect would be more appropriate if age and clinic location could be adjusted for. This adjustment is especially important when the data are imbalanced: e.g. the patients in the treatment group are generally older than those in the control group. Without adjustment, treatment effects may be masked or exaggerated by confounding effects. The researchers involved in the β-Blocker Heart Attack Trial study may also be interested in finding the interventions for prolonging a patient's life as described in Section 1. In either case, the previous proposed class of statistics can not be applied directly to achieve our goals. We will in this section resolve this problem by generalizing the above class of statistics to accommodate the situations where finite strata are present.

Suppose there are S strata. For each stratum i, a sample of survival data $\{T_{ij}, C_{ij}, Z_{ij}(t)\}$ are observed, where j refers to the subject, $j = 1...n_i$, T_{ij} and C_{ij} are the times to failure and censoring respectively, and $Z_{ij}(t)$ are $p \times 1$ vector covariate processes with $t \leq T_{ij} \wedge C_{ij}$. The observed failure counting process and the at-risk process for subject j in stratum i are denoted by N_{ij} and Y_{ij}. Let $n = \sum_i n_i$,

$$\overline{N}_i(t) = \sum_{j=1}^{n_i} N_{ij}(t), \quad \overline{N}(t) = \sum_{i=1}^{S} N_i(t),$$

and

$$\overline{Y}_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t), \quad \overline{Y}(t) = \sum_{i=1}^{S} Y_i(t),$$

and define the filtrations

$$\mathcal{F}_t^n = \sigma\{N_{ij}(s),\ Y_{ij}(s+),\ Z_{ij}(s+),\ i = 1 \ldots S,\ j = 1 \ldots n_i,\ s \leq t\}, \qquad n = 1, 2, \ldots.$$

We again assume the general random censorship model, that the data in different stratum are independent of each other, and that the cumulative hazard of failure times at time t, $\Lambda_{ij}^n(t)$, depend only on $Z_{ij}(t)$, $i = 1 \ldots S$, $j = 1 \ldots n_i$.

Define $X^n(f,t) = \left\{X_1^n(f,t), ..., X_p^n(f,t)\right\}^T$ with the kth element

$$X_k^n(f,t) \quad = \quad n^{-1/2} \int_0^t \sum_{i=1}^{S} f\{\boldsymbol{b}_i^n(s)\} \sum_{j=1}^{n_i} Y_{ij}(s)[Z_{ijk}(s) - \overline{Z}_{ik}(s)] dN_{ij}(s), \tag{5.1}$$

where $\overline{Z}_{ik}(s) = \overline{Y}_i^{-1}(s) \sum_j Y_{ij}(s) Z_{ijk}(s)$, $f \in H$, and $\boldsymbol{b}_i^n = (b_{i1}^n, ..., b_{ir}^n)^T$, for some $r < \infty$, $b_{il} \in \mathbf{B}(\mathcal{F}^n)$, for all $l = 1...r$, $k = 1...p$, and $i = 1...p$.

Let Λ_{iz} denote the cumulative hazard in stratum i for the covariate path $z \in \mathcal{Z}$. We are interested in testing the null hypothesis $H_0 : \Lambda_{iz}^n = \Lambda_i$ over $[0, \infty)$ for all $z \in \mathcal{Z}$ and $i = 1...S$, against contiguous alternatives of the form

$$d\Lambda_i^n(t \mid z(t)) = \frac{\exp\left[\frac{\beta^T z(t)}{\sqrt{n}} g(t)\right] d\Lambda_i(t)}{1 + (\exp\left[\frac{\beta^T z(t)}{\sqrt{n}} g(t)\right] - 1)\Delta\Lambda_i(t)}, \tag{5.2}$$

for some finite $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)^T$, some baseline hazard $\Lambda_i(t)$, and some real function g such that g is bounded by $G < \infty$.

THEOREM 5 *Under the general random censorship model and the "local alternative" sequence (5.2), suppose that* $n_i \longrightarrow \infty$ *for all* i, *and that the following conditions hold :*

1. *There exist* S *functions* $\pi_i : [0,\infty) \mapsto [0,1]$ *such that*

$$\sup_{t\in[0,\infty)} \left| \frac{\overline{Y}_i(t)}{n_i} - \pi_i(t) \right| \xrightarrow{P} 0 \quad \textit{as } n \longrightarrow \infty.$$

2. *The covariate processes are adapted, bounded, and left-continuous with right-hand limits.*

Set $\mathcal{I}_i = sup\{t : \pi_i(t) > 0\}$ *and* $u_i = sup\ \mathcal{I}_i$, $i = 1...S$.

