Published in final edited form as:

Biometrics. 2013 June ; 69(2): 328–337. doi:10.1111/biom.12016.

Mark-specific Hazard Ratio Model with Multivariate Continuous Marks: An Application to Vaccine Efficacy

M. Juraska¹ and P. B. Gilbert^{1,2}

M. Juraska: mjuraska@u.washington.edu

¹Department of Biostatistics, University of Washington, Seattle, Washington, 98195, U.S.A.

²Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109, U.S.A.

Summary

In randomized placebo-controlled preventive HIV vaccine efficacy trials, an objective is to evaluate the relationship between vaccine efficacy to prevent infection and genetic distances of the exposing HIV strains to the multiple HIV sequences included in the vaccine construct, where the set of genetic distances is considered as the continuous multivariate ‘mark’ observed in infected subjects only. This research develops a multivariate mark-specific hazard ratio model in the competing risks failure time analysis framework for the assessment of mark-specific vaccine efficacy. It allows improved efficiency of estimation by employing the semiparametric method of maximum profile likelihood estimation in the vaccine-to-placebo mark density ratio model. The model also enables the use of a more efficient estimation method for the overall log hazard ratio in the Cox model. Additionally, we propose testing procedures to evaluate two relevant hypotheses concerning mark-specific vaccine efficacy. The asymptotic properties and finite-sample performance of the inferential procedures are investigated. Finally, we apply the proposed methods to data collected in the Thai RV144 HIV vaccine efficacy trial.

Keywords

Competing risks; Competing risks; Cox regression model; Density ratio model; Failure time data; Genetic data; Mark variable; Mark-specific vaccine efficacy

1. Introduction

In studies with competing risks data, commonly a time-to-event is evaluated, and data from those who experience the event – termed “mark” data – are collected. The analysis of such data often involves the comparison of hazard rates of mark-specific failure between two treatment groups. We propose estimation and testing procedures for the continuous mark specific hazard ratio model that improves efficiency compared to alternative approaches and accommodates a multivariate mark.

The extensive genetic diversity of HIV poses a central challenge to development of an efficacious preventive HIV vaccine. Efficacy trials of such vaccines randomize HIV-uninfected volunteers to receive vaccine or placebo, and the primary objective assesses vaccine efficacy (VE), typically defined as one minus the vaccine-to-placebo hazard ratio of

Correspondence to: M. Juraska, mjuraska@u.washington.edu.

Supplementary Materials

Web Appendices referenced in Sections 2–5 are available with this paper at the Biometrics website on Wiley Online Library.

HIV acquisition (Halloran, Struchiner, and Longini, 1997). A secondary objective assesses VE to prevent acquisition with particular genotypes of HIV, where we expect that VE will be greatest against HIVs genetically similar to the HIV strains included in the vaccine, and will be lower against HIVs more distant from the vaccine strains (Gilbert, McKeague, and Sun, 2008) (henceforth GMS). It is important to maximize efficiency for assessing genotype-specific vaccine efficacy, to avoid discontinuation of product development owing to a false negative error.

Maximizing biological relevance and statistical power for assessing genotype-specific vaccine efficacy requires careful definition of the mark variable measuring genetic divergence of an exposing HIV to the vaccine strains. In general, viral vaccines stimulate protective immune responses to the exact protein sequences of the viruses inside the tested vaccine, but fail to stimulate protective immune responses to viruses with certain differences in their genetic sequence compared to the vaccine strains. Therefore, the genetic distance variable is defined in a way to reflect a biological model of how well the vaccine can induce immune responses that ‘cross-react’ with viruses different from the vaccine strains; under the model for cross-reactivity the vaccine is more likely to stimulate a protective immune response to HIVs with smaller distances to the vaccine strains. Thus, the statistical assessment of if and how VE varies with viral genetic distance empirically tests the fidelity of the distance for measuring cross-reactive protection. Because of challenges involved in modeling cross-reactivity, multiple immunologically meaningful genetic distances have been considered, based on (i) different conceptual definitions, (ii) different HIV sequence regions, (iii) different methods to predict which HIV amino acids stimulate antibody responses, and (iv) different reference sequences inside the vaccine.

From 2003 to 2009, a randomized Phase III preventive vaccine efficacy trial (Rerks-Ngarm et al., 2009), the RV144 ‘Thai trial’, was conducted to evaluate the combination of the recombinant canarypox vector vaccine (ALVAC-HIV [vCP1521]), expressing the HIV genes *gag*, *pro*, and the gp120 portion of *env*, and the recombinant gp120 subunit vaccine (AIDSVAX B/E). HIV-uninfected volunteers in Thailand, primarily at heterosexual risk for HIV infection, were randomly allocated to receive the vaccine ($n_1 = 8197$) or placebo ($n_0 = 8198$), and followed for 42 months for the primary endpoint of HIV infection. HIV was diagnosed in 125 subjects (51 vaccine, 74 placebo), and, based on the Cox model, overall vaccine efficacy was estimated to be 31.2% (95% CI, 1.1 to 52.1; $p=0.04$). This result generated great interest to understand how vaccine protection may have depended on the genetic divergence of infecting HIVs from the HIV vaccine strains. Full genome HIV sequences were measured from 121 of the 125 infected subjects, with between 2 and 13 sequences sampled per subject. In Section 5, we describe the immunologically relevant genetic distance marks that we analyze.

GMS developed a nonparametric estimator for the univariate mark-specific hazard ratio and procedures for testing whether the mark-specific hazard ratio is (i) unity or (ii) independent of the mark. Problems (i) and (ii) are equivalent to the null hypothesis of zero vaccine efficacy against any exposing virus and the null hypothesis that vaccine efficacy does not depend on the viral divergence, respectively; both null hypotheses are also considered in Section 3. Sun, Gilbert, and McKeague (2009) developed the mark-specific proportional hazards model, which, in the case of a univariate mark, provides an alternative approach to inference about the mark-specific vaccine efficacy defined in (2).

We propose a more efficient method of estimation and hypothesis testing for the mark-specific hazard ratio in the framework of competing risks failure time analysis that accommodates multivariate marks. Our approach makes use of the maximum profile likelihood estimation method in the semiparametric density ratio/biased sampling model

developed by Qin (1998). In parallel with Qin (1998), a similar method of maximum profile partial likelihood estimation was derived by Gilbert, Lele, and Vardi (1999) and Gilbert (2000) for semiparametric biased sampling models with K possibly biased samples.

1.1 Notation

Let T denote the time to failure and $\mathbf{V} \in \mathbb{R}^s$, a continuous, possibly multivariate, mark variable. Without loss of generality, the support of each component of $\mathbf{V}$ is taken to be $[0, 1]$. Let C be the time to censoring. The observed right-censored failure time is $X = \min(T, C)$ with the failure indicator $\delta = I(T \leq C)$. The mark $\mathbf{V}$ is always observed if $\delta = 1$; otherwise it is unobservable. Let Z denote the indicator of assignment to the treatment group (in vaccine trials, $Z = 1$ indicates vaccine and $Z = 0$ indicates placebo). Let $(Z_i, X_i, \delta_i, \mathbf{V}_i)$, $i = 1, \dots, n$, be i.i.d. replicates of $(Z, X, \delta, \mathbf{V})$. The observed data consist of the observations $(Z_i, X_i, \mathbf{V}_i)$ for individuals with $\delta_i = 1$ and the observations (Z_i, X_i) for those with $\delta_i = 0$.

