Omnibus Tests for Comparison of Competing Risks with Adjustment for Covariate Effects

Ian W. McKeague

Department of Statistics, Florida State University, Tallahassee, Florida 32306, U.S.A.

e-mail: mckeague@stat.fsu.edu

and

Peter B. Gilbert

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

e-mail: pgilbert@hsph.harvard.edu

and

Phyllis J. Kanki

Harvard AIDS Institute and Department of Immunology and Infectious Diseases

Harvard University, Boston, Massachusetts 02115, U.S.A.

e-mail: pkanki@hsph.harvard.edu

SUMMARY. This paper develops omnibus tests for comparing cause-specific hazard rates and cumulative incidence functions at specified covariate levels. Confidence bands for the difference and the ratio of two conditional cumulative incidence functions are also constructed. The omnibus test is formulated in terms of a test process given by a weighted difference of estimates of cumulative cause-specific hazard rates under Cox proportional hazards models. A simulation procedure is devised for sampling from the null distribution of the test process, leading to graphical and numerical techniques for detecting significant differences in the risks. The approach is applied to a cohort study of type-specific HIV infection rates.

KEY WORDS: Cause-specific hazard rates; Cox proportional hazards model; Cumulative incidence function; Dependent competing risks; Human immunodeficiency virus.

1. Introduction

In longitudinal studies, where individuals are subject to failure from a number of competing risks and the eventual failure can be attributed to precisely one of the risks, it can be of interest to compare the various hazards when they are adjusted for covariate effects. Comparisons of this type are useful, for example, in the design of vaccines for the human immunodeficiency virus (HIV).

The global HIV pandemic is characterized by the circulation of many HIV genotypes (Louwagie et al., 1993). In order for an HIV vaccine to be efficacious in a particular geographic region, it may be necessary to match the genotypes of the HIV antigens contained in the vaccine to the local HIV genotypes that pose the greatest risk of HIV infection (Heyward, Osmanov, and Esparza, 1992; Moore and Anderson, 1994). For instance, in the hope of obtaining broad vaccine protection, the ongoing HIV vaccine efficacy trial in Bangkok is testing a vaccine matched for the local circulating HIV type 1 (HIV-1) subtypes B and E (Berman, 1998). In general, when preparing for an HIV vaccine efficacy trial in a particular region, it is important to compare the infection rates of locally circulating HIV genotypes, to guide prioritization of HIV antigen types for inclusion in the tested vaccine. Moreover, to make comparisons of genotype-specific infection rates for this purpose interpretable, it is important to adjust the comparison for risk factors of HIV infection.

The methods developed in this paper are motivated by a prospective study of a cohort of female prostitutes in Senegal, which forms a candidate population for a new HIV vaccine trial. Two types of HIV (HIV-1 and HIV-2) are known to circulate in Senegal, and the methods are applied to compare their infection rates, with adjustment for risk factors such as age and the frequency of sexual contacts. In addition, the methods are applied to compare covariate-adjusted male-to-female per-contact transmission probabilities of the two virus types. This is accomplished with our testing procedure by adjusting for the differing prevalences of HIV-1 and HIV-2 in the exposing male partner population entered as a time-dependent covariate. Understanding differences in transmissibilities between viral genotypes may be useful for forecasting the evolution of the global HIV pandemic and for improving the design of HIV vaccines and treatments.

We consider a competing risks framework in which individuals are at risk from k types of failure and covariate measurements on each individual are available. Within this framework, we provide graphical and numerical methods for comparing any two of the k cause-specific hazard rates, or cumulative incidence functions, at a specified covariate level. We use the standard formulation of the competing risks model which assumes the existence of a latent failure time T_j corresponding to each failure type $j = 1, \ldots, k$. The observed time of failure is given by $X = \min_j T_j$, which may be right censored. When X is uncensored, the cause of failure $\delta \in \{1, \ldots, k\}$ is observed along with a p-vector Z representing the covariate information. For simplicity, we restrict attention to a time-independent covariate, but a time-dependent covariate $Z(t)$ can be handled with only minor changes. The conditional distributions of the latent failure times given the covariate are in general not identifiable from data on (X, δ, Z), even in the absence of censoring. Instead, statistical interest focuses on the conditional cumulative incidence function $F_j(t|z) = P(X \leq t, \delta = j|Z = z)$, and the conditional cause-specific hazard rate $\lambda_j(t|z) = \lim_{\epsilon \to 0} P(t \leq X < t + \epsilon, \delta = j|X \geq t, Z = z)/\epsilon$, both of which are estimable from the competing risks data, see Prentice et al. (1978).

We use Cox proportional hazards models to specify each $\lambda_j(t|z)$ and assume that the censoring is conditionally independent of the latent failure times given Z. Under these assumptions we develop a graphical method, along with a formal procedure, for testing the null hypothesis

$$H_0: \lambda_1(t|z_0) = \lambda_2(t|z_0),\ \ 0 \le t \le \tau$$

where z_0 is a specified p-vector of covariate levels, and $[0, \tau]$ is the time interval of interest. The following alternative hypotheses will be considered:

$$\begin{aligned} H_1: \lambda_1(t|z_0) &\ne \lambda_2(t|z_0),\ \text{for some}\ \ 0 \le t \le \tau \\ H_2: F_1(t|z_0) &\le F_2(t|z_0),\ \ 0 \le t \le \tau \\ H_3: \lambda_1(t|z_0) &\le \lambda_2(t|z_0),\ \ 0 \le t \le \tau \end{aligned}$$

with strict inequality for some $t \in [0, \tau]$ in H_2 and H_3. The hypothesis H_0 is equivalent to equality of the corresponding cumulative incidence functions over $[0, \tau]$, because $F_j(t|z) = \int_0^t \lambda_j(u|z) S_X(u|z)\, du$, where $S_X(t|z) = P(X > t|Z = z)$ is the conditional survival function of X. The hypotheses H_2 and H_3 are ordered alternatives expressing the notion that cause 2 is "more serious" than cause 1, with H_3 being the more restrictive alternative. H_2 is appropriate for a comparison in terms of absolute risk, and H_3 in terms of risk intensity. We are interested in developing omnibus tests that are consistent against all departures from H_0 in the directions of H_1, H_2 and H_3.

There is a large literature on the problem of testing the equality of two cause-specific hazard rates, but at least one of the following assumptions has been required: independent competing risks, no covariates, no censoring, or parametric models. Nonparametric tests that allow censoring and dependent competing risks, but no covariates, have been considered by Aly, Kochar and McKeague (1994), Sun and Tiwari (1995), Lam (1988), Hu and Tsai (1999), Luo and Turnbull (1999) and Sun (2000), among others. Cox proportional hazards models have been widely used in the competing risks context, see, e.g., Prentice et al. (1978), Holt (1978), Larson (1984), Andersen, Hansen and Keiding (1991), Lunn and McNeil (1995), and Cheng, Fine and Wei (1998). As far as we know, however, omnibus tests for the comparison of two cause-specific hazard rates have not been studied in this context.

The paper is organized as follows. In Section 2 we develop test statistics for detecting departures from H_0 in the direction of the alternatives H_1, H_2 and H_3. In Section 3 we derive confidence bands for the difference and ratio of two conditional cumulative incidence functions. Section 4 describes the results of a simulation study assessing the accuracy and power of the proposed tests. The application to data on HIV infection rates is presented in Section 5. Theoretical results are placed in the Appendix.

