STATISTICAL ANALYSIS OF MARK-SPECIFIC PROPORTIONAL HAZARDS MODEL

by

Mei Li

A dissertation submitted to the faculty of
the University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2011

Approved by:

Dr. Yanqing Sun

Dr. Jiancheng Jiang

Dr. Zhou Weihua

Dr. Arun Ravindran

ABSTRACT

MEI LI. Statistical analysis of mark-specific proportional hazards model.
(Under the direction of DR. YANQING SUN)

Competing risks occur frequently in survival analysis, and in some cases, the competing risks are not discrete. In this dissertation, we develop some statistical inferences to analyze continuous competing risks.

In Chapter 2, inspired by the HIV vaccine trials, we extend the modeling of mark-specific hazards function to multivariate marks to better fit the HIV data. We develop the partial likelihood based parametric procedure to estimate the coefficents. The asymptotic properties of the proposed estimators are derived. We propose some tests to examine a variety of null hypotheses to understand how relevant the two distances are for protection. Finite sample performances of the proposed methods, are examined through extensive simulations and are shown satisfying. The methods are applied to STEP data to evaluate the vaccine efficacy and its dependence on the multivariate marks. A goodness of fit procedure is also developed. The test statistics are constructed based on the score function and the generalized weighted martingale residuals. The performance of tests are also examined through simulations. And the tests are used to check adequacy of the multivariate mark-specific proportional hazard model for STEP data.

In Chapter 3, we develop a goodness of fit procedure for the stratified mark-specific proportional hazard model with continuous marks. Coefficents are estimated through partial likelihood based kernel smoothing method. The asymptotic properties of the proposed estimators are derived. We also construct confidence bands for vaccine efficacy. We focus on the goodness of fit test of the model. The test statistics are constructed based on the generalized weighted martingale residuals. The finite sample properties of proposed tests are examined through simulations.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my advisor, Professor Yanqing Sun, for her thoughtful and insightful directions through years. Her inspiring, encouraging, and personal guidance have provided me with a deeper understanding of the research topic, and made the completion of this research possible.

I also would like to thank the members of my committee, Professor Jiancheng Jiang, Professor Weihua Zhou, and Professor Ravindran Arun, for their help and support.

I would like to give thanks to Professor Peter Gilbert for motivating the research topics and for giving his expertise throughout the study.

I acknowledge that this research is partially supported by NSF grants DMS-0604576 and NIH grant R37 AI054165-09. I am grateful for the Department of Mathematics and S-tatistics at the University of North Carolina at Charlotte for providing me with an excellent work environment and for supporting me during the past years.

Finally, I want to thank my husband, Dichao Peng, my parents, and my friends for their continuous encouragement and support.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

In the traditional survival analysis, we often assume the subject will only experience one type of failure. However, in many clinical trials, the subject may experience more than one type of event, and the occurrence of one type of event hinders the occurrence of other types of events (Pintilie, 2006). For example, patients who undergo bone marrow transplant are followed in order to observe acute graft versus host disease (GVHD). At the end of the study, patients will either developed acute GVHD, or did not, or died from other causes. So death without acute GVHD is the competing risk for the acute GVHD (Wang, 2010). In demographic mortality studies, when evaluating the efficacy of heart transplants, one may want to treat deaths due to heart failure differently from deaths due to other causes, such as accident and cancer (Lee, 2003). As discussed above, competing risks occur frequently in medical research.

1.1 Competing Risks Data

Let T be the failure time that may be subject to censoring, Z the covariate vector, and $J \in \{1, 2, ..., m\}$ the type or cause of failure. A typical right censored competing risks data can be represented by $\{X_i, \delta_i, \delta_i J_i, Z_i\}$, $i = 1, \ldots, n$. Where $X_i = min(T_i, C_i)$, $\delta_i = I(T_i \leq C_i)$ and C_i is the censoring time. When $\delta_i = 1$, the cause of failure is observed; Otherwise, it is undefined and unknown. A basic estimable quantity based on the competing risks data is the cause-specific hazard functions, defined by

$$\lambda_j(t; z) = \lim_{\Delta t \to 0} P(t < T < t + \Delta t, J = j | T > t; Z = z)/\Delta t, \tag{1.1}$$

for $J \in \{1, 2, \ldots, m\}$. And the overall hazard function is

$$\lambda(t; z) = \lim_{\Delta t \to 0} P(t < T < t + \Delta t | T > t; Z = z)/\Delta t.$$

Prentice et al. (1978) introduced a Cox regression framework for the analysis of failure time data by cause-specific hazard function in the presence of finitely many competing risks. The regression function is defined as

$$\lambda_j(t|z) = \lambda_{0j}(t)\exp(\beta_j^T z(t)), \quad j = 1, 2, \ldots, m, \tag{1.2}$$

where $\lambda_{0j}(t)$ is an unspecified baseline function, and β_j is a column vectors of cause-specific regression coefficients. A lot of work has been done on the discrete cases, Kuk (1992), Aly, Kocha, and Mckeague (1994), Lunn and McNeil (1995), Sun and Tiwari (1995), Lam (1998), Hu and Tisa (1999), Luo and Tunbull (1999), Sun (2001), and Scheike, Zhang and Gerds (2008).

1.2 Competing Risks Data with Continuous Marks

Many important applications of competing risks methodologies involve continuous causes-of-failure (marks). For example, in a AIDS clinical trial of drug regimens for treating HIV infection, the time to treatment failure (typically defined by levels of viral load rising above a threshold (Gilbert et al., 2001)) can decrease with increases in a distance measure describing the extent of drug-selected HIV genetic evolution within a patient between baseline and the time of failure. Detecting such an association can help in designing anti-HIV treatments that overcome the problem of drug resistance, which represents one of the greatest barriers to achieving durably efficacious treatment of HIV infection (Hirsch et al., 2000; Yeni et al., 2002). In this example, the mark V is a measure of the accumulated HIV genetic resistance resulting from exposure to an antiretroviral treatment, which is measured only on subjects who fail treatment, at the time of treatment failure. A second example of interest is a prospective cohort study of a population at risk for acquiring HIV infection. In this application, T is the time from cohort entry until HIV infection, and V is the value of a metric measuring genotypic or phenotypic dissimilarity of the HIV virus that infects a study participant from a reference HIV strain. For example, V could be Hamming's genetic distance and the reference strain could be the prototype virus contained in an HIV vaccine that is under development for field testing in the cohort population.

Other examples of continuous mark variables include lifetime medical cost or a quality of life score associated with survival time (Olschewski and Schumacher, 1990). The grouping of continuous mark data into discrete marks is unsatisfactory because that amounts to a coarsening of the data and the results will depend on the way the groups are defined. Huang and Louis (1998) first considered a continuous mark and developed the nonparametric maximum likelihood estimator of the joint distribution of the failure time and the mark.

Denote the time to endpoint as T and the mark variable V, the observable random variables are $(X, \delta, \delta V)$, where $X = \min\{T, C\}$, $\delta = I(T \leq C)$, and C is a censoring random variable. When the failure time T is observed, $\delta = 1$ and the mark V is also observed, whereas if T is censored, the mark is unknown. Gilbert, Mckeague and Sun (2004) introduced the mark-specific hazard function

$$\lambda(t, v) = \lim_{h_1, h_2 \to 0} P\{T \in [t, t + h_1), V \in [v, v + h_2) | T \geq t\}/h_1 h_2, \tag{1.3}$$

with t ranging over a fixed interval $[0, \tau]$, where τ is the end of follow-up. The mark-specific function is the natural analog of its discrete counterpart, with similar interpretation. In particular, $\lambda(t, v)$ is the instantaneous risk of failure by a cause V in a small interval $[v, v + h_2)$ in the presence of all other causes. Just as the cause-specific hazard functions are the basic estimable quantities when the mark variable is discrete, the mark-specific hazard function $\lambda(t, v)$ is estimable from the available data and forms the basis for inference when the mark variable is continuous. Indeed, the likelihood function under the competing risks data with continuous mark has a similar form and is derived as follows. Assume that the continuous mark variable V has a known bounded support; rescaling V if necessary, this support is taken to be $[0, 1]$. Let $f(t, v)$ be the joint density of (T, V), and $S_T(t)$ be survival function of T. Then $\lambda(t, v) = f(t, v)/S_T(t)$ and $\lambda(t) = \int_0^1 \lambda(t, v)\, dv$ is the overall hazard function of T. The likelihood function given n i.i.d. observations $(X_i, \delta_i, \delta_i V_i)$, $i = 1, \ldots, n$

from the above model can be expressed in terms of the mark-specific hazard rate as

$$\prod_o f(X_i, V_i) \prod_c S_T(X_i) = \prod_o \lambda(X_i, V_i) \prod_{i=1}^{n} \exp\left\{ - \int_0^1 \int_0^{X_i} \lambda(s, v)\, ds\, dv \right\}, \tag{1.4}$$

where $\prod_o$ denotes the product over the observed failure times, $\prod_c$ denotes the product over right censored failure times, and each product only applies to the expression immediately in front. Gilbert et al. (2004) developed nonparametric testing procedures for

$$H_0 \colon \ \lambda(t, v) \text{ does not depend on } v \text{ for } t \in [0, \tau]$$

against the following alternative hypotheses:

$$\begin{aligned} H_1 \colon \ \lambda(t, v_1) &\leq \lambda(t, v_2) \text{ for all } v_1 \leq v_2, \ t \in [0, \tau]; \\ H_2 \colon \ \lambda(t, v_1) &\neq \lambda(t, v_2) \text{ for some } v_1 \leq v_2, \ t \in [0, \tau] \end{aligned}$$

with strict inequalities for some t, v_1, v_2 in H_1.

In the AIDS clinical trial example, the test of H_0 versus the monotone alternative H_1 assesses whether the absolute (instantaneous) risk of treatment failure increases with the level of acquired drug resistance. The test is useful for evaluating if V is a clinically relevant measure of a treatment's resistance cost. Knowledge of clinically meaningful genetic resistance cost metrics would be helpful for identifying combination drug regimens that do not select for drug resistant virus, and thus provide long-lasting treatment efficacy. In the second example mentioned above, the test of H_0 versus the two-sided hypothesis H_2 assesses whether the HIV metric V is associated with the instantaneous risk of HIV infection. Finding evidence for H_2 may suggest that the metric V can be used to guide selection of the types of HIV antigens to include in HIV vaccines (Gilbert et al., 2001). For example, if H_0 is rejected and the infection risk appears particularly high for $v > 0.7$, then it may behoove vaccine researchers to insert HIV antigens characterized by $v > 0.7$. Carrying out the test for multiple metrics in multiple genes could help identify the metric(s) that

optimize the breadth of expected protective coverage of the vaccine. This application is important because the broad genotypic and phenotypic diversity of HIV poses one of the greatest challenges to developing an effective AIDS vaccine (UNAIDS, 2001).

Let $\lambda_1(t,v)$ and $\lambda_2(t,v)$ denote the mark-specific hazard functions for the vaccine group and placebo group, respectively. The mark-specific vaccine efficacy is defined as $VE(t,v) = 1 - \lambda_1(t,v)/\lambda_2(t,v)$. Assuming the HIV vaccine trial is a randomized double-blinding trial and since HIV infection is a rare event in HIV vaccine efficacy trials, $VE(t,v)$ approximately measure the vaccine effect to reduce susceptibility to HIV acquisition given exposure to strain v at time t; See Gilbert, Mckeague and Sun (2008). Gilbert, Mckeague and Sun (2008) developed some nonparametric test procedures to evaluate the mark-specific HIV vaccine efficacy, the the mark of interest is a measure of genetic distance between the HIV sequence sampled from a volunteer infected in the trial and the HIV sequence represented in the tested vaccine construct. Specifically, they considered testing the null hypothesis

$$H_0^0\colon\ VE(t,v) = 0 \text{ for } (t,v) \in [0,\tau] \times [0,1]$$

against the following alternative hypotheses:

$$H_1^0\colon\ VE(t,v) \geq 0 \text{ for all } \ (t,v) \in [0,\tau] \times [0,1];$$
$$H_2^0\colon\ VE(t,v) \neq 0 \text{ for some } \ (t,v) \in [0,\tau] \times [0,1]$$

with strict inequality for some $(t,v) \in [0,\tau] \times [0,1]$ in H_1^0. Testing H_0^0 evaluates whether there is any vaccine efficacy against any HIV strain. If H_0^0 is rejected, then it is of interest to assess if vaccine efficacy varies with strain distance. They have developed test for

$$H_0\colon\ VE(t,v) \text{ does not depend on } v \text{ for } t \in [0,\tau]$$

against the following alternative hypotheses:

$$H_1\colon\ VE(t,v_1) \geq VE(t,v_2) \text{ for all } v_1 \leq v_2,\ t \in [0,\tau];$$
$$H_2\colon\ VE(t,v_1) \neq VE(t,v_2) \text{ for some } v_1 \leq v_2,\ t \in [0,\tau]$$

with strict inequality for some t, v_1, v_2 in H_1. Testing H_0 versus H_1 assesses whether vaccine efficacy decreases with HIV sequence divergence. Detecting that the vaccine protects against some strains but not others, and quantifying the relationship between vaccine efficacy and viral divergence, is useful for guiding vaccine deployment decisions and for designing new vaccines that provide greater breadth of protection.

To evaluate HIV vaccine efficacy with adjustments for covariates and develop more efficient statistical methods, Sun, Gilbert and Mckeague (2009) studied the mark-specific proportional hazards model with continuous marks defined as:

$$\lambda(t,v|z(t)) = \lambda_0(t,v)\exp(\beta(v)^T z(t)), \tag{1.5}$$

where the baseline hazard function $\lambda_0(t,v)$ is an unspecified function of (t,v) and the p-dimensional regression parameter $\beta(v)$ are unknown continuous functions of the marks v.Let $Z(t) = (Z_1, Z_2(t))$ where Z_1 is the treatment(vaccine) group indicator and $Z_2(t)$ are possibly time-dependent covariates. The covariate $Z_2(t)$ adjusted mark-specific vaccine efficacy is then $VE(v) = 1 - exp(\beta_1(v))$, where $\beta_1(v)$ is the coefficient for Z_1. Sun, Gilbert, and McKeague (2009) developed some estimations and hypothesis tests for $VE(v)$ based univariate continuous marks. Model (1.5) can provide more powerful tests of mark-specific vaccine efficacy. Furthermore, ignoring the mark variable and studying vaccine efficacy using the standard Cox model, as is widely practiced in vaccine trials for many infectious diseases, can give misleading results since the ordinary Cox model averages the mark-specific vaccine efficacy over its range, and important vaccine effects may be missed.

1.3 Multivariate Continuous Marks

Despite the research progress, the previous work did not account for multivariate marks. This is a serious limitation given that all of the candidate HIV vaccines tested in HIV vaccine efficacy trials have contained multiple antigens/immunogens, with rational to attempt to elicit multiple types of immune response that recognize and block different types of HIV viruses. The greater the number of virus types that can be recognized and killed by vaccine-induced immune responses, the greater the potential overall vaccine efficacy. Many HIV vaccine candidates have multiple "antigens/immunogens" that are designed to elicit certain kinds of immune responses. In the first two efficacy trials, the HIV vaccine construct contained two envelope (env) gene antigens, based on two distinct strains of HIV, such that a 2-dimensional mark variable is of interest (Flynn et al., 2005; Pitisitthithum et al., 2006). The vaccine construct evaluated in the third and fourth efficacy trials contained gag, pol, and nef gene antigens, making a 3-dimensional mark variable of interest (Buchbinder et al., 2008; Gray et al., 2009). Lastly, the fifth and most recent efficacy trial tested a vaccine that contained gag, pol, and nef gene antigens, as well as three distinct env gene antigens, making a 6-dimensional mark variable of interest (Rerks-Ngarm et al., 2009).

The previous work dealt with the multivariate mark issue by collapsing the multiple distances into a univariate mark– the minimum of the distances to each vaccine antigen. This approach is reasonable under the belief that the only thing that matters for protection is the nearness of the exposing HIV to the closest antigen represented inside of the vaccine (e.g., Gilbert, McKeague and Sun, 2008). However, there are many ways in which this assumption may fail. For one example, based on host genetics (e.g., HLA type or Fc-γ-receptor type), one subgroup may be protected through immune responses that recognize HIV peptides that are similar to HIV peptides represented in antigen 1, whereas another subgroup may be protected through immune responses that are similar to HIV peptides represented in antigen 2; in this case the vaccine efficacy depends on both individual distances and less so on the minimum distance. For a second example, there are many ways to define a genetic distance in a putatively immunologically relevant way (several distances

were used in the sieve analysis of the Buchbinder et al., (2008), efficacy trial reported by Rolland et al., (2011), and if two distances are used such that the first considers many HIV sites irrelevant for protection whereas the second sagely restricts attention to key HIV sites that are contained in epitopes that cause protection, then the first distance could be shorter even though vaccine efficacy only depends on the second. Therefore, a more general approach to assessing and modeling how vaccine efficacy depends on multiple genetic distances is needed, that does not pre-assume a particular way to collapse the multivariate distances into a univariate distance. Outside of the survival analysis field, Gilbert (2000) studied such a general approach with multivariate marks, based on a semiparametric biased sampling model. However, this method is limited by the fact that the model conditions on infection, so that conditional odds ratios but not prospective relative risks of infection can be estimated, and by the fact that the model treats HIV infection as a binary outcome, not accounting for the time to HIV infection.

Sun, Gilbert and McKeague (2009) studied the mark-specific proportional hazard model

$$\lambda(t, v|z) = \lambda_0(t, v)e^{\beta^T(v)z(t)}, \tag{1.6}$$

for evaluating the mark-specific vaccine efficacy. Note under the proportional hazard model, the ratio of hazard function of any two individuals is a constant independent of time. This assumption may not always be met in practice, and can be relaxed through stratification (Lee, 2003). Here we propose to study the stratified mark-specific proportional hazard model to multivariate marks where the baseline hazard function can vary with stratum. For simplicity, we consider a two-dimensional mark variable $v = (v_1, v_2)$. The methods for general multivariate marks follow similar outlines. Note that the mark-specific relative risk function $\beta(v) = \beta(v_1, v_2)$ is an unspecified p-dimensional function depending on (v_1, v_2). Estimation of $\beta(v)$ without any structural assumptions is possible following the procedure for one-dimensional mark variable by Sun, Gilbert and McKeague (2009). However this would require a very large sample size due to curse of dimensionality. In this paper, we

propose estimation procedure under a parametric structure for $\beta(v)$. Consider the model

$$\beta(v) = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_{12} v_1 v_2, \tag{1.7}$$

which is the first order Taylor approximation of $\beta(v)$ plus an interaction term which will be used to investigate whether there is a confounding effect between the two marks. Here each of $\beta_0, \beta_1, \beta_2$, and β_{12} is a p-dimensional vector. Thus $\beta(v)$ is completely specified by the $4 \times p$ parameters denoted by $\bar{\beta} = (\beta_0^T, \beta_1^T, \beta_2^T, \beta_{12}^T)^T$. We develop an estimation procedure for the stratified mark-specific proportional hazard model under model (2.2) for $\beta(v)$ and testing procedures for the hypothesis relevant to the HIV vaccine efficacy evaluation. We consider the following null hypotheses to understand how relevant the two marks are for protection: $H_{10} : \beta_1 = \beta_2 = \beta_{12} = 0$; $H_{20} : \beta_{12} = 0$; $H_{30} : \beta_1 = \beta_{12} = 0$ and $H_{40} : \beta_2 = \beta_{12} = 0$. The null hypothesis H_{10} indicates that the relative risks does not depend on marks; H_{20} implies that the distances v_1 and v_2 do not have confounding effects on relative risks; The null hypothesis H_{30} implies that relative risks are not affected by v_1; While H_{40} implies that relative risks are not affected by v_2. Note that the β's given in the hypotheses are p-dimensional vectors, these tests can also be carried out for a single component corresponding to a given covariate to examine how the covariate effect is modified the marks or strain distances.

1.4 Overview

In Chapter 2, we propose the parametric estimator of $\beta(v)$ based on maximum partial likelihood method. The asymptotic properties of the proposed estimator are derived. The likelihood based hypotheses tests such as the likelihood ratio test, Wald test and score test are proposed to examine a variety of null hypotheses to understand how relevant the two marks are for protection against HIV infection. Extensive simulations are conducted to examine the finite sample performance of the proposed methods. An application to STEP data is discussed to show the usage of the methodology. We also proposed some former statistical tests to check goodness of fit of the multivariate mark-specific proportional hazard model based on the generalized weighted martingale residuals. The proposed

goodness-of-fit tests are justified and examined through simulations and also applied to STEP data.

CHAPTER 2: PROPORTIONAL HAZARDS MODEL WITH MULTIVARIATE CONTINUOUS MARKS

In this chapter, we develop the estimation and hypothesis testing procedures for the mark-specific proportional hazards model. We consider a parametric model for the regression coefficients $\beta(v)$. The asymptotic properties of the estimators are studied. The test statistics are constructed based on the asymptotic distributions of these estimators. We conduct a simulation study of the proposed estimation and hypothesis testing procedures. An application in the STEP data is used to illustrate the proposed methods.