1.2 Assumptions

We assume C is conditionally independent of both T and $\mathbf{V}$ given Z , that is, $C \perp T|Z$ and $C \perp \mathbf{V}|Z$, and assume $T \perp \mathbf{V}|Z$. These assumptions ensure identifiability of the Euclidean parameters in the density ratio model introduced in Section 2. In the HIV vaccine trial setting, the assumption $T \perp \mathbf{V}|Z$ entails that, for example, vaccine recipients with infection time $T = 6$ months have the same distribution of the mark $\mathbf{V}$ as vaccine recipients with infection time $T = 2.5$ years, which may approximately hold given a limited shift in the HIV sequence distribution over the period of 2 years. An analogous statement comparing vaccine recipients with infection times, say, 50 years apart would be clearly incorrect due to the genetic shift of HIV. Hence, the fact that HIV vaccine efficacy trials only last between 3–5 years is crucial for the assumption to be (approximately) met. In Section 3.1, we develop a Kolmogorov–Smirnov-type test for assessing validity of the assumption $T \perp \mathbf{V}|Z$, and, in Section 4.3, we demonstrate some robustness properties of the proposed inferential procedures to violation of the assumption $T \perp \mathbf{V}|Z$.

1.3 Estimands

We define the conditional multivariate mark-specific hazard function as

$$\lambda(t, \mathbf{v}|Z=z) = \lim_{h_1, h_{21}, \dots, h_{2s} \searrow 0} \frac{P(T \in [t, t+h_1), \mathbf{V} \in \prod_{i=1}^s [v_i, v_i+h_{2i}) | T \geq t, Z=z)}{h_1 h_{21} \cdots h_{2s}}, \quad (1)$$

which is a natural generalization of the cause-specific hazard function in the presence of finitely many causes of failure (Prentice et al., 1978). GMS defined the mark-specific vaccine efficacy as

$$VE(t, \mathbf{v}) = 1 - \lambda(t, \mathbf{v}|Z=1)/\lambda(t, \mathbf{v}|Z=0). \quad (2)$$

As described in GMS, $VE(t, \mathbf{v})$ has an approximate interpretation as the multiplicative vaccine effect to reduce susceptibility to HIV infection given exposure to strain $\mathbf{v}$ at time t .

The remainder of this article is organized as follows. In Section 2, we introduce a semiparametric model for the mark-specific hazard ratio in (2), discuss identifiability of the Euclidean model parameters, and state the asymptotic properties of the proposed estimator for the mark-specific hazard ratio. Hypothesis testing procedures are introduced in Section 3. We summarize our findings from the simulation experiment in Section 4 and apply the proposed inferential procedures to data from the Thai trial in Section 5. Proofs of main results are available in the Web Appendix.

2. Estimation method

The mark-specific hazard function factors as

$$\lambda(t, \mathbf{v}|Z=z) = f(\mathbf{v}|T=t, Z=z) \lambda(t|Z=z), \quad (3)$$

where $f(\cdot|T=t, Z=z)$ is the conditional density function of $\mathbf{V}$ given $T=t$ and $Z=z$, and $\lambda(\cdot|Z=z)$ is the ordinary marginal hazard function not involving the mark. Consequently, the mark-specific hazard ratio can be written as

$$\frac{\lambda(t, \mathbf{v}|Z=1)}{\lambda(t, \mathbf{v}|Z=0)} = \frac{f(\mathbf{v}|T=t, Z=1)}{f(\mathbf{v}|T=t, Z=0)} \frac{\lambda(t|Z=1)}{\lambda(t|Z=0)}. \quad (4)$$

The factorization in (4) is advantageous because the two ratios can be estimated separately. For the mark density ratio, we consider the semiparametric density ratio model (Qin, 1998)

$$f(\mathbf{v}|T=t, Z=1)/f(\mathbf{v}|T=t, Z=0) = g(\mathbf{v}, \boldsymbol{\phi}) \quad (5)$$

where $\boldsymbol{\phi} \in \mathbb{R}^d$ is the parameter of interest and $g(\mathbf{v}, \boldsymbol{\phi})$ is a known time-independent weight function, continuously differentiable in $\boldsymbol{\phi}$. The nuisance parameter distribution $f(\mathbf{v}|T=t, Z=0)$ is time-independent following the assumption $T \perp \mathbf{V}|Z$, and is treated nonparametrically. Due to the equality of the conditional density functions

$$f(\mathbf{v}|T=t, Z=z) = f(\mathbf{v}|T=t, Z=z, \delta=1), \quad (6)$$

which occurs because of the assumption $(T, C) \perp \mathbf{V}|Z$, the parameter $\boldsymbol{\phi}$ in (5) is estimable using cases with an observed failure only. Using Bayes rule, model (5) implies that

$$\frac{g(\mathbf{v}_1, \boldsymbol{\phi})}{g(\mathbf{v}_2, \boldsymbol{\phi})} = \frac{f(\mathbf{v}_1|Z=1)}{f(\mathbf{v}_1|Z=0)} \left\{ \frac{f(\mathbf{v}_2|Z=1)}{f(\mathbf{v}_2|Z=0)} \right\}^{-1} = \frac{P(Z=1|\mathbf{v}_1)}{P(Z=0|\mathbf{v}_1)} \left\{ \frac{P(Z=1|\mathbf{v}_2)}{P(Z=0|\mathbf{v}_2)} \right\}^{-1}$$

for marks $\mathbf{v}_1, \mathbf{v}_2 \in [0, 1]^s$. Thus $g(\mathbf{v}_1, \boldsymbol{\phi})/g(\mathbf{v}_2, \boldsymbol{\phi})$ is the odds of being assigned vaccine for an individual infected with strain $\mathbf{v}_1$, relative to that for an individual infected with strain $\mathbf{v}_2$.

A common choice of the weight function $g(\mathbf{v}, \boldsymbol{\phi})$ is $\exp\{\alpha + g(\tilde{\mathbf{v}}, \boldsymbol{\beta})\}$ where $\boldsymbol{\phi} = (\alpha, \boldsymbol{\beta}^T)^T$ and $g(\tilde{\mathbf{v}}, \boldsymbol{\beta})$ is a polynomial function. In this case, the following proposition characterizes the necessary and sufficient condition for identifiability of $\boldsymbol{\phi}$ (see Web Appendix A for a proof).

Proposition 1: In the semiparametric density ratio model

$$f(\mathbf{v}|T=t, Z=1)/f(\mathbf{v}|T=t, Z=0) = e^{\alpha(t) + \tilde{g}(\mathbf{v}, \boldsymbol{\beta})},$$

T and $\mathbf{V}$ are conditionally independent given Z if and only if $\alpha(t)$ is constant.