2. Test procedure

2.1. *Preliminaries*

Let C denote the censoring time. The competing risks model data are assumed to be given by n independent replicates of $(\tilde{X}, \tilde{\delta}, Z)$, where $\tilde{X} = \min(X, C)$, $\tilde{\delta} = \delta I(X \leq C)$ and $I(\cdot)$ is the indicator function. It is also assumed that C is conditionally independent of $T_1, \ldots, T_k$ given Z. The latent failure times T_j do not have to be independent, but we do require that $P(T_j = T_l) = 0$ for $j \neq l$. The cause of failure $\delta = j$ when $X = T_j$.

The conditional cause-specific hazard rates $\lambda_j(t|z)$ are specified by separate Cox proportional hazards models [Cox (1972), Andersen et al. (1991)]:

$$\lambda_j(t|z) = \lambda_{0j}(t) \exp(\beta_j' z),$$

where $\lambda_{0j}(\cdot)$ is an unspecified baseline hazard function and β_j is a p-vector of regression parameters for the jth cause of failure. The partial likelihood score function for β_j is

$$U_j(\beta) = \sum_{i=1}^{n} \Delta_{ji}(Z_i - \bar{Z}(\beta, \tilde{X}_i)),$$

where $\Delta_{ji} = I(\tilde{\delta}_i = j)$ and

$$\bar{Z}(\beta, t) = \sum_{i=1}^{n} Y_i(t) \exp(\beta' Z_i) Z_i \Big/ \sum_{i=1}^{n} Y_i(t) \exp(\beta' Z_i), \quad Y_i(t) = I(\tilde{X}_i \geq t).$$

The denominator of $\bar{Z}(\beta,t)$ is denoted $S^{(0)}(\beta,t)$ in the sequel. The maximum partial likelihood estimator $\hat{\beta}_j$ is the solution to the estimating equation $U_j(\beta) = 0$. Under some mild regularity conditions (Andersen et al., 1993, Chapter VII), $\mathcal{I}_j^{-1/2}(\hat{\beta}_j)(\hat{\beta}_j - \beta_j)$ is asymptotically zero-mean normal with identity covariance matrix, where $\mathcal{I}_j(\beta)$ is minus the derivative matrix of $U_j(\beta)$. The $\hat{\beta}_j$ are asymptotically independent. The counting process $N_{ji}(t) = \Delta_{ji} I(\tilde{X}_i \leq t)$ records observed failures of type j. The jth cumulative cause-specific hazard function $\Lambda_j(t|z) = \int_0^t \lambda_j(u|z)\,du$ can be estimated by

$$\hat{\Lambda}_j(t|z) = \hat{\Lambda}_{0j}(t)\exp(\hat{\beta}_j' z),$$

where the leading term is the Nelson–Aalen type estimator

$$\hat{\Lambda}_{0j}(t) = \int_0^t \frac{\sum_{i=1}^n dN_{ji}(u)}{S^{(0)}(\hat{\beta}_j, u)}.$$

These estimators are special cases of estimators studied by Andersen et al. (1991) in connection with proportional hazards models for the transition intensities of non-homogeneous Markov chains.

2.2. *Test process*

Our approach is based on a comparison of the cumulative cause-specific hazard estimators at the specified covariate level z_0, using a general predictable locally bounded non-negative weight processes $W(\cdot)$:

$$L(t) = \int_0^t W(u)(\hat{\Lambda}_2(du|z_0) - \hat{\Lambda}_1(du|z_0)).$$

The weight process provides a flexible way to control the relative importance attached to differences in the cause-specific hazards at different times, and is useful for controlling instability in the tails. Some examples of weight processes are discussed in Section 2.5.

A plot of the test process $L(t)$ is helpful in looking for possible departures from H_0, with a tendency for large absolute values under H_1, large positive values under H_2, and an increasing trend under H_3. This can be seen from the identity

$F_j(t|z) = \int_0^t \lambda_j(u|z) S_X(u|z)\, du$. However, these plots can be difficult to interpret due to fluctuations in the test process which occur even under the hypothesis of equal cause-specific hazards.

In the Appendix we show that $L(t)$ converges in distribution under H_0 to a zero-mean Gaussian process provided $W(t)/\sqrt{n}$ converges uniformly in probability over $[0,\tau]$ to a bounded function. Yet the limiting covariance is complicated, so this result does not immediately translate into a workable test procedure. For that purpose, we develop a simulation method for approximately sampling from the null distribution of $L(t)$, adapting the procedure presented in Lin, Wei and Ying (1993) for checking adequacy of the Cox proportional hazards model. This procedure was also used by Cheng, Fine and Wei (1998) to obtain confidence intervals and bands for the predicted conditional cumulative incidence function $F_j(t|z_0)$.

2.3. *Sampling from the null distribution of the test process*

First we need a representation of the test process in terms of the basic martingales

$$M_{ji}(t) = N_{ji}(t) - \int_0^t Y_i(u) \exp(\beta_j' Z_i) \lambda_{0j}(u)\, du.$$

Using Taylor series expansions of $\exp(\hat\beta_j' z_0)$ and $U_j(\hat\beta_j)$ around β_j, the process $L(t)$ is seen, under H_0, to be asymptotically equivalent to the process $L_2(t) - L_1(t)$, where

$$L_j(t) = \psi(t, \hat\beta_j)' \mathcal{I}_j^{-1}(\hat\beta_j) \sum_{i=1}^n \int_0^\infty (Z_i - \bar Z(\hat\beta_j, u))\, dM_{ji}(u) + \sum_{i=1}^n \int_0^t \frac{W(u) \exp(\hat\beta_j' z_0)}{S^{(0)}(\hat\beta_j, u)}\, dM_{ji}(u)$$

and

$$\psi(t, \beta) = \tfrac{1}{2} \int_0^t W(u)(z_0 - \bar Z(\beta, u))(\hat\Lambda_1(du|z_0) + \hat\Lambda_2(du|z_0)).$$

The process $L^*(t)$ used to simulate $L(t)$ is defined by replacing $M_{ji}(u)$ in the first and second terms of $L_j(t)$ by $G_{ji} N_{ji}(u)$ and $\tilde G_{ji} N_{ji}(u)$, respectively, where $\{G_{ji}, \tilde G_{ji}: j = 1,2; i = 1,\ldots,n\}$ are independent standard normal variables. Realizations of $L^*(t)$ are approximate draws from the null distribution of the test process. More specifically, under H_0 the conditional distribution of $L^*(t)$ given the observed data is the same in the limit as the unconditional distribution of $L(t)$, see the Appendix. The method works essentially because $M_{ji}(t)$ has mean zero and variance $E\{N_{ji}(t)\}$.

An alternative version of $L^*(t)$, with the same asymptotic properties, is obtained by replacing $M_{ji}(u)$ by $G_{ji}N_{ji}(u)$ in both terms of $L_j(t)$; this is closer to the method proposed by Lin, Wei and Ying (1993). However, we prefer the first method because it better reflects the asymptotic independence of the two terms in $L_j(t)$; see the proof of Theorem 1 in the Appendix for more details.

2.4. *Test statistics*

For formal procedures, the following test statistics are suggested for detecting departures from H_0 in the direction of H_1, H_2, H_3:

$$D_1 = \sup_{0\le t\le\tau} |L(t)|, \quad D_2 = \sup_{0\le t\le\tau} L(t), \quad D_3 = \sup_{0\le s<t\le\tau} \{L(t) - L(s)\},$$

respectively.