3. *For any* $t \in \mathcal{I}_i$,

   (i) $\Lambda_i(t) < \infty$.

   (ii) There exist left-continuous functions v^i_{kl}, $k,\ l = 1,...,p$, *with right-hand limits such that for all* $t \in \mathcal{I}_i$,

$$\sup_{s\in[0,t]} \left| \overline{Y}_i(s)^{-1} \sum_{j=1}^{n_i} Y_{ij}(s) \left[Z_{ijk}(s) - \overline{Z}_{ik}(s)\right] \left[Z_{ijl}(s) - \overline{Z}_{il}(s)\right] - v^i_{kl}(s) \right| \xrightarrow{P} 0$$

   as $n_i \longrightarrow \infty$, *and* $v^i_{kl}(s)$ *are zero outside of* $\mathcal{I}_i$.

Then,

(a) $X^n(\cdot,\cdot)$, *with components defined in (5.1), converges HJD-weakly in the uniform topology on* $\boldsymbol{A}(\mathbf{H}, (\mathbf{D}[0,\infty])^p)$ *to a multivariate Gaussian process* $X(\cdot,\cdot)$ *with mean function*

$$\boldsymbol{\mu}(f,t) = \int_0^t \sum_{i=1}^S f\{\boldsymbol{b}_i(s)\} \left[\boldsymbol{v}^i(s)\boldsymbol{\beta}\right] g(s) d\Lambda_i(s), \tag{5.3}$$

and covariance function

$$\boldsymbol{V}_{fg}(s,t) = \int_0^{s\wedge t} \sum_{i=1}^S f\{\boldsymbol{b}_i(x)\} g\{\boldsymbol{b}_i(x)\} \boldsymbol{v}^i(s) \pi_i(s) \left[1 - \Delta\Lambda_i(s)\right] d\Lambda(s), \tag{5.4}$$

for all $f, g \in \mathbf{H}$ *and all* $s,t \in [0,\infty]$, *where* $\boldsymbol{v}^i$ *is a* $p \times p$ *matrix with* (k,l)*th element* v^i_{kl};

(b) For any $t \in [0,\infty]$, the collection

$$\left\{X^n(f,t),\ \sup_{s\in[0,t]} X_k^n(f,s),\ \inf_{s\in[0,t]} X_k^n(f,s), k=1,\ldots,p,\ \sup_{s\in[0,t]} [X^n(f,s)]^T [X^n(f,s)] : t \in T\right\}$$

consists of measurable random variables which jointly converge weakly in the uniform topology over all $f \in \mathbf{H}$ to the corresponding collection

$$\left\{X(f,t),\ \sup_{s\in[0,t]} X_k(f,s),\ \inf_{s\in[0,t]} X_k(f,s), k=1,\ldots,p,\ \sup_{s\in[0,t]} X^T(f,s)X(f,s) : t \in T\right\};$$

(c) The covariance estimator

$$\boldsymbol{V}_{fg}^n(s,t) \;=\; n^{-1}\int_0^{s\wedge t} \sum_{i=1}^{S} f\{\boldsymbol{b}_i^n(\cdot)\}g\{\boldsymbol{b}_i^n(\cdot)\} \sum_{j=1}^{n_i} Y_{ij}[Z_{ij}-\overline{Z}_i][Z_{ij}-\overline{Z}_i]^T \frac{\overline{Y}_i - \Delta\overline{N}_i}{\overline{Y}_i - 1}\frac{d\overline{N}_i}{\overline{Y}_i}$$

is uniformly consistent for $\boldsymbol{V}_{fg}$ over all $f,g \in \mathbf{H}$ and all $s,t \in [0,\infty]$.

Theorem 5 can be shown by the same arguments used to develop Theorem 3.