2.1 Model and Data Description

Consider a two dimensional mark variable $v = (v_1, v_2)$ stratified multivariate mark-specific proportional hazard model

$$\lambda_k(t, v|z) = \lambda_{0k}(t, v)e^{\beta^T(v)z_k(t)}, \qquad k = 1, 2, \ldots, K, \tag{2.1}$$

where K is the number of the strata, $\lambda_{0k}(t, v)$ is an unspecified baseline mark-specific hazards function, $\beta(v)$ is p-dimensional regression function of v given by

$$\beta(v) = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_{12} v_1 v_2. \tag{2.2}$$

Here each of β_0, β_1, β_2, and β_{12} is a p-dimensional vector. Let $z = (z_1, z_2(t))^T$, where z_1 is the vaccine group indicator, $z_1 = 1$ for the vaccine group and $z_1 = 0$ for the placebo group, while $z_2(t)$ is other possibly time-dependent covariates. Under the mark-specific proportional hazards model (2.1), the mark-specific vaccine efficacy can be expressed as $VE(v) = 1 - \lambda(t, v|z_1 = 1)/\lambda(t, v|z_1 = 0) = 1 - exp(\beta_1(v))$, where $\beta_1(v)$ is the coefficient corresponding z_1. Let T_k be the failure time for an individual in the kth stratum, V_k be

the mark observation at failure and Z_k be the associated p-dimensional covariate. Assume that T_k, V_k, Z_k follow model (2.1). Under right censoring, the observed random variables are $(X_k, \delta_k, \delta_k V_k, Z_k)$, where $X_k = min(T_k, C_k)$, $\delta_k = I(T_k \leq C_k)$, and C_k is the censoring random variable, which is assumed to be independent of (T_k, V_k) given Z_k. The mark variable V_k is observed if the corresponding failure time uncensored. We consider statistical inferences for (2.1) and (2.2) with the observations $(X_{ki}, \delta_{ki}, \delta_{ki} V_{ki}, Z_{ki})$, $i = 1, \ldots, n_k$, which are independent identically distributed (iid) replicates of $(X_k, \delta_k, \delta_k V_k, Z_k)$, for $k = 1, 2, \ldots, K$. We assume that the end of follow-up time is τ, i.e., the failure time beyond τ are all censored.

2.2 Estimation

Under model (2.2), $\beta(v)$ is completely specified by the $4 \times p$ parameters denoted by $\bar{\beta} = (\beta_0^T, \beta_1^T, \beta_2^T, \beta_{12}^T)^T$. Let $\bar{v} = (1, v_1, v_2, v_1 v_2)^T$ and $\tilde{Z}_{ki}(t, v) = Z_{ki}(t) \otimes \bar{v}$ with $\otimes$ being the Kronecker product. Similar to Kalbfleisch and Prentice (1980) for competing risks model with finite number of causes, the log-partial likelihood function for $\bar{\beta}$ can be expressed as

$$\begin{aligned} l(\bar{\beta}) &= \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau \big[\beta^T(v) Z_{ki}(s) - \log\big\{\sum_{l=1}^{K} \sum_{j=1}^{n_k} Y_{lj}(s) \exp(\beta^T(v) Z_{lj}(s))\big\}\big] N_{ki}(ds, dv) \\ &= \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau \big[\bar{\beta}^T \tilde{Z}_{ki}(s, v) - \log\big\{\sum_{l=1}^{K} \sum_{j=1}^{n_k} Y_{lj}(s) \exp(\bar{\beta}^T \tilde{Z}_{lj}(s, v))\big\}\big] N_{ki}(ds, dv), \end{aligned} \tag{2.3}$$

where $N_{ki}(t, v) = I(X_{ki} \leq t, V_{ki} \leq v, \delta_{ki} = 1, k \in K)$ is the marked counting process of the ith individual in stratum k and $Y_{ki}(t) = I(X_{ki} \geq t)$ is the at risk process. For the multivariate mark variable V_{ki}, the relation $V_{ki} \leq v$ in $N_{ki}(t, v)$ means that the inequality holds for each component of the multivariate marks. Here we note that $\bar{\beta}^T \tilde{Z}_{ki}(s, v) = (\beta_0^T, \beta_1^T, \beta_2^T, \beta_{12}^T)(Z_{ki}^T(s), Z_{ki}^T(s) v_1, Z_{ki}^T(s) v_2, Z_{ki}^T(s) v_1 v_2)^T = \beta^T(v) Z_{ki}(s)$. The partial likelihood estimator $\hat{\bar{\beta}}$ for $\bar{\beta}$ is obtained by maximizing $l(\bar{\beta})$. Let

$$S_k^{(j)}(t, v, \bar{\beta}) = n_k^{-1} \sum_{i=1}^{n_k} Y_{ki}(t) exp\{\bar{\beta}^T \tilde{Z}_{ki}(t, v)\} \tilde{Z}_{ki}(t, v)^{\otimes j}. \tag{2.4}$$

Taking the derivative of $l(\bar{\beta})$ with respect to $\bar{\beta}$, the score function can be written as

$$U(\bar{\beta}) = \frac{\partial l(\bar{\beta})}{\partial \bar{\beta}} = \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau [\tilde{Z}_{ki}(t,v) - \frac{S_k^{(1)}(t,v,\bar{\beta})}{S_k^{(0)}(t,v,\bar{\beta})}] N_{ki}(dt,dv). \tag{2.5}$$

The information matrix is given by the derivative of $U(\bar{\beta})$ with respect to $\bar{\beta}$

$$I(\bar{\beta}) = \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau [\frac{S_k^{(2)}(t,v,\bar{\beta})}{S_k^{(0)}(t,v,\bar{\beta})} - (\frac{S_k^{(1)}(t,v,\bar{\beta})}{S_k^{(0)}(t,v,\bar{\beta})})^{\otimes 2}] N_{ki}(dt,dv). \tag{2.6}$$

The maximum partial likelihood estimator $\hat{\bar{\beta}}$ for $\bar{\beta}$ is a solution to $U(\bar{\beta}) = 0$. Under condition A, the matrix $I(\bar{\beta})$ is positive definite with probability 1 as $n_k \to \infty$ for $1 \leq k \leq K$. Thus $\hat{\bar{\beta}}$ exists almost surely for large sample sizes. The baseline function $\lambda_{0k}(t,v)$ can be estimated by smoothing the increments of the following estimator of the doubly cumulative baseline function $\Lambda_{0k}(t,v) = \int_0^t \int_0^v \lambda_{0k}(s,u) ds du$ as

$$\hat{\Lambda}_{0k}(t,v) = \int_0^t \int_0^v \frac{N_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,u,\hat{\bar{\beta}})}, \tag{2.7}$$

where $N_{k\cdot}(t,v) = \sum_{i=1}^{n_k} N_{ki}(t,v)$. A kernel estimator of $\lambda_{0k}(t,v)$ is given by $\hat{\lambda}_{0k}(t,v) = \int_0^\tau \int_0^1 K_{h_1}^{(1)}(t-s) K_{h_2}^{(1)}(v-u) \hat{\Lambda}_{0k}(ds,du)$, where $K_{h_1}^{(1)}(x) = K^{(1)}(x/h_1)/h_1$ and $K_{h_2}^{(2)}(x) = K^{(2)}(x/h_2)/h_2$ with $K^{(1)}(\cdot)$ and $K^{(2)}(\cdot)$ be the kernel functions and h_1 and h_2 the bandwidths.

2.3 Asymptotic Properties

We make use of the following regularity conditions.

Condition A

(A.1) The covariate process $Z_k(t)$ is left continuous with bounded variation and satisfies the moment condition $E[\|Z_k(t)\|^4 \exp(M\|Z_k(t)\|)] < \infty$, where $\|\cdot\|$ is the Euclidean norm and M is a constant such that $(\beta_0, \beta_1, \beta_2, \beta_{12}) \in (-M, M)^p$ for $1 \leq k \leq K$.

(A.2) For $j = 0, 1, 2$, let $s_k^{(j)}(t,v,\bar{\beta}) = E(S_k^{(j)}(t,v,\bar{\beta}))$. Let $\bar{\beta}_0$ be the true value of $\bar{\beta}$ under the model (2.1) and (2.2). $\lambda_{0k}(t,v)$ is continuous on $[0,\tau] \times [0,1]^2$. Each component

of $s_k^{(j)}(t,v,\bar{\beta})$ is continuous on $[0,\tau]\times[0,1]^2\times\mathcal{B}$ where $\mathcal{B}$ is an neighborhood of β. And $s_k^{(0)}(t,v,\bar{\beta})>0$ on $[0,\tau]\times[0,1]^2\times\mathcal{B}$ for $1\le k\le K$. $n_k/n\to p_k$ as $n_k\to\infty$. The matrix $\Sigma(\bar{\beta})=\sum_{k=1}^{K}\sum_{i=1}^{n_k}p_k\int_0^1\int_0^\tau[\frac{s_k^{(2)}(t,u,\bar{\beta})}{s_k^{(0)}(t,u,\bar{\beta})}-(\frac{s_k^{(1)}(t,u,\bar{\beta})}{s_k^{(0)}(t,u,\bar{\beta})})^{\otimes 2}]\ \lambda_{0k}(t,v)s_k^{(0)}(t,u,\bar{\beta})dtdu$ is positive definite.

The following theorems state the asymptotic consistency and asymptotic normality for $\hat{\bar{\beta}}$.

Theorem 2.1. *Under conditions (A.1)-(A.2), $\hat{\bar{\beta}}$ converges in probability to $\bar{\beta}_0$ as $n_k\to\infty$ for $1\le k\le K$.*

Theorem 2.2. *Under conditions (A.1)–(A.2), $n^{1/2}(\hat{\bar{\beta}}-\bar{\beta})\xrightarrow{\mathcal{D}}N(0,\Sigma^{-1}(\bar{\beta}_0))$ as $n_k\to\infty$ for $1\le k\le K$. $\Sigma(\bar{\beta}_0)$ can be consistently estimated by $n^{-1}I(\hat{\bar{\beta}})$.*

Following theorem 2.2, the large sample $100(1-\alpha)\%$ confidence interval for the jth component $\bar{\beta}_j$ of $\bar{\beta}$ is

$$\hat{\bar{\beta}}_j\pm n^{-\frac{1}{2}}z_{\alpha/2}\hat{\sigma}_j,$$

where $\hat{\sigma}_j$ is the jth diagonal element of $n^{-1}I(\hat{\bar{\beta}})$ and $Z_{\alpha/2}$ is the $1-\alpha/2$ quantile of a standard normal distribution.

2.4 Hypothesis Testing

In this section, we develop tests for testing the following hypothesis:

$H_{10}:\beta_1=\beta_2=\beta_{12}=0$;

$H_{20}:\beta_{12}=0$;

$H_{30}:\beta_2=\beta_{12}=0$;

and $H_{40}:\beta_1=\beta_{12}=0$. We propose three test statistics including the likehood ratio test, Wald test and score test. Note that the logarithm of partial likelihood function of β under (2.1) and (2.2) is $l(\bar{\beta})=\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau\big[\bar{\beta}^T\tilde{Z}_{ki}(s,v)-log\big\{\sum_{l=1}^{K}\sum_{j=1}^{n_k}Y_{lj}(s)*exp(\bar{\beta}^T\tilde{Z}_{lj}(s,v))\big\}\big]N_{ki}(ds,dv)$. Let $\hat{\bar{\beta}}$ be the maximum partial likelihood estimator maximizing $l(\bar{\beta})$. Let $\hat{\bar{\beta}}_{H_0}$ be the estimator of $\bar{\beta}$ under H_{10}, $\bar{\beta}_{H_0}$ is the maximizer of $l(\bar{\beta})$ under the restriction $\beta_1=\beta_2=\beta_{12}=0$ for testing H_{10}. The likelihood ratio test (LRT) statistic is

$$T_l=2\{l(\hat{\bar{\beta}})-l(\hat{\bar{\beta}}_{H_0})\}.$$

Routine analysis following Serfling (1980) shows that under H_{10}, T_l converges in distribution to a chi-square distribution with $3p$ degrees of freedom.

The Wald test is given by

$$T_w = (\hat{\bar{\beta}} - \hat{\bar{\beta}}_{H_0})^T [I(\hat{\bar{\beta}})](\hat{\bar{\beta}} - \hat{\bar{\beta}}_{H_0}),$$

where information matrix $I(\hat{\bar{\beta}})$ is defined in (2.6).

The score test statistic is given by

$$T_s = U^T(\bar{\beta}_{H_0}) I(\hat{\bar{\beta}}_{H_0})^{-1} U(\bar{\beta}_{H_0}).$$

where the score function $U(\bar{\beta}_{H_0})$ and information matrix $I(\hat{\bar{\beta}})$ is defined in (2.5) and (2.6) respectively. Under H_{10}, both T_w and T_s converge in distribution to a chi-square distribution with $3p$ degrees of freedom. Under H_{20}, T_l, T_w, and T_s converge in distribution to a chi-square distribution with p degrees of freedom. The asymptotic distributions of these test statistics under H_{30} and H_{40} is chi-square with $2p$ degrees of freedom. The likelihood ratio test rejects the null hypothesis H_{10} if $T_l > \chi^2_{3p,\alpha}$, the upper α quantile of the chi-square distribution with 3p degrees of freedom. The corresponding critical values for testing H_{20}, H_{30}, and H_{40} are $\chi^2_{p,\alpha}$, $\chi^2_{2p,\alpha}$ and $\chi^2_{2p,\alpha}$, respectively. Similar decision rules hold for the Wald test (T_w) and score test (T_s).

2.5 Goodness-of-fit Tests

The estimation and testing procedures developed in section 2.3 and 2.4 are developed under model (2.1) and (2.2). The validity of these prcedures depends on goodness of fit of the multivariate mark-specific proportional hazard model. This section develops some goodness of fit tests of model (2.1) and (2.2). Similar to Lin, Wei and Ying (1993) and Spiekerman and Lin (1996), we derive the model checking test statistics based on the

martingale residuals is defined as

$$\hat{M}_{ki}(t,v) = \int_0^t \int_a^v N_{ki}(ds,du) - Y_{ki}(s)\exp(\hat{\tilde{\beta}}^T \tilde{Z}_{ki}(s,u))\hat{\Lambda}_{0k}(ds,du). \tag{2.8}$$

Where $\hat{\tilde{\beta}}$ is the maximum partial likelihood estimator given in section 2.2, and $\hat{\Lambda}_{0k}(t,k)$ is defined in (2.7). $\hat{M}_{ki}(t,v)$ may be interpreted as the difference at time t between the observed and the predicted number of events with marks less than v for the ith subject in kth stratum. Thus the martingale residuals are informative about the model misspecification. It is easy to check that

$$\begin{aligned}
\sum_{k=1}^{K}\sum_{i=1}^{n_k} \hat{M}_{ki}(t,v) &= \sum_{k=1}^{K}\sum_{i=1}^{n_k} \int_0^t \int_a^v \left[N_{ki}(ds,du) - Y_{ki}(s)\exp(\hat{\tilde{\beta}}^T \tilde{Z}_{ki}(s,u)) \frac{N_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,u,\hat{\tilde{\beta}})} \right] \\
&= \sum_{k=1}^{K}\sum_{i=1}^{n_k} N_{ki}(t,v) - \sum_{k=1}^{K}\sum_{i=1}^{n_k} \int_0^t \int_a^v Y_{ki}(s)\exp(\hat{\tilde{\beta}}^T \tilde{Z}_{ki}(s,u)) \times \\
&\quad \frac{N_{k\cdot}(ds,du)}{\sum_{j=1}^{n_k} Y_{kj}(s) exp(\hat{\tilde{\beta}}^T \tilde{Z}_{kj}(s,u))} \\
&= 0.
\end{aligned}$$

For $1 \le k \le K$, let

$$W_k(t,v,z) = n^{-1/2} \sum_{i=1}^{n_k} g_k(Z_{ki},z)\hat{M}_{ki}(t,v), \tag{2.9}$$

here $g_k(Z_{ki},z)$ is a $1 \times r$-vector of known bounded functions of Z_{ki} and z. For example, one may take $g_k(Z_{ki},z) = f_k(Z_{ki})I(Z_{ki} \le z)$, where $f_k(\cdot)$ is a known function and $I(Z_{ki} \le z) = (I(Z_{1ki} \le z_1), \ldots, I(Z_{pki} \le z_p))$, in which case $r = p$. We construct goodness of fit test statistics based on the test process $W(t,v,z) = (W_1(t,v,z), \ldots, W_K(t,v,z))$. If model (2.1) and (2.2) hold, the process $W(t,v,z)$ fluctuates randomly about zero. Various test statistics can be constructed by selecting different weight functions $f_k(\cdot)$ and using different functionals of the process $W(t,v,z)$. Here we propose the supremum test statistics to test the overall fit of the model:

$$S = \sup_{1 \le k \le K} \sup_{t,v,z} |W_k(t,v,z)|. \tag{2.10}$$

The distribution of $W(t,v,z)$ and S can be approximated by using the Guassian multiplier method (Lin, Wei and Ying, 1993) as we shall drive next.

Let $S_{kg}^{(0)}(t,v,z,\bar{\beta}) = n_k^{-1}\sum_{i=1}^{n_k} Y_{ki}(t)\exp\{\bar{\beta}^T\tilde{Z}_{ki}(t,v)\}g_k(Z_{ki},z)$ and $S_{kg}^{(1)}(t,v,z,\bar{\beta}) = n_k^{-1}\sum_{i=1}^{n_k} Y_{ki}(t)\exp\{\bar{\beta}^T\tilde{Z}_{ki}(t,v)\}\tilde{Z}_{ki}(t,v)\otimes g_k(Z_{ki},z)$, where $A\otimes B$ is the Kronecker product of matrices A and B. Let $s_{kg}^{(0)}(t,v,z,\bar{\beta}) = E(S_{kg}^{(0)}(t,v,z,\bar{\beta}))$ and $s_{kg}^{(1)}(t,v,z,\bar{\beta}) = E(S_{kg}^{(1)}(t,v,z,\bar{\beta}))$. From the proof given in Appendix I, we have the following decomposition.

Theorem 2.3.

$$
\begin{aligned}
&W_k(t,v,z)\\
&= n^{-1/2}\sum_{l=1}^{K}\sum_{i=1}^{n_l}\int_0^\tau\int_0^v I(l=k)I(s\le t)\left[g_l(Z_{li},z) - \frac{s_{lg}^{(0)}(s,u,z,\bar{\beta})}{s_l^{(0)}(s,u,\bar{\beta})}\right]M_{li}(ds,du)\\
&+n^{-1/2}(R_k(t,v,z))^T(\Sigma(\bar{\beta}))^{-1}\sum_{l=1}^{K}\sum_{i=1}^{n_l}\left\{\int_0^\tau\int_0^1\left[\tilde{Z}_{li}(s,u) - \frac{s_l^{(1)}(s,u,\bar{\beta})}{s_l^{(0)}(s,u,^-beta)}\right]M_{li}(ds,du)\right\}^T\\
&+o_p(1).
\end{aligned}
\tag{2.11}
$$

where

$$
R_k(t,v,z) = (n_k/n)\int_0^t\int_a^v\left(\frac{s_k^{(1)}(s,x,\bar{\beta})\otimes s_{kg}^{(0)}(s,x,z,\bar{\beta})}{s_k^{(0)}(s,x,\bar{\beta})} - s_{kg}^{(1)}(s,x,z,\bar{\beta})\right)\lambda_{0k}(s,x)\,dsdx.
$$

The expression (2.11) shows that the process $W_k(t,v,z)$ is asymptotically equivalent to the sum of iid terms involving the integrations with respect to $M_{li}(s,u)$. The empirical process theorem (Donsker theorem) can be used to show that the process $W(t,v,z) = (W_1(t,v,z),\ldots,W_K(t,v,z))$ converges weakly to a multi-dimensional Gaussian random field. Let $\{\xi_{li}, i=1,\ldots,n_l, l=1,\ldots,K\}$ be iid standard normal random variables. Using the Gaussian multiplier technique of Lin, Wei and Ying (1993), the distribution of $W(t,v,z)$ can be approximated by the distribution of $W^*(t,v,z) = (W_1^*(t,v,z),\ldots, W_K^*(t,v,z))$,

where for $1 \leq k \leq K$,

$$
\begin{aligned}
& W_k^*(t,v,z) \\
& = n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \int_0^\tau \int_0^v I(l=k) I(s \leq t) \left[g_l(Z_{li}, z) - \frac{S_{lg}^{(0)}(s,u,z,\hat{\bar{\beta}})}{S_l^{(0)}(s,u,z,\hat{\bar{\beta}})} \right] \xi_{li} \, N_{li}(ds, du) \\
& + n^{-1/2} (\hat{R}_k(t,v,z))^T (\Sigma(\bar{\beta}))^{-1} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \left\{ \int_0^\tau \int_0^1 \left[\tilde{Z}_{li}(s,u) - \frac{S_l^{(1)}(s,u,\bar{\beta})}{S_l^{(0)}(s,u,\bar{\beta})} \right] \xi_{li} N_{li}(ds, du) \right\}^T,
\end{aligned}
\tag{2.12}
$$

where

$$
\hat{R}_k(t,v,z) = (n_k/n) \int_0^t \int_0^v \left(\frac{S_k^{(1)}(s,x,\hat{\bar{\beta}}) \otimes S_{kg}^{(0)}(s,x,z,\hat{\bar{\beta}})}{S_k^{(0)}(s,x,\hat{\bar{\beta}})} - S_{kg}^{(1)}(s,x,z,\hat{\bar{\beta}}) \right) d\hat{\Lambda}_{0k}(s,x).
$$

Let

$$
S_0^* = \sup_{1 \leq k \leq K} \sup_{t,v,z} |W_k^*(t,v,z)|. \tag{2.13}
$$

The distribution of S can be approximated by the empirical distribution of S^* obtained repeatedly generating a large number of sets of iid random variables ξ_{li} while holding the observed data fixed. We reject the model (2.1) and (2.2) at significance level α if S is greater than the upper α quantile of S^*.

2.6 Simulation Study

2.6.1 Estimation and Hypothesis Testing Procedures

In this section, we use simulation to check the proposed estimation and hypothesis test under model (2.1) and (2.2). The simulations are set up to mimic the STEP data. The STEP data includes 1900 randomized men. It is a stratified random sample, about 1100 in the "$Ad5 \leq 200$" stratum and about 800 in the "$Ad5 > 200$" stratum. The strata refer to whether a volunteer has pre-existing immunity to the Adenovirus serotype 5 vector that was used in the vaccine; 200 is the titer of neutralization (≤ 200 is low, > 200 is high). The study enrolled people to the $Ad5 \leq 200$ stratum first, and later amended the study to include an additional enrolment of $Ad5 > 200$. For $Ad5 \leq 200$ there were 52 total

infections, pretty even in vaccine and placebo (28 vaccine; 24 placebo; annual incidences 4.2%, 3.5%). For $Ad5 > 200$ there were 30 total infections, 21 in vaccine and 9 in placebo (annual incidence 5.4% and 2.3%).