2.5. *Choice of weight process*

The simplest choice is $W_1(t) = \sqrt{n}$, which reduces the test process to the normalized difference of the estimated cumulative cause-specific hazard functions:

$$L(t) = \sqrt{n}(\hat{\Lambda}_2(t|z_0) - \hat{\Lambda}_1(t|z_0)).$$

This may be a good choice for the graphical procedure, where ease of interpretation is important. The variance of $L(t)$ increases with t, however, so for a formal test it is preferable to use a decreasing weight process that gives less weight to the tail, such as

$$W_2(t) = \sum_{i=1}^{n} I(\tilde{X}_i \ge t\}/\sqrt{n}.$$

A more sophisticated weight process that achieves a similar effect is

$$W_3(t) = \hat{S}_X(t-|z_0)^{1/2}\left(\frac{\exp(\hat{\beta}_1' z_0)}{S^{(0)}(\hat{\beta}_1, t)} + \frac{\exp(\hat{\beta}_2' z_0)}{S^{(0)}(\hat{\beta}_2, t)}\right)^{-1/2},$$

where $\hat{S}_X(t|z_0)$ is the Cox-model based estimator of the conditional survival function of X, see Andersen et al. (1993, p. 509). In this case the asymptotic distribution of $L(t)$ is of a relatively simple form; the variance function $V(t)$ simplifies to the conditional cumulative incidence function $F_1(t|z_0)$, see Theorem 1 in the Appendix. For $k = 2$ and no covariates, $W_3(t)$ reduces to a weight process considered by Aly et al. (1994, Section 3.1) and makes the test statistics D_1, D_2, D_3 asymptotically distribution-free.

Another relatively simple choice is $W_4(t) = \sqrt{n}\hat{S}_X(t - |z_0)$, which gives the normalized difference of the estimated cumulative incidence functions:

$$L(t) = \sqrt{n}(\hat{F}_2(t|z_0) - \hat{F}_1(t|z_0)),$$

where $\hat{F}_j(t|z_0) = \int_0^t \hat{S}_X(u - |z_0)\hat{\Lambda}_j(du|z_0)$.

2.6. *Remarks*

Within the Cox model setting and under mild conditions on the weight process, the tests based on D_1, D_2 and D_3 are consistent against any departure from H_0 in the direction of their respective alternatives; see the remark following the proof of Theorem 1 in the Appendix.

When the covariate is one-dimensional it is feasible to use a nonparametric model for $\lambda_j(t|z)$. In that case, a nonparametric estimator (McKeague and Utikal, 1990) of $\Lambda_j(t|z_0)$ should be used in place of $\hat{\Lambda}_j(t|z_0)$ in $L(t)$. This alternative version of $L(t)$ can be shown to converge in distribution to a time-changed Wiener process, and provides the basis for an asymptotically distribution-free test of H_0. However, the nonparametric approach fails in general, due to the curse of dimensionality. An alternative approach for covariates with a moderate number of components would be to use an additive risk model in place of the Cox model, as in Shen and Cheng (1999) who developed simultaneous confidence bands for cause-specific cumulative incidence functions.

In some applications it may be of interest to test H_0 simultaneously over a range of values $\mathcal{Z}$ of the covariate level z_0. Our approach extends readily to this setting with the test process $L(t) = L(t, z_0)$ and its simulated version $L^*(t) = L^*(t, z_0)$ now indexed by $(t, z_0) \in [0, \tau] \times \mathcal{Z}$, and the test statistics D_1, D_2, D_3 replaced by their maximal values over $z_0 \in \mathcal{Z}$.

Hu and Tsai (1999) recently considered the problem of finding optimal weight functions for a class of linear rank tests of H_0 in the absence of covariates. The resulting test statistics are optimal against local Lehmann alternatives. A Lehmann alternative in our setting means $\lambda_2(t|z) = \lambda_{01}(t) \exp(\gamma + \beta_2' z)$ for some constant γ, so H_0 (H_3) is equivalent to $\gamma + (\beta_2 - \beta_1)' z_0 = 0$ (> 0). Thus H_0 can be tested by fitting a model of Holt (1978) because γ, β_1 and β_2 are regression parameters in a cause-specific hazard model having common baseline hazards. However, the class of Lehmann alternatives is too restrictive for many applications, see the discussion of the relative merits of log-rank and Kolmogorov–Smirnov type tests in Andersen et al. (1993, pp. 390–395); in the HIV application, for example, there is no reason to believe that the HIV-1 and HIV-2 cause-specific hazard rates are proportional to one another. The proposed omnibus tests should be more powerful against the types of alternatives that are likely to occur in practice.

In some applications, it may be plausible that a covariate has the same effect on the two cause-specific hazards, i.e., the regression coefficients are the same in the both Cox models [Lunn and McNeil's (1995) duplication method B provides a simple way of testing this]. Such structure can be exploited to improve efficiency. This can be done by reformulating the two Cox models in terms of failure-type-specific covariates and a single vector of regression coefficients (cf. Andersen et al., 1993, p. 478). The second sampling method described in subsection 2.3 is then used to sample from the asymptotic null distribution of the test process. The processes $L_j(t)$ are asymptotically dependent in this case, and the statement of Theorem 1 needs to be modified accordingly.

3. Confidence bands

In this section we construct confidence bands for the difference and the ratio of two conditional cumulative incidence functions.

A consistent estimator of the difference $\delta(t) = F_2(t|z_0) - F_1(t|z_0)$ can be obtained from $L(t)$ using $W_4(t)$ as the weight process: $\hat{\delta}(t) \equiv \hat{F}_2(t|z_0) - \hat{F}_1(t|z_0) = L(t)/\sqrt{n}$. From the proof of Theorem 1 in the Appendix, it follows that the process $\sqrt{n}(\hat{\delta}(t)-\delta(t))$ converges in distribution to the (null) limiting distribution of $L(t)$. The earlier Monte Carlo procedure based on $L^*(t)$ can be used to estimate an upper $\alpha/2$-quantile, $c_{\alpha/2}(t)$, of the limiting distribution of $|L(t)|$. An approximate $100(1-\alpha)\%$ pointwise confidence band for $\delta(t)$ is then given by $\hat{\delta}(t) \pm c_{\alpha/2}(t)n^{-1/2}$. A simultaneous band for $\delta(t)$ can be found by suitably scaling the pointwise band: $\hat{\delta}(t) \pm ac_{\alpha/2}(t)n^{-1/2}$, $t \in [0, \tau]$, where $a \geq 1$; here the Monte Carlo procedure is used to adjust the constant a to furnish the desired $100(1-\alpha)\%$ simultaneous confidence level. Such bands are illustrated in Figure 5 below. Alternatively, the transformation approach of Cheng, Fine and Wei (1998, p.221), henceforth CFW, could be adapted for this purpose.

A simultaneous band for the ratio $\rho(t) = F_1(t|z_0)/F_2(t|z_0)$ can be based on the estimate $\hat{\rho}(t) = \hat{F}_1(t|z_0)/\hat{F}_2(t|z_0)$. Approximate draws from the joint asymptotic distribution of $\sqrt{n}(\hat{F}_j(t|z_0) - F_j(t|z_0))$, $j = 1, 2$, are obtained from the simulated processes $U_j^*(t)$, $j = 1, 2$, given by expression (2) of CFW. Using the functional delta method (Andersen et al., 1993), it can be shown that the distribution of the process $\sqrt{n}(\hat{\rho}(t) - \rho(t))$ coincides in the limit with conditional distribution of

$$U^*(t) = U_1^*(t)/\hat{F}_2(t|z_0) + U_2^*(t)\hat{F}_1(t|z_0)/\hat{F}_2(t|z_0)^2$$

given the data, where $0 < t_1 < t < t_2$ and it is assumed that $F_2(t_1|z_0) > 0$ and $t_2 < \inf\{t\colon P(\tilde{X} \geq t) = 0\}$. A simultaneous confidence band for $\rho(t)$ can then be constructed as in CFW (p. 221), with the obvious modifications: $U^*(t)$ replaces $\hat{U}_k(t; z_0)$

and $\hat{\rho}(t)$ replaces $\hat{F}_k(t|z_0)$.