REMARK 3 *We can also apply the test procedures and the Monte Carlo approach proposed in Section 4 to these processes and develop the results of Theorem 4 following along the lines of the proof given in Section 4.*

6. Discussion. We generalized the class of single-covariate nonparametric test procedures proposed by Jones and Crowley (1989) to multivariate-covariate situations with finite strata. These statistics can, in clinical trials, be applied to investigate the treatment effects or s-sample trend after adjusting for possible confounding factors, as well as to explore the potential interventions after influential factors are controlled for. Considering these stochastic processes indexed by both the time scale and the weight function, we showed that certain large families of these processes converge weakly in the uniform topology to multivariate Gaussian processes also doubly indexed by both time and weight function. This result permits us to develop more powerful test procedures than the current ones for versatile alternatives. Via simulation studies for the two-sample problems in Kosorok and Lin (1997), we have shown that our newly proposed statistics can increase power under a variety of settings. However, the general behavior of these statistics with covariates present and the criterion for an appropriate choice of weight functions in practice haven't been fully explored. Further studies need to be performed for a better understanding of these issues.

The simulation study in Kosorok and Lin also evaluates the performance of a Monte Carlo approach closely related to the one proposed in the present paper. This approach seems quite effective for moderate sample sizes. There are other martingale-type statistics, such as the weighted Kaplan-Meier statistics of Pepe and Fleming (1989), to which the theory in the present paper could be potentially applied. Further research on the relative efficiencies of these martingale-type statistics can give us guidelines on the choice of the type of statistics to use in different situations and would thus be very beneficial in practice.

APPENDIX

Proof of the finite dimensional weak convergence of $X^n(\cdot,\cdot)$ in Theorem 3. The kth element of $X^n(f,t)$ can be written as

$$\begin{aligned} X_k^n(f,t) &= n^{-1/2}\int_0^t f\{\boldsymbol{b}^n(s)\}\sum_{j=1}^n Y_j(s)[Z_{jk}(s)-\overline{Z}_k(s)]dM_j(s) + \\ & \quad n^{-1/2}\int_0^t f\{\boldsymbol{b}^n(s)\}\sum_{j=1}^n Y_j(s)[Z_{jk}(s)-\overline{Z}_k(s)]d\Lambda_j^n(s), \\ &\equiv X_{\circ k}^n(f,t)+\mu_k^n(f,t). \end{aligned} \tag{A.1}$$

Note that

$$\mu_k^n(t) = n^{-1/2}\int_0^t f\{\boldsymbol{b}^n(s)\}\sum_{j=1}^n Y_j(s)\left[Z_{jk}(s)-\overline{Z}_k(s)\right]\left[d\Lambda_j^n(s)-d\Lambda(s)\right],$$

and
$$d\Lambda_j^n(s)-d\Lambda(s) = \frac{\exp\{\frac{\beta^T Z_j(s)g(s)}{\sqrt{n}}\}-1}{1+\left(\exp\{\frac{\beta^T Z_j(s)g(s)}{\sqrt{n}}\}-1\right)\Delta\Lambda(s)}\left[1-\Delta\Lambda(s)\right]d\Lambda(s).$$

Since each b_l^n, $l=1,\ldots,r$, is in $\mathbf{B}(\mathcal{F}^n)$ and f is continuous on $[0,1]^r$,

$$\sup_{s\in[0,t]} \mid f\{\boldsymbol{b}^n(s)\}-f\{\boldsymbol{b}(s)\}\mid \xrightarrow{P} 0, \tag{A.2}$$

for all $t \in \mathcal{I}$. The boundedness of Z_j and g yields

$$\lim_{n\to\infty} \sup_{s\in[0,\infty)} \left| n^{1/2}\left[d\Lambda_j^n(s) - d\Lambda(s)\right] - \boldsymbol{\beta}^T Z_j(s) g(s)\left[1 - \Delta\Lambda(s)\right] d\Lambda(s) \right| = 0.$$

Therefore $\mu^n(f,t)$ converges uniformly on $[0,\infty]$ in probability to $\mu(f,t)$, since

$$\sum_{j=1}^{n} Y_j\left[Z_{jk} - \overline{Z}_k\right] Z_{jl} = \sum_{j=1}^{n} Y_j\left[Z_{jk} - \overline{Z}_k\right]\left[Z_{jl} - \overline{Z}_l\right].$$

The finite dimensional convergence of X^n can be established by verifying the finite dimensional weak convergence of $X_\circ^n = (X_{\circ 1}^n, ..., X_{\circ p}^n)^T$, where $X_{\circ k}^n$ are as given in (A.1), to Gaussian processes with mean $\mathbf{0}$ and covariance function $\boldsymbol{V}_{fg}$, where $f, g \in \mathrm{H}$. Since the distribution of failure times may not be absolutely continuous, an adaption of the time-transformed method used in the proof of Theorem 4.2.1 in Gill (1980) will be utilized in this proof.