We consider two strata $K = 2$ corresponding to two levels of Ad5 strata neutralization. Set the baseline hazard $\lambda_{0k}(t, v)$ to be constants for $k = 1$ and 2, with $\lambda_{01} = 0.4$ and $\lambda_{02} = 0.6$. The simulations are constructed for $n_k = 250$, 400 and 500 in each stratum. We set the covariate z as Bernoulli random variable with 0.5 probability of success where $z = 1$ corresponds to the vaccine group and $z = 0$ for the placebo group. We generate the censoring times from an exponential distribution, independent of (T, V), and the follow-up time is $\tau = 2.0$. The following selections of $\beta(v)$ specified in terms of $(\beta_0, \beta_1, \beta_2, \beta_{12})$ are chosen to examine the proposed estimation and hypothesis testing procedures.

$M_{10} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, 0., 0., 0.);$

$M_{11} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, .9, 0., 0.);$

$M_{12} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, .9, .8, 0.);$

$M_{13} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, .9, .8, .6).$

$M_{20} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-3.5, .3, .1, 0.);$

$M_{21} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-3.5, .3, .1, 5.0);$

$M_{22} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-3.5, .3, .1, 5.5);$

$M_{23} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-3.5, .3, .1, 6.0).$

$M_{30} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, 1.2, 0., 0.);$

$M_{31} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, 0.8, 1.0, 0.);$

$M_{32} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, 0.5, 1.2, 0.);$

$M_{33} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, 0.6, 0.8, 1.0);$

$M_{34} : (\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, 0.4, 0.9, 1.2).$

Here M_{10} is a model under the null hypothesis H_{10} that no β except β_0 is nonzero. Under M_{11}, the vaccine efficacy depends on only the first mark. Under M_{12}, the vaccine efficacy depends on both marks but not the intersection of these marks. Under M_{13}, the vaccine efficacy depends on both marks and the intersection of these marks. M_{20} is a model under the null hypothesis H_{20} that all βs except β_{12} is nonzero. And M_{30} is a

model under the null hypothesis H_{30} that no beta except β_1 is nonzero. The average vaccine efficacy defined as $AVE = \int_0^1 \int_0^1 (1 - e^{\beta(v)}) dv_1 dv_2$ is around 50%. Table 2.1, 2.2 and 2.3 list the bias and the standard deviation of the estimators (SSE), the average of the estimated standard deviation (ESE) and the coverage probability (CP) for β for total sample size $n = n_1 + n_2 = 500$, 800, and 1000. Those tables show that the biases of the estimators are small, ESE approximates SSE well and the coverage probability are very close to their nominal level of 0.95. Table 2.4, 2.5 and 2.6 summarize the empirical sizes and powers of the likelihood ratio test, Wald test and the score test for testing H_{10}, H_{20} and H_{30} at the significance level $\alpha = 0.05$. Those tables show that the empirical sizes of all these tests are close to 0.05. The likelihood ratio test has better power than Wald and score test. The departure from H_{10} increases as the simulation model moves from M_{11} to M_{13}, corresponding power also increases. Similarly the power for testing H_{20} increases in the direction M_{21} to M_{23}. And the power for testing H_{30} also increases in the direction M_{31} to M_{34}. And naturally, the power increases as the sample size increase. The coverage probabilities for β_0, β_1, β_2 and β_{12} are also listed to demonstrate that the proposed maximum partial likelihood methods work very well.

2.6.2 Goodness-of-fit Procedure

In this section, we conduct a simulation study to check the finite sample performance of the proposed testing procedure. The size of the test is examined using the following simple mark-specific proportional hazards model:

$$M_{40} : \lambda_k(t, v|z) = \lambda_{0k} e^{(\beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_{12} v_1 v_2)^T z_k}, \quad k = 1, 2, \tag{2.14}$$

where $\lambda_{01} = 0.4$, $\lambda_{02} = 0.6$ and $(\beta_0, \beta_1, \beta_2, \beta_{12}) = (-1.65, .4, .9, 1.2)$. z_{ki} takes value 0 or 1 as a treatment indicator. We generate covariates Z_{ki} from Bernoulli distribution with $P(Z_{ki} = 1) = 0.5$. The censoring times are generated from an exponential distribution, independent of (T_{ki}, V_{ki}).

Table 2.11 shows the empirical sizes of the test for total sample sizes of $n = n_1 + n_2 =$ 100, 200 and 300. The empirical sizes are calculated based on 1000 simulations and 500

Gaussian multiplier samples. And the emperical size of all those tests are shown to be very close to 0.05.

To evaluate the power of proposed test, consider the model

$$\lambda_k(t, v|z) = \lambda_{0k}(t, v)e^{(-at^b+0.4v_1+0.6v_2)z}, \quad k = 1, 2, \tag{2.15}$$

where $\lambda_{0k} = abkt^{(b-1)}e^{at^b-0.4v_1-0.6v_2}$ for $k = 1$ and 2. Again we take Z_{ki} as a Bernoulli random variable with $P(Z_{ki} = 1) = 0.5$. Model (2.15) is not a mark-specific proportional hazard model since the hazard ratio $\lambda_k(t, v|Z_{ki} = 1)/\lambda_k(t, v|Z_{ki} = 0) = e^{at^b-0.4v_1-0.6v_2}$ changes with time. We consider the following choices of (a,b):

$(M_{41}) : (a, b) = (0.20, 0.30)$;

$(M_{42}) : (a, b) = (0.25, 0.40)$;

$(M_{43}) : (a, b) = (0.30, 0.50)$. As a and b increases, the hazard ratio under (2.15) increases faster with t, which represents an increases departure from the null hypothesis. For each of the above selected (a, b), random right censoring times are generated from an exponential distribution, independent of (T_{ki}, V_{ki}). sample sizes of $n = n_1 + n_2 = 100$, $n = 200$ and 300 are studied. The empirical power of the test at the significance level 0.05 under (2.15) for different choice of (a, b) are given in Table 2.11. Each entry of the table is based on 1000 simulations and 500 Gaussian multiplier samples. The power of the test increases with the increasing of (a, b), and also increases with sample size. The limited simulation study demonstrates the validity of the proposed goodness of fit testing procedure. The test provides a valuable tool to check the adequacy of the stratified mark-specific proportional hazard model (2.1).

2.7 Application to the STEP Data

2.7.1 Estimation Procedure

We now illustrate our model with an analysis of the STEP data. The 'Step' trial randomized 1836 HIV negative men to receive either the Merck Adenovirus 5 vaccine (MRKAd5) or placebo, and was conducted in North and South America (Buchbinder et

al., 2008). Women were also enrolled, but because only 1 became HIV infected, essentially all of the information about vaccine efficacy is restricted to men. Of the 1836 men 88 acquired HIV infection, of which 66 had between 5 and 14 HIV sequences measured. The 22 men with no sequence data are excluded from the analysis. The randomization was stratified by whether a volunteer has pre-existing immunity to the Adenovirus serotype 5 vector that was used in the vaccine, defined by an Ad5 neutralization titer below versus above 200 (sample sizes 1058 and 778, respectively), such that the method is implemented accounting for these strata. one. In this data set, we focus on men only, because only one woman got infected.

We define the mark as the HIV genetic distance for a subject whose HIV infection was diagnosed in the acute phase; diagnosis in the acute phase means that the HIV-specific PCR test is positive but the ELISA antibody test is negative. About 50% of the infections in STEP were diagnosed in this way. The genetic distance is defined as the percent mismatch of amino acids (infecting strain compared to the strain inside the vaccine). Those amino acids are the HIV Gag protein that are in cytotoxic T lymphocyte epitopes and are recognized by at least 10% of vaccine recipients at the Week 8 visit.

For $Ad5 \leq 200$ there were 54 total infections, 29 infections in the vaccine group (7 of which with missing marks) and 25 infections in the placebo group (with 8 missing marks). The annual incidences were 5.3% for the vaccine and 4.7% for the placebo. For $Ad5 > 200$ there were 33 total infections (7 of them with missing marks), 24 infections in the vaccine (with 7 missing marks) and 9 infections in placebo. The annual incidences were 6.1% for the vaccine and 2.3% for the placebo.

Two factors motivate the need to consider multiple genetic distances for the sieve analysis of the Step data. First, the MRKAd5 vaccine contained three HIV genes: Gag, Pol, and Nef. As a control it is also of interest to consider genetic distances to all of the HIV vaccine genes combined not included in the vaccine, env-rev-tat-vif-vpr-vpu, for which the HXB2 reference strain is used. Second, as described in greater detail in the clinical paper (Rolland et al., 2011), two different bioinformatics methods, NetMHC (Buus, 2003), and Epipred (Heckerman, 2007) were used to predict for each infected subject, based on their

HLA alleles, the set of HIV peptides that could potentially constitute T cell epitopes and hence could potentially cause vaccine-protection. NetMHC predicts binding of peptides to 4-digit HLA alleles; the software discriminates on the basis of quantitative peptide MHC binding data and discerns strong and weak binders. In contrast, Epipred identifies known and potential HIV-1 CTL epitope motifs using 2-digit HLA information. In addition to the known HLA-restricted epitopes previously reported at the Los Alamos National Laboratory HIV database (HIVDB), we accepted all epitope motifs with a posterior probability of $>$ 0.8. HLA-specific epitopes were predicted in the MRKAd5 HIV-1 Gag-Pol-Nef vaccine sequences and in all proteins from HXB2 (available at the HIVDB, http:www.hiv.lanl.gov). Using Epipred, the first step in computing the distance between a subject's sequences and the reference sequence is to compute the nonparametric maximum likelihood estimate (NPMLE) of the number of peptides shared between the reference sequence and the subject's sequences, defined as the sum of estimated epitope-probabilities across all 8, 9, 10, 11-mers in the reference sequence that are exactly matched in all of the subject's sequences. Then, the distance is the NPMLE of the percent of peptides mismatched in at least one of the subject's sequences, defined as one minus the ratio of the NPMLE of the number of shared peptides (computed in the first step) and the NPMLE of the number of peptides in the reference sequence. Because the NetMHC software returns results of non-binder, weak binder, or strong binder, we defined the distance as the estimated percent of epitopes mismatched in at least one of a subject's sequences, the latter defined as the number of weak or strong binding 8, 9, 10, or 11-mers in the reference sequence that mismatch the corresponding peptide in at least one of the subject's sequences.

We consider distances defined using the reference HIV regions Gag, Pol, Nef, Gag-Pol-Nef, env-rev-tat-vif-vpr-vpu, as well as using the two bioinformatics prediction methods. In the following table, we choose different combinations of reference region and bioinformatics method as bivariate marks (v_1, v_2), where (except for the last control case) we always include at least one Gag distance, given the hypothesis that this gene is most important for protection. We analyze the data with the mark specific proportional hazard model (2.1) where the covariate z is the indicator for the treatment, with $z = 1$ for the vaccine group

and 0 for placebo group, and $\beta(v)$ takes the form (2.2).

Model	mark	type	ref	gene
S_1	v_1	Epipred	Step	Gag
	v_2	Epipred	Step	Pol
S_2	v_1	Epipred	Step	Gag
	v_2	Epipred	Step	Nef
S_3	v_1	Epipred	Step	Gag
	v_2	Epipred	HXB2	env-rev-tat-vif-vpr-vpu
S_4	v_1	netMHC	Step	Gag
	v_2	netMHC	Step	Pol
S_5	v_1	netMHC	Step	Gag
	v_2	netMHC	Step	Nef
S_6	v_1	netMHC	Step	Gag
	v_2	netMHC	HXB2	env-rev-tat-vif-vpr-vpu
S_7	v_1	Epipred	Step	Gag
	v_2	netMHC	Step	Gag
S_8	v_1	Epipred	HXB2	env-rev-tat-vif-vpr-vpu
	v_2	netMHC	HXB2	env-rev-tat-vif-vpr-vpu

In the tables 2.7, 2.8, we compare the result using original and standardized marks. Here we standardize mark by subtract the minimum one in that group, then divide the range (which equals the maximum minus minimum mark). Since the support of the original marks (v_1, v_2) is not $(0, 1)$, we modify the average vaccine efficacy (AVE) as

$$AVE^* = \frac{\int_{V_1^m}^{V_1^M} \int_{V_2^m}^{V_2^M} (1 - e^{\beta}(v)) dv_1 dv_2}{(V_1^M - V_1^m)(V_2^M - V_2^m)},$$

here V_i^m and V_i^M are the minimum and maximum of mark V_i, $i = 1, 2$. Table 2.7 shows that most AVE^* are negative, that because there are more failures in vaccine group than in placebo group. And the large negative AVEs for S4-S7 are not valid, because from the

figures 2.1 to 2.8, we can see the failure time's range for vaccine and placebo group are quite different. In that case the model (2.1) and (2.2) are not appropriate for these data. In addition, we use likelihood ratio test, Wald test and score test to test $H_{10} : \beta_1 = \beta_2 = \beta_{12} = 0$, which indicates that the relative risks does not depend on marks; $H_{20} : \beta_{12} = 0$, which means the distances v_1 and v_2 do not have confounding effects on relative risks; $H_{30} : \beta_2 = \beta_{12} = 0$, implies the relative risks are not affected by v_1. Table 2.9 and 2.10 show the p-value of likelihood ratio test, Wald test and score test for H_{10}, H_{20} and H_{30}. From the goodness-of-fit we checked later, only S2 and S8 fit the proposed model. And under the selection of S8, we may reject H_{10} and H_{30} under nominal level 0.05, and can also reject H_{20} under nominal level 0.10.

2.7.2 Goodness-of-fit Procedure

Next we check the goodness of fit of model (2.1) and (2.2) for the real date STEP based on the previous analysis which are carried out, the test is conducted at the significance level $\alpha = 0.05$, and by setting $f_k(z) = 1$. Following the procedure given in section 2.5, we check if $S = sup_{1\leq k\leq K} sup_{t,v,z}|W_k(t, v, z)|$ falls below 0.95 quantile of $S^* = sup_{1\leq k\leq K} sup_{t,v,z}|W_k^*(t, v, z)|$. The asymptotic distribution of S^* can be approximated by repeatedly generating sets of independent standard normal random variables while holding the observed data fixed. We generate 500 copies of S^* and calculate the p-value as the percentage of values of S^* greater than S. Table 2.12 list the p-value of the test statistic S for the STEP data for different combinations of the marks. It shows that except S2 and S8, other selections of marks do not fit our proposed model. For S4, S5, S6 and S7, this cause from the fact that the range of V1 being too short for the placebo compared to the vaccine group. In the future, it is worth to look at the analysis for S4-S7 based on a subset where V_1 has the same range for both the vaccine and placebo and V_2 has the same range for both the vaccine and placebo.

2.8 Complements

In this section, we give the derivations of the main results presented in previous sections of this chapter. First we present the following result that for proving the asymptotic normality of $\hat{\tilde{\beta}}(v)$ and provides important insight into the constructions of the confidence bands and test statistics that follow. Let

$$\tilde{W}_A(v) = n^{-1/2} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^v \int_0^\tau A(u) \left[\tilde{Z}_{ki}(s,u) - \frac{s_k^{(1)}(s,u,\bar{\beta})}{s_k^{(0)}(s,u,\bar{\beta})} \right] M_{ki}(ds,du), \tag{2.16}$$

where $A(u)$ is a deterministic $p \times p$ matrix with bounded components.

Lemma 2.1. *Assume that each component of the $p \times p$ matrix $A(v), v \in [a,b]$, is continuous. Under conditions (A.1)–(A.2), $\tilde{W}_A(v)$ converges weakly to a p-dimensional mean-zero Gaussian martingale, $W_A(v)$, with continuous sample paths on $v \in [a,b]$. The covariance matrix of $W_A(v)$ is given by $\mathrm{Cov}(W_A(v)) = \int_a^v A(u)\Sigma(u)A(u)\,du$. The estimator of $\Sigma(u)$ is given by*

$$\hat{\Sigma}_{\hat{A}}(v) = n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_a^v \int_0^\tau \hat{A}(u) J_{kn}(t, \hat{\beta}(u)) \hat{A}^T(u)\, N_{ki}(dt,du), \tag{2.17}$$

where $\hat{A}(v)$ is a consistent estimator of $A(v)$ uniformly in $v \in [a,b] \subset [0,1]$. It can be shown that $\hat{\Sigma}_A(v)$ is a consistent estimator of $\mathrm{Cov}(W_A(v))$.

Proof of Lemma 2.1.

It is easy to check that the conditions of Lemma 1 of Sun and Wu (2005) are satisfied under Condition A. It follows that $\tilde{W}_A(v)$ converges weakly to a vector of continuous mean-zero Gaussian random processes, $W_A(v)$, $v \in [0,1]^2$. Now we show that $W_A(v)$ has independent increments. Let $w_{ki}(v) = \int_a^v \int_0^\tau A(u)\big[\tilde{Z}_{ki}(t,u) - s_k^{(1)}(t,u,\bar{\beta})/s_k^{(0)}(t,u,\bar{\beta})\big] M_{ki}(dt,du)$. Then $\tilde{W}_A(v) = n^{-1/2}\sum_{i=1}^n w_{ki}(v)$. For $0 \le v_1 \le v_2 \le 1$, the covariance matrix of $W_A(v_1)$ and $W_A(v_2) - W_A(v_1)$ is equal to $E\{w_{ki}(v_1)(w_{ki}(v_2) - w_{ki}(v_1))^T\}$. Since $M_{ki}(t,v_1)$ and $M_{ki}(t,v_2) - M_{ki}(t,v_1)$, $0 \le t \le \tau$, are orthogonal square integrable martingales, it follows

that $E\{w_{ki}(v_1)(w_{ki}(v_2)-w_{ki}(v_1))^T\}=0$. Hence $W_A(v)$, $v\in[0,1]$, is a vector of mean-zero Gaussian random processes with independent increments.

Further, the covariance matrix of $W_A(v)$ is equal to

$$\begin{aligned}
& E\{w_i(v)(w_i(v))^T\} \\
= \; & E\Big\{\int_0^v\int_0^\tau A(u)\Big[Z_{ki}(t)-\frac{s_k^{(1)}(t,u,\bar\beta)}{s_k^{(0)}(t,u,\bar\beta)}\Big]^{\otimes 2}A(u)\,N_{ki}(dt,du)\Big\} \\
= \; & E\Big\{\int_0^v\int_0^\tau A(u)\Big[Z_{ki}(t)-\frac{s_k^{(1)}(t,u,\bar\beta)}{s_k^{(0)}(t,u,\bar\beta)}\Big]^{\otimes 2}A(u)Y_{ki}(t)exp(\bar\beta^T\tilde Z_{ki}(t))\lambda_{0k}(t,u) \\
= \; & \int_0^v A(u)\Sigma(u)A(u)\,du.
\end{aligned}$$

This completes the proof of Lemma 2.1. *Q.E.D.*

Lemma 2.2. *Under conditions (A.1)-(A.2), $S_k^{(j)}(t,v,\bar\beta)$ converge to $s_k^{(j)}(t,v,\bar\beta)$ in probability uniformly in $(t,v)\in[0,\tau]\times[0,1]^2\times\ \mathcal{B}^{4p}$ as $n_k\to\infty$, for $j=0,1,2$ and $1\le k\le K$.*

Proof of Lemma 2.2.

Let $S_{ki}^{(j)}(t,v,\bar\beta)=Y_{ki}(t)exp\{\bar\beta^T\tilde Z_{ki}(t,v)\}\tilde Z_{ki}(t,v)^{\otimes j}$, then $S_k^{(j)}=n_k^{-1}\sum_{i=1}^{n_k}S_{ki}^{(j)}(t,v,\bar\beta)$. We show the lemma for $j=0$. The proofs for $j=1$ and 2 follow similarly. Let $\omega_{ki}=(X_{ki},Z_{ki})$, and $Z_{ki}\in[-B,B]^p$ for some $B>0$. ω_{ki}, $i=1,\ldots,n_k$, is a random sample from a probability distribution P_k on a measurable space $(\mathcal{X}_k,\mathcal{A}_k)$, where $\mathcal{X}_k=[0,\tau]\times[-B,B]^p$ and $\mathcal{A}_k$ is its Borel σ-field. Let $\mathcal{F}$ be the class of all coordinate projections $f_{t,v,\bar\beta}(\omega_{ki}):\mathcal{X}_k\longrightarrow R$, where $f_{t,v,\bar\beta}=S_{ki}(t,v,\bar\beta)$, for $(t,v,\bar\beta)\in[0,\tau]\times[0,1]^2\times\mathcal{B}$. Then $S_k(t,v,\bar\beta)=n_k^{-1}\sum_{i=1}^{n_k}f_{t,v,\bar\beta}(\omega_{ki})$. Let $\|f_{t,v,\bar\beta}\|_{P,r}=(P_k|f_{t,v,\bar\beta}|^r)^{1/r}=(E|S_{ki}(t,v,\bar\beta)|^r)^{1/r}$ be $L_r(P_k)$-norm of $f_{t,v,\bar\beta}$. Next, we show that $\mathcal{F}$ is Glivenko-Cantelli. Since $Z_{ki}(\cdot)$ is of bounded variation, for simplicity we assume that $Z_{ki}(\cdot)$ is an nonnegative monotone increasing process. In general, $Z_{ki}(\cdot)$ can be expressed as the difference of two nonnegative monotone increasing processes plus a constant. In this case, the class of functions of interest, $\mathcal{F}$ is the product of several Glivenko-Cantelli (Donsker) classes, therefore.