4. Simulation study

We designed simulations to study the performance of the proposed test procedure under various scenarios relevant to the application in the next section. Key questions to be addressed are: Does the nominal size accurately match observed levels at moderate sample sizes? Which weight function gives the best performance in terms of power? How does a trend in the baseline hazard function, say due to heterogeneity in biological susceptibility to infection, affect the size and power?

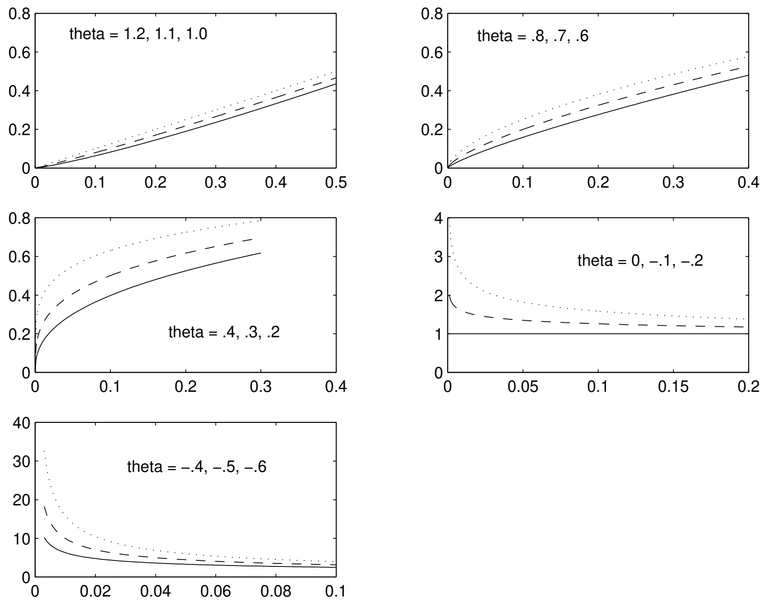

Figure 1: Weibull baseline hazard functions used for the simulation study, plotted over the relevant interval $[0, \tau]$ in each case. Values of θ_1 correspond to the solid line, and values of θ_2 to the other lines.

We considered a $p = 2$ dimensional covariate Z, having independent components uniformly distributed on $[0, 1]$, and specified $z_0 = (0.5, 0.5)$. The $k = 2$ latent failure times were taken to be conditionally independent with conditional cause-specific hazard functions given by Cox models of the form $\lambda_j(t|z) = t^{\theta_j} \exp(z_1 + z_2)$, for various choices of $\theta_j > -1$, $j = 1, 2$. Note that the conditional distribution of T_j given Z is Weibull with shape parameter $\theta_j + 1$. The baseline hazard is decreasing and convex if $-1 < \theta_j < 0$, constant if $\theta_j = 0$, increasing and concave if $0 < \theta_j \leq 1$, and increasing and convex if $1 \leq \theta_j$. The various baseline hazards being compared are plotted over $[0, \tau]$ in Figure 1.

The censoring C was taken to be exponentially distributed, with decay parameter adjusted so 15–30% of the observations were censored in $[0, \tau]$. The size and power of the test based on D_2 at the nominal 0.05 level were estimated from 1000 independent samples, with critical values obtained in each sample from 1000 realizations of $L^*(t)$.

Table 1.

Observed levels and powers of test for equality of conditional cause-specific hazard rates based on D_2 at nominal level 0.05

		$n = 100$				$n = 200$			
θ_1	θ_2	W_1	W_2	W_3	W_4	W_1	W_2	W_3	W_4
1.2	1.2	0.056	0.052	0.057	0.054	0.046	0.046	0.045	0.041
	1.1	0.105	0.112	0.115	0.103	0.149	0.158	0.158	0.153
	1.0	0.180	0.200	0.204	0.186	0.298	0.321	0.319	0.303
0.8	0.8	0.063	0.061	0.070	0.057	0.057	0.059	0.059	0.064
	0.7	0.144	0.159	0.166	0.150	0.158	0.169	0.169	0.158
	0.6	0.282	0.296	0.300	0.291	0.430	0.479	0.479	0.462
0.4	0.4	0.054	0.046	0.052	0.048	0.052	0.054	0.052	0.054
	0.3	0.166	0.172	0.184	0.167	0.257	0.301	0.303	0.284
	0.2	0.375	0.446	0.448	0.407	0.642	0.722	0.717	0.679
0.0	0.0	0.050	0.052	0.059	0.052	0.059	0.061	0.061	0.059
	−0.1	0.247	0.333	0.330	0.280	0.393	0.505	0.507	0.451
	−0.2	0.654	0.815	0.814	0.739	0.898	0.981	0.980	0.949
−0.4	−0.4	0.052	0.061	0.065	0.059	0.057	0.046	0.045	0.050
	−0.5	0.441	0.824	0.830	0.675	0.644	0.982	0.982	0.901
	−0.6	0.866	1.000	1.000	0.999	0.968	1.000	1.000	1.000

The results, reported in Table 1, show that the observed levels for D_2 quite accurately match the 0.05 nominal level of the test, and trends in the baseline hazard functions have no appreciable effect on accuracy. Similar results, not reported here, were found for the tests based on D_1 and D_3. The weight functions W_2 and W_3 gave better performance than W_1 and W_4 in terms of power; this was expected because W_1 and W_4 do not down-weight observations in the tail, where there tends to be a sharp increase in the variance of the cumulative baseline hazard estimate.

5. Application: Comparison of HIV-1 and HIV-2 infection rates

In 1985, the Senegalese government in collaboration with the Harvard School of Public Health began a prospective cohort study in self-identified female prostitutes in Dakar, Senegal (Kanki et al., 1990; Kanki et al., 1992, Kanki, 1999). Cohort participants were followed through regular visits to health clinics for varying time-intervals between February 7, 1985 and November 1, 1999. At each clinic visit, women were tested for infection with human immunodeficiency virus types 1 (HIV-1) and 2 (HIV-2). Covariates measuring risk for HIV infection were also collected, including nationality, age at cohort entry, calendar date at entry, years of registered prostitution at entry, average number of sexual partners per week, extent of condom use, and infection with sexually transmitted diseases (STDs) other than HIV. The last three covariates are time-dependent. This prospective study of 1948 initially HIV uninfected women provides excellent data for comparing type-specific hazard rates of the two viral competing risks of infection. For interpretability, it is important to adjust the comparison for risk factors.

In addition, we apply the tests to assess if the male-to-female per-sexual contact transmission probability (the "infectivity") of HIV-1 differs from that of HIV-2. This can be done by adjusting the comparison of HIV-1 and HIV-2 crude hazard rates for differences in the HIV-1 and HIV-2 prevalence rates in the exposing male partner population. To accomplish this within the framework of the tests developed here, we include the log ratio of HIV-2 versus HIV-1 prevalence in the infected male partner population as a time-dependent covariate. For the ith woman t years into follow-up, this ratio was calculated as the HIV-2 versus HIV-1 prevalence ratio in all female sex workers under follow-up at the calendar time corresponding to t. These calculations included sex workers HIV infected and uninfected at entry, totaling 3141 women. By fixing the level of the log partner prevalence ratio covariate at zero, the statistics D_1 and D_3 test the null hypothesis of equal HIV-1 and HIV-2 infectivity versus 2-sided and 1-sided alternatives, respectively (Gilbert et al., submitted manuscript, provide a detailed argument for this). Studies showing that HIV-1 has a shorter asymptomatic period (Marlink et al., 1994), a higher viral load (Albert et al., 1990; Popper et al., 1999), and a higher perinatal transmission rate (Poulsen et al., 1992; Abbot et al., 1994; Adjorlo-Johnson et al., 1994) support the hypothesis that HIV-1 is more heterosexually infectious than HIV-2.