Following Qin (1998), the profile score functions for the parameter of interest $\boldsymbol{\phi}$ and the Lagrange multiplier $\lambda \in [0, 1]$ (that arises by profiling out the nuisance parameter) are

$$U_{\boldsymbol{\phi}}(\boldsymbol{\phi}, \lambda) = \sum_{i=1}^n \delta_i \left(-\frac{\lambda \dot{g}(\mathbf{V}_i, \boldsymbol{\phi})}{1 + \lambda(g(\mathbf{V}_i, \boldsymbol{\phi}) - 1)} + Z_i \frac{\dot{g}(\mathbf{V}_i, \boldsymbol{\phi})}{g(\mathbf{V}_i, \boldsymbol{\phi})} \right) \quad U_{\lambda}(\boldsymbol{\phi}, \lambda) = -\sum_{i=1}^n \delta_i \frac{g(\mathbf{V}_i, \boldsymbol{\phi}) - 1}{1 + \lambda(g(\mathbf{V}_i, \boldsymbol{\phi}) - 1)}$$

where $\hat{g}(\mathbf{v}, \boldsymbol{\phi}) = g(\mathbf{v}, \boldsymbol{\phi}) / \boldsymbol{\phi}$. The maximum profile likelihood estimator $(\hat{\boldsymbol{\phi}}^T, \hat{\boldsymbol{\lambda}}^T)^T$ for $(\boldsymbol{\phi}^T, \boldsymbol{\lambda}^T)^T$ is defined as the solution to the system $\mathbf{U}_{\boldsymbol{\phi}}(\boldsymbol{\phi}, \boldsymbol{\lambda}) = \mathbf{0}$ and $U_{\boldsymbol{\lambda}}(\boldsymbol{\phi}, \boldsymbol{\lambda}) = 0$.

For the marginal hazard ratio in (4), we propose to use the Cox regression model

$$\lambda(t|Z=1)/\lambda(t|Z=0) = e^{\gamma}. \quad (7)$$

The parameter γ may be estimated with the standard maximum partial likelihood estimator (MPLE), or, alternatively, with the more efficient method of Lu and Tsiatis (2008) that leverages auxiliary data predictive of the failure time (we implemented the Lu and Tsiatis method in the R `speff2trial` package). If additional, possibly time-dependent, covariates $\mathbf{Z}^*(t) = (Z_1^*(t), \dots, Z_p^*(t))^T$ are measured, then the Cox model (7) can be extended to $\lambda(t|Z=1, \mathbf{Z}^*(t) = \mathbf{z}^*(t)) = \lambda(t|Z=0, \mathbf{Z}^*(t) = \mathbf{0})e^{\gamma + \gamma^{*T} \mathbf{z}^*(t)}$, and any estimation method for $(\gamma, \gamma^{*T})^T$ available for the Cox model can be employed. For example, using $\mathbf{Z}^*(t) = (t, Zt)^T$ specifies $\lambda(t|Z=1, t)/\lambda(t|Z=0, t) = e^{\gamma + \gamma_2^* t}$, allowing the overall treatment effect to change over time.

2.1 Asymptotic properties of the proposed estimator

Qin (1998) showed that $(\hat{\boldsymbol{\phi}}^T, \hat{\boldsymbol{\lambda}}^T)^T$ is a consistent and asymptotically normal (CAN) estimator if the proportions of failures in the placebo and treatment group converge to positive constants. The MPLE $\hat{\gamma}$ and the Lu and Tsiatis (2008) estimator have also been shown to be CAN. The present goal is to characterize the joint asymptotic distribution of $(\hat{\boldsymbol{\phi}}^T, \hat{\boldsymbol{\lambda}}^T, \hat{\gamma})^T$.

We restrict attention to the density ratio model (5) with weight function $g(\mathbf{v}, \boldsymbol{\phi}) = e^{\alpha + g(\tilde{\mathbf{v}}, \boldsymbol{\beta})}$, where $\boldsymbol{\phi} = (\alpha, \boldsymbol{\beta}^T)^T$ and $g(\tilde{\mathbf{v}}, \boldsymbol{\beta})$ is a polynomial function in $\tilde{\mathbf{v}}$, and the marginal Cox model adjusted for the covariates $(Z, \mathbf{Z}^{*T}(t))^T$. Although Theorem 1 applies in this setting, for expositional simplicity the results are presented for $g(\tilde{\mathbf{v}}, \boldsymbol{\beta})$ a quadratic form in $\tilde{\mathbf{v}}$ and the marginal Cox model (7) with sole covariate the treatment group indicator Z . In this special case, the mark-specific hazard ratio model in (4) takes the form $e^{\alpha + g(\tilde{\mathbf{v}}, \boldsymbol{\beta}) + \gamma}$. Let $\boldsymbol{\theta} = (\boldsymbol{\phi}^T, \boldsymbol{\lambda}, \gamma)^T$ denote the vector of Euclidean parameters, and let the random map $\boldsymbol{\Psi}_n(\boldsymbol{\theta})$ denote the set of all estimating functions for $\boldsymbol{\theta}$. The estimator $\hat{\boldsymbol{\theta}}_n$ for $\boldsymbol{\theta}$ is obtained as the solution to $\boldsymbol{\Psi}_n(\boldsymbol{\theta}) = \mathbf{0}$.

Next, define the random processes

$$\eta_{n,\gamma}(t) = \bar{Z}_n(t; \gamma) = \frac{\frac{1}{n} \sum_{k=1}^n Z_k I(X_k \geq t) e^{\gamma Z_k}}{\frac{1}{n} \sum_{k=1}^n I(X_k \geq t) e^{\gamma Z_k}} = \frac{\xi_{n,\gamma}(t)}{\zeta_{n,\gamma}(t)}, t \in [0, \tau], \quad \eta_{0,\gamma}(t) = \frac{E[Z I(X \geq t) e^{\gamma Z}]}{E[I(X \geq t) e^{\gamma Z}]} = \frac{\xi_{0,\gamma}(t)}{\zeta_{0,\gamma}(t)}, t \in [0, \tau],$$

where τ is selected such that $P(T > \tau) > 0$. We will use the notation $\eta_{\cdot,\gamma}$ or $\eta_{\cdot,\boldsymbol{\theta}}$ to denote the processes indexed by the component γ or by the full parameter vector $\boldsymbol{\theta}$. For $(x, \Delta, \mathbf{v}, z) \in [0, \tau] \times \{0, 1\} \times [0, 1]^s \times \{0, 1\}$, define the map $\boldsymbol{\varphi} = (\varphi_1^T, \varphi_2, \varphi_3)^T$ as

$$\boldsymbol{\varphi}(\boldsymbol{\theta}, \eta_{m,\boldsymbol{\theta}}) = \Delta \left\{ \left(-\frac{\lambda \dot{g}(\mathbf{v}, \boldsymbol{\phi})}{1 + \lambda(g(\mathbf{v}, \boldsymbol{\phi}) - 1)} + z \frac{\dot{g}(\mathbf{v}, \boldsymbol{\phi})}{g(\mathbf{v}, \boldsymbol{\phi})} \right)^T, -\frac{g(\mathbf{v}, \boldsymbol{\phi}) - 1}{1 + \lambda(g(\mathbf{v}, \boldsymbol{\phi}) - 1)}, z - \eta_{m,\boldsymbol{\theta}}(x) \right\}^T. \quad (8)$$

The dependence of $\boldsymbol{\phi}$ on $(x, \Delta, \mathbf{v}, z)$ is suppressed in the notation. Let $\boldsymbol{\Psi}$ be the limit of $\boldsymbol{\Psi}_n$ as $n \rightarrow \infty$ (and also $m \rightarrow \infty$) and let $\boldsymbol{\Psi}(\boldsymbol{\theta}_0) = \mathbf{0}$. Let $\mathbb{G}_P$ denote the zero-mean P -Brownian bridge process and define the class of functions $f_{\boldsymbol{\theta},r}(x, z) = z^r I(x - t) e^{\gamma z}$ for $r = 0, 1$. Theorem 1 characterizes the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ (see Web Appendix B for a proof).