Let $\{t_h\}$, $\{v_j\}$ and $\{\bar\beta_m\}$ be the grid points of finite partitions of the intervals $[0,\tau]$, $[0,1]^2$, and $\mathcal{B}$, respectively. Let $\{t_{h'},t_h\}$, $\{v_{j'},v_j\}$ and $\{\bar\beta_{m'},\bar\beta_m\}$ be the grid points on the

opposite ends of a hyper-cubic of the partitions such that $0 \le t_h - t_{h'} \le \epsilon$, $0 \le v_j - v_{j'} \le \epsilon$ and $0 \le \bar{\beta}_m - \bar{\beta}_{m'} \le \epsilon$ for $\epsilon > 0$. Define the bracketing functions $l_{h'j'm'} = S_i^{(0)}(t_{h'}, v_{j'}, \bar{\beta}_{m'})$ and $u_{hjm} = S_i^{(0)}(t_h, v_j, \bar{\beta}_m)$. Then for any $f_{t,v,\bar{\beta}} \in \mathcal{F}$, there is a bracket $[l_{h'j'm'}, u_{hjm}]$ such that $f_{t,v,\bar{\beta}} \in [l_{h'j'm'}, u_{hjm}]$.

$$\begin{aligned}
\|u_{hjm} - l_{h'j'm'}\|_{P,2} &\le \|S_{ki}^{(0)}(t_h, v_j, \bar{\beta}_m) - S_{ki}^{(0)}(t_{h'}, v_{j'}, \bar{\beta}_{m'})\|_{P,2} \\
&= \|Y_{ki}(t_h) exp\{\bar{\beta}_m^T \tilde{Z}_{ki}(t_h, v_j)\} \\
&\quad - Y_{ki}(t_{h'}) exp\{\bar{\beta}_{m'}^T \tilde{Z}_{ki}(t_{h'}, v_{j'})\}\|_{p,2} \\
&\le [C_1\|t_h - t_{h'}\| + C_2\|v_j - v_{j'}\| + C_3\|\bar{\beta}_m - \bar{\beta}_{m'}\|)]^{1/2} \\
&\le C\epsilon^{1/2},
\end{aligned}$$

where C_1, C_2, C_3 and C are some positive constants. Hence, the bracketing number $N_{[]}(\epsilon^{1/2}, \mathcal{F}, L_2(P_k))$ is of the polynomial order $(1/\epsilon)^{p+3}$. Thus $N_{[]}(\epsilon, \mathcal{F}, L_2(P_k))$ is of the order $(1/\epsilon)^{2(p+3)}$. By the Glivenko–Cantelli Theorem (Theorem 19.4 of van der Vaart), $S_k^{(j)}(t, v, \bar{\beta})$ converges in probabitlity uniformly to $s_k^{(j)}(t, v, \bar{\beta})$ for $(t, v, \bar{\beta}) \in [0, \tau] \times [0, 1]^2 \times B$.

Proof of Theorem 2.1.

Consider

$$\begin{aligned}
X(\bar{\beta}) &= n^{-1}(l(\bar{\beta}) - l(\bar{\beta}_0)) \\
&= n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau \left[(\bar{\beta} - \bar{\beta}_0)^T \tilde{Z}_{ki}(s, u) - log\big[\frac{S_k^{(0)}(s, u, \bar{\beta})}{S_k^{(0)}(s, u, \bar{\beta}_0)}\big] \right] N_{ki}(ds, du)
\end{aligned}$$

Note that under condition $(A.2)$,

$$\begin{aligned}
\partial^2 X(\bar{\beta})/\partial\bar{\beta}^2 &= -n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau \left[\frac{S_k^{(0)}(s, u, \bar{\beta})}{S_k^{(0)}(s, u, \bar{\beta}_0)} - (\frac{S_k^{(0)}(s, u, \bar{\beta})}{S_k^{(0)}(s, u, \bar{\beta}_0)})^{\otimes 2} \right] N_{ki}(ds, du) \\
&= -n^{-1} I(\bar{\beta}).
\end{aligned}$$

converges in probability to:

$$-\sum_{k=1}^{k} P_k \int_0^1 \int_0^\tau \frac{s_k^{(2)}(s,u,\bar{\beta})}{s_k^{(0)}(s,u,\bar{\beta})} - \frac{s_k^{(1)}(s,u,\bar{\beta})^{\otimes 2}}{s_k^{(0)}(s,u,\bar{\beta})} s_k^{(0)}(s,u,\bar{\beta}_0)\lambda_{0ki}(ds,du)dsdu,$$

uniformly in $\bar{\beta} \in \mathcal{B}$ by lemma 2.2 and by the uniform convergence of $n_k^{-1}\sum_{i=1}^{n_k} N_{ki}(t,v) \xrightarrow{P}$ $\int_0^t \int_0^v \lambda_{0k}(s,u) S_k^{(0)}(s,u,\bar{\beta}_0)dsdu$ in $(t,v) \in [0,\tau] \times [0,1]^2$ (Gilbert al., 2004). The limiting matrix function is a minus positive definite matrix at $\bar{\beta} = \bar{\beta}_0$, Hence, $X(\beta)$ converges in probability to a function of $\bar{\beta}$ which is concave with a unique maximum at $\bar{\beta}_0$. Since $\bar{\beta}$ is the maximizer of $X(\beta)$, we have $\hat{\bar{\beta}}$ converges in probability to $\bar{\beta}_0$ as $n \to \infty$ by Var der Vaart (1998).

Proof of Theorem 2.2.

Note that $U(\hat{\bar{\beta}}) - U(\bar{\beta}_0) = I(\bar{\beta}^*)(\hat{\bar{\beta}} - \bar{\beta}_0)$, where $\bar{\beta}^*$ is on the line segment between $\hat{\bar{\beta}}$ and $\bar{\beta}_0$. By the uniform convergence of $n^{-1}I(\bar{\beta}) \xrightarrow{P} -\Sigma(\bar{\beta})$ in probability in $\beta \in \mathcal{B}$ and the consistency of $\hat{\bar{\beta}}$ to $\bar{\beta}_0$, we have

$$\begin{aligned} n^{1/2}(\hat{\bar{\beta}} - \bar{\beta}) &= -(I(\bar{\beta}^*)/n)^{-1} n^{-1/2} U(\bar{\beta}_0) \\ &= (\Sigma(\bar{\beta}_0))^{-1} n^{-1/2} U(\bar{\beta}_0) + o_p(1). \end{aligned}$$

It remains to show that: $n^{-1/2}U(\bar{\beta}_0) \xrightarrow{\mathcal{D}} N(0, \Sigma(\bar{\beta}_0))$.
let $M_{ki}(t,v) = N_{ki}(t,v) - \int_0^v \int_0^t Y_{ki}(s)\lambda_{ki}(s,u)dsdu$, then

$$\begin{aligned} n^{-1/2}U(\bar{\beta}) &= n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau [\tilde{Z}_{ki}(s,u) - \frac{S_k^{(1)}(t,u,\bar{\beta})}{S_k^{(0)}(t,u,\bar{\beta})}]N_{ki}(ds,du) \\ &= n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau \left[\tilde{Z}_{ki}(s,u) - \frac{S_k^{(1)}(t,u,\bar{\beta})}{S_k^{(0)}(t,u,\bar{\beta})}\right] M_{ki}(ds,du) + \\ &\quad n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau \left[\tilde{Z}_{ki}(s,u) - \frac{S_k^{(1)}(t,u,\bar{\beta})}{S_k^{(0)}(t,u,\bar{\beta})}\right] exp(\bar{\beta}^T Z_{ki}(s,u)) * \\ &\quad Y_{ki}(s)\lambda_{0k}(s,u)dsdu \\ &= n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau \left[\tilde{Z}_{ki}(s,u) - \frac{S_k^{(1)}(t,u,\bar{\beta})}{S_k^{(0)}(t,u,\bar{\beta})}\right] M_{ki}(ds,du) \end{aligned}$$

$$
\begin{aligned}
&= n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau\left[\tilde{Z}_{ki}(s,u)-\frac{s_k^{(1)}(t,u,\bar{\beta})}{s_k^{(0)}t,u,\bar{\beta})}\right]M_{ki}(ds,du)+o_p(1)\\
&= \tilde{W}_I(1)+o_p(1).
\end{aligned}
$$

By lemma 2 of Gilber, Mckeague and Sun (2006). From Lemma 2.1 , $\tilde{W}_I(1)$ and thus $n^{-1/2}U(\bar{\beta})$ converges to $N(0,\Sigma(\bar{\beta}))$, This completes the proof. *Q.E.D.*

Proof of Theorem 2.3.

Consider the following decomposition:

$$
\begin{aligned}
\hat{M}_{ki}(t,v) &= \int_0^t\int_0^v[N_{ki}(ds,du)-Y_{ki}(s)\exp(\hat{\bar{\beta}}^T\tilde{Z}_{ki}(s,u))\hat{\Lambda}_{0k}(ds,du)]\\
&= M_{ki}(t,v)+\int_0^t\int_0^v Y_{ki}(s)\exp((\bar{\beta}^T\tilde{Z}_{ki}(s,u))\Lambda_{0k}(ds,du)\\
&\quad-\int_0^t\int_0^v Y_{ki}(s)\exp(\hat{\bar{\beta}}^T\tilde{Z}_{ki}(s,u))\hat{\Lambda}_{0k}(ds,du)\\
&= M_{ki}(t,v)-\int_0^t\int_0^v Y_{ki}(s)\exp(\hat{\bar{\beta}}^T\tilde{Z}_{ki}(s,u))[\hat{\Lambda}_{0k}(ds,du)-\Lambda_{0k}(ds,du)]\\
&\quad-\int_0^t\int_0^v Y_{ki}(s)[\exp(\hat{\bar{\beta}}^T\tilde{Z}_{ki}(s,u))-\exp(\bar{\beta}^T\tilde{Z}_{ki}(s,u)]\Lambda_{0k}(ds,du)]. \quad (2.18)
\end{aligned}
$$

Note that

$$
\begin{aligned}
\hat{\Lambda}_{0k}(t,v)-\Lambda_{0k}(t,v) &= \int_0^t\int_0^v\frac{N_{k\cdot}(ds,du)}{n_kS_k^{(0)}(s,u,\hat{\bar{\beta}})}-\Lambda_{0k}(t,v)\\
&= \int_0^t\int_0^v\left[\frac{1}{n_kS_k^{(0)}(s,u,\hat{\bar{\beta}})}-\frac{1}{n_kS_k^{(0)}(s,u,\bar{\beta})}\right]N_{k\cdot}(ds,du)\\
&\quad+\int_0^t\int_0^v\frac{M_{k\cdot}(ds,du)}{n_kS_k^{(0)}(s,u,\bar{\beta})}+o_p(n_k^{-1/2})\\
&= \int_0^t\int_0^v\frac{(S_k^{(1)}(s,u,\bar{\beta}))^T(\bar{\beta}-\hat{\bar{\beta}})}{n_kS_k^{(0)}(s,u,\hat{\bar{\beta}})S_k^{(0)}(s,u,\bar{\beta})}N_{k\cdot}(ds,du)\\
&\quad+\int_0^t\int_0^v\frac{M_{k\cdot}(ds,du)}{n_kS_k^{(0)}(s,u,\bar{\beta})}+o_p(n_k^{-1/2})
\end{aligned}
$$

$$= \int_0^t \int_0^v \frac{(S_k^{(1)}(s,u,\bar{\beta}))^T(\bar{\beta}-\hat{\bar{\beta}})\lambda_{0k}(s,u)}{S_k^{(0)}(s,u,\hat{\bar{\beta}})}\,dsdu + \int_0^t \int_0^v \frac{M_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,u,\bar{\beta})} + o_p(n_k^{-1/2}). \tag{2.19}$$

Where $M_{k\cdot}(t,v) = \sum_{i=1^k} M_{ki}(t,v)$. Plugging (2.19) into (2.18), we have

$$\begin{aligned} W_k(t,v,z) &= n^{-1/2}\sum_{i=1}^{n_k} g_k(Z_{ki},z)M_{ki}(t,v) \\ &\quad + n_k n^{-1/2}\int_0^t\int_0^v \frac{S_{kg}^{(0)}(s,u,z,\bar{\beta})[(S_k^{(1)}(s,u,\bar{\beta}))^T(\hat{\bar{\beta}}-\bar{\beta})]}{S_k^{(0)}(s,u,\bar{\beta})}\lambda_{0k}(s,u)\,dsdu \\ &\quad - n^{-1/2}\int_0^t\int_0^v \frac{S_{kg}^{(0)}(s,u,z,\bar{\beta})}{S_k^{(0)}(s,u,\bar{\beta})}\,M_{k\cdot}(ds,du) \\ &\quad - n_k n^{-1/2}\int_0^t\int_0^v (\hat{\bar{\beta}}-\bar{\beta})^T S_{kg}^{(1)}(s,u,z,\bar{\beta})\lambda_{0k}(s,u)\,dsdu + o_p(1). \end{aligned} \tag{2.20}$$

From the proof of Theorem 2.2, we have $n^{1/2}(\hat{\bar{\beta}}-\bar{\beta}) = (\Sigma(\bar{\beta}))^{-1}\tilde{W}_I(1) + o_p(1)$, then the second term of (2.20) equals to

$$(n_k/n)\left\{\int_0^t\int_0^v \frac{[S_k^{(1)}(s,u,\bar{\beta})\otimes S_{kg}^{(0)}(s,u,z,\bar{\beta})]^T\lambda_{0k}(s,u)}{S_k^{(0)}(s,u,\bar{\beta})}dsdu\right\}\Sigma^{-1}(\bar{\beta})\tilde{W}_I(1) + o_p(1)$$

Similarly, the fourth term of (2.20) is equal to

$$(n_k/n)\left\{\int_0^t\int_a^v [S_{kg}^{(1)}(s,u,z,\bar{\beta})]^T\lambda_{0k}(s,u)\,dsdu\right\}\Sigma^{-1}(\bar{\beta})\tilde{W}_I(1) + o_p(1).$$

Bringing the above expressions into (2.20) , we have

$$
\begin{aligned}
&W_k(t,v,z)\\
&= n^{-1/2}\sum_{i=1}^{n_k}\int_0^t\int_0^v\left[g_k(Z_{ki},z)-\frac{S_{kg}^{(0)}(s,u,z,\bar\beta)}{S_k^{(0)}(s,u,z,\bar\beta)}\right]M_{ki}(ds,du)+\\
&\frac{n_k}{n}\int_0^t\int_0^v\left(\frac{S_k^{(1)}(s,u,\bar\beta)\otimes S_{kg}^{(0)}(s,u,\bar\beta)}{S_k^{(0)}(s,u,\bar\beta)}-S_{kg}^{(1)}(s,u,z,\bar\beta)\right)^T\lambda_{0k}(s,u)\,dsdu(\Sigma(\bar\beta))^{-1}\tilde W_I(1)\\
&+o_p(n_k^{1/2})\\
&= n^{-1/2}\sum_{i=1}^{n_k}\int_0^t\int_0^v\left[g_k(Z_{ki},z)-\frac{S_{kg}^{(0)}(s,u,z,\bar\beta)}{S_k^{(0)}(s,u,\bar\beta)}\right]M_{ki}(ds,du)+n^{-1/2}(R_k(t,v,z))^T*\\
&(\Sigma(\bar\beta))^{-1}\sum_{l=1}^{K}\sum_{i=1}^{n_l}\left\{\int_0^\tau\int_0^1\left[\tilde Z_{li}(s,u)-\frac{S_l^{(1)}(s,u,\bar\beta)}{S_l^{(0)}(s,u,\bar\beta)}\right]M_{li}(ds,du)\right\}^T+o_p(1)\\
&= n^{-1/2}\sum_{l=1}^{K}\sum_{i=1}^{n_l}\int_0^\tau\int_0^v I(l=k)I(s\le t)\left[g_l(Z_{li},z)-\frac{S_{lg}^{(0)}(s,u,z,\bar\beta)}{S_l^{(0)}(s,u,\bar\beta)}\right]M_{li}(ds,du)\\
&+n^{-1/2}(R_k(t,v,z))^T(\Sigma(\bar\beta))^{-1}\sum_{l=1}^{K}\sum_{i=1}^{n_l}\left\{\int_0^\tau\int_0^1\left[\tilde Z_{li}(s,u)-\frac{S_l^{(1)}(s,u,\bar\beta)}{S_l^{(0)}(s,u,\bar\beta)}\right]M_{li}(ds,du)\right\}^T\\
&+o_p(1).
\end{aligned}
\tag{2.21}
$$

By the uniform convergence of $S_k^{(0)}(s,u,\bar\beta)$, $S_k^{(1)}(s,u,\bar\beta)$, $S_{kg}^{(0)}(s,u,z,\bar\beta)$, and $S_{kg}^{(1)}(s,u,z,\bar\beta)$ to $s_k^{(0)}(s,u,\bar\beta)$, $s_k^{(1)}(s,u,\bar\beta)$, $s_{kg}^{(0)}(s,u,z,\bar\beta)$, and $s_{kg}^{(1)}(s,u,z,\bar\beta)$ in $(s,u)\in[0,\tau]\times[0,1]^2$ in probability, respectively. And by the weak convergence of $n_k^{-1/2}\sum_{i=1}^{n_k}\int_0^t\int_0^v\frac{S_{kg}^{(0)}(s,u,z,\bar\beta)}{S_k^{(0)}(s,u,\bar\beta)}M_{ki}(ds,du)$ and $n_k^{-1/2}\sum_{i=1}^{n_k}\int_0^\tau\int_0^1\frac{S_k^{(1)}(s,u,\bar\beta)}{S_k^{(0)}(s,u,\bar\beta)}M_{ki}(ds,du)$ for $k=1,\dots K$, the terms $S_k^{(0)}(s,u,\bar\beta)$, $S_k^{(1)}(s,u,\bar\beta)$, $S_{kg}^{(0)}(s,u,z,\bar\beta)$, and $S_{kg}^{(1)}(s,u,z,\bar\beta)$ can be replaced by their expected values respectively. This complete the proof. *Q.E.D.*

Table 2.1: Summary statistics for the estimators $\hat{\tilde{\beta}}$ and coverage probabilities of the 95% simultaneous confidence intervals for $(\beta_0, \beta_1, \beta_2, \beta_{12})$ under model M_{10}.

(n_1, n_2)	Coefficient	Bias	SSE	ESE	CP
$(250, 250)$	β_0	-0.0846	0.9893	0.9150	.947
	β_1	0.0387	1.7113	1.5916	.953
	β_2	0.0648	1.7045	1.5952	.960
	β_{12}	0.0696	2.917	2.7694	.955
$(400, 400)$	β_0	-0.0451	0.7190	.7046	.960
	β_1	0.0410	1.2005	1.2230	.966
	β_2	0.0160	1.2336	1.2222	.963
	β_{12}	-0.0382	2.0698	2.1219	.965
$(500, 500)$	β_0	-0.04244	0.6217	0.6247	.954
	β_1	0.0502	1.0950	1.0813	.954
	β_2	0.0047	1.0686	1.0860	.952
	β_{12}	0.0099	1.8662	1.8778	.956

Table 2.2: Summary statistics for the estimators $\hat{\tilde{\beta}}$ and coverage probabilities of the 95% simultaneous confidence intervals for $(\beta_0, \beta_1, \beta_2, \beta_{12})$ under model M_{20}.

(n_1, n_2)	Coefficient	Bias	SSE	ESE	CP
$(250, 250)$	β_0	-0.0210	0.7824	0.7884	.964
	β_1	0.0174	1.2611	1.2736	.968
	β_2	0.0100	1.2865	1.2824	.954
	β_{12}	0.0005	2.0566	2.0802	.958
$(400, 400)$	β_0	-0.0340	0.6226	0.6171	.957
	β_1	0.0211	0.9909	0.9945	.962
	β_2	0.0308	1.00038	1.0020	.962
	β_{12}	0.0080	1.6225	1.6216	.963
$(500, 500)$	β_0	-0.0338	0.5483	0.5475	.957
	β_1	0.0377	0.8769	0.8829	.960
	β_2	0.0492	0.9084	0.8905	.950
	β_{12}	-0.0437	1.4557	1.4412	.950

Table 2.3: Summary statistics for the estimators $\hat{\tilde{\beta}}$ and coverage probabilities of the 95% simultaneous confidence intervals for $(\beta_0, \beta_1, \beta_2, \beta_{12})$ under model M_{40}.

(n_1, n_2)	Coefficient	Bias	SSE	ESE	CP
$(250, 250)$	β_0	-0.0428	0.8601	0.8208	.957
	β_1	0.0686	1.33300	1.2942	.957
	β_2	-0.0021	1.4785	1.4268	.960
	β_{12}	-0.0177	2.2907	2.468	.961
$(400, 400)$	β_0	-0.0179	0.6560	0.6356	.946
	β_1	0.0043	1.0254	1.0026	.946
	β_2	-0.0703	1.1276	1.1045	.950
	β_{12}	0.0816	1.7516	1.7407	.954
$(500, 500)$	β_0	-0.0221	0.5761	0.5655	.950
	β_1	0.0379	0.9161	0.8921	.951
	β_2	0.0007	1.0051	0.9813	.943
	β_{12}	-0.0091	1.5954	1.5482	.945

Table 2.4: For testing $H_{10} : \beta_1 = \beta_2 = \beta_{12} = 0$. The empirical size and power of likelihood ratio test, Wald test and score test at the significance level 0.05.