During the follow-up period 199 prostitutes became infected with HIV, 127 with type 1 only, 66 with type 2 only, and 6 with both types. The time-to-infection was calculated as the time from entry into the cohort to the midpoint between the last HIV seronegative visit date and the first seropositive visit date. Among the 6 dual-infected women, 3 had first seropositive test seropositive for both viruses; these women were assumed to have a simultaneous HIV-1 and HIV-2 seroconversion date. Since these cases violated an assumption of the test, they were not used in the analysis. As shown in Figure 2, in which no adjustment was made for covariate effects, the cumulative incidence of HIV-1 exceeded that of HIV-2, and the infection hazard of each type peaked after 5 years and then waned.

Most women in the initially HIV uninfected cohort were Senegalese (73.2%) or Ghanaian (14.0%), with average age 30.4 years (range 19 to 56 years), average 2.7 years of registered prostitution (range 0 to 26 years), and average date of cohort entry 5.16 years after the study start date February 7, 1985 (range 0 to 14.45 years). The mean number of partners per week was 6.8 (interquartile range 3.5 to 8.0). Condom use was entered as a binary covariate with levels always and sometimes or never, with

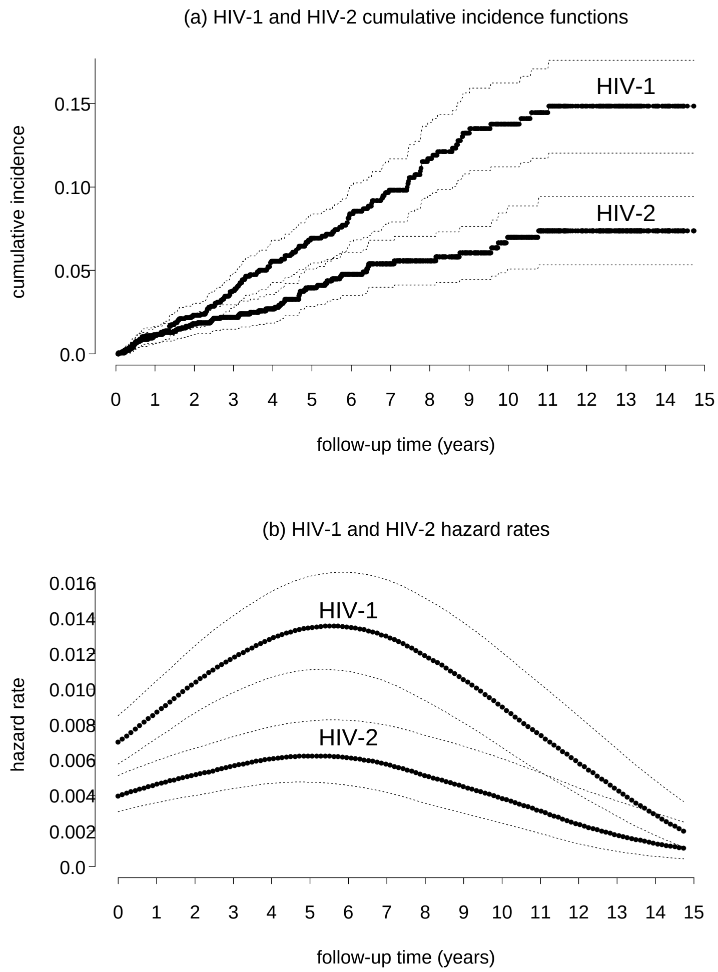

Figure 2: Senegal cohort study. No adjustment for covariates. (a) Estimated HIV-1 and HIV-2 cumulative incidence functions with 95% pointwise confidence limits. (b) Nonparametric Epanechnikov kernel estimates of the HIV-1 and HIV-2 hazard rates from smoothed estimates of the cumulative cause-specific hazard functions (Ramlau-Hansen, 1983); normal approximation 95% pointwise confidence limits calculated by transforming the symmetric confidence limits (Andersen et al. 1993, p. 249).

76.1% reporting sometimes or never, and STDs was entered as a binary covariate with levels no STD or at least one STD, with 50.9% testing positive for at least one STD during follow-up. Missing values for the time-dependent covariates were filled in by nearest observed values within a subject's longitudinal profile of clinic visits.

At the 0.05 significance level, univariable Cox models identified four risk factors for HIV-1 infection: non-Senegalese nationality, Ghanaian nationality, later date of cohort entry, and older at cohort entry (Table 2). The same variables were significant univariable risk factors for HIV-2 infection, with hazard ratios of comparable magnitude to those for HIV-1, with exception that an earlier date of cohort entry predicted an increased risk of HIV-2 infection. The interaction between Senegalese nationality and the date of cohort entry was also significant, with a later date of entry predicting an increased risk of HIV-2 infection for Senegalese sex workers and a decreased risk for non-Senegalese sex workers. The only covariate that predicted a significantly differential HIV-1 versus HIV-2 infection risk, as assessed by a re-coded Cox model (Method B of Lunn and McNeil, 1995), was the calendar date of cohort entry. Therefore, it would be reasonable to apply the tests that assume each covariate other than the date of entry has the same effect on the HIV-1 and HIV-2 hazard rates. However, to minimize assumptions, we applied the tests that allow each covariate to have a different regression relationship with the HIV types.

First, we applied the tests to compare cause-specific hazard rates adjusted for the significant risk-factor covariates without adjusting for the relative HIV-2 versus HIV-1 partner prevalence. The covariates were Senegalese nationality, date of cohort entry, the interaction of Senegalese nationality with date of cohort entry, and age at cohort entry (Ghanaian nationality was excluded because it was highly correlated with Senegalese nationality). Tests of H_0 versus H_2 and H_3 applied at covariate level $z_0 =$ (Senegalese, 5.16 years, 5.16 years, 30.4 years) indicated a significantly greater type 1 hazard rate and cumulative incidence function (Figure 3).

Table 2.

Significant risk factors in Cox models for HIV-1 infection, for HIV-2 infection, and for HIV-1 versus HIV-2 infection of 1948 female prostitutes participating in the Senegal cohort study, from 1985 to 1999

	Covariate	estimate[a] (s.e.)	p-value	p-value (PH)[b]
HIV-1	Senegalese	−0.470 (0.218)	0.030	0.17
	Ghanaian	0.619 (0.263)	0.017	0.075
	date of entry[c]	0.107 (0.027)	< 0.001	0.21
	age at entry[d]	0.026 (0.014)	0.054	0.028
HIV-2	Senegalese	−0.595 (0.281)	0.032	0.99
	Ghanaian	0.737 (0.333)	0.024	0.93
	date of entry	−0.127 (0.045)	0.004	0.30
	Senegalese × date of entry	0.265 (0.115)	0.021	0.63
	age at entry	0.054 (0.017)	0.0017	0.17
$\frac{\text{HIV-1}}{\text{HIV-2}}$[e]	date of entry	0.234 (0.053)	< 0.001	0.96

[a] Estimated coefficient in the univariable Cox model.
[b] p-value for test of proportional hazards (PH), based on the method of Grambsch and Therneau (1994).
[c] The covariate calendar date of cohort entry was calculated as the years since February 7, 1985, the date the first participant enrolled in the cohort study.
[d] Age in years at cohort entry.
[e] Fit by the duplication method B of Lunn and McNeil (1995).

Second, we applied the tests to compare the infectivity of HIV-1 and HIV-2 by also adjusting for the log ratio of HIV-2 versus HIV-1 prevalence in partners over time. The tests applied at covariate level $z_0 =$ (Senegalese, 5.16 years, 5.16 years, 30.4 years, 0) showed significant differences (p-values < 0.01, see Figure 4). The tests repeated for many other levels of the covariates also indicated significant differences.