For any collection of m bounded left-continuous step functions on $[0,\infty]$, $\{c_h, h = 1...m\}$, $m < \infty$, let

$$C^n(t) = \sum_{h=1}^{m} c_h(t) f_h\{\boldsymbol{b}^n(t)\}, \quad \text{and} \quad C(t) = \sum_{h=1}^{m} c_h(t) f_h\{\boldsymbol{b}(t)\}.$$

By (A.2), for all $t \in \mathcal{I}$,

(A.3) $$\sup_{s\in[0,t]} \mid C^n(s) - C(s) \mid \xrightarrow{P} 0.$$

Define a p-dimensional process $\widetilde{\boldsymbol{U}}^n$ with the kth elements

$$\widetilde{U}_k^n(t) = n^{-1/2} \int_0^t C^n(s) \sum_{j=1}^{n} Y_j(s)\left[Z_{jk}(s) - \overline{Z}_k(s)\right] dM_j(s),$$

and $H_{jk}(s) = n^{-1/2} C^n(s) Y_j(s)[Z_{jk}(s) - \overline{Z}_k(s)]$, for $k = 1...p$ and $j = 1...n$.

We can enumerate all the discontinuities of Λ_j^n and Λ, for all $n \geq 1$, in a single sequence $t_1, t_2, ...$, say. Choose $\delta_h > 0$, $h = 1, 2, ...$, such that $\sum_{h=1}^{\infty} \delta_h < \infty$. Define the time transformation $\phi^* : [0,\infty] \mapsto [0,\infty]$ by

$$\phi^*(t) = t + \sum_{h:t_h \leq t} \delta_h.$$

Let $\mathcal{I}^* = [0, \phi^*(u-))$ if $u \notin \mathcal{I}$ and $\mathcal{I}^* = [0, \phi^*(u-)]$ if $u \in \mathcal{I}$.

The processes N_j^*, Y_j^*, M_j^*, H_{jk}^* are defined as follows. Firstly, if $t^* = \phi(t)$ for some t, we let $N_j^*(t^*) = N_j(t)$, $Y_j^*(t^*) = Y_j(t)$, and $H_{jk}^*(t^*) = H_{jk}(t)$. Next, we define N_j^* on the intervals $[\phi^*(t_h-), \phi^*(t_h))$ by letting N_j^*, conditional on $Y_j(t_h)$, make a single jump at the point R_{jh} with probability $Y_j(t_h)\Delta\Lambda_j^n(t_h)$, where R_{jh} is an independent variable uniformly distributed on $(\phi^*(t_h-), \phi^*(t_h))$. Also, for $t^* \in [\phi^*(t_h-), \phi^*(t_h))$, $h = 1, 2, \ldots$, we define $Y_j^*(t^*) = Y_j(t_h)$, $H_{jk}^*(t^*) = H_{jk}(t_h)$, and

$$M_j^*(t^*) = M_j(t_h-) + N_j^*(t^*) - N_j(t_h-) - Y_j(t_h)I_{\{R_{jh} \le t^*\}}\Delta\Lambda_j^n(t_h).$$

Define

$$\mathcal{E}_{t^*}^{*n} \equiv \sigma\left\{R_{jh}, \text{for all } h : \phi^*(t_h-) \le t^*; N_j^*(s^*), s^* \le t^*; j = 1 \ldots n\right\}$$

and

$$\mathcal{F}_{t^*}^{*n} \equiv \begin{cases} \sigma\left\{\mathcal{F}_t^n, \mathcal{E}_{t^*}^{*n}\right\}, & \text{if } t^* = \phi^*(t), \\ \sigma\left\{\mathcal{F}_{t-}^n, \mathcal{E}_{t^*}^{*n}\right\}, & \text{if } \phi^*(t-) \le t^* < \phi^*(t). \end{cases}$$

We can see that $M_j^*(t^*)$, $j = 1 \ldots n$, are square integrable $\mathcal{F}_{t^*}^{*n}$-martingales, with $H_{jk}^*(t^*)$ and $Y_j^*(t^*)$ being $\mathcal{F}_{t^*}^{*n}$-predictable.