Model	$n_1(Trt, Plb)$	$n_2(Trt, Plb)$	AVE	LRT	Wald	Score	$CP(\beta_0, \beta_1, \beta_2, \beta_{12})$
M_{10}	250(11,46)	250(16,60)	80.8	6.7	5.2	5.7	94.7,95.3,96.0,95.5
	400(18,73)	400(26,97)	80.8	4.4	4.0	4.5	96.0,96.6,96.3,96.5
	500(22,92)	500(33,121)	80.8	4.8	4.2	4.4	95.4,95.4,95.2,95.6
M_{11}	250(17,46)	250(25,60)	68.9	56.5	55.8	55.9	95.1, 95.2, 95.1, 95.7
	400(28,73)	400(40,97)	68.9	64.8	63.8	64.7	96.0, 96.4, 96.0, 95.6
	500 (35,92)	500(50,121)	68.9	74.0	72.4	72.8	94.5, 94.7, 95.8, 95.1
M_{12}	250(25,46)	250(35,60)	52.3	56.4	53.2	54.9	95.5,95.7,94.9,95.6
	400(41,74)	400(51,97)	52.3	74.1	73.2	73.8	95.8,95.7,95.8,95.3
	500(51,92)	500(72,121)	52.3	82.4	81.4	81.8	95.1,96.2,95.2,95.6
M_{13}	250(30,46)	250(42,60)	41.4	76.4	72.9	75.9	93.5,93.2,94.0,94.0
	400(49,74)	400(67,97)	41.4	91.7	91.2	91.5	95.3,95.3,94.6,95.0
	500(61,93)	500(84,121)	41.4	97.7	97.7	97.7	94.7,95.0,94.2,94.5

Table 2.5: For testing $H_{20} : \beta_{12} = 0$. The empirical size and power of likelihood ratio test, Wald test and score test at the significance level 0.05.

Model	$n_1(Trt, Plb)$	$n_2(Trt, Plb)$	AVE	LRT	Wald	Score	$CP(\beta_0, \beta_1, \beta_2, \beta_{12})$
M_{20}	250(25,46)	250(36,60)	52.3	4.9	4.2	4.6	96.4,96.8,95.4,95.8
	400(41,74)	400(57.96)	52.3	3.9	3.8	3.8	95.7,96.2,96.2,96.3
	500(51,92)	500(71,120)	52.3	5.0	5.0	5.0	95.7,96.0,95.0,95.0
M_{21}	250(17,45)	250(25,60)	68.9	31.0	35.3	33.5	95.5, 96.6, 96.7, 96.2
	400(28,74)	400(40,96)	68.9	38.8	42.2	41.5	94.6, 95.0, 94.5, 94.0
	500 (35,93)	500(50,121)	68.9	50.8	54.8	53.4	95.9, 96.1, 94.8, 95.5
M_{22}	250(23,46)	250(32,60)	57.9	34.0	39.3	38.2	96.5, 97.3, 96.5, 96.7
	400(37,74)	400(52,97)	57.9	51.0	55.2	54.2	95.4, 95.3, 95.8, 95.6
	500 (46,92)	500(65,121)	57.9	60.4	63.5	62.8	95.5, 96.0, 95.0, 96.3
M_{23}	250(30,46)	250(41,60)	42.1	40.0	43.3	42.0	95.7, 96.4, 95.4, 95.4
	400(48,74)	400(67,97)	42.1	53.7	57.6	56.8	94.7, 94.6, 95.9, 95.1
	500 (61,92)	500(83,121)	42.1	67.0	69.8	69.0	95.8, 95.8, 95.7, 95.5

Table 2.6: For testing $H_{30} : \beta_2 = \beta_{12} = 0$, the empirical size and power of likelihood ratio test, Wald test and score test at the significance level 0.05.

Model	$n_1(Trt, Plb)$	$n_2(Trt, Plb)$	AVE	LRT	Wald	Score	$CP(\beta_0, \beta_1, \beta_2, \beta_{12})$
M_{30}	250(20,46)	250(29,60)	62.8	5.6	4.8	5.5	95.7,95.7,96.0,96.1
	400(33,74)	400(47.97)	62.8	3.9	3.6	3.9	94.6,94.6,95.0,95.4,
	500(41,92)	500(58,120)	62.8	5.3	5.0	5.2	95.0,95.1,94.3,94.5
M_{31}	250(27,46)	250(37,60)	49.4	57.0	56.0	56.6	94.6,95.1,94.5,94.1
	400(43,74)	400(60,96)	49.4	70.9	69.7	70.0	96.1 95.8,96.3,95.4
	500(54,92)	500(75,121)	49.4	79.6	79.1	79.5	95.3,95.2,96.3,96.1
M_{32}	250(25,46)	250(36,60)	51.8	66.9	64.3	66.1	96.2,95.8,96.1,95.7
	400(41,74)	400(58,97)	51.8	83.5	82.8	83.0	95.5 95.2,95.0,95.1
	500(51,92)	500(72,120)	51.8	89.2	88.6	88.8	95.5 95.8,95.8 95.3
M_{33}	250(29,46)	250(41,60)	43.3	62.8	60.2	61.8	94.6,95.0,95.3,94.5
	400(48,74)	400(65,96)	43.3	83.0	82.1	82.6	93.6,93.7,94.6,94.2
	500(59,92)	500(82,120)	43.3	90.8	90.3	90.4	95.6,95.3,95.5,95.1
M_{34}	250(30,46)	250(41,60)	42.3	74.6	73.3	74.2	95.2,95.6,95.1,95.0
	400(48,74)	400(66,97)	42.3	91.3	90.9	91.2	95.4,95.5,95.6,96.2
	500(60,92)	500(83,121)	42.3	96.0	95.7	95.9	94.2,95.3,94.4,94.7

Table 2.7: Estimation of of $(\beta_0, \beta_1, \beta_2, \beta_{12})$ under orignal marks.

Model	$(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_{12})$	$SE(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_{12})$	(V_1^m, V_1^M)	(V_2^m, V_2^M)	AVE^*
S_1	$(-9.2287, 21.7608, 17.7119, -39.0028)$	$(13.2853, 30.2153, 33.4435, 75.5529)$	$(0.2792, .5754)$	$(.2672, .6848)$	-0.9815
S_2	$(-23.7200, 48.2382, 29.6228, -57.4511)$	$(14.8561, 33.8216, 19.6736, 44.5532)$	$(0.2792, .5754)$	$(.5096, .8695)$	-1.1422
S_3	$(-19.5149, 37.3522, 24.1250, -43.6239)$	$(10.3192, 20.4953, 12.9958, 24.2905)$	$(0.2792, .5754)$	$(.6269, .8240)$	-0.9627
S_4	$(-4.0509, 16.0184, 4.0596, -17.8728)$	$(1.7197, 4.8840, 3.6920, 8.3948)$	$(0.0, .8571)$	$(.0435, .7143)$	-47.4516
S_5	$(-3.9356, 13.7176, 4.6342, -14.9112)$	$(1.4871, 4.3390, 1.9393, 4.6546)$	$(0.0, .8571)$	$(0.0, 1.0)$	-1841.0429
S_6	$(-8.15875, 27.4792, 9.0838, -29.8365)$	$(3.5335, 8.8494, 4.8093, 11.0204)$	$(0.0, .8571)$	$(.2222, 1.0)$	-164.6004
S_7	$(1.8977, -3.0317, 17.5899, -14.3385)$	$(3.2152, 8.2121, 5.6859, 10.2312)$	$(.2792, .5754)$	$(0.0, .8571)$	-81.0033
S_8	$(24.9929, -30.2899, -62.9896, 79.8365)$	$(21.1750, 28.8837, 36.8384, 49.5279)$	$(.6269, .8240)$	$(.2222, 1.0)$	-1.3893

Table 2.8: Estimator of $(\beta_0, \beta_1, \beta_2, \beta_{12})$ under standardized marks.

Model	$(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_{12})$	$SE(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_{12})$	AVE
S_1	$(-0.7464, 3.7829, 2.3479, -4.5770)$	$(0.9278, 2.7257, 2.8261, 4.2928)$	-2.8249
S_2	$(0.4664, -0.9358, -0.7815, 1.7421)$	$(1.1345, 2.9339, 2.3066, 3.9144)$	-0.0558
S_3	$(0.2235, -4.6530, -1.8821, 6.7805)$	$(0.8561, 7.0811, 2.7491, 8.5654)$	0.6664
S_4	$(0.2092, 1.74509, -2.9374, 2.6182)$	$(0.9606, 3.0765, 2.2719, 4.1130)$	-0.9981
S_5	$(1.0782, -0.6448, -3.0065, 2.6467)$	$(0.9411, 2.0458, 2.0963, 3.0184)$	-0.0571
S_6	$(0.7397, -0.0650, -3.0347, 3.2543)$	$(1.4456, 5.8498, 2.8475, 6.4918)$	-0.1963
S_7	$(-1.0940, 0.0876, 8.1077, -4.7879)$	$(0.9302, 1.9413, 6.4748, 6.7672)$	-32.4863
S_8	$(0.6571, -0.2429, -9.9050, 10.2485)$	$(1.1332, 3.6885, 16.5549, 19.7318)$	0.4835

Table 2.9: The p-value of the likelihood ratio test, Wald test and score test for H_{10}, H_{20} and H_{30} with various selections of original marks.

Model	TEST	LRT	Wald	Score
S_1	H_{10}	0.4000	0.4242	0.4070
	H_{20}	0.5984	0.5948	0.5991
	H_{30}	0.8600	0.8593	0.8605
S_2	H_{10}	0.1246	0.1777	0.1336
	H_{20}	0.3408	0.3750	0.3476
	H_{30}	0.1990	0.2427	0.2078
S_3	H_{10}	0.2168	0.2534	0.2401
	H_{20}	0.5161	0.5162	0.4934
	H_{30}	0.4053	0.4004	0.4022
S_4	H_{10}	0.0001	0.0050	0.0010
	H_{20}	0.7047	0.7020	0.7033
	H_{30}	0.7928	0.7928	0.7932
S_5	H_{10}	0.0001	0.0123	0.0021
	H_{20}	0.1908	0.1900	0.1900
	H_{30}	0.2526	0.2515	0.2511
S_6	H_{10}	0.0001	0.0052	0.0007
	H_{20}	0.3272	0.3008	0.2970
	H_{30}	0.5934	0.5695	0.5638
S_7	H_{10}	0.0001	0.0050	0.0007
	H_{20}	0.5204	0.5250	0.5250
	H_{30}	0.0001	0.0023	0.0005
S_8	H_{10}	0.0383	0.07148	0.0490
	H_{20}	0.0710	0.1044	0.0932
	H_{30}	0.0416	0.0846	0.0526

Table 2.10: The p-values of the likelihood ratio test, Wald test and score test for H_{10}, H_{20} and H_{30} with various selections of standardized marks.

Model	TEST	LRT	Wald	Score
S_1	H_{10}	0.4000	0.4249	0.4051
	H_{20}	0.5985	0.5959	0.5944
	H_{30}	0.8601	0.8599	0.8590
S_2	H_{10}	0.1245	0.1730	0.1298
	H_{20}	0.3407	0.3615	0.3510
	H_{30}	0.1990	0.2357	0.2114
S_3	H_{10}	0.2168	0.2591	0.2365
	H_{20}	0.5161	0.5162	0.5137
	H_{30}	0.4053	0.4149	0.4043
S_4	H_{10}	0.0001	0.0051	0.0010
	H_{20}	0.7046	0.7024	0.7034
	H_{30}	0.7927	0.7929	0.7934
S_5	H_{10}	0.0001	0.0123	0.0021
	H_{20}	0.1908	0.1904	0.1899
	H_{30}	0.2526	0.2519	0.2513
S_6	H_{10}	0.0001	0.0052	0.0007
	H_{20}	0.3272	0.3011	0.2955
	H_{30}	0.5933	0.5699	0.5628
S_7	H_{10}	0.0001	0.0050	0.0007
	H_{20}	0.5201	0.5265	0.5286
	H_{30}	0.0001	0.0023	0.0005
S_8	H_{10}	0.0383	0.0728	0.0476
	H_{20}	0.0710	0.1039	0.0850
	H_{30}	0.0416	0.0852	0.0504

Table 2.11: The empirical size and power of goodness-of-fit test at the significance level 0.05.

Model	$n_1(Trt, Plb)$	$n_2(Trt, Plb)$	Size/Power
M_{40}	50(5,9)	50(8,12)	6.2
	100(12,18)	100(16,24)	5.2
	150(18,27)	250(25,36)	6.0
M_{41}	50(4,8)	50(8,13)	49.80
	100(9,16)	100(16,27)	79.00
	150(14,24)	150(25,40)	87.40
M_{42}	50(5,10)	50(10,15)	61.00
	100(11,20)	100(20,31)	87.00
	150(17,30)	150(31,47)	92.40
M_{43}	50(7,12)	50(11,17)	69.40
	100(13,24)	100(23,35)	92.40
	150(21,36)	150(36,52)	97.00

Table 2.12: Summary results of goodness-of-fit test of mark-specific PH model for STEP data with various selections of standardized marks.

Mark Selection	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
P-value	0.1500	.3900	.0900	.1400	.0900	.0400	.0900	.5600

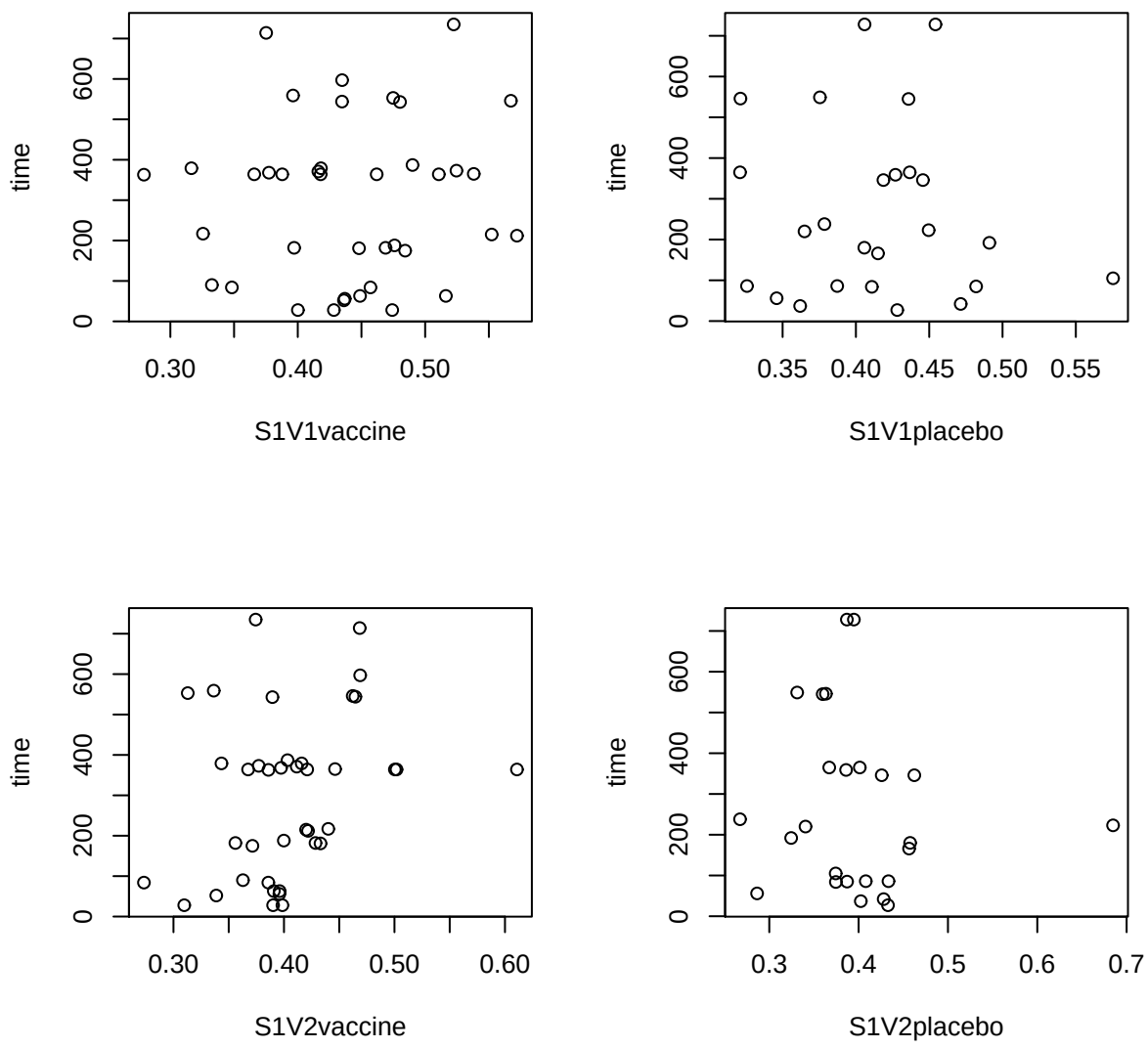

Figure 2.1: Survival time for vaccine and placebo group under model S1 for mark1 and mark2

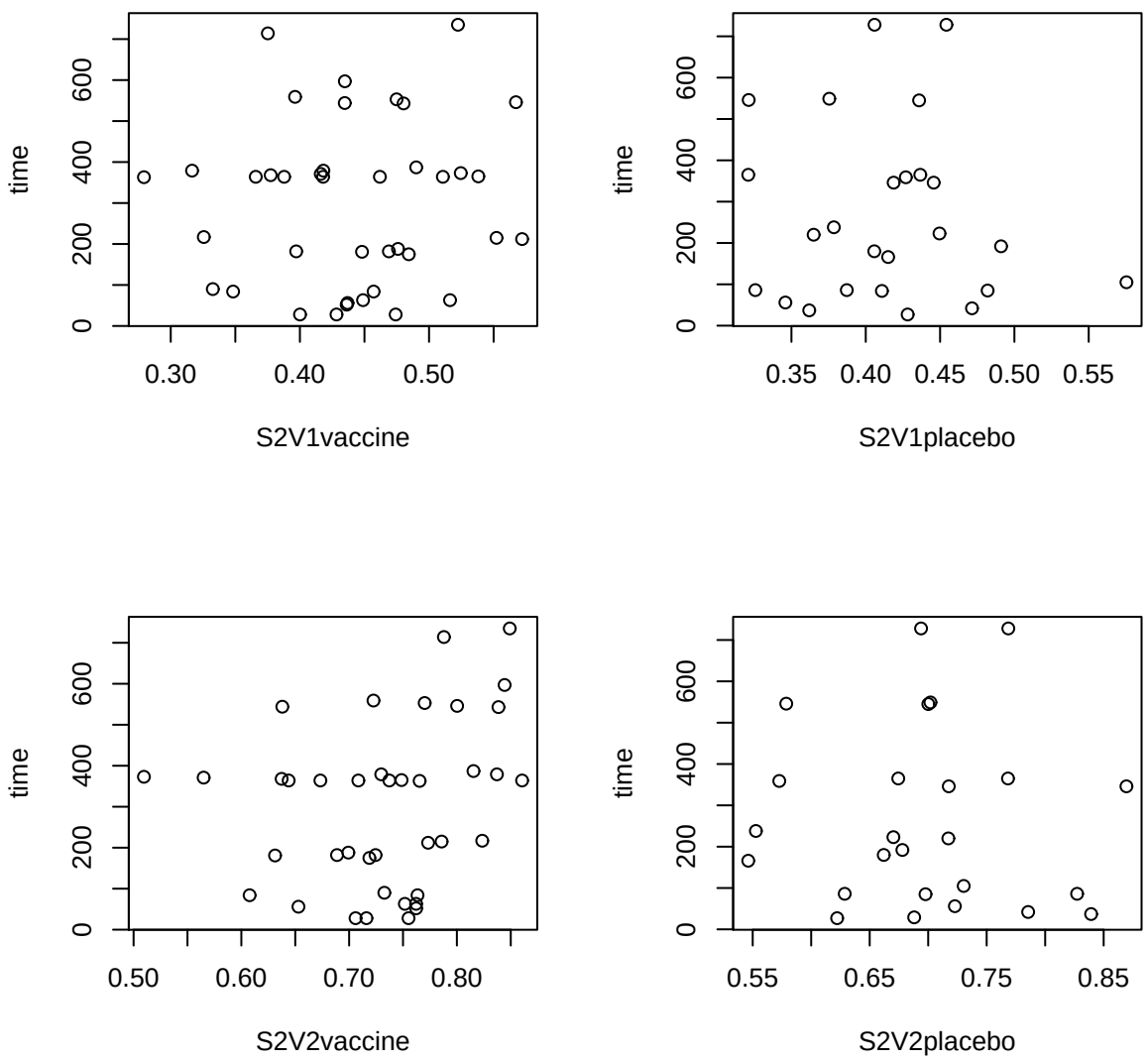

Figure 2.2: Survival time for vaccine and placebo group under model S2 for mark1 and mark2