Theorem 1: Let $\hat{\theta}_n$ be the solution to $\Psi_n(\theta) = \mathbf{0}$ and let $\Psi(\theta_0) = \mathbf{0}$. Then

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} -\dot{\Psi}_{\theta_0}^{-1} \mathbf{Z}$$

where

$$\mathbf{Z} = \mathbb{G}_P \{ \varphi_1^T(\theta_0, \eta_{0,\theta_0}), \varphi_2(\theta_0, \eta_{0,\theta_0}), \varphi_3(\theta_0, \eta_{0,\theta_0}) + p_\delta l_{\theta_0} \}^T$$

with the map $\varphi = (\varphi_1^T, \varphi_2, \varphi_3)^T$ defined in (8), $p_\delta = P(\delta = 1)$, and

$$l_{\theta_0} = l_{\theta_0}(x, z) = \int \zeta_{0,\theta_0}^{-1}(t) (f_{\theta_0,t,1}(x, z) - \eta_{0,\theta_0}(t) f_{\theta_0,t,0}(x, z)) dF(t | \delta = 1).$$

Here $F(t | \delta = 1)$ is the conditional cdf of X given $\delta = 1$, and $\dot{\Psi}_{\theta_0}$ is the continuously invertible derivative of the map $\theta \mapsto \Psi(\theta)$ at θ_0 and has matrix form

$$\dot{\Psi}_{\theta_0} = \begin{pmatrix} \dot{\Psi}_{11} & \dot{\Psi}_{12} & 0 \\ \dot{\Psi}_{12}^T & \dot{\Psi}_{22} & 0 \\ 0^T & 0 & \dot{\Psi}_{33} \end{pmatrix}$$

with entries

$$\begin{aligned} \dot{\Psi}_{11} &= \int \Delta \left\{ - \left(\frac{\rho_1}{1 + \rho_1(g(v, \phi_0) - 1)} - \frac{z}{g(v, \phi_0)} \right) \ddot{g}(v, \phi_0) + \left(\frac{\rho_1^2}{[1 + \rho_1(g(v, \phi_0) - 1)]^2} - \frac{z^2}{g^2(v, \phi_0)} \right) \dot{g}(v, \phi_0) \dot{g}^T(v, \phi_0) \right\} dP(x, \Delta, \Delta v, z) \\ \dot{\Psi}_{12} &= \int \frac{-\Delta \dot{g}(v, \phi_0)}{[1 + \rho_1(g(v, \phi_0) - 1)]^2} dP(x, \Delta, \Delta v, z) \\ \dot{\Psi}_{22} &= \int \frac{\Delta(g(v, \phi_0) - 1)^2}{[1 + \rho_1(g(v, \phi_0) - 1)]^2} dP(x, \Delta, \Delta v, z) \\ \dot{\Psi}_{33} &= \int -\Delta \eta_{0,\theta_0}(x) (1 - \eta_{0,\theta_0}(x)) dP(x, \Delta, \Delta v, z) \end{aligned}$$

where $\ddot{g}(v, \phi) = \frac{\partial^2 g(v, \phi)}{\partial \phi \partial \phi^T}$.

Denote the column vector of functions

$\tilde{\varphi}(\theta, \eta_{0,\theta}) = (\varphi_1^T(\theta, \eta_{0,\theta}), \varphi_2(\theta, \eta_{0,\theta}), \varphi_3(\theta, \eta_{0,\theta}) + p_\delta l_\theta)^T$. The following corollary describes the asymptotic variance of $\sqrt{n}(\hat{\theta}_n - \theta_0)$.

Corollary 1: The asymptotic random vector $\dot{\Psi}_{\theta_0}^{-1} \mathbf{Z}$ in Theorem 1 is normally distributed with zero mean and covariance matrix $\Gamma = \dot{\Psi}_{\theta_0}^{-1} \Omega \dot{\Psi}_{\theta_0}^{-1}$ where

$$\Omega = P \tilde{\varphi}(\theta_0, \eta_{0,\theta_0}) \tilde{\varphi}^T(\theta_0, \eta_{0,\theta_0}) - P \tilde{\varphi}(\theta_0, \eta_{0,\theta_0}) P \tilde{\varphi}^T(\theta_0, \eta_{0,\theta_0}).$$

Let $\hat{\Gamma}_n$ denote the empirical estimator for Γ obtained by replacing P by the empirical probability measure $\mathbb{P}_n$, θ_0 by $\hat{\theta}_n$, and η_{0,θ_0} by $\eta_{n,\hat{\theta}_n}$ in the definition of Γ . Corollary 1 leads to the construction of Wald confidence intervals (pointwise in v) for the components of θ , and, subsequently, for the parameter $V E(v) = 1 - e^{\alpha + g(\hat{v}, \beta) + \gamma}$.

3. Hypothesis testing

For illustrating the testing procedures, we consider the simplified weight function $g(\mathbf{v}, \boldsymbol{\varphi}) = e^{\alpha + \boldsymbol{\beta}^T \mathbf{v}}$, $\boldsymbol{\varphi} = (\alpha, \boldsymbol{\beta}^T)^T$, which leads to the mark-specific vaccine efficacy function $VE(\mathbf{v}) = 1 - e^{\alpha + \boldsymbol{\beta}^T \mathbf{v} + \gamma}$. We develop likelihood ratio and Wald tests to evaluate the null hypothesis

$$H_0^0: VE(\mathbf{v}) = 0 \text{ for all } \mathbf{v} \in [0, 1]^s, \quad (9)$$

which states that the vaccine provides no protection against infection with any HIV strain. If H_0^0 is rejected, the question arises as to whether vaccine efficacy depends on the viral divergence; thus, we develop the likelihood ratio and Wald tests for the null hypothesis

$$H_0: VE(\mathbf{v}) \equiv VE \text{ for all } \mathbf{v} \in [0, 1]^s. \quad (10)$$