Let

$$\widetilde{U}_k^{*n}(t^*) = \int_0^{t^*} \sum_{j=1}^n H_{jk}^*(s)dM_j^*(s)$$

and obtain the predictable covariations $\left\langle M_j^*, M_{j'}^*\right\rangle(\cdot) = 0$, for $j \ne j'$; $\left\langle M_j^*, M_j^*\right\rangle(t^*) = \langle M_j, M_j\rangle(t)$, for $t^* = \phi^*(t)$; and

$$\left\langle M_j^*, M_j^*\right\rangle(t^*) = \langle M_j, M_j\rangle(t_h-) + Y_j(t_h)I_{\{R_{jh} \le t^*\}}\left[1 - \Delta\Lambda_j^n(t_h)\right]\Delta\Lambda_j^n(t_h),$$

for $\phi^*(t_h-) \le t^* < \phi^*(t_h)$. $\widetilde{U}_k^*$ is then a square integrable $\mathcal{F}^{*n}$-martingale with predictable covariation

$$\left\langle \widetilde{U}_k^{*n}, \widetilde{U}_{k'}^{*n}\right\rangle(t^*) = \begin{cases} \int_0^t \sum_{j=1}^n H_{jk}(s)H_{jk'}(s)d\langle M_j, M_j\rangle(s), \text{ if } t^* = \phi^*(t), \\ \int_0^{t_h-} \sum_{j=1}^n H_{jk}(s)H_{jk'}(s)d\langle M_j, M_j\rangle(s) \\ \quad + \sum_{j=1}^n H_{jk}(t_h)H_{jk'}(t_h)Y_j(t_h)I_{\{t^* \le R_{jh}\}}\left[1 - \Delta\Lambda_j^n(t_h)\right]\Delta\Lambda_j^n(t_h), \\ \qquad \text{if } \phi^*(t_h-) \le t^* < \phi^*(t_h). \end{cases}$$

By (A.3) and condition 2.(ii), we have

$$\text{(A.4)} \qquad \sup_{s\in[0,t]} \left| \sum_{j=1}^{n} H_{jk}(s)H_{jk'}(s) - C^2(s)\pi(s)v_{kk'}(s) \right| \xrightarrow{P} 0,$$

for all $k, k' = 1, ..., p$, and $t \in \mathcal{I}$. Therefore, for every $t^* \in \mathcal{I}^*$,

$$\text{(A.5)} \quad \left\langle \widetilde{U}_k^{*n}, \widetilde{U}_{k'}^{*n} \right\rangle (t^*) \xrightarrow{P} \begin{cases} \int_0^t C^2(s)\pi(s)v_{kk'}(s)\,[1-\Delta\Lambda(s)]\,d\Lambda(s), \text{ if } t^* = \phi^*(t), \\ \int_0^{t_h-} C^2(s)\pi(s)v_{kk'}(s)\,[1-\Delta\Lambda(s)]\,d\Lambda(s) + \\ \qquad h(t_h)\frac{t^*-\phi^*(t_h-)}{\delta_h}\,[1-\Delta\Lambda(t_h)]\,\Delta\Lambda(t_h), \\ \qquad \text{if } \phi^*(t_h-) \le t^* < \phi^*(t_h), \end{cases}$$

If $u \notin \mathcal{I}$, $\lim_{t\uparrow u} \pi(t) = 0$. For any $w > t$,

$$\begin{aligned} \text{(A.6)} \qquad & \int_t^w d\left\langle \widetilde{U}_k^n, \widetilde{U}_{k'}^n \right\rangle (s) \\ &= n^{-1} \int_t^w [C^n(.)]^2 \sum_{j=1}^{n} Y_j \left[Z_{jk} - \overline{Z}_k\right]\left[Z_{jk'} - \overline{Z}_{k'}\right]\left(1 - \Delta\Lambda_j^n\right) d\Lambda_j^n \\ &\le \kappa \left[n^{-1} \int_t^w \sum_{j=1}^{n} Y_j(s)\,d\Lambda_j^n(s) \right], \end{aligned}$$

for some $\kappa < \infty$. Let $\pi_j^n(s) = \mathrm{P}\{Y_j(s) = 1 \mid Z_j(s)\}$. Then

$$\begin{aligned} \text{(A.7)} \qquad & \mathrm{E}\left[n^{-1} \int_t^u \sum_{j=1}^{n} Y_j(s)\,d\Lambda_j^n(s) \right] \\ &\le n^{-1} \int_t^\infty \sum_{j=1}^{n} \pi_j^n(s)\,d\Lambda_j^n(s) \\ &\le n^{-1} \sum_{j=1}^{n} \mathrm{P}\{C_j \ge t \mid Z_j(s)\} \int_t^\infty S_j^n(s)\,d\Lambda_j^n(s) \\ &= n^{-1} \sum_{j=1}^{n} \pi_j^n(t) \quad \xrightarrow{n\to\infty} \quad \pi(t) \quad \xrightarrow{t\uparrow u} 0, \end{aligned}$$