From a public health perspective, it is of interest to estimate the difference between the HIV-1 and HIV-2 cumulative incidence functions for various covariate subgroups.

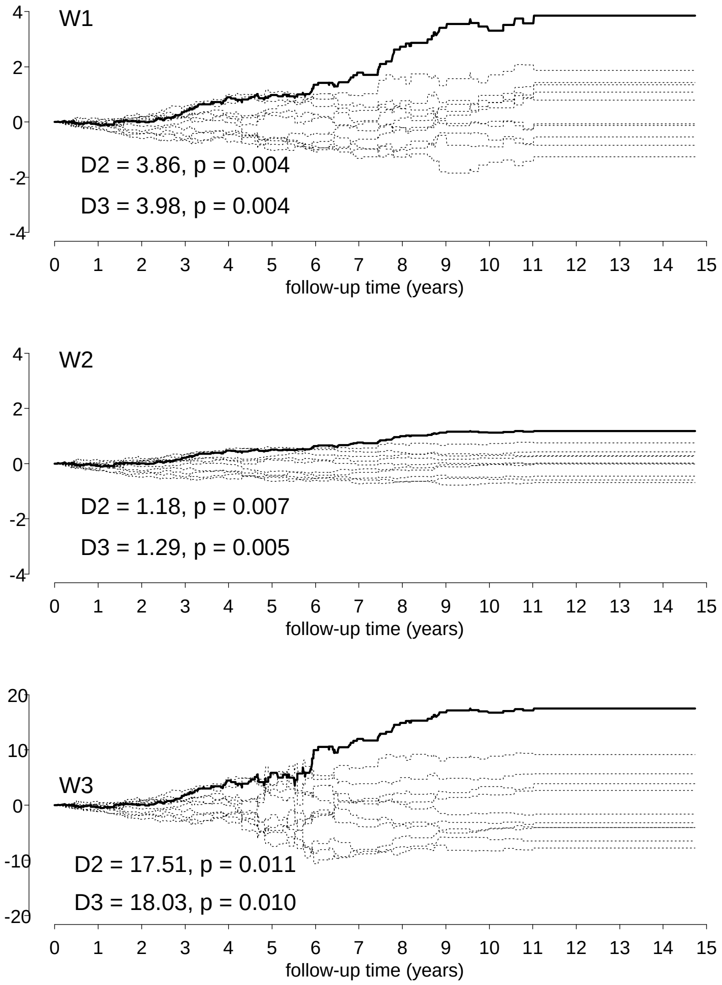

Figure 3: Senegal cohort study. The test process $L(t)$ (solid line) and ten realizations of $L^*(t)$ (dashed lines) for the weight processes $W_1(t)$, $W_2(t)$ and $W_3(t)$, with adjustment for covariates nationality (Senegalese versus else), date of cohort entry, the interaction of nationality and date of cohort entry, and the age at cohort entry at level $z_0 =$ (Senegalese, 5.16 years, 5.16 years, 30.4 years).

The confidence bands displayed in Figure 5 were computed via the first procedure described in Section 3, for the same covariate subgroups as in Figures 3 and 4. Based on the 90% simultaneous bands, we find the probablity of HIV-1 infection exceeds that of HIV-2 by up to 0.18 over the follow-up period, cf. Figure 2(a).

To check the proportional hazards assumption for each HIV type, we applied

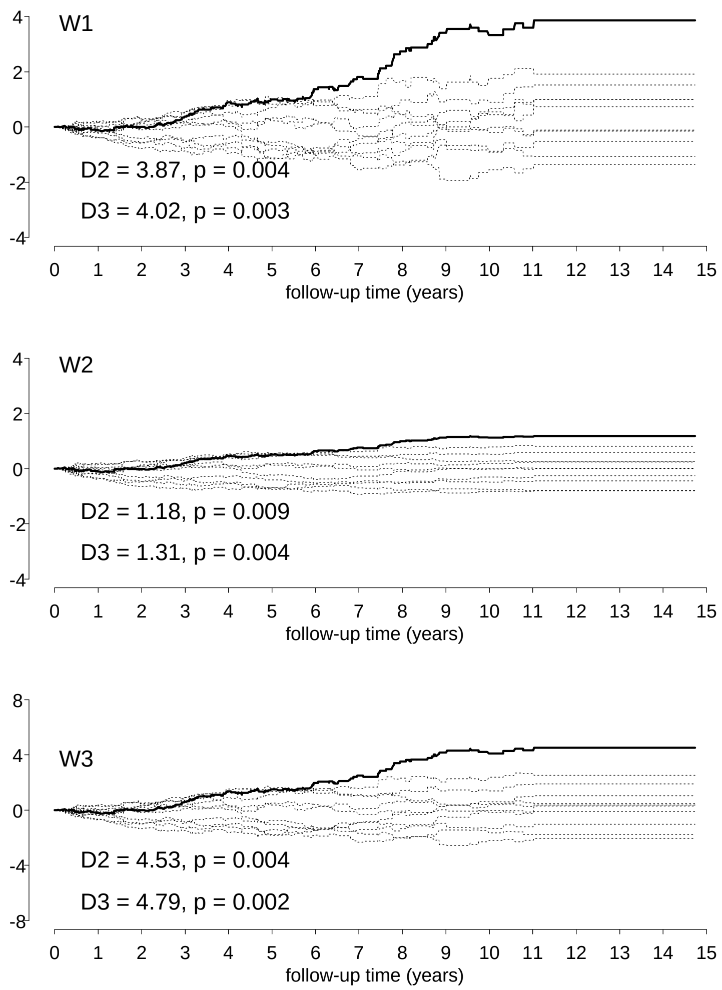

Figure 4: Senegal cohort study. The test process $L(t)$ (solid line) and ten realizations of $L^*(t)$ (dashed lines) for the weight processes $W_1(t)$, $W_2(t)$ and $W_3(t)$, with adjustment for covariates nationality (Senegalese versus else), date of cohort entry, the interaction of nationality and date of cohort entry, the age at cohort entry, and the log HIV-2/HIV-1 partner prevalence at level $z_0 =$ (Senegalese, 5.16 years, 5.16 years, 30.4 years, 0).

Grambsch and Therneau's (1994) diagnostic test based on re-scaled Schoenfeld residuals. The test did not reveal any serious departures from proportional hazards except for the relationship between the HIV-1 hazard and the age at cohort entry (p-value 0.028, Table 2). The diagnostic plot suggested that the relationship was approximately cubic, and quadratic and cubic age terms were marginally significant in a Cox model.