Under models (5) and (7), H_0^0 is equivalent to $\{\tilde{H}_0^0: \boldsymbol{\beta} = \mathbf{0} \text{ and } \gamma = 0\}$. The likelihood ratio test of $\tilde{H}_0^0$ against $\{\tilde{H}_1^0: \boldsymbol{\beta} \neq \mathbf{0} \text{ or } \gamma \neq 0\}$ uses Simes (1986) procedure, in which the profile likelihood ratio test statistic for $\boldsymbol{\beta}$ in model (5) and the partial likelihood ratio test statistic for γ in model (7) are evaluated separately. P-values $p_{\boldsymbol{\beta}}$ and p_{γ} are obtained based on the fact that the likelihood ratio statistics are asymptotically χ_s^2 and χ_1^2 under $\tilde{H}_0^0$, respectively. Simes' procedure rejects H_0^0 if either $\max(p_{\boldsymbol{\beta}}, p_{\gamma}) \leq \alpha$ or $\min(p_{\boldsymbol{\beta}}, p_{\gamma}) \leq \alpha/2$ where α is the nominal familywise level of significance. The Wald test of H_0^0 versus H_1^0 is based on the statistic $n(\hat{\boldsymbol{\beta}}_n^T, \hat{\gamma}_n) \hat{\boldsymbol{\Gamma}}_{n, \boldsymbol{\beta} \gamma}^{-1} (\hat{\boldsymbol{\beta}}_n^T, \hat{\gamma}_n)^T$ where $\hat{\boldsymbol{\Gamma}}_{n, \boldsymbol{\beta} \gamma}$ is the submatrix of $\hat{\boldsymbol{\Gamma}}_n$ pertaining to the components $\boldsymbol{\beta}$ and γ . Under H_0^0 , the Wald test statistic is asymptotically χ_{s+1}^2 . We additionally propose a weighted one-sided Wald-type test of H_0^0 based on the Z-statistic

$$Z = \left(\sum_{i=1}^s \frac{\hat{\beta}_{n,i}}{\widehat{\text{var}} \hat{\beta}_{n,i}} - \frac{\hat{\gamma}_n}{\widehat{\text{var}} \hat{\gamma}_n} \right) / \sqrt{\widehat{\text{var}} \left(\sum_{i=1}^s \frac{\hat{\beta}_{n,i}}{\widehat{\text{var}} \hat{\beta}_{n,i}} - \frac{\hat{\gamma}_n}{\widehat{\text{var}} \hat{\gamma}_n} \right)}, \quad (11)$$

which is designed to increase power to detect alternative hypotheses where both the

marginal vaccine efficacy $VE = 1 - \frac{\lambda(t|Z=1)}{\lambda(t|Z=0)} = 1 - e^{\gamma} > 0$ and $VE(\mathbf{v})$ declines with all of the components of $\mathbf{V}$. The null hypothesis H_0 is equivalent to $\{H_0: \boldsymbol{\beta} = \mathbf{0}\}$, and thus the density ratio model (5) alone serves to construct the likelihood ratio and Wald test statistics for H_0 .

3.1 Diagnostic test for $T \perp \mathbf{V} | Z$

By Proposition 1, the $T \perp \mathbf{V} | Z$ assumption is necessary for parameter identifiability in the time-independent density ratio model. For a given treatment group $Z = z$, we propose a diagnostic test of the null hypothesis $\{K_0: T \perp \mathbf{V} | Z = z\}$ based on the statistic

$$D_z = \sup_{t, \mathbf{v}} |\hat{F}_{T\mathbf{V}}(t, \mathbf{v} | z) - \hat{F}_T(t | z) \hat{F}_{\mathbf{V}}(\mathbf{v} | z)|, \quad (12)$$

where $\hat{F}_{T\mathbf{V}}(t, \mathbf{v} | z)$ is the nonparametric maximum likelihood estimator of the joint cdf of $(T, \mathbf{V})$ for group z developed by Huang and Louis (1998), $\hat{F}_T(t | z)$ is one minus the Kaplan-Meier estimator of the conditional survival function of T , and $\hat{F}_{\mathbf{V}}(\mathbf{v} | z)$ is the empirical distribution function of the observed values of $\mathbf{V}$. The critical values for the distribution of (12) under K_0 can be calculated using a bootstrap algorithm as follows:

1. Draw an independent sample (X_i^*, δ_i^*) from the original time-on-study data (X_i, δ_i) for group $Z = z$ subjects, with replacement.
2. Draw a sample V_i^* from the original mark data V_i of group $Z = z$ subjects who became infected ($\delta = 1$), with replacement.
3. Compute D_z based on the bootstrap data $(X_i^*, \delta_i^*, \delta_i^* V_i^*)$ for subjects i in group $Z = z$.
4. Repeat Steps (1)–(3) B times.
5. Estimate the critical value as the α -quantile of D_z across the B replicates.

The bootstrap test has the nominal size asymptotically and is consistent against all alternatives (Romano, 1988). We test the overall hypothesis $\{K_0 : T \perp V | Z\}$ using Simes' procedure applied to the tests of K_0 performed separately for the two groups $Z = 1$ and $Z = 0$.

4. Simulation study

4.1 Assessment of the proposed methods under model validity

The simulation setup mimics typical placebo-controlled HIV vaccine efficacy trials. Initially, we investigate finite-sample performance of the $VE(v)$ estimation and testing procedure in the model $VE(v) = 1 - e^{\alpha + \beta v + \gamma}$, $v \in [0, 1]$, i.e., the data are generated following models (5) with $g(v, \alpha, \beta) = e^{\alpha + \beta v}$ and (7). We compare the performance of the estimator to that of the semiparametric estimator of Sun, Gilbert, and McKeague (SGM) (2009), and compare the performance of the Wald and likelihood ratio tests of H_0^0 and H_0 to those of GMS.

We specify the failure times T to be exponential with rates λ and λe^γ in the placebo and vaccine groups. The rate $\lambda = \log(0.85)/(-3)$ is chosen so that 15% of placebo recipients are infected by 3 years. We specify the censoring times C to be $\text{Uniform}(0, 15)$ in each group which implies a 20% chance of censoring by 3 years. The observed time on study X is defined as $\min(T, C, 3)$. The vaccine-to-placebo assignment ratio is 1:1. A univariate mark V for placebo and vaccine recipients is generated from distributions with density functions $f(v|Z=0) = \{2e^{-2v}/(1 - e^{-2})\} I(0 \leq v \leq 1)$ and $f(v|Z=1) = f(v|Z=0)e^{\alpha + \beta v}$, where, for a given value of β , the value of the parameter $\alpha = \alpha(\beta)$ is defined as the solution to $\int_0^1 f(v|Z=0)e^{\alpha + \beta v} dv = 1$. Only the mark values for subjects with $T \leq 3$ and $\delta = 1$ would be observed in a real study and hence are used in the analysis.

We consider sample sizes $n = 556, 741$, and 1481 per arm so that the expected numbers of observed placebo infections by year 3 are $N_{pI} = 75, 100$, and 200 , respectively. The simulation scenarios are characterized by the following values of (β, γ) : $(0, 0)$, $(0.5, -0.8)$, $(1.2, -0.2)$, and $(2.1, -1.3)$. The true $VE(v)$ functions are depicted in Web Figure 1; in subsequent tables and figures, the simulation scenarios are labeled by the pertaining $VE(0)$ and $VE(1)$ values. For all simulations the nominal significance level is taken to be 5% for two-sided tests and 2.5% for one-sided tests. The results are based on 10^4 replicated data sets.

Figures 1, 2, and Web Figure 2 characterize finite-sample bias of estimates and asymptotic standard error estimators of $\widehat{VE}(v)$ as well as coverage probabilities of 95% pointwise Wald confidence intervals for $VE(v)$. In each simulation scenario, bias is the smallest in the central support region of the mark and slightly increases in the right tail. As expected, bias decreases with an increasing number of events. Web Figure 4 compares bias of the proposed and the SGM estimator. Asymptotic and empirical standard error estimates are consistently

in good accordance with one another, and Web Figure 5 illustrates a substantial efficiency gain relative to the SGM method in this setting. Coverage probabilities for all mark values approach the nominal confidence level with an increasing number of failures.