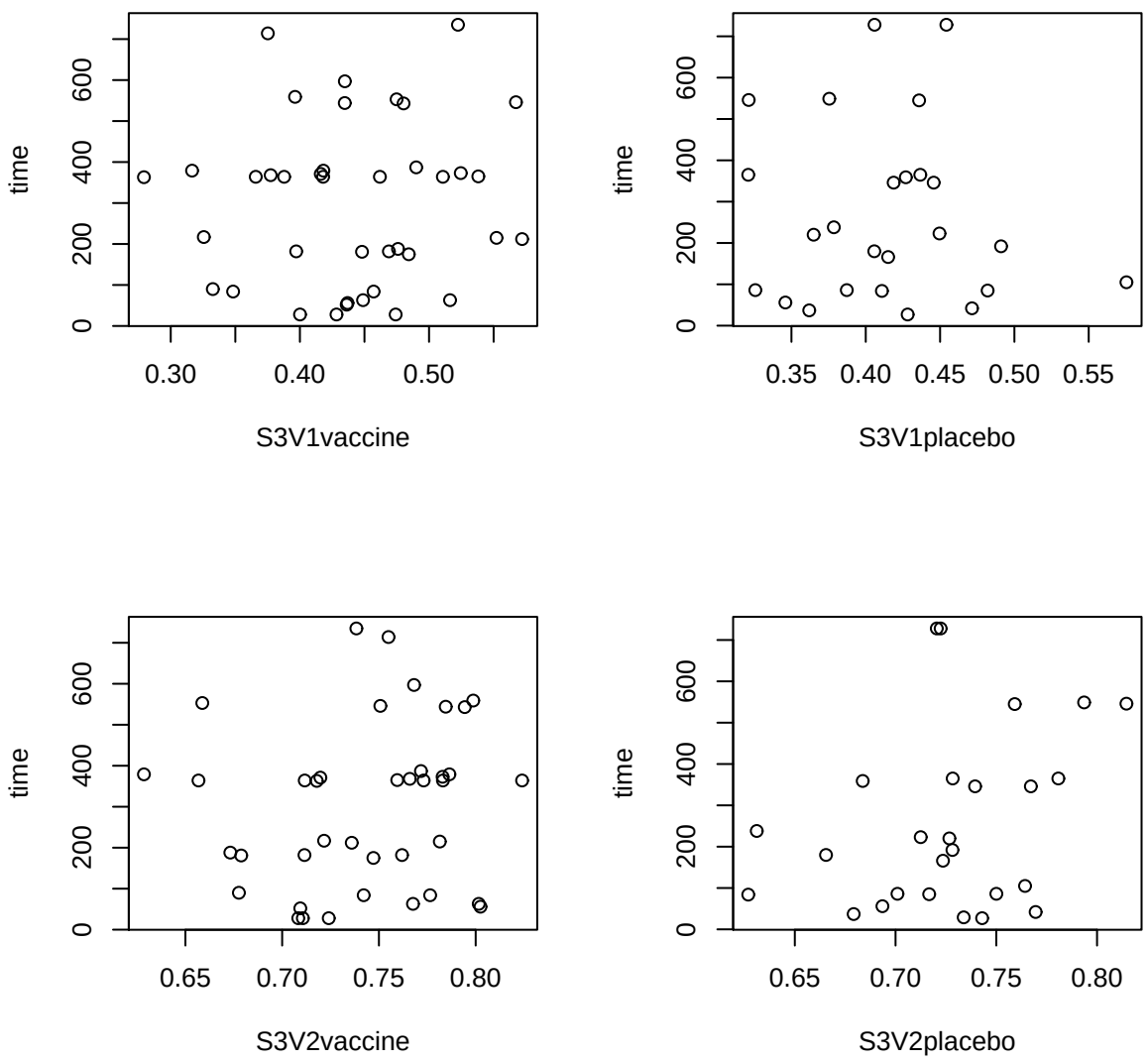

Figure 2.3: Survival time for vaccine and placebo group under model S3 for mark1 and mark2

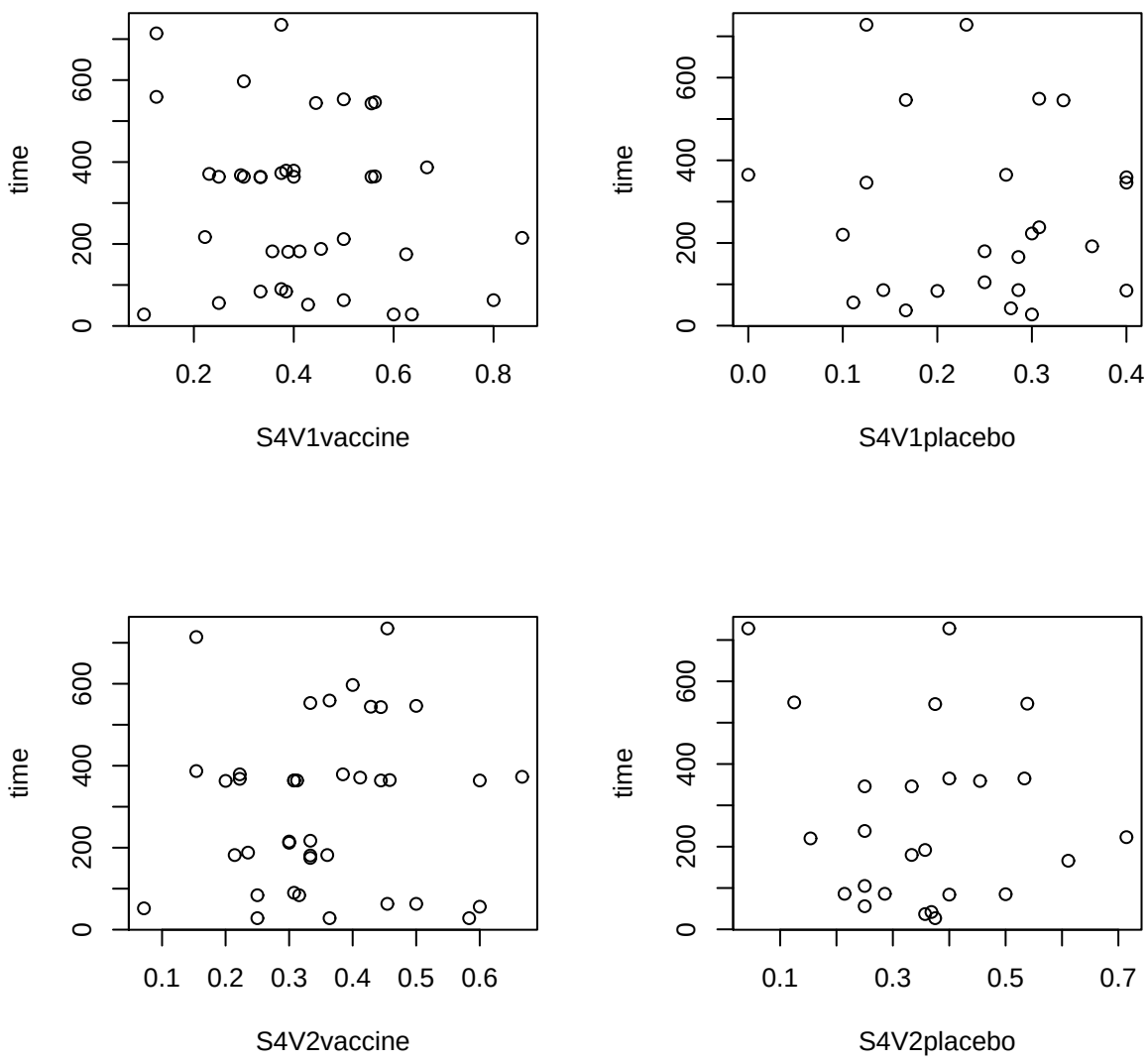

Figure 2.4: Survival time for vaccine and placebo group under model S4 for mark1 and mark2

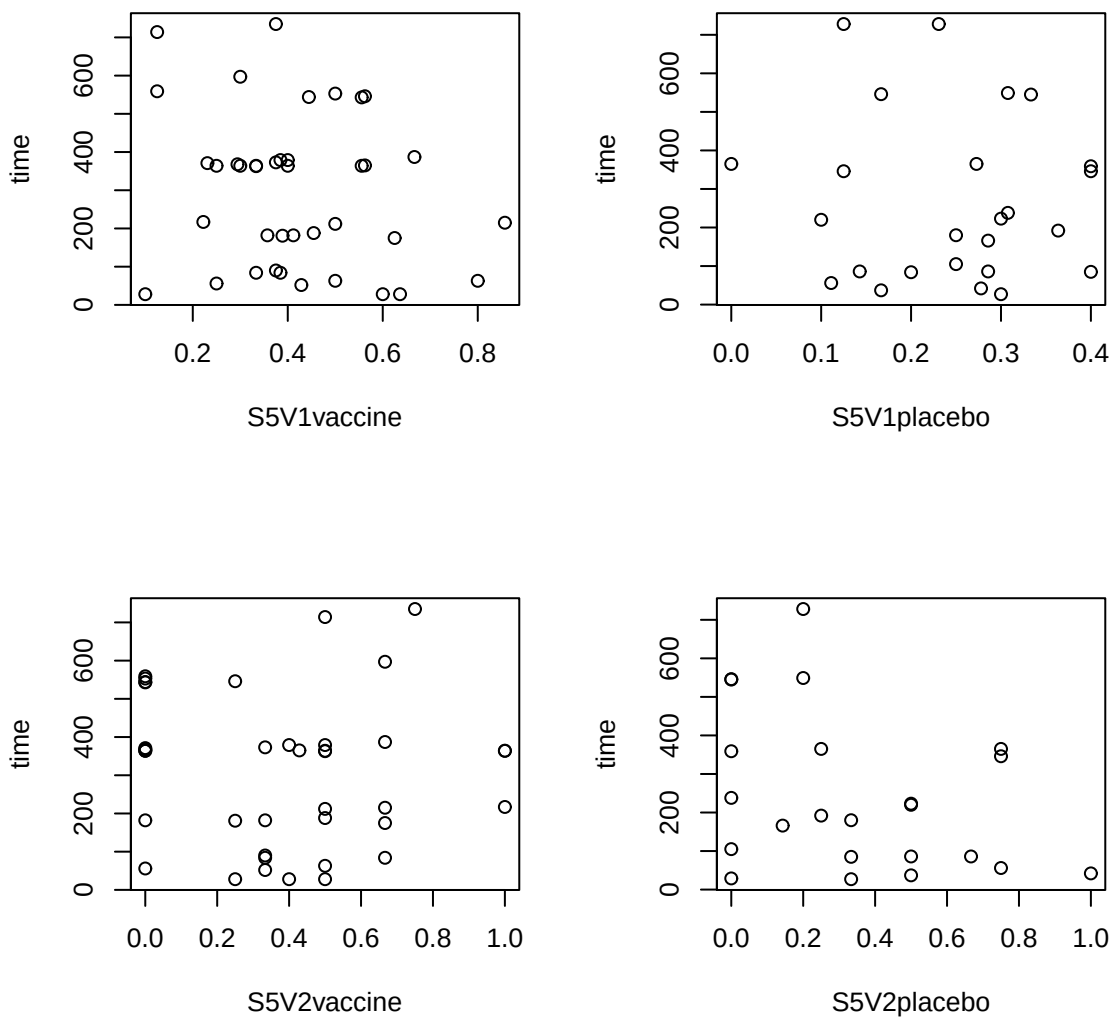

Figure 2.5: Survival time for vaccine and placebo group under model S5 for mark1 and mark2

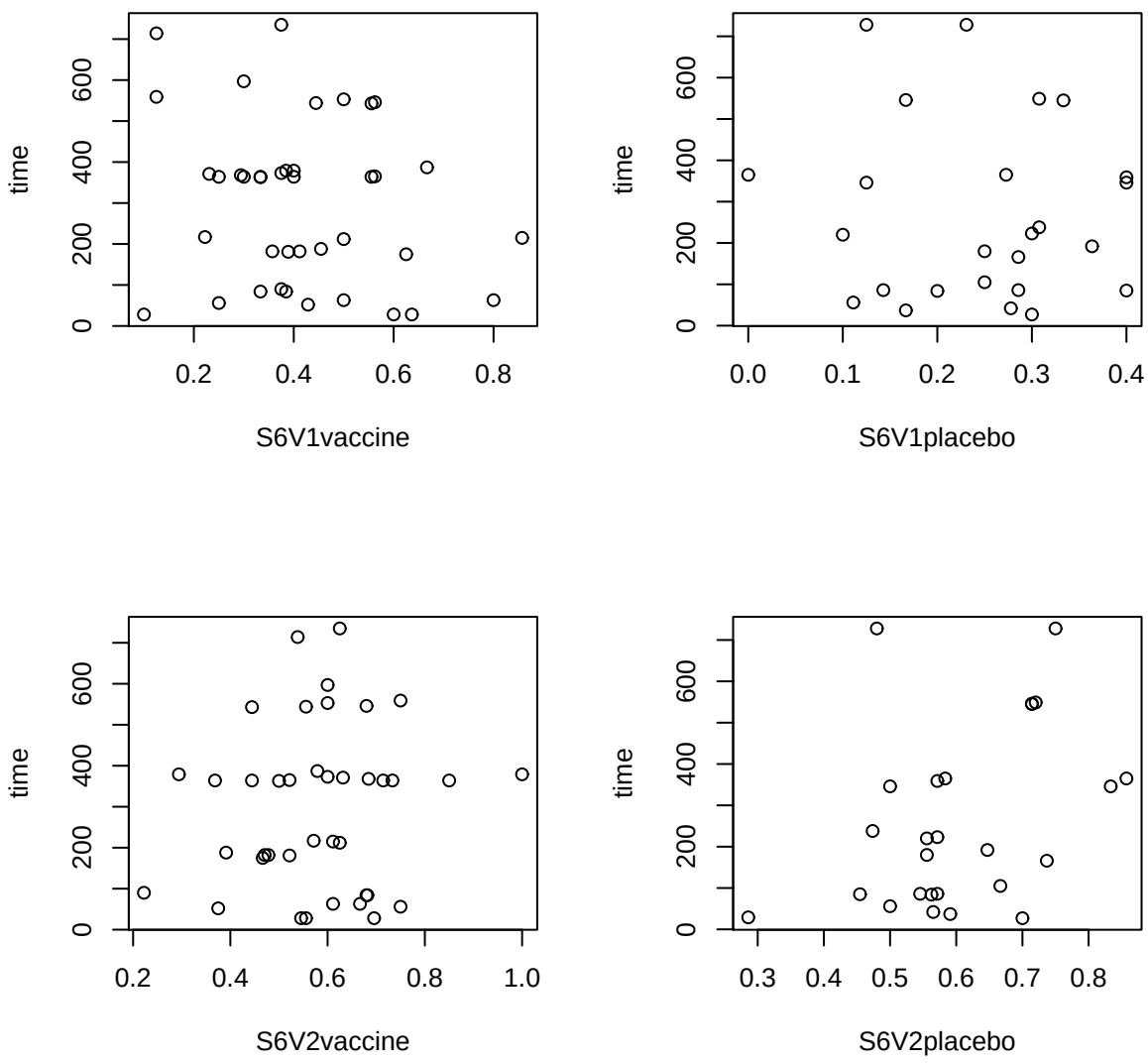

Figure 2.6: Survival time for vaccine and placebo group under model S6 for mark1 and mark2

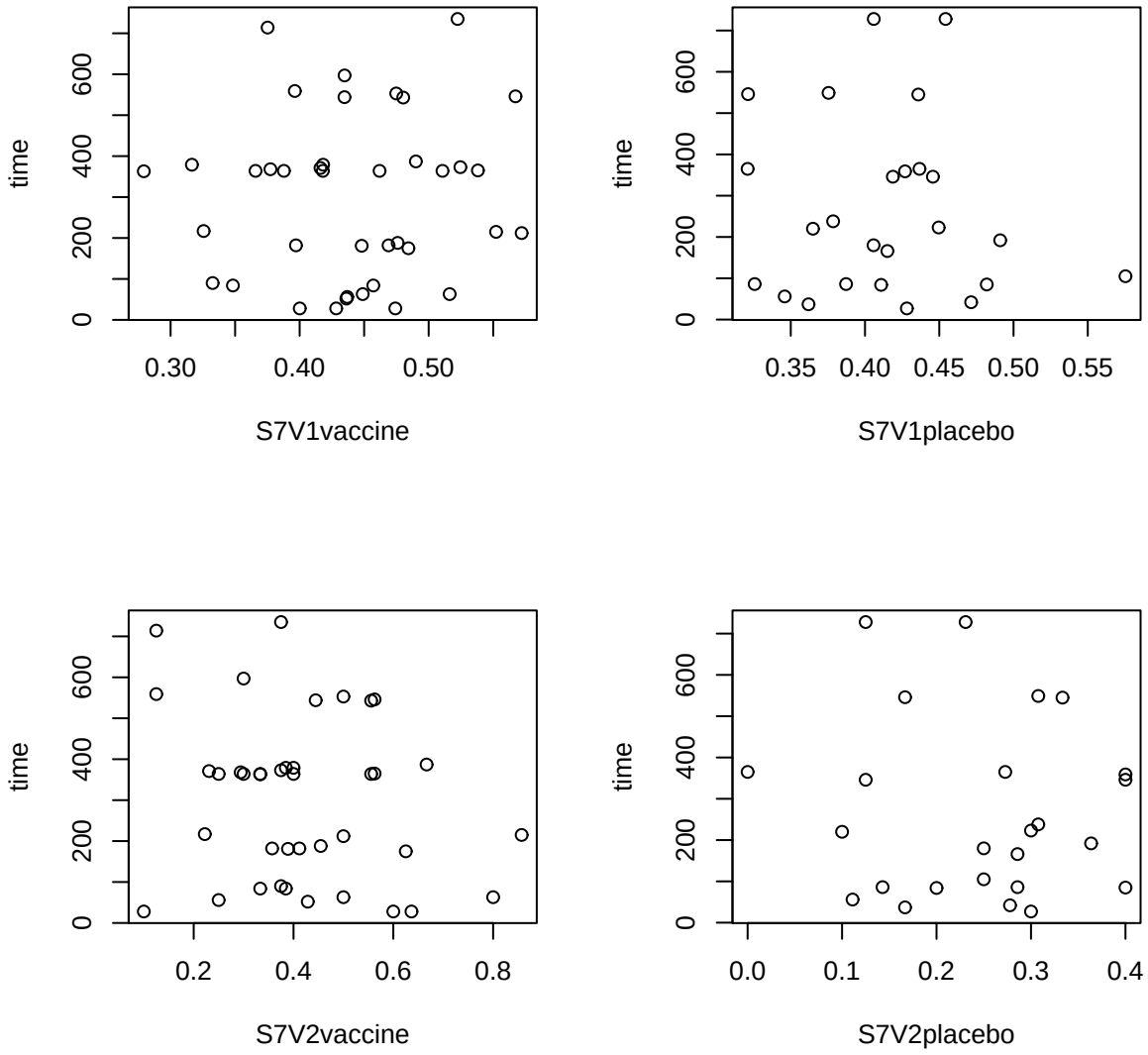

Figure 2.7: Survival time for vaccine and placebo group under model S7 for mark1 and mark2

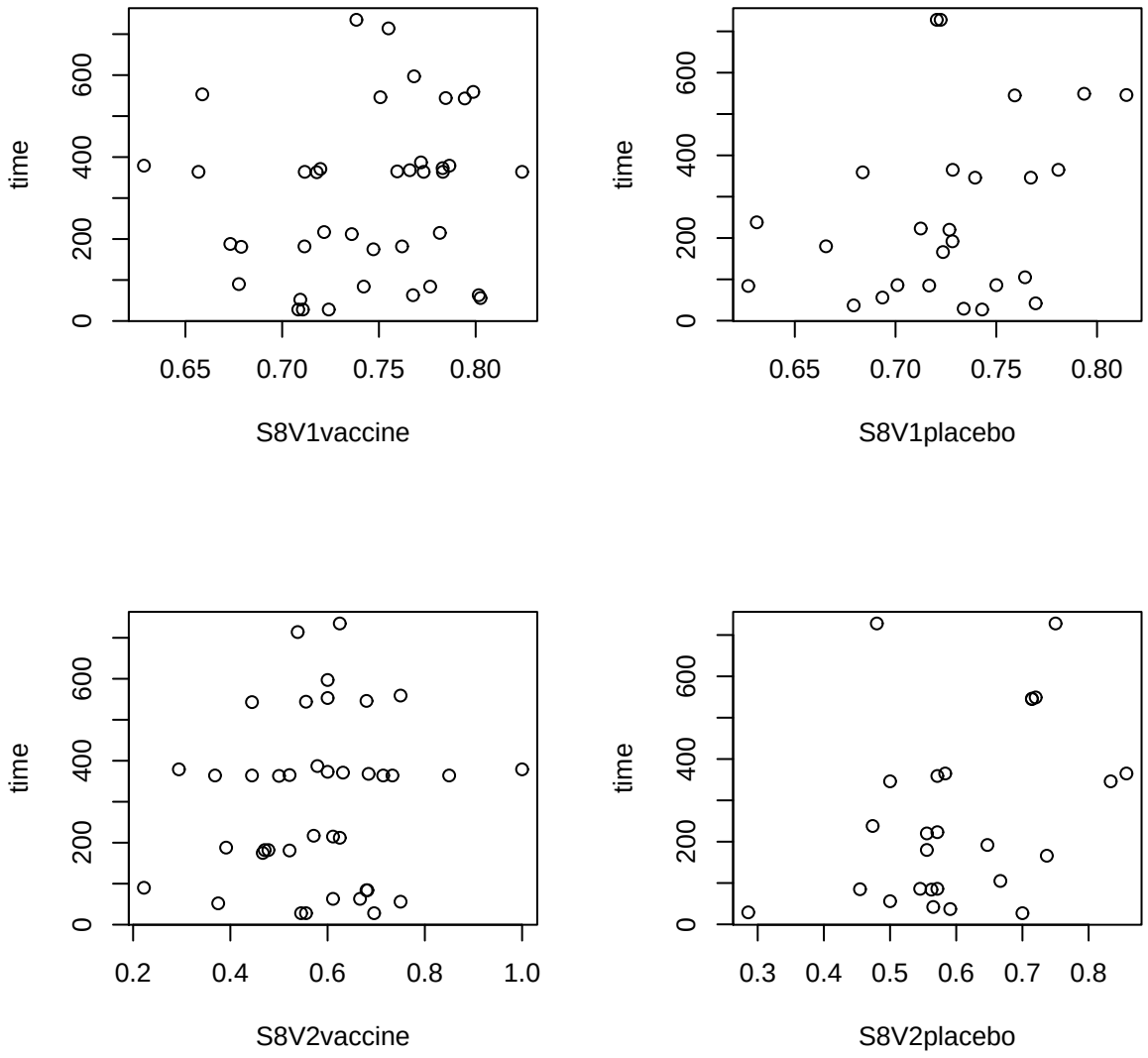

Figure 2.8: Survival time for vaccine and placebo group under model S8 for mark1 and mark2

CHAPTER 3: GOODNESS-OF-FIT OF STRATIFIED PROPORTIONAL HAZARDS MODELS WITH CONTINUOUS MARKS

3.1 Introduction

The proportional hazard model (Cox model) has been wildly used in survival analysis. Motivated by the need to evaluate HIV vaccine efficacy trials, Sun, Gilbert and McKeague (2009) studied the mark-specific proportional hazards (PH) model with continuous marks under the competing risk setting. In case of the Cox model, the model assumption may fail in three ways: (i) the time invariance of the hazard ratio does not hold; (ii) the functional forms of individual covariates in the exponent of the model are misspecified; (iii) the exponential form of the link function for the hazard ratio is inappropriate. The model misspecification can have detrimental effects on the validity and efficiency of the partial likelihood inference (Lagakos & Schoenfeld, 1984; Struthers & Kalbfleisch, 1986; Lagakos, 1988; Lin & Wei, 1989). Lin, Wei and Ying (1993) and Spiekerman & Lin (1996) developed some formal martingale residual based goodness of fit procedures to check the validity of the Cox model.