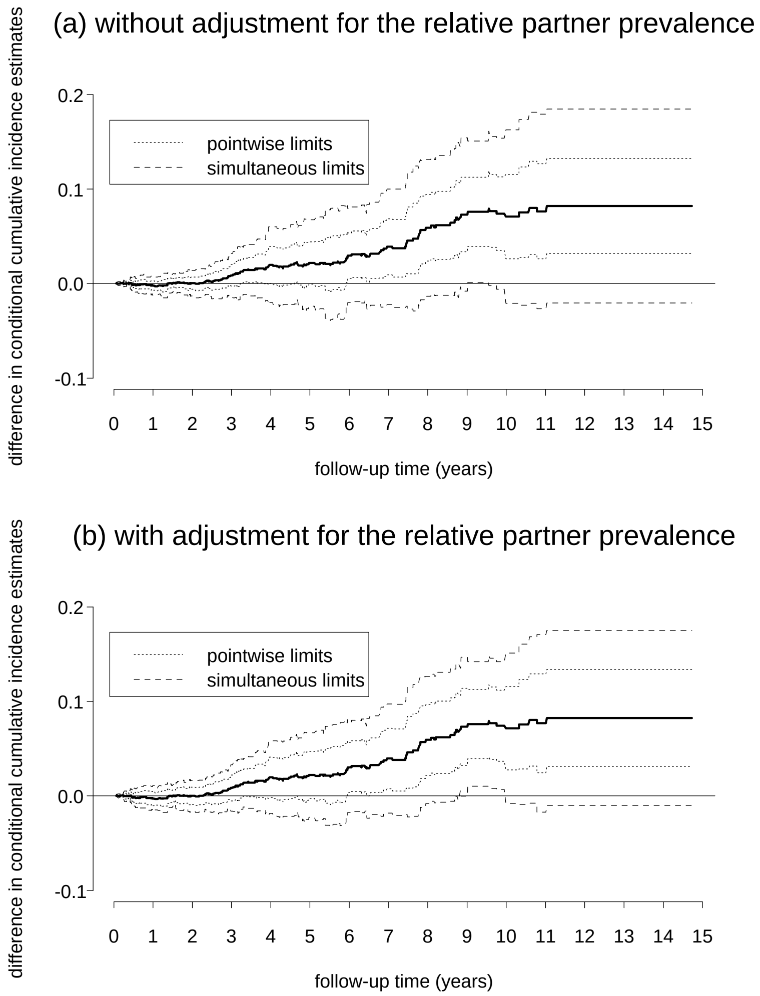

Figure 5: 95% pointwise and 90% simultaneous confidence bands for the difference in conditional HIV type-specific cumulative incidence functions. (a) conditions on the set of covariates listed in the legend to Figure 3, and (b) conditions on the set of covariates listed in the legend to Figure 4.

The tests that also adjusted for the quadratic and cubic age covariates gave comparable results as the earlier tests.

These analyses show that, under adjustment for covariates that measure exposure and susceptibility to HIV-1 and to HIV-2 infection, the risk of infection, and the male-to-female infectivity, was greater for HIV-1 than for HIV-2. Gilbert et al. (2000, manuscript submitted to Amer. J. Epidem.) present a complete analysis of the Senegal cohort data, with purpose to compare the infectivity of the virus types. The finding of higher HIV-1 infectivity has important implications for epidemiological modeling and for the design of HIV vaccines.

Acknowledgements

The authors are grateful to Professor Souleymane Mboup and the Harvard–Senegal collaboration for the provision of data from their long term prospective HIV study. They also thank the associate editor and the referees for many helpful comments. This research was partially supported by NSF grant 9971784 (McKeague), NIH grant AI24643 (Gilbert) and U.S. Army Medical Research grants 17-C-7072, 17-C-0138, 17-95-5005 and P01 A1 30795 (Kanki).

Appendix

The theorem below is established under standard regularity conditions that can be found in a form suitable for our setting in Andersen et al. (1993, Conditions VII.2.1 and VII.2.2). In addition to these conditions, we suppose there exists a bounded non-negative function $w(t)$ such that

$$\sup_{0\le t\le\tau} |n^{-1/2}W(t) - w(t)| \quad \to \quad 0 \tag{A.1}$$

in probability.

Let $s^{(0)}(\beta,t)$, $\bar{z}(\beta,t)$ and Σ_j denote the limits of $S^{(0)}(\beta,t)/n$, $\bar{Z}(\beta,t)$, and the matrix $\mathcal{I}_j(\beta_j)/n$, respectively, and define

$$\tilde{\psi}(t,\beta) = \tfrac{1}{2}\int_0^t w(u)(z_0 - \bar{z}(\beta,u))(\Lambda_1(du|z_0) + \Lambda_2(du|z_0)).$$

Theorem 1. *(a) The limiting distribution of the test process $L(t)$ under H_0 is that of the zero-mean Gaussian process*

$$\tilde{\psi}(t,\beta_1)'\xi_1 + \tilde{\psi}(t,\beta_2)'\xi_2 + B(V(t)), \tag{A.2}$$

where the three terms are independent, ξ_j is distributed $N(0,\Sigma_j^{-1})$, $B(\cdot)$ is a standard Wiener process and

$$V(t) = \tfrac{1}{2}\int_0^t w^2(u)\left(\frac{\exp(\beta_1' z_0)}{s^{(0)}(\beta_1,u)} + \frac{\exp(\beta_2' z_0)}{s^{(0)}(\beta_2,u)}\right)(\Lambda_1(du|z_0) + \Lambda_2(du|z_0)).$$

(b) The limit of the conditional distribution of $L^*(t)$ given the observed data coincides with the limiting distribution of $L(t)$ under H_0.

Proof. (a) Under H_0, we can decompose the test process as

$$\begin{aligned} L(t) \quad = \quad & \int_0^t W(u)(\exp(\hat{\beta}_2' z_0) - \exp(\beta_2' z_0))\hat{\Lambda}_{02}(du) \\ & + \int_0^t W(u)\exp(\beta_2' z_0)(\hat{\Lambda}_{02}(du) - \Lambda_{02}(du)) \\ & - \int_0^t W(u)(\exp(\hat{\beta}_1' z_0) - \exp(\beta_1' z_0))\hat{\Lambda}_{01}(du) \\ & - \int_0^t W(u)\exp(\beta_1' z_0)(\hat{\Lambda}_{01}(du) - \Lambda_{01}(du)). \end{aligned}$$

The condition (A.1) on $W(t)$ can be used to show that under H_0 the process $\psi(t,\beta)/\sqrt{n}$ converges to $\tilde{\psi}(t,\beta)$ uniformly in probability over $[0,\tau]$. Using a Taylor series expansion of $\exp(\hat{\beta}_2' z_0)$ around β_2, the sum of the first two terms in the above decomposition of $L(t)$ is seen to be asymptotically equivalent to

$$\sqrt{n}(\hat{\beta}_2 - \beta_2)'\tilde{\psi}(t,\beta_2) + \int_0^t w(u)\exp(\beta_2' z_0)\, dQ_2(u), \tag{A.3}$$

where

$$Q_j(t) = \sqrt{n}(\hat{\Lambda}_{0j}(t) - \Lambda_{0j}(t)) + \sqrt{n}(\hat{\beta}_j - \beta_j)' \int_0^t \bar{z}(\beta_j,u)\Lambda_{0j}(du).$$

The sum of the last two terms in the decomposition of $L(t)$ has a similar representation. By a theorem of Andersen et al. (1993, p. 504), the two terms in (A.3) are asymptotically independent; the limiting distribution of the first term is that of $\tilde{\psi}(t, \beta_2)'\xi_2$, and the limiting distribution of the second term is that of a zero-mean Gaussian martingale with variance function $V_2(t)$, where

$$V_j(t) = \int_0^t w^2(u) \frac{\exp(2\beta_j' z_0)}{s^{(0)}(\beta_j, u)} \lambda_{0j}(u)\, du.$$

Due to orthogonality of the basic martingales M_{1i}, M_{2i} corresponding to failure types 1 and 2, which do not occur simultaneously, the first two terms in $L(t)$ are asymptotically independent from the last two. The sum of the two independent Gaussian martingale parts is equal in distribution to a continuous Gaussian martingale, or time-changed Wiener process, with variance function $V_1(t) + V_2(t) = V(t)$ under H_0. This completes the proof of (a).

For (b), first note that $L^*(t)$ consists of a sum of four terms that are conditionally independent given the data. Also, as we have seen above, the corresponding terms in $L(t)$ are asymptotically independent. Thus it suffices to show that the conditional distribution of each term in $L^*(t)$ given the data is the same in the limit as the unconditional distribution of the corresponding term in $L(t)$. This follows directly by applying the argument of Lin, Wei and Ying (1993, Appendix 1) to each of the four terms.