Tables 1 and 2 summarize the observed sizes and powers of the Wald and likelihood ratio tests of H_0^0 and H_0 , and include a comparison with the alternative non- and semiparametric tests of GMS. The sizes of all considered tests are in good agreement with the nominal significance level. As for two-sided testing of H_0^0 , powers of the Wald and likelihood ratio test are comparable, and both tests outperform the $\hat{U}_1^3$ and $\hat{U}_1^4$ tests of GMS, with power gains up to 19%, for all specified values of the model parameters. The one-sided weighted Wald-type test (11) of H_0^0 is designed to be sensitive to alternative hypotheses with positive marginal vaccine efficacy and $VE(v)$ declining in v , for which it is conjectured to be more powerful than the standard log-rank test that assesses the marginal vaccine efficacy only; this conjecture is verified in the results with power gain reaching up to 27%.

Powers of the Wald and likelihood ratio tests of H_0 are considerably higher (up to 36%) than those of GMS's $\hat{U}_2^{np}$ and $\hat{U}_2^{sp}$ tests. We conjecture that the power gain can be ascribed to the additional assumptions made by our methods compared to those of GMS. Web Table 1 supports nominal size of the diagnostic test (12) of $\{K_0 : T \perp V|Z\}$.

4.2 Simulation scenarios with a bivariate mark variable

We additionally study simulation scenarios involving a bivariate mark $V = (V_1, V_2)$, where V_1 and V_2 are generated as independent univariate marks from distributions described in Section 4.1. The point and interval estimator for bivariate mark-specific vaccine efficacy exhibits similar properties as in scenarios with a univariate mark (Web Appendix D). For bivariate marks, the likelihood ratio tests of H_0^0 and H_0 are slightly more powerful than the respective Wald tests, and the diagnostic test (12) of the null hypothesis $\{T \perp (V_1, V_2)|Z\}$ has nominal size.

4.3 Robustness analysis of the proposed methods under model mis-specification

Assuming the model $VE(v) = 1 - e^{\alpha + \beta v + \gamma}$, we examine robustness of the $VE(v)$ estimator and of the Wald and likelihood ratio tests of H_0^0 and H_0 to violation of the $T \perp V|Z$ assumption. To this end, we generate samples (T, V) with correlation ρ_{TV} . For an infected placebo recipient, we first generate $T = t$, and second generate V from the conditional density function

$$f(v|T=t, Z=0; c) = \{cte^{-ctv} / (1 - e^{-ct})\} I(0 \leq v \leq 1), c > 0,$$

where c governs the magnitude of ρ_{TV} . Using Proposition 1, for an infected vaccine recipient with $T = t$, we generate V from the conditional density function $f(v|T = t, Z = 1; c) = f(v|T = t, Z = 0; c)e^{\alpha(t) + \beta v}$, where $\alpha(t) = \alpha(t, \beta, c)$ is the solution to

$\int_0^1 f(v|T=t, Z=0; c)e^{\alpha(t) + \beta v} dv = 1$. The data generating mechanism implies that the true $VE(t, v) = 1 - e^{\alpha(t) + \beta v + \gamma}$. The values of c are chosen such that ρ_{TV} varies over 0.1–0.5. In each scenario, power to reject the null hypothesis $\{T \perp V|Z\}$ using the diagnostic test (12) is evaluated.

Web Figure 3 depicts the Monte Carlo average deviation $\widehat{VE}(v) - \overline{VE}(v)$ where $\overline{VE}(v) = \int_0^3 VE(t, v) f(t) dt / \int_0^3 f(t) dt$ and $f(t)$ is the density function of the failure time T .

For scenarios where the magnitude of correlation is detected with moderate power, the deviation is minimal. Such scenarios are of particular interest because, in these cases, violation of the assumption $T \perp V|Z$ may frequently remain undetected. For scenarios with power to detect departures from $T \perp V|Z$ greater than 90%, the deviation tends to increase for mark values in the right tail of the distribution, although this is fairly innocuous because these are the scenarios where the method would fail the diagnostics and hence would not be used.

Web Tables 2 and 3 indicate that the likelihood ratio and Wald tests of H_0^0 and H_0 retain correct size under violation of $T \perp V|Z$. This robustness property holds for any violation of this assumption, because under either null hypothesis H_0^0 and H_0 , $\alpha(t) \equiv \alpha = 0$ for all $t > 0$, i.e., the estim and $VE(v)$ in the time-independent model coincides with the true $VE(t, v)$ characterizing the data generating mechanism regardless of the value of ρ_{TV} . Web Tables 2 and 3 show that power of the likelihood ratio and Wald tests is minimally sensitive to correlation between T and V and tends to be higher than that of GMS's tests.

Finally, we performed a robustness analysis with respect to violation of the proportional marginal hazards assumption in which we considered distributions of $T|Z = 0$ and $T|Z = 1$ characterized by the hazard functions $\lambda(t|Z = 0) \equiv \lambda$ and $\lambda(t|Z = 1) = \lambda e^{\gamma_0 + \gamma_1 t}$; thus γ_1 governs the amount of variation in the marginal hazard ratio over time. We observed minimal differences in the patterns of average deviation $\widehat{VE}(v) - \overline{VE}(v)$, standard error estimates, and power relative to those presented for correctly specified models (results not reported). We conjecture that the similar performance occurs because $VE(t, v) = 1 - e^{\alpha + \beta v + \gamma_0 + \gamma_1 t}$ does not include an interaction of V with T , and hence $VE(t, v)$ and $VE(v)$ have the same amount of change in v .

5. Application

5.1 Genetic distance definition

In the RV144 Thai trial, full length HIV genomes were measured from 121 of the 125 infected subjects; 3 are missing data because their viral load was too low for the Sanger sequencing technology to work, and 1 dropped out. The 4 subjects with missing data are excluded from the analysis. The recent RV144 'immune correlates' study showed that vaccine recipients with high levels of IgG antibodies that bind to the V1/V2 region of the HIV Envelope gp120 protein had a significantly lower rate of HIV infection (Haynes et al., 2012). To help discern the potential role of V1/V2 antibody epitopes as an immunologic cause of the observed partial vaccine efficacy, we focus the genetic distance definitions on the V1/V2 region (85 amino acids long). Three HIV gp120 sequences were included in the vaccine: 92TH023 in ALVAC canarypox vector, and A244, MN in AIDSVAX gp120. 92TH023 and A244 are subtype E-viruses whereas MN is subtype B, and 110 (91%) of the 121 subjects were infected with subtype E viruses. The E viruses are closer genetically to the infecting sequences than MN, and thus are more likely to stimulate protective immune responses. Accordingly, our analysis focuses on the 92TH023 and A244 insert sequences, and excludes the 11 subjects with non-E infecting sequences.

To maximize biological relevance and statistical power of the analysis, V1/V2 distances were computed only including the subset of the 85 amino acids that were predicted to be part of antibody epitopes that were stimulating vaccine-induced responses. Web Appendix E describes this process, which yielded $q = 9$ important amino acid sites that were used for defining the distances.