In this chapter, we study the stratified mark-specific proportional hazard model with continuous marks. The estimation procedure and derive the asymptotic properties of the estimator are developed following Sun, Gilbert and McKeague (2009). Confidence bands of vaccine efficacy are also constructed. The main contribution of this chapter is developing a goodness of fit procedure for the stratified mark-specific proportional hazard model with continuous marks. The finite sample performance of the proposed tests are examined by simulations.

3.2 Estimation

Suppose that a study population is divided into K strata. The stratified mark-specific proportional hazards model postulates that the conditional mark-specific hazard function

for an individual with covariate $z(\cdot)$ in the kth stratum is

$$\lambda_k(t, v|z(t)) = \lambda_{0k}(t, v)\exp\left\{\beta(v)^T z(t)\right\}, \quad k = 1, \ldots, K \tag{3.1}$$

where $\lambda_{0k}(t, v)$ is unspecified baseline hazard functions and $\beta(v)$ is the unknown p-dimensional regression function of v. Sun, Gilbert and Mckeague (2009) developed some estimation and hypothesis testing procedures for (3.1) in the case of $K = 1$. Sun and Gilbert (2011) studied model (3.1) based on missing marks. In the following we present an estimation method for model (3.1) parallel to Sun, Gilbert and Mckeague (2009). But the focus of this chapter is on goodness of fit of model (3.1) presented in section 3.5.

Let n_k be the number of observations in the kth stratum, $k = 1, \ldots, K$. The total sample size is then $n = \sum_{k=1}^{K} n_k$. For each k, let $(X_{ki}, \delta_{ki}, \delta_{ki}V_{ki}, Z_{ki})$, $i = 1, \ldots, n_k$ be i.i.d. replicates of $(X_k, \delta_k, \delta V_k, Z_k)$.

For a p-dimensional covariate Z_k in the kth stratum, let T_k be the failure time and V_k the mark variable observable upon failure. We assume that (T_k, V_k, Z_k) follows model (3.1). The mark variable V_k is assumed to have a known and bounded support; rescaling V_k if necessary, this support is taken without loss of generality to be $[0, 1]$. The observed random variables are $(X_k, \delta_k, \delta_k V_k, Z_k)$, where $X_k = min(T_k, C_k)$, $\delta_k = I(T_k \leq C_k)$ and C_k is a censoring random variable. The mark is observed whenever the corresponding failure time is uncensored. The censoring time is assumed to be conditionally independent of (T_k, V_k) given Z_k. The estimation of model (3.1) using the observations $(X_{ki}, \delta_{ki}, \delta_{ki}v_{ki}, Z_{ki})$ for $i = 1, \ldots, n_k$, $k = 1, \ldots, K$ can be based on a localized version of the log partial likelihood function for $\beta = \beta(v)$ at a fixed v:

$$l(v, \beta) = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau K_h(u-v)\left[\beta^T Z_{ki}(t) - \log\left(\sum_{j=1}^{n_k} Y_{kj}(t)e^{\beta^T Z_{kj}(t)}\right)\right] N_{ki}(dt, du), \tag{3.2}$$

where $K_h(x) = K(x/h)/h$, $K(\cdot)$ is a kernel function with support $[-1, 1]$, τ is the end of the follow-up period and $h = h_n$ is a bandwidth. Here $N_{ki}(t, v) = I(X_{ki} \leq t, \delta_{ki} = 1, V_{ki} \leq v)$ is the marked counting process with a jump at an uncensored failure times X_{ki} and the

associated mark V_{ki}, and $Y_{ki}(t) = I(X_{ki} \geq t)$. The log partial likelihood function (3.2) resembles that of Kalbfleisch and Prentice (1980) in the case of discrete marks, except that it borrows strength from observations having marks in the neighborhood of v. The kernel function is designed to give greater weight to observations with marks near v than those further away. The local maximum partial likelihood estimator of $\beta(v)$ is a maximizer $\hat{\beta}(v)$ of (3.2). For $\beta \in \mathbb{R}^p$, $t \geq 0$, let

$$S_k^{(j)}(t,\beta) = n_k^{-1} \sum_{i=1}^{n_k} Y_{ki}(t) \exp\{\beta^T Z_{ki}(t)\} Z_{ki}(t)^{\otimes j}, \tag{3.3}$$

where for any $z \in \mathbb{R}^p$, we denote $z^{\otimes 0} = 1$, $z^{\otimes 1} = z$ and $z^{\otimes 2} = zz^T$. Taking the derivative of $l(v,\beta)$ with respect to β gives the score function

$$U(v,\beta) = \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau K_h(u-v) \left[Z_{ki}(t) - \frac{S_k^{(1)}(t,\beta)}{S_k^{(0)}(t,\beta)} \right] N_{ki}(dt,du). \tag{3.4}$$

The second derivative of $l(v,\beta)$ with respect to β yields

$$l''_\beta(v,\beta) = -\sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^1 \int_0^\tau K_h(u-v) J_{kn}(t,\beta)\, N_{ki}(dt,du),$$

where $J_{kn}(t,\beta) = \frac{S_k^{(2)}(t,\beta)}{S_k^{(0)}(t,\beta)} - \left(\frac{S_k^{(1)}(t,\beta)}{S_k^{(0)}(t,\beta)} \right)^{\otimes 2}$. The maximum partial likelihood estimator is a solution to $U(v,\hat{\beta}(v)) = 0$. $\hat{\beta}(v)$ can be obtained using the Newton–Raphson algorithm: Set $\beta^{(j+1)}(v) = \beta^{(j)}(v) - \{l''_\beta(v,\beta^{(j)})\}^{-1} U(v,\beta^{(j)})$, until $\beta(v)$ convergence. The baseline function $\lambda_{0k}(t,v)$ can also be estimated, by smoothing the increments of the following estimator of the doubly cumulative baseline function $\Lambda_{0k}(t,v) = \int_0^t \int_0^v \lambda_{0k}(s,u)\, dsdu$:

$$\hat{\Lambda}_{0k}(t,v) = \int_0^t \int_0^v \frac{N_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,\hat{\beta}(u))}, \tag{3.5}$$

where $N_{k\cdot}(t,v) = \sum_{i=1}^{n_k} N_{ki}(t,v)$. A kernel estimator of $\lambda_{0k}(t,v)$ is given by $\hat{\lambda}_{0k}(t,v) = \int_0^\tau \int_0^1 K_{h_1}^{(1)}(t-s) K_{h_2}^{(1)}(v-u) \hat{\Lambda}_{0k}(ds,du)$, where $K_{h_1}^{(1)}(x) = K^{(1)}(x/h_1)/h_1$ and $K_{h_2}^{(2)}(x) = K^{(2)}(x/h_2)/h_2$ with $K^{(1)}(\cdot)$ and $K^{(2)}(\cdot)$ be the kernel functions and h_1 and h_2

the bandwidths.

3.3 Asymptotic Results

Define $s_k^{(j)}(t,\beta) = ES_k^{(j)}(t,\beta)$ and

$$J_k(t,\beta) = \frac{s_k^{(2)}(t,\beta)}{s_k^{(0)}(t,\beta)} - \left(\frac{s_k^{(1)}(t,\beta)}{s_k^{(0)}(t,\beta)}\right)^{\otimes 2}.$$

We make use of the following regularity conditions; not all of these conditions are required for the proof of each theorem, nor are they the minimum required set of conditions.

Condition A

(A.1) $\beta(v)$ has componentwise continuous second derivatives on $[0,1]$. For each $k = 1,\ldots,K$, the second partial derivative of $\lambda_{0k}(t,v)$ with respect to v exists and is continuous on $[0,\tau]\times[0,1]$. The covariate process $Z_k(t)$ has paths that are left continuous and of bounded variation, and satisfies the moment condition $E[\|Z_k(t)\|^4 \exp(2M\|Z_k(t)\|)] < \infty$, where M is a constant such that $(v,\beta(v)) \in [0,1]\times(-M,M)^p$ for all v and $\|A\| = \max_{k,\,l}|a_{kl}|$ for a matrix $A = (a_{kl})$.

(A.2) Each component of $s_k^{(j)}(t,\theta)$ is continuous on $[0,\tau]\times[-M,M]^p$, and $\sup_{t\in[0,\,\tau],\ \theta\in[-M,\,M]^p} \|S^{(j)}(t,\theta) - s^{(j)}(t,\theta)\| = O_p(n^{-1/2})$, for $j = 0,1,2$.

(A.3) $s_k^{(0)}(t,\theta) > 0$ on $[0,\tau]\times[-M,M]^p$ and the matrix $\Sigma(v) = \sum_{k=1}^K p_k\Sigma_k(v)$ is positive definite, where $\Sigma_k(v) = \sum_{k=1}^K \int_0^\tau J_k(t,\beta(v))\lambda_{0k}(t,v)s_k^{(0)}(t,\beta(v))\,dt$, $p_k = \lim_{n\to\infty} n_k/n$ and $0 < p_k < 1$.

(A.4) $E(N_{ki}(dt,dv)|\mathcal{F}_{t-}) = E(N_{ki}(dt,dv)|Y_{ki}(t),Z_{ki}(t))$, where $(\mathcal{F}_t)$ is the (right-continuous) filtration generated by the processes $(N_{ki},Y_{ki},Z_{ki}), i = 1,\ldots,n_k,\ k = 1,\ldots,K$.

(A.5) The kernel function $K(\cdot)$ is symmetric with support $[-1,1]$ and of bounded variation. The bandwidth satisfies $nh^2 \to \infty$ and $nh^5 \to 0$ as $n \to \infty$.

Note that the condition (A.2) holds under the condition (A.1) given some additional moment conditions on $Z(t) - Z(s)$ and $\exp(b^TZ(t)) - \exp(b^TZ(s))$. If $Z(t) = Z$, not depending on t, then (A.2) holds by the Donsker Theorem (Theorem 19.5 of van der Vaart,

1998). The condition (A.4) assumes that the mark-specific instantaneous failure rate at time t given the observed information up to time t only depends on the failure status and the current covariate value. Under (A.4) and by the definition (1.3), $E(N_{ki}(dt,dv)|\mathcal{F}_{t-}) = Y_{ki}(t)\lambda(t,v|Z_{ki}(t))\,dtdv$. Let $M_{ki}(t,u) = \int_0^t \int_0^u [N_{ki}(ds,dx) - Y_{ki}(s)\lambda(s,x|Z_{ki}(s))\,dsdx]$. It follows by Aalan and Johansen (1978) that $M_{ki}(\cdot,v_1)$ and $M_{ki}(\cdot,v_2) - M_{ki}(\cdot,v_1)$ are orthogonal square integrable martingales with respect to $\mathcal{F}_t$ for any $0 \le v_1 \le v_2 \le 1$. To avoid the problems at the boundaries $v = 0, 1$, we shall study the asymptotic properties of $\hat{\beta}(v)$ for the interior values of $v \in [a,b] \subset (0,1)$.

First we present the following result that is essential for proving the asymptotic normality of $\hat{\beta}(v)$ and provides important insight into the constructions of the confidence bands and test statistics that follow. Let

$$\tilde{W}_A(v) = n^{-1/2} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_a^v \int_0^\tau A(u) \left[Z_{ki}(t) - \frac{s_k^{(1)}(t,\beta(u))}{s_k^{(0)}(t,\beta(u))} \right] M_{ki}(dt,du), \tag{3.6}$$

where $A(u)$ is a deterministic $p \times p$ matrix with bounded components.

Theorem 3.1. *Assume that each component of the $p \times p$ matrix $A(v), v \in [a,b]$, is continuous. Under conditions (A.1)–(A.4), $\tilde{W}_A(v)$ converges weakly to a p-dimensional mean-zero Gaussian martingale, $W_A(v)$, with continuous sample paths on $v \in [a,b]$ as $n \to \infty$. The covariance matrix of $W_A(v)$ is given by* $\mathrm{Cov}(W_A(v)) = \int_a^v A(u)\Sigma(u)A(u)\,du$.

Let

$$\hat{\Sigma}_{\hat{A}}(v) = n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_a^v \int_0^\tau \hat{A}(u) J_{kn}(t,\hat{\beta}(u)) \hat{A}^T(u)\, N_{ki}(dt,du), \tag{3.7}$$

where $\hat{A}(v)$ is a consistent estimator of $A(v)$ uniformly in $v \in [a,b] \subset [0,1]$. It can be shown that $\hat{\Sigma}_A(v)$ is a consistent estimator of $\mathrm{Cov}(W_A(v))$.

The consistency and asymptotic normality of $\hat{\beta}(v)$ are established in the next two theorems.

Theorem 3.2. *Under conditions (A.1)–(A.5), $\hat{\beta}(v)$ converges to $\beta(v)$ uniformly in $v \in [a,b] \subset (0,1)$ as $n \to \infty$.*

Theorem 3.3. *Under conditions (A.1)–(A.5),* $(nh)^{1/2}(\hat{\beta}(v)-\beta(v)) \xrightarrow{\mathcal{D}} N(0, \nu_0\Sigma^{-1}(v))$ *for* $v \in [a,b] \in (0,1)$ *as* $n \to \infty$*, where* $v_o = \int_{-1}^{1} K^2(u)du$.

The proof of Theorem 3.3 uses a Taylor expansion of the score function, leading to $\hat{\beta}(v) - \beta(v) = -(l''_\beta(v, \beta^*(v)))^{-1}U(v, \beta(v))$, where $\beta^*(v)$ is on the line segment between $\hat{\beta}(v)$ and $\beta(v)$. The asymptotic variance of $n^{-1/2}h^{1/2}U(v, \beta(v))$ is shown to be $\nu_0\Sigma(v)$, which is the in probability limit of $\tilde{\Sigma}_n(\beta(v)) = n^{-1}h\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau (K_h(u-v))^2 J_{kn}(t,\beta(v))\, N_{ki}(dt,du)$. It can also be shown that $\hat{\Sigma}(v) \equiv -l''_\beta(v,\hat{\beta}(v))/n \xrightarrow{P} \Sigma(v)$ as $n \to \infty$. Thus, the asymptotic variance of $(nh)^{1/2}(\hat{\beta}(v)-\beta(v))$ can be estimated by $\hat{\Sigma}_1(v) = (l''_\beta(v,\hat{\beta}(v))/n)^{-1}\ \tilde{\Sigma}_n(\hat{\beta}(v))(l''_\beta(v,\hat{\beta}(v))/n)^{-1}$. An alternative estimator is $\hat{\Sigma}_2(v) = -\nu_0(l''_\beta(v,\hat{\beta}(v))/n)^{-1}$. It is easy to check that $\nu_0 = 3/5$ for Epanechnikov's kernel $K(x) = \frac{3}{4}(1-x^2),\ -1 < x < 1$. Simulations indicate that the two estimators have similar finite sample performance.

Theorem 3.3 will lead to the construction of pointwise confidence intervals for the vaccine efficacy. Simultaneous inference over $v \in [a,b]$ will be possible in terms of the estimate $\hat{B}(v) = \int_a^v \hat{\beta}(u)\, du$ of the cumulative regression coefficient $B(v) = \int_a^v \beta(u)\, du$. We have the following weak convergence result for $\hat{B}(v)$.

Theorem 3.4. *Under conditions (A.1)–(A.5),* $n^{1/2}(\hat{B}(v) - B(v))$ *converges weakly to a p-dimensional mean-zero Gaussian martingale,* $W_{\Sigma^{-1}}(v)$*, with continuous sample paths on* $v \in [a,b]$ *as* $n \to \infty$*. The covariance matrix of* $W_{\Sigma^{-1}}(v)$ *is* $\int_a^v \Sigma(u)^{-1}\, du$*, which can be consistently estimated by* $\hat{\Sigma}_{\hat{A}}(v)$ *defined by (3.7) with* $A(v) = (\Sigma(v))^{-1}$ *and* $\hat{A}(v) = (\hat{\Sigma}(v))^{-1}$.

3.4 Confidence Bands for Vaccine Efficacy

In the context of the vaccine trial application, let $z(t) = (z_1, z_2^T(t))^T$, where z_1 is the treatment group (1=vaccine; 0=placebo) and z_2 are other related explanatory variables. Let $\beta(v) = (\beta_1(v), \beta_2^T(v))^T$. Then the vaccine efficacy can be expressed as $\mathrm{VE}(v) = 1 - \exp(\beta_1(v))$. The estimated vaccine efficacy is $\widehat{\mathrm{VE}}(v) = 1 - \exp(\hat{\beta}_1(v))$. By Theorem 3.3 and the delta method, $(nh)^{1/2}(\widehat{\mathrm{VE}}(v) - \mathrm{VE}(v)) \xrightarrow{\mathcal{D}} N(0, \nu_0\sigma_1^2(v)\exp(2\beta_1(v)))$ for $v \in [a,b]$, where $\sigma_1^2(v)$ is the first element on the diagonal of $\Sigma^{-1}(v)$. Let $\hat{\sigma}^2_{\beta_1}(v)$ be the first element on

the diagonal of $\hat{\Sigma}_1(v)$. By the discussions on the consistent estimators for the asymptotic variance following Theorem 3.3, $\hat{\sigma}^2_{\beta_1}(v)$ is a consistent estimator for $\nu_0\sigma_1^2(v)$. A pointwise $100(1-\alpha)\%$ confidence band for $\mathrm{VE}(v)$ is given by

$$\widehat{\mathrm{VE}}(v) \pm (nh)^{-1/2} z_{\alpha/2} \hat{\sigma}_{\beta_1}(v) \exp(\hat{\beta}_1(v)), \qquad a \le v \le b, \tag{3.8}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of the standard normal distribution.

3.5 Goodness-of-fit Tests

Similar to Lin, Wei and Ying (1993), we derive the model checking test statistics based on the martingale residuals, which, in our case, is defined as

$$\hat{M}_{ki}(t,v) = \int_0^t \int_a^v [N_{ki}(ds,du) - Y_{ki}(s)\exp((\hat{\beta}(u))^T Z_{ki})\hat{\Lambda}_{0k}(ds,du)]. \tag{3.9}$$

$\hat{M}_{ki}(t,v)$ may be interpreted as the difference at time t between the observed and the predicted number of events with marks less than v for the ith subject in kth stratum. Thus the martingale residuals are informative about the model misspecification. It can be checked that $n^{-1/2}\sum_{k=1}^K \sum_{i=1}^{n_k} \hat{M}_{ki}(t,v) = o_p(1)$. This property is similar to that for the martingale residuals of the standard Cox model, where the sum of all martingale residuals is exactly zero. The difference here is caused by the kernel smoothing in a neighborhood of v.

Consider the test process $W(t,v,z) = (W_1(t,v,z), \ldots, W_K(t,v,z))$, where for $1 \le k \le K$,

$$W_k(t,v,z) = n^{-1/2} \sum_{i=1}^{n_k} g_k(Z_{ki},z)\hat{M}_{ki}(t,v), \tag{3.10}$$

and $g_k(Z_{ki},z)$ is a $1 \times r$-vector of known bounded functions of Z_{ki} and z. For example, one may take $g_k(Z_{ki},z) = f_k(Z_{ki})I(Z_{ki} \le z)$, where $f_k(\cdot)$ is a know function, $I(Z_{ki} \le z) = (I(Z_{1ki} \le z_1), \ldots, I(Z_{pki} \le z_p))$, and$Z_{jki}$ is the jth element of Z_{ki}. In this case $r = p$. If model (3.1) holds, the process $W(t,v,z)$ fluctuates randomly about zero. The distribution of $W(t,v,z)$ can be approximated using the Guassian multiplier method as we describe

next. Various test statistics can be constructed by selecting different weight functions $f_k(\cdot)$ and using different functionals of the process $W(t, v, z)$.

Let $S_{kg}^{(0)}(t, z, \beta) = n_k^{-1} \sum_{i=1}^{n_k} Y_{ki}(t) \exp\{\beta^T Z_{ki}(t)\} g_k(Z_{ki}, z)$ and $S_{kg}^{(1)}(t, z, \beta) = n_k^{-1}$ $\sum_{i=1}^{n_k} Y_{ki}(t) \exp\{\beta^T Z_{ki}(t)\} Z_{ki}(t) \otimes g_k(Z_{ki}, z)$, where $A \otimes B$ is the Kronecker product of matrices A and B.