where S_j^n is the corresponding survival function of Λ_j^n. Also,

$$\int_t^u [C^n(s)]^2 v_{kk'}(s)\pi(s)\,[1-\Delta\Lambda(s)]\,d\Lambda(s) \;\le\; \kappa \int_t^u \pi(s)\,d\Lambda(s),$$

and

$$\text{(A.8)} \qquad \int_t^u \pi(s)\,d\Lambda(s) \;=\; \int_t^u (1 - L(s-))\,S(s)\,d\Lambda(s) \;\le\; \pi(t) \quad \xrightarrow{t\uparrow u} 0,$$

where S is the corresponding survival function of Λ and L is a distribution function such that $\pi(s) = (1 - L(s-))S(s)$.

When $u < \infty$, by the fact that $\pi(u) = 0$ and by (A.6) and (A.7),

$$\int_u^\infty d\left\langle \widetilde{U}_k^n, \widetilde{U}_{k'}^n \right\rangle(t) \;\xrightarrow{\text{P}}\; 0, \qquad \text{for all } k,\ k' = 1, ..., p.$$

Also $v_{kk'}(t) = 0$ for any $t \geq u$ and $k, k' = 1, ..., p$; therefore

$$\int_u^\infty C^2(s)\pi(s)v_{kk'}(s)\,[1 - \Delta\Lambda(s)]\,d\Lambda(s) \;=\; 0.$$

Since for any $w^* > t^*$

$$\left\langle \widetilde{U}_k^{*n}, \widetilde{U}_{k'}^{*n} \right\rangle(w^*) - \left\langle \widetilde{U}_k^{*n}, \widetilde{U}_{k'}^{*n} \right\rangle(t^*) \;\leq\; \left\langle \widetilde{U}_k^n, \widetilde{U}_{k'}^n \right\rangle(w) - \left\langle \widetilde{U}_k^n, \widetilde{U}_{k'}^n \right\rangle(t-),$$

where w and t are such that $\phi^*(w-) \leq w^* < \phi^*(w)$ and $\phi^*(t-) \leq t^* < \phi^*(t)$, (A.5) holds for $t^* \in [0, \infty]$.

For all $k, k' = 1...p$ and $\epsilon > 0$, let

$$\widetilde{U}_k^{*n\epsilon} = \int_0^{t^*} \sum_{j=1}^n H_{jk}^*(s)\; I_{\{|H_{jk}^*(s)|>\epsilon\}}\; dM_j^*(s).$$

Then, almost surely,

$$\sup_{t^*\in[0,\infty]} \left|\Delta[\widetilde{U}_k^{*n}(t^*) - \widetilde{U}_k^{*n\epsilon}(t^*)]\right| \leq \epsilon \sup_{t^*\in[0,\infty]} \left|\sum_{j=1}^n \Delta M_j^*(t^*)\right| \leq \epsilon,$$

and if $\phi^*(t-) \leq t^* < \phi^*(t)$,

$$\left\langle \widetilde{U}_k^{*n\epsilon}, \widetilde{U}_{k'}^{*n\epsilon} \right\rangle(t^*) \;\leq\; \left(\sup_{j,k}\ \sup_{s\in[0,\infty)} I_{\{|H_{jk}(s)|>\epsilon\}}\right) \left\langle \widetilde{U}_k^{*n}, \widetilde{U}_{k'}^{*n}(t) \right\rangle \;\xrightarrow{\text{P}}\; 0,$$

since $\sup_{j,k} \sup_{s\in[0,\infty)} |H_{jk}(s)| \leq n^{-1/2}\kappa^*$ for some $\kappa^* < \infty$.