For each of the two insert reference sequences and each sequence from a subject, the genetic distance was computed using the Blosom90 scoring matrix, adjusted such that the addition or subtraction of an N-linked glycosylation site is scored 3-fold or 1.5-fold higher than the largest Blosom90 score depending on whether the mismatch is in position 2 of the aminoacid triple comparing the N-linked glycosylation site. First, a distance was computed as

$$d(seq_{\text{subject}}, seq_{\text{reference}}) = \sum_{i=1}^q \frac{2s(a_i^s, a_i^r)}{s(a_i^s, a_i^s) + s(a_i^r, a_i^r)},$$

where a_i^s is the amino acid at site i in the subject's sequence, a_i^r is the amino acid at site i in the reference sequence, and $s(x, y)$ is the score from the modified Blosom90 matrix. Second, the distances were re-scaled to values between 0 and 1 (2 identical sequences have distance 0). Between 2 and 13 sequences (total 1030 sequences) were measured per infected subject, and the final distance was defined as the subject-specific median distance. This yields a data set of a bivariate mark (V_1, V_2) measured from 110 infected subjects, with 49 and 52 unique mark values in V_1 and V_2 , respectively. Figure 3 shows the distributions of V_1 and V_2 .

5.2 Inference about mark-specific vaccine efficacy

Point and 95% pointwise confidence interval estimates of $VE(v)$ are plotted in Figures 4a–b. The weighted Wald test of $H_0: VE(v) \equiv 0$ yields p -values of 0.016 and 0.024 for the 92TH023 and A244 distances. The likelihood ratio test of $H_0: VE(v) = VE$ yields $p = 0.52$ for the 92TH023 and $p = 0.83$ for the A244 distance. These p -values are consistent with the broad confidence bands in Figures 4a and 4b. The Thai trial had low power to detect a change in $VE(v)$ with viral divergence. The assumption of conditional independence between T and V given Z was verified using the Kolmogorov–Smirnov-type test, with p -values of 0.88 and 0.98 for the placebo and vaccine groups, based on 1000 bootstrap iterations. The Qin and Zhang (1997) goodness-of-fit test does not reject the validity of the mark density ratio model for either the 92TH023 ($p = 0.49$) or the A244 distance ($p = 0.51$). Moreover, a test of proportional hazards assumed in (7) (Grambsch and Therneau, 1994) does not reject ($p = 0.21$).

Figure 4c displays $\widehat{VE}(v_1, v_2)$. The weighted Wald test of H_0 yields $p = 0.019$. The likelihood ratio test of H_0 yields $p = 0.81$. The p -value for the test of $T \perp (V_1, V_2) | Z$ is 0.64 and 0.89 in the placebo and vaccine groups again supporting the appropriateness of the method. The goodness-of-fit test does not reject the validity of the bivariate-mark density ratio model ($p = 0.46$).

6. Discussion

This article develops inferential procedures for the mark-specific hazard ratio model with a multivariate continuous mark variable. We show that the mark-specific hazard ratio factors as the mark density ratio and the ordinary marginal hazard ratio. This factorization is appealing because it enables the two components to be estimated separately. The semiparametric method of maximum profile likelihood estimation in the mark density ratio model yields improved efficiency of mark-specific hazard ratio estimation relative to current alternative approaches. We derive the joint asymptotic distribution of the estimators for the mark density ratio and the marginal hazard ratio. This result allows estimation of mark-specific vaccine efficacy ($VE(v)$) with pointwise confidence intervals and testing of

hypotheses of interest about $VE(\mathbf{v})$. The proposed weighted Wald-type test is recommended for testing $VE(\mathbf{v}) \equiv 0$ for all $\mathbf{v}$, and the likelihood ratio test for testing $VE(\mathbf{v})$ invariant in $\mathbf{v}$.

Due to growing complexity in HIV vaccine design, it becomes increasingly desirable to evaluate $VE(\mathbf{v})$ as a function of a multivariate mark representing a set of genetic distances. Our approach allows flexible semiparametric modeling of a multivariate mark variable, and the estimation method does not impose practical constraints on the mark dimension.

A limitation of the methods is that they do not apply if the $T \perp V|Z$ assumption fails; in this scenario, an alternative approach may be used that is based on a different factorization of the mark-specific hazard function,

$$\lambda(t, \mathbf{v}|Z=z) = f(t, \mathbf{v}|Z=z)S^{-1}(t|Z=z),$$

with $f(t, \mathbf{v}|Z=z)$ the conditional joint density function of $(T, \mathbf{V}_T)^T$ given $Z=z$, and $S(t|Z=z)$ the conditional marginal survival function of T given $Z=z$. This approach eliminates the $T \perp V|Z$ assumption as well as the proportional marginal hazards assumption in exchange for flexible parametric modeling of the failure time in the density ratio model. It would be of interest to develop estimation and testing methods for this approach.

For a univariate mark variable, the present method is an alternative to the SGM method, which estimates the same $VE(\mathbf{v})$ estimand and tests the same null hypotheses. The following issues are relevant for the choice of method. First, the present method makes two assumptions not made by SGM: $T \perp V|Z$ and a parametric form for the mark density ratio. Therefore, where the goodness-of-fit tests reject either of these assumptions, the SGM method is more appropriate. However, if the data support both assumptions, then the present method may be preferred for its substantially improved efficiency, a key issue given the concern for false negative errors. Second, SGM assumes a mark-specific proportional hazards model, such that $VE(t, \mathbf{v})$ is independent of t . The present method partially relaxes this assumption via the factorization $1 - VE(t, \mathbf{v}) = g(\mathbf{v}, \varphi) \{1 - VE(t)\}$, which allows flexible modeling of $VE(t)$, for example via the Cox model with time-dependent covariates, which allows $VE(t, \mathbf{v})$ to vary with time. Third, the SGM method uses kernel smoothing to estimate $VE(\mathbf{v})$ over the range of mark values, thereby providing more flexible estimation. However, issues of band width selection and complicated behavior of the estimator in the tails of the mark variable make the SGM method more complicated to implement.

Although the presented methods are motivated by a specific scientific application, they provide a general approach to the analysis of survival data in the presence of a continuous, possibly multivariate, mark variable. The R code implementing the proposed methods is available upon request and is posted at the second author's webpage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the reviewers and Jon Wellner for helpful suggestions, and thank the participants, investigators, and sponsors of the RV144 Thai trial, including the U.S. Military HIV Research Program (MHRP); U.S. Army Medical Research and Materiel Command; NIAID; U.S. and Thai Components, Armed Forces Research Institute of Medical Science Ministry of Public Health, Thailand; Mahidol University; SanofiPasteur; and Global Solutions for Infectious Diseases. This research was supported by NIH NIAID grant 2 R37 AI054165-10.