Remark. The tests based on D_1 and D_3 are consistent against their respective Cox model alternatives, provided $w(t)$ is bounded away from zero on $[0, \tau]$. The same holds for D_2 provided, in addition, that $w(t)/S_X(t|z_0)$ is a decreasing function of $t \in [0, \tau]$; the weight function $W_4(t)$ furnishes such an example. This can be seen by extending the proof of Theorem 1 to obtain, under any Cox model departure from H_0, that $L(t) = \eta(t) + \sqrt{n}R(t)$, where $\eta(t)$ converges in distribution to a process of the form

(A.2) and $R(t)$ converges uniformly in probability on $[0, \tau]$ to

$$r(t) = \int_0^t w(u)(\Lambda_2(du|z_0) - \Lambda_1(du|z_0)) = \int_0^t w(u) S_X(u|z_0)^{-1}(F_2(du|z_0) - F_1(du|z_0)).$$

Also note that $r(t) \neq 0$ for some t under H_1; $r(t) > 0$ for some t under H_2; and $r(t) - r(s) > 0$ for some $s < t$ under H_3, cf. Aly et al. (1994, Appendix).

References

Abbott, R.C., Ndour-Sarr, A., Diouf, A., Ndoye, T., Ba, D., Tall, N., Mbaye, N., Kb, F., Diadhiou, F., Essex, M., Marlink, R., Mboup, S., and Kanki, P. (1994). Risk determinants for HIV infection and adverse obstetrical outcomes in pregnant women in Dakar, Senegal. *Journal of AIDS* **7**, 711-717.

Adjorlolo-Johnson, G., DeCock, K., Ekpini, E., Vetter, K.M., Sibailly, T., Brattegaard, K., Yavo, D., Doorly, R., Whitaker, J.P., and Kestens, L. (1994). Prospective comparison of mother-to-child transmission of HIV-1 and HIV-2 in Abidjan, Ivory Coast. *Journal of the American Medical Association* **272**, 462-466.

Albert, J., Naucler, A., Bottiger, B., Broliden, P., Albino, P., Ouattara, S., Bjorkegren, C., Valentin, A., Biberfeld, G., Fenyo, E. (1990). Replicative capacity of HIV-2, like HIV-1, correlates with severity of immunodeficiency. *AIDS* **4**, 291-295.

Aly, E.-E., Kochar, S. C. and McKeague, I. W. (1994). Some tests for comparing cumulative incidence functions and cause-specific hazard rates. *J. Amer. Statist. Assoc.* **89**, 994–9.

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

Andersen, P. K., Hansen, L. S. and Keiding, N. (1991). Non- and semi-parametric estimation of transition probabilities from censored observation of a non-homogeneous Markov process. *Scand. J. Statist.* **18**, 153–67.

Benichou, J. and Gail, M. H. (1990). Estimates of absolute cause-specific risk in cohort studies. *Biometrics* **46**, 813–26.

Berman, P.W. (1998). Development of bivalent rgp120 vaccines to prevent HIV type

1 infection. *AIDS Research and Human Retroviruses* **14 Suppl 3**, S277-S289.

Cheng, S. C., Fine, J. P. and Wei, L. J. (1998). Prediction of cumulative incidence function under the proportional hazards model. *Biometrics* **54**, 219–228.

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

Grambsch, P. and Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**, 515-526.

Heyward, W.L., Osmanov, S. and Esparza, J. (1992). Establishment of WHO-sponsored field sites for HIV vaccine evaluation in developing countries. *Antibiotics and Chemotherapy* **48 13**, 9-144.

Holt, J. D. (1978). Competing risk analyses with special reference to matched pair experiments. *Biometrika* **65**, 159–166.

Hu, X. S. and Tsai, W. Y. (1999). Linear rank tests for competing risks model. *Statist. Sinica* **9**, 971–983.

Kanki, P. J., Marlink, R. G., Mboup, S. and Essex, M. (1990). Epidemiology of HIV-2 in prostitutes in Senegal (abstract). *AIDS Res. and Hum. Retroviruses* **6**, 76.

Kanki, P. J., Mboup, S., Marlink, R.J., Travers K., Gueye, A., Boye, C., Sankalé, J.-L., Donnelly, C., Leisenring, W., Siby, T., Thior, I., Dia, M., Gueye, E.-H., Hsieh, C., N'Doye, I. and Essex, M. (1992). Prevalence and risk determinants of Human Immunodeficiency Virus type 2 (HIV-2) and Human Immunodeficiency Virus type 1 (HIV-1) in West African female prostitutes. *Amer. J. Epidem.* **136**, 895–907.

Kanki, P. J. (1999). Human immunodeficiency virus type 2 (HIV-2). *AIDS Reviews* **1**, 101–8.

Lam, K. F. (1998). A class of tests for the equality of k cause-specific hazard rates in a competing risks model. *Biometrika* **85**, 179–188.

Larson, M. G. (1984). Covariate analysis of competing-risks data with log-linear models *Biometrics* **40**, 459–69.

Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80**, 557–72.

Louwagie, J., McCutchan, F.E., Peeters, M., Brennan, T.P., Sanders, B.E., Eddy, G.A., van der Groen, G., Fransen, K., Gershy-Damet, G.-M., Deleys, R., and Burke, D.S. (1993). Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS* **7**, 769-780.

Lunn, M. and McNeil, D. (1995). Applying Cox regression to competing risks. *Biometrics* **51**, 524–32.

Luo, X. and Turnbull, B. (1999). Comparing two treatments with multiple competing risks endpoints. *Statist. Sinica* **9**, 985–997.

Marlink, R., Kanki, P., Thior, I., Travers, K., Eisen, G., Siby, T., Traore, I., C.C., H., MC., D., Gueye, E.H., Hellinger, J., Gueye, N.A., Sankal, J.L., Ndoye, I., Mboup, S., and Essex, M. (1994). Reduced rate of disease development with HIV-2 compared to HIV-1. *Science* **265**, 1587-1590.

McKeague, I. W. and Utikal, K. (1990). Inference for a nonlinear counting process regression model. *Ann. Statist.* **18**, 1172–87.

Moore, J.P. and Anderson, R. (1994). The WHO and why of HIV vaccine trials. *Nature* **372**, 313-314.

Popper, S.J., Dieng-Sarr, A., Travers, K.U., Guye-Ndiaye, A., Mboup, S., Essex, M., and Kanki, P.J. (1999). Lower HIV-2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. *Journal of Infectious Diseases* **180**, 1116-1121.

Poulsen, A.G., Kvinesdal, B.B., Aaby, P., Lisse, I.M., Gottschau, A., K., M., Dias, F., and Lauritzen, E. (1992). Lack of evidence of vertical transmission of human immunodeficiency virus type 2 in a sample of the general population in Bissau. *Journal of AIDS* **5**, 25-30.

Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T. and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics* **34**, 541–54.

Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* **11**, 453–66.

Shen, Y. and Cheng, S. C. (1999). Confidence bands for cumulative incidence curves under the additive risk model. *Biometrics* **55**, 1093–1100.

Sun, Y. (2000). Generalized nonparametric test procedures for comparing multiple cause-specific hazard rates. *J. Nonparametr. Statist.*, in press.

Sun, Y. and Tiwari, R. C. (1995). Comparing cause-specific hazard rates of a competing risks model with censored data. *IMS Lecture Notes-Monograph Series* **27**, 255–70. *Analysis of Censored Data*, edited by H. L. Koul and J. V. Deshpandé.