Rebolledo's theorem combined with the Cramer-Wold device now yields weak convergence in the Skorohod topology on $(\mathbf{D}[0,\infty])^{mp}$ of collections of $\{X^{*n}(f), f \in \mathbf{h}\}$, where $\mathbf{h}= \{f_1, ..., f_m\}$ and the kth element of $X^{*n}(f)$ is

$$n^{-1/2} \int_0^{(\cdot)} f^*\{\boldsymbol{b}^n(s)\} \sum_{j=1}^n Y_j^*(s) \left[Z_{jk}^*(s) - \overline{Z}_k^*(s)\right] dM_j^*(s),$$

where f^* and Z^* are defined corresponding to H^*, to m Gaussian processes with mean $\mathbf{0}$ and covariance function

$$\boldsymbol{V}^*_{fg}(t^*) \equiv \begin{cases} \boldsymbol{V}_{fg}(t), \text{ if } \phi^*(t) = t^*, \\ \boldsymbol{V}_{fg}(t_h-) + \frac{t^*-\phi^*(t_h-)}{\delta_h}\Delta\boldsymbol{V}_{fg}(t_h), \text{ if } \phi^*(t_h-) \le t^* < \phi^*(t_h), \end{cases}$$

for $f, g \in \mathbf{h}$.

Since the above limiting process is continuous, Theorem 1.3.6 of van der Vaart and Wellner (1996) yields HJD-weak convergence in the uniform topology on $(\mathbf{D}[0,\infty])^{mp}$; and, since the inverse transformation of ϕ^* is continuous in the uniform metric, reapplication of van der Vaart and Wellner's Theorem 1.3.6 yields HJD-weak convergence in the uniform topology on $(\mathbf{D}[0,\infty])^{mp}$ of collections of $\{X^n_\circ(f), f \in \mathbf{h}\}$, to m p-variate Gaussian processes with mean $\mathbf{0}$ and covariance function $\boldsymbol{V}_{fg}(t)$, for $f, g \in \mathbf{h}$.

The uniform consistency of the variance estimators on $\mathcal{I}$ is established if we can show that for any $t \in \mathcal{I}$,

$$\text{(A.9)} \qquad \sup_{s\in[0,t]} \left| \int_0^s \frac{d\overline{N}(x)}{\overline{Y}(x)} - \Lambda(s) \right| \xrightarrow{P} 0.$$

Define

$$\begin{aligned} M^n_{**}(t) &= \int_0^t I_{\{\overline{Y}(s)>0\}} \left\{ \frac{d\overline{N}(s)}{\overline{Y}(s)} - d\Lambda(s) \right\} \\ &= \int_0^t \frac{I_{\{\overline{Y}(s)>0\}}}{\overline{Y}(s)} \sum_{j=1}^n dM^n_j(s) + \int_0^t \frac{I_{\{\overline{Y}(s)>0\}}}{\overline{Y}(s)} \sum_{j=1}^n Y_j(s) \left[d\Lambda^n_j(s) - d\Lambda(s) \right]. \end{aligned}$$

The second term converges uniformly to 0 in probability as $n \longrightarrow \infty$. The first term is a square integrable martingale with predictable variation process

$$\int_0^t \frac{I_{\{\overline{Y}>0\}}}{\overline{Y}^2} \sum_{j=1}^n \left(1 - \Delta\Lambda^n_j\right) d\Lambda^n_j \le \int_0^t \frac{I_{\{\overline{Y}>0\}}}{\overline{Y}^2} \sum_{j=1}^n \left(d\Lambda^n_j - d\Lambda\right) + \int_0^t \frac{I_{\{\overline{Y}>0\}}}{\overline{Y}} \frac{n}{\overline{Y}} d\Lambda \xrightarrow{P} 0,$$

since $\sup_t |d\Lambda^n_j(t) - d\Lambda(t)| \longrightarrow 0$, $\Lambda(t) < \infty$ and $\overline{Y}(t)/n \xrightarrow{P} \pi(t) > 0$, $\forall t \in \mathcal{I}$. Thus, $\langle M^n_{**}, M^n_{**}\rangle(t) \xrightarrow{P} 0$, $\forall t \in \mathcal{I}$. By Lenglart's inequality (Lenglart, 1977), (A.9) holds for all $t \in \mathcal{I}$.

For $u \notin \mathcal{I}$,

$$\sup_{h,h',k,k'} \left| V^n_{hk,h'k'}(u,u) - V^n_{hk,h'k'}(t,t) \right| \le 4M^2 n^{-1} \int_t^u \sum_{j=1}^n dN_j \xrightarrow{P} 0 \text{ as } t \uparrow u,$$

since by (A.7)

$$\mathrm{P}\left\{ \left| n^{-1} \int_t^u \sum_{j=1}^n dN_j \right| > \epsilon \right\} \le \frac{1}{\epsilon} \mathrm{E}\left[n^{-1} \int_t^u \sum_{j=1}^n dN_j \right] \le \frac{1}{\epsilon} \mathrm{E}\left[n^{-1} \int_t^u \sum_{j=1}^n Y_j d\Lambda_j^n \right] \xrightarrow{t \uparrow u} 0,$$

for any $\epsilon > 0$. Also, by (A.8),

$$\sup_{h,h',k,k'} \left| V_{hk,h'k'}(u,u) - V_{hk,h'k'}(t,t) \right| \le 4M^2 n^{-1} \int_t^u \pi(s) d\Lambda(s) \longrightarrow 0, \quad \text{as } t \uparrow u.$$