References

- Gilbert P. Large sample theory of maximum likelihood estimates in semiparametric biased sampling models. *Annals of Statistics*. 2000; 28:151–194.
- Gilbert P, Lele S, Vardi Y. Maximum likelihood estimation in semiparametric selection bias models with application to AIDS vaccine trials. *Biometrika*. 1999; 86:27–43.
- Gilbert P, McKeague I, Sun Y. The 2-sample problem for failure rates depending on a continuous mark: an application to vaccine efficacy. *Biostatistics*. 2008; 9:263–276. [PubMed: 17704528]
- Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994; 81:515–526.
- Halloran M, Struchiner C, Longini I. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *American Journal of Epidemiology*. 1997; 146:789–803. [PubMed: 9384199]
- Haynes B, Gilbert P, McElrath M, Zolla-Pazner S, Tomaras G, Alam S, Evans D, Montefiori D, Karnasuta D, Sutthent R, Liao H, DeVico A, Lewis G, Williams C, Fong Y, Janes H, DeCamp A, Huang Y, Rao M, Billings E, Karasavvas N, Robb M, Ngaay V, de Souza M, Paris R, Ferrari G, Bailer R, Soderberg K, Andrews C, Berman P, Frahm N, De Rosa S, Alpert M, Yates N, Shen X, Koup R, Pitisuttithum P, Kaewkungwal J, Nitayaphan S, Rerks-Ngarm S, Michael N, Kim J. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *New England Journal of Medicine*. 2012; 366:1275–1286. [PubMed: 22475592]
- Huang Y, Louis T. Nonparametric estimation of the joint distribution of survival time and mark variables. *Biometrika*. 1998; 85:785–798.
- Lu X, Tsiatis A. Improving the efficiency of the log-rank test using auxiliary covariates. *Biometrika*. 2008; 95:679–694.
- Prentice R, Kalbfleisch J, Peterson A, Flournoy N, Farewell V, Breslow N. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978; 34:541–554. [PubMed: 373811]
- Qin J. Inferences for case-control and semiparametric two-sample density ratio models. *Biometrika*. 1998; 85:619–630.
- Qin J, Zhang B. A goodness-of-fit test for logistic regression models based on case-control data. *Biometrika*. 1997; 84:609–618.
- Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, Prensri N, Namwat C, de Souza M, Adams E, Benenson M, Gurunathan S, Tartaglia J, McNeil J, Francis D, Stablein D, Bix D, Chunsuttiwat S, Khamboonruang C, Thongcharoen P, Robb M, Michael N, Kunasol P, Kim J. for the MOPH-TAVEG Investigators. Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. *New England Journal of Medicine*. 2009; 361:2209–2220. [PubMed: 19843557]
- Romano J. A bootstrap revival of some nonparametric distance tests. *Journal of the American Statistical Association*. 1988; 83:698–708.
- Simes R. An improved Bonferroni procedure for multiple tests of significance. *Biometrika*. 1986; 73:751–754.
- Sun Y, Gilbert P, McKeague I. Proportional hazards models with continuous marks. *Ann. Stat.* 2009; 37:394–426. [PubMed: 19838313]

Figure 1.

Finite-sample bias of estimation of $VE(v) = 1 - e^{\alpha + \beta v + \gamma}$ with a univariate mark. N_{pl} and N_{vl} denote the expected numbers of observed placebo and vaccine infections, respectively. This figure appears in color in the electronic version of this article.

Figure 2.

Asymptotic versus empirical (dotted) standard error estimates of $\widehat{VE}(v)$ with a univariate mark. N_{pl} and N_{vl} denote the expected numbers of observed placebo and vaccine infections, respectively. This figure appears in color in the electronic version of this article.

Figure 3.

RV144 trial: distribution of the distances to the 92TH023 and A244 vaccine insert sequence by vaccine/placebo group. This figure appears in color in the electronic version of this article.

(a)

(b)

(c)

Figure 4.

Estimated mark-specific vaccine efficacy for (a) the univariate distance to the 92TH023 insert sequence, (b) the univariate distance to the A244 insert sequence, and (c) the bivariate distance with components defined in (a) and (b). 95% pointwise confidence bands are displayed in (a) and (b). This figure appears in color in the electronic version of this article.

Powers of tests of the null hypothesis $\{H_0^0 : V E(v) = 0 \text{ for all } v \in [0, 1]\}$ against (i) the alternative hypothesis $\{H_0^1 : \text{non } H_0^0\}$, and (ii) the alternative hypothesis $\{H_2^0 : V E > \text{ and } V E(v) \text{ is a decreasing function in } v\}$. As for testing H_0^0 versus H_1^0 , the proposed Wald and likelihood ratio tests are compared with the alternative nonparametric tests $\hat{U}_3^1$ and $\hat{U}_4^1$ (Gilbert, McKeague, and Sun, 2008) based on comparing nonparametric maximum likelihood estimates of the cumulative mark-specific hazard functions $\Lambda(t, v|Z=k) = \int_0^v r^k \lambda(s, u|Z=k) ds du$, for the vaccine and placebo arms. Furthermore, power of the one-sided weighted Wald-type test (11) of H_0^0 versus H_2^0 is compared to that of the two-sided log-rank test of equal failure distributions in the vaccine and placebo arm. The nominal significance level is taken to be 5% for each two-sided test and 2.5% for each one-sided test.

Table 1

VE(0)	VE(1)	N_{pl}	N_{ω}	$H_0^0 \text{ vs } H_1^0$		$H_0^0 \text{ vs } H_2^0$	
				Wald	LRatio	$\hat{U}_3^1$	$\hat{U}_4^1$
0.0	0.0	200	200	0.04	0.04	0.04	0.04
		100	100	0.05	0.04	0.06	0.06
		75	75	0.05	0.05	0.06	0.04
0.6	0.4	200	94	1.00	1.00	1.00	1.00
		100	47	0.99	0.99	0.79	0.80
		75	35	0.97	0.97	0.98	0.97
0.5	-0.7	200	167	0.93	0.90	0.50	0.89
		100	83	0.63	0.58	0.28	0.57
		75	63	0.53	0.51	0.24	0.48
0.9	0.1	200	58	1.00	1.00	0.98	0.98
		100	29	1.00	1.00	0.80	0.80
		75	22	1.00	1.00	1.00	1.00

Table 2

Powers of two-sided tests of the null hypothesis $\{H_0 : VE(v) \equiv VE \text{ for all } v \in [0,1]\}$. The proposed Wald and likelihood ratio tests are compared with the alternative non- and semiparametric tests $\hat{U}_2^{np}$ and $\hat{U}_2^{sp}$ (Gilbert, McKeague, and Sun, 2008) based on comparing the nonparametric maximum likelihood estimate $\hat{\Lambda}(t, v|Z=1) - \hat{\Lambda}(t, v|Z=0)$ with a non- and semiparametric estimate of $\Lambda(t, v|Z=1) - \Lambda(t, v|Z=0)$ under H_0 . For each test, the nominal significance level is set to 5%.

VE(0)	VE(1)	N_{PI}	N_{ω}	Wald	LRatio	$\hat{U}_2^{np}$	$\hat{U}_2^{sp}$
0.0	0.0	200	200	0.04	0.04	0.04	0.05
		100	100	0.04	0.05	0.04	0.04
		75	75	0.06	0.06	0.03	0.04
0.6	0.4	200	94	0.18	0.18	0.09	0.09
		100	47	0.12	0.12	0.10	0.07
		75	35	0.12	0.12	0.10	0.07
0.5	-0.7	200	167	0.89	0.89	0.66	0.65
		100	83	0.60	0.61	0.36	0.35
		75	63	0.53	0.54	0.24	0.26
0.9	0.1	200	58	0.97	0.97	0.78	0.76
		100	29	0.79	0.80	0.44	0.36
		75	22	0.67	0.66	0.36	0.32