For $u < \infty$, similar arguments can also be used to establish that

$$\sup_{h,h',k,k'} \left| V^n_{hk,h'k'}(\infty,\infty) - V^n_{hk,h'k'}(u,u) \right| \xrightarrow{P} 0, \text{ and}$$
$$\sup_{h,h',k,k'} \left| V_{hk,h'k'}(\infty,\infty) - V_{hk,h'k'}(u,u) \right| \longrightarrow 0.$$

Consistency now follows by the Skorohod-measurability of the sup-norm and Theorem 1 (c).□

REFERENCES

β-Blocker Heart Attack Trial Research Group (1982). A randomized trial of propranolol in patients with acute myocardial infarction. *Journal of the American Medical Association* **247** 1707–1714.

BROWN, B. W., JR., HOLLANDER, M., and KORWAR, R. M. (1974). Nonparametric tests of independence of censored data, with applications to heart transplant studies. In *Reliability and Biometry, Statistical Analysis of Lifelength*, F. Proschan and R. J. Serfling, 327–354. Philadelphia: Society for Industrial and Applied Mathematics.

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

DOLIVO, F. (1974). Counting processes and integrated conditional rates: a martingale approach with application to detection. Technical Report, College of Engineering, Univ. Michigan.

ETHIER, S. N., and KURTZ, T. G. (1986). *Markov Processes.* New York: John Wiley and Sons.

FLEMING, T. R., and HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis.* New York: John Wiley and Sons.

FLEMING, T. R., HARRINGTON, D. P., and O'SULLIVAN, M. (1987). Supremum versions of the log-rank and generalized Wilcoxon statistics. *Journal of the American Statistical Association* **82**, 312–320.

GASTWIRTH, J. L. (1985). The use of maximin efficiency robust tests in combining contingency tables and survival analysis. *Journal of the American Statistical Association* **80**, 380–384.

GEHAN, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52** 203–223.

GILL, R. D. (1980). *Censoring and Stochastic Integrals.* volume 124 of *Mathematical Centre Tracts.* Amsterdam: Mathematisch Centrum.

HARRINGTON, D. P., AND FLEMING, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69**, 133–143.

JONES, M. P. and CROWLEY, J. (1989). A general class of nonparametric tests for survival analysis. *Biometrics* **45** 157–170.

JONES, M. P. and CROWLEY, J. (1990). Asymptotic properties of a general class of nonparametric tests for survival analysis. *Ann. Statistics* **18** 1203–1220.

KOSOROK, M. R. (1997). Families of Weighted Log-rank Statistics as Function-indexed Stochastic Processes (version 1b). Technical Report 98, Department of Biostatistics, University of Wisconsin, Madison, Wisconsin.

KOSOROK, M. R. and LIN, C. (1997). The versatility of function-indexed weighted log-rank statistics (version 2a). Technical Report 105, Department of Biostatistics, University of Wisconsin, Madison, Wisconsin.

O'BRIEN, P. C. (1978). A nonparametric test for association with censored data. *Biometrics* **34** 243–250.

PEPE, M. S. and FLEMING, T. R. (1989). Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics* **45** 497–507.

PETO, R. and PETO, J. (1972). Asymptotically efficient rank-invariant test procedures (with discussion). *Journal of the Royal Statistical Society, Series A* **135** 185–206.

PRENTICE, R. L. (1978). Linear rank tests with right-censored data. *Biometrika* **65** 167–170.

TARONE, R. E. (1975). Tests for trend in life table analysis. *Biometrika* **62** 679–682.

TARONE, R. E. (1981). On the distribution of the maximum of the log-rank statistic and the modified Wilcoxon statistic. *Biometrics* **37**, 79–85.

TARONE, R. E. and WARE, J. (1977). On distribution-free test for equality of survival distributions. *Biometrika* **64** 156-160.

VAN DER VAART, A. W., and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes.* New York: Springer.