Theorem 3.5. *Assuming conditions* $(A.1) - (A.5)$*, we have for* $1 \le k \le K$*,*

$$
\begin{aligned}
& W_k(t, v, z) \\
& = n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \int_0^\tau \int_a^v I(l = k) I(s \le t) \left[g_l(Z_{li}, z) - \frac{S_{lg}^{(0)}(s, z, \beta(u))}{S_l^{(0)}(s, \beta(u))} \right] M_{li}(ds, du) \\
& + n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \left\{ \int_0^\tau \int_a^v (R_k(t, u, z))^T (\Sigma(u))^{-1} \left[Z_{li}(s) - \frac{S_l^{(1)}(s, \beta(u))}{S_l^{(0)}(s, \beta(u))} \right] M_{li}(ds, du) \right\}^T \\
& + o_p(1)
\end{aligned}
\tag{3.11}
$$

where

$$
R_k(t, u, z) = (n_k/n) \int_0^t \left(\frac{S_k^{(1)}(s, \beta(u)) \otimes S_{kg}^{(0)}(s, z, \beta(u))}{S_k^{(0)}(s, \beta(u))} - S_{kg}^{(1)}(s, z, \beta(u)) \right) \lambda_{0k}(s, u)\, ds.
$$

The process $W(t, v, z) = (W_1(t, v, z), \ldots, W_k(t, v, z))$ converges weakly to a k dimensional mean zero Gaussion random process on $[0, \tau] \times [a, b] \times R^p$ as $n \to \infty$. Let $\{\xi_{li}, i = 1, \ldots, n_l, l = 1, \ldots, K\}$ be iid standard normal random variables. Using the Gaussian multiplier technique of Lin, Wei and Ying (1993), the distribution of $W(t, v, z)$ can be approximated by the distribution of $W^*(t, v, z) = (W_1^*(t, v, z), \ldots, W_K^*(t, v, z))$, where for

$1 \le k \le K$,

$$
\begin{aligned}
&W_k^*(t,v,z) \\
&= n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \int_0^{\tau} \int_a^{v} I(l=k) I(s \le t) \left[g_l(Z_{li}, z) - \frac{S_{lg}^{(0)}(s, z, \hat{\beta}(u))}{S_l^{(0)}(s, \hat{\beta}(u))} \right] \xi_{li} \, N_{li}(ds, du) \\
&+ n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \left\{ \int_0^{\tau} \int_a^{v} (\hat{R}_k(t,u,z))^T (\hat{\Sigma}(u))^{-1} \left[Z_{li}(s) - \frac{S_l^{(1)}(s, \hat{\beta}(u))}{S_l^{(0)}(s, \hat{\beta}(u))} \right] \xi_{li} N_{li}(ds, du) \right\}^T,
\end{aligned}
\tag{3.12}
$$

where

$$
\hat{R}_k(t,u,z) = (n_k/n) \int_0^t \left(\frac{S_k^{(1)}(s, \hat{\beta}(u)) \otimes S_{kg}^{(0)}(s, z, \hat{\beta}(u))}{S_k^{(0)}(s, \hat{\beta}(u))} - S_{kg}^{(1)}(s, z, \hat{\beta}(u)) \right) d\hat{\Lambda}_{u,0k}(s),
$$

and

$$
\hat{\Lambda}_{u,0k}(t) = \int_0^t \int_0^1 K_h(u-x) \frac{dN_{k\cdot}(ds, dx)}{n_k S_k^{(0)}(s, \hat{\beta}(x))} \tag{3.13}
$$

is the estimator for the mark-specific cumulative baseline function $\Lambda_{u,0k}(t) = \int_0^t \lambda_{0k}(s,u)\, ds$.

Various test statistics based on the functionals of the process $W(t,v,z)$ can be constructed to check the lack of fit model (3.1). Let $(g_k^{(1)}(x,z), \ldots, g_k^{(r)}(x,z))$ and $(W_k^{(1)}(t,v,z), \ldots, W_k^{(r)}(t,v,z))$ be the r components of $g_k(x,z)$ and $W_k(t,v,z)$, respectively. We consider the following supremum test statistic to test the overall fit of the model:

$$
T = \sup_{1 \le k \le K} \sup_{1 \le j \le r} \sup_{(t,v,z) \in \mathcal{C}} |W_k^{(j)}(t,v,z)|, \tag{3.14}
$$

where $\mathcal{C} = [0,\tau] \times [a,b] \times \mathbb{R}^p$ and $[a_k, b_k] \in (0,1)$. Let $T^* = \sup_{1 \le k \le K} \sup_{1 \le j \le r} \sup_{(t,v,z) \in \mathcal{C}} |W_k^{*(j)}(t,v,z)|$, where $W_k^{*(j)}(t,v,z)$ is the jth component of $W_k^*(t,v,z)$. The distribution of T can be approximated by repeatedly generating iid sets of standard normal random variables $\{\xi_{li}\}$. We reject model (3.1) at significance level α if T is greater than the upper α quantile of T^*. When $g_k(Z_{ki}, z) = Z_{ki}$, T is the supremum test based on the score process. The selection $g_k(Z_{ki}, z) = I(Z_{ki} \le z)$, where all components of Z other than the jth component are set to be ∞, can be used to test the functional form of the jth

component of covariate Z. On the other hand, one can take $g_k(Z_{ki}, z) = (Z_{ki}, I(Z_{ki} \leq z))$ to test the overall fit of model (3.1).

3.6 Simulation Study

In this section, we conduct a simulation study to check the finite sample performance of the proposed testing procedure. The size of the test is examined using the following simple mark-specific proportional hazards model:

$$\lambda_k(t, v|z) = \exp\{\gamma v + (\alpha + \beta v)z\}, \quad t \geq 0,\ 0 \leq v \leq 1, \tag{3.15}$$

where α, β and γ are constants, z takes value 0 or 1 as a treatment indicator. We consider the simple case where $k = 1$. The mark-specific baseline function is $\lambda_{0k}(t, v) = \exp(\gamma v)$. We generate covariates Z_{ki} from Bernoulli distribution with $P(Z_{ki} = 1) = 0.5$. The censoring times are generated from an exponential distribution, independent of (T_{ki}, V_{ki}), with the censoring rates ranging from 20% to 30%. Set the follow-up time $\tau = 2.0$. The censoring rate before τ is around 10%. We set the interval of analyses for v as $[a, b] = [.1, .9]$. The observed failure times with marks outside the interval $[a, b]$ can also be used since the smoothing at v takes the cases with marks in its h-neighborhood. The Epanechnikov kernel $K(x) = .75(1 - x^2)I\{|x| \leq 1\}$ is used throughout. Table 3.1 shows the empirical sizes of the test under different choice of α, β and γ, for sample sizes of $n = 200$, $n = 350$ and 500 and bandwidths $h = 0.2,\ 0.25,\ 0.3$. The empirical sizes are calculated based on 1000 simulations and 500 Gaussian multiplier samples. They are very close to 0.05.

To evaluate the power of proposed test, consider the model

$$\lambda_k(t, v|z_{ki}(t)) = \lambda_{0k}(t, v)exp\{(\beta v - ct)z_{ki}\}, \tag{3.16}$$

where $\lambda_{0k}(t, v) = k \exp(ct - \beta v)$ for $k = 1, \ldots, K$. Again we take Z_{ki} as a Bernoulli random variable with $P(Z_{ki} = 0.5)$. For simplicity here we consider only one failure cause $K = 1$. Model (3.16) is not a mark-specific proportional hazard model since the hazard ratio $\lambda_k(t, v|Z_{ki} = 1)/\lambda_k(t, v|Z_{ki} = 0) = exp\{ct - \beta v\}$ changes with time. We set $\beta = 0.1$

and consider $c = 0.8$, 1.0, 1.2 and 1.4. As c increases, the hazard ratio under (3.16) increases faster with t, which represents an increases departure from the null hypothesis. For each of the above selected c, random right censoring times are generated from an exponential distribution, independent of (T_{ki}, V_{ki}), to yield around 30% of censoring. Again we set $\tau = 2.0$ and $[a, b] = [0.1, 0.9]$. sample sizes of $n = 200$, $n = 350$ and 500 are studied. The empirical power of the test at the significance level 0.05 under (3.16) for $c = 0.8$, 1.0, 1.2 and 1.4, $n = 300$, 500, 800, and $h = 0.20$, 0.25 and 0.30 are given in Table 3.2. Each entry of the table is based on 1000 simulations and 500 Gaussian multiplier samples. The power of the test increases with c, and also increases with sample size. The limited simulation study demonstrates the validity of the proposed goodness of fit testing procedure. The test provides a valuable tool to check the adequacy of the mark-specific proportional hazard model (3.1).

3.7 Complements

In this section, we give the derivations of the main results presented in previous sections of this chapter. Proof of Theorem 3.1.

It is easy to check that the conditions of Lemma 1 of Sun and Wu (2005) are satisfied under Condition A. It follows that $\tilde{W}_A(v)$ converges weakly to a vector of continuous mean-zero Gaussian random processes, $W_A(v)$, $v \in [a, b]$. Now we show that $W_A(v)$ has independent increments. Let $w_{ki}(t, v) = \int_a^v \int_0^t A(u)\big[Z_{ki}(t) - s_k^{(1)}(t, \beta(u))/s_k^{(0)}(t, \beta(u))\big] M_{ki}(dt, du)$. Then $\tilde{W}_A(v) = n^{-1/2} \sum_{k=1}^K \sum_{i=1}^{n_k} w_{ki}(\tau, v)$. For $a \le v_1 \le v_2 \le b$, the covariance matrix of $W_A(v_1)$ and $W_A(v_2) - W_A(v_1)$ is equal to $\sum_{k=1}^K \frac{n_k}{n} E\{w_{ki}(\tau, v_1)(w_{ki}(\tau, v_2) - w_{ki}(\tau, v_1))^T\}$. Since $M_{ki}(t, v_1)$ and $M_{ki}(t, v_2) - M_{ki}(t, v_1)$, $0 \le t \le \tau$, are orthogonal square integrable martingales, it follows that $w_{ki}(t, v_1)$ and $w_{ki}(t, v_2) - w_{ki}(t, v_1)$, $0 \le t \le \tau$, orthogonal square integrable martingales. Hence $E\{w_{ki}(\tau, v_1)(w_{ki}(\tau, v_2) - w_{ki}(\tau, v_1))^T\} = 0$. So $W_A(v)$, $v \in [a, b]$, is a vector of mean-zero Gaussian random processes with independent increments.

Further, the covariance matrix of $W_A(v)$ is equal to

$$
\begin{aligned}
& \sum_{k=1}^K P_k E\{w_{ki}(\tau, v)(w_i(\tau, v))^T\} \\
= & \sum_{k=1}^K P_k E\bigg\{ \int_a^v \int_0^\tau A(u)\bigg[Z_{ki}(t) - \frac{s_k^{(1)}(t, \beta(u))}{s_k^{(0)}(t, \beta(u))}\bigg]^{\otimes 2} A(u)\, N_{ki}(dt, du)\bigg\} \\
= & E\bigg\{ \int_a^v \int_0^\tau A(u)\bigg[Z_{ki}(t) - \frac{s_k^{(1)}(t, \beta(u))}{s_k^{(0)}(t, \beta(u))}\bigg]^{\otimes 2} A(u) y_k(t|Z_{ki}(t))\lambda_k(t, u|Z_{ki}(t))\, dtdu\bigg\} \\
= & \int_a^v A(u) E\bigg\{ \int_0^\tau \bigg[Z_{ki}(t) - \frac{s_k^{(1)}(t, \beta(u))}{s_k^{(0)}(t, \beta(u))}\bigg]^{\otimes 2} y_k(t|Z_{ki}(t))\lambda_k(t, u|Z_{ki}(t))\, dt\bigg\} A(u)\, du \\
= & \int_a^v A(u)\Sigma(u)A(u)\, du.
\end{aligned}
$$

This completes the proof of Theorem 3.1. *Q.E.D.*

Proof of Theorem 3.2.

We shall prove Theorem 3.2 by verifying the conditions of Lemma 1 of Sun, Gilbert, Mckeague (2009).

Let

$$\begin{aligned}
\eta_n(u,\theta) &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^u\int_0^\tau \left[\theta^T Z_{ki}(t) - \log(S_k^{(0)}(t,\theta))\right] N_{ki}(dt,du) \\
\xi_n(u,\theta) &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^u\int_0^\tau \left[\theta^T Z_{ki}(t) - \log(s_k^{(0)}(t,\theta))\right] N_{ki}(dt,du) \\
Q_n(v,\theta) &= n^{-1}l(v,\theta) + n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\log n_k \int_0^1 K_h(u-v)\, N_{ki}(\tau,du).
\end{aligned}$$

Then by Condition A, $\eta_n(v,\theta) = \xi_n(v,\theta) + O_p(n^{-1/2})$ and

$$Q_n(v,\theta) = \int_0^1 K_h(u-v)\,\eta_n(du,\theta) = \int_0^1 K_h(u-v)\,\xi_n(du,\theta) + O_p(n^{-1/2}h^{-1}),$$

uniformly in $(v,\theta) \in [0,1]\times[-M,M]$, for $M>0$. Let

$$Q(v,\theta) = \sum_{k=1}^{K} P_k E\bigg[\int_0^\tau \left[\theta^T Z_{ki}(t) - \log(s_k^{(0)}(t,\theta))\right]\lambda_{k0}(t,v)\exp(\beta^T(v)Z_{ki}(t))Y_{ki}(t)\,dt\bigg].$$

Following similar steps of the proof of Theorem 1 of Sun, Gilbert and Mckeague (2009), $\beta(v)$ is the well separated point of maximum of $Q(v,\theta)$ for $v\in[0,1]$ uniformly in $(v,\theta)\in [a,b]\times[-M,M]$, and $Q_n(v,\hat{\beta}(v)) \geq Q_n(v,\beta(v))$. *Q.E.D.*

Proof of Theorem 3.3.

In the proof of this theorem, we set $\beta = \beta(v)$ for simplicity. Note that under Condition A, using a second order Taylor expansion for $\lambda_k(t,u|Z_{ki}(t))$ in the neighborhood of v, we have

$$\begin{aligned}
&n^{-1/2}\bigg|\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau K_h(u-v)\bigg[Z_{ki}(t) - \frac{S_k^{(1)}(t,\beta)}{S_k^{(0)}(t,\beta)}\bigg]Y_{ki}(t)[\lambda_k(t,v|Z_{ki}(t)) - \\
&\lambda_k(t,u|Z_i(t))]\,dtdu\bigg| = O_p(n^{1/2}h^2),
\end{aligned}$$

uniformly in $v \in [0,1]$. It follows that

$$
\begin{aligned}
& n^{-1/2}U(v,\beta) \\
& = n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau K_h(u-v)\left[Z_{ki}(t)-\frac{S_k^{(1)}(t,\beta)}{S_k^{(0)}(t,\beta)}\right][N_{ki}(dt,du)- \\
& Y_{ki}(t)\lambda_k(t,v|Z_{ki}(t))\,dtdu] \\
& = n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau K_h(u-v)\left[Z_{ki}(t)-\frac{S_k^{(1)}(t,\beta)}{S_k^{(0)}(t,\beta)}\right][N_{ki}(dt,du)- \\
& Y_{ki}(t)\lambda_k(t,u|Z_{ki}(t))\,dtdu] + O_p(n^{1/2}h^2),
\end{aligned}
$$

uniformly in $v \in [0,1]$.

Next, we show that for each v, $n^{-1/2}h^{1/2}U(v,\beta)$ converges weakly to a normal distribution. By Lemma 2 of Sun, Gilbert and Mckeague (2009), $n^{-1/2}M_k(t,v)$ converges weakly to a mean-zero Gaussian process for $1 \le k \le K$. By Condition A, $\|S_k^{(j)}(t,\beta)-s_k^{(j)}(t,\beta)\| = o_p(n^{-1/2+\delta})$, uniformly in t for $j=0,1$ and $1 \le k \le K$, for some $0<\delta<1/2$. Note that $n^{-1/2+\delta}h^{-1/2} = o(1)$ for $\delta = 1/4$ as $nh^2 \to \infty$. We have $h^{1/2}K_h(u-v)\|S_k^{(j)}(t,\beta)-s_k^{(j)}(t,\beta)\|$ goes to zero in probability for $1 \le k \le K$. Applying Lemma 2 of Gilbert, McKeague and Sun (2006), we have

$$
\begin{aligned}
n^{-1/2}h^{1/2}U(v,\beta) &= n^{-1/2}h^{1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau K_h(u-v)\left[Z_{ki}(t)-\frac{s_k^{(1)}(t,\beta)}{s_k^{(0)}(t,\beta)}\right] \\
&\quad [N_{ki}(dt,du)-Y_{ki}(t)\lambda_k(t,u|Z_{ki}(t))\,dtdu] + O_p(n^{1/2}h^{5/2}) + o_p(1) \\
&= n^{-1/2}h^{1/2}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\int_0^1\int_0^\tau K_h(u-v)\left[Z_{ki}(t)-\frac{s_k^{(1)}(t,\beta(u))}{s_k^{(0)}(t,\beta(u))}\right] \\
&\quad [N_{ki}(dt,du)-Y_k(t)\lambda_k(t,u|Z_i(t))\,dtdu] + O_p(n^{1/2}h^{5/2}) + o_p(1) \\
&= h^{1/2}\int_0^1 K_h(u-v)\,\tilde{W}_I(du) + O_p(n^{1/2}h^{5/2}) + o_p(1), \qquad (3.17)
\end{aligned}
$$

where $\tilde{W}_I(v)$ is defined in (3.6) with $A = I$ and $a = 0$.

Since $\tilde{W}_I(v) \xrightarrow{\mathcal{D}} W_I(v)$ by Theorem 3.1. By the almost sure representation theorem (Shorack and Wellner, 1986), there exist $\tilde{W}_I^*(v)$ and $W_I^*(v)$ on some probability space that have the same distributions and sample paths as $\tilde{W}_I(v)$ and $W_I(v)$, respectively, such that $\tilde{W}_I^*(v) \xrightarrow{\text{a.s.}} W_I^*(v)$ uniformly in $v \in [0,1]$. Hence $\int_0^1 K_h(u-v)\,\tilde{W}_I^*(du) = \int_0^1 K_h(u-$

$v)\, W_I^*(du) + O_p(n^{-1/2}h^{-1})$ by integration by parts since $K(\cdot)$ has bounded variation. It follows that

$$h^{1/2}\int_0^1 K_h(u-v)\,\tilde{W}_I(du) \stackrel{D}{=} h^{1/2}\int_0^1 K_h(u-v)\,\tilde{W}_I^*(du)$$
$$= h^{1/2}\int_0^1 K_h(u-v)\,W_I^*(du) + O_p(n^{-1/2}h^{-1/2}). \tag{3.18}$$

Since $W_I^*(v)$ is a Gaussian martingale with covariance matrix of $\int_0^v \Sigma(u)\,du$, $h^{1/2}\int_0^1 K_h(u-v)\,W_I^*(du)$ is a mean zero Gaussian random vector with covariance matrix equal to $h\int_0^1 K_h^2(u-v)\Sigma(u)\,du \to \nu_0\Sigma(v)$ as $h \to 0$. Hence, $h^{1/2}\int_0^1 K_h(u-v)\,\tilde{W}_I(du) \xrightarrow{\mathcal{D}} N(0,\nu_0\Sigma(v))$ as $h \to 0$, $nh \to \infty$. By the Slutsky theorem, $n^{-1/2}h^{1/2}U(v,\beta)$ converges weakly to $N(0,\nu_0\Sigma(v))$ as $nh^2 \to \infty$ and $nh^5 \to 0$.

Note that $U(v,\hat{\beta}) - U(v,\beta) = l''_\beta(v,\beta^*(v))(\hat{\beta}(v)-\beta(v))$, where $\beta^*(v)$ is on the line segment between $\hat{\beta}(v)$ and $\beta(v)$. By Condition A and the uniform consistency of $\hat{\beta}(v)$ on $v \in [a,b] \subset (0,1)$, we have $n^{-1}l''_\beta(v,\beta^*(v)) = -\Sigma(v) + o_p(1)$, uniformly in $v \in [a,b]$ for $0 < \delta < 1/2$. Hence,

$$n^{1/2}h^{1/2}(\hat{\beta}(v)-\beta(v)) = -(l''_\beta(v,\beta^*(v))/n)^{-1}n^{-1/2}h^{1/2}U(v,\beta)$$
$$= (\Sigma(v))^{-1}n^{-1/2}h^{1/2}U(v,\beta) + o_p(1), \tag{3.19}$$

uniformly in $v \in [a,b]$. It follows that $(nh)^{1/2}(\hat{\beta}(v)-\beta(v)) \xrightarrow{\mathcal{D}} N(0,\nu_0\Sigma(v)^{-1})$ as $nh^2 \to \infty$ and $nh^5 \to 0$. *Q.E.D.*

Proof of Theorem 3.4.

From (3.17) and the first line of (3.19), we have, for $v \in [a,b]$,

$$\int_a^v n^{1/2}(\hat{\beta}(u)-\beta(u))\,du = -\int_a^v (\Sigma(u))^{-1}\int_0^1 K_h(x-u)\tilde{W}_I(dx)du + o_p(1).$$

Exchanging the order of integration and by the compact support of the kernel function

$K(\cdot)$ on $[-1,1]$, we have

$$\begin{aligned}
&\int_a^v n^{1/2}(\hat{\beta}(u)-\beta(u))\,du = -\int_0^1 \left[\int_a^v (\Sigma(u))^{-1}K_h(x-u)\,du\right]\tilde{W}_I(dx)+o_p(1)\\
&= -\int_{a+h}^{v-h}\left[\int_a^v (\Sigma(u))^{-1}K_h(x-u)\,du\right]\tilde{W}_I(dx)\\
&-\int_{a-h}^{a+h}\left[\int_a^v (\Sigma(u))^{-1}K_h(x-u)\,du\right]\tilde{W}_I(dx)\\
&-\int_{v-h}^{v+h}\left[\int_a^v (\Sigma(u))^{-1}K_h(x-u)\,du\right]\tilde{W}_I(dx)+o_p(1). \qquad (3.20)
\end{aligned}$$

By Theorem 3.1, $\tilde{W}_I(x)$ converges weakly to a mean-zero Gaussian process with continuous paths. Under the assumption (A.4), $\int_a^v(\Sigma(u))^{-1}K_h(x-u)\,du$ has bounded variation and converges uniformly to $\Sigma(x)^{-1}$ for $x\in(a+h,v-h)$. By Lemma 2 of Gilbert, McKeague and Sun (2006), the first term in (3.20) is equal to $-\int_a^v(\Sigma(x))^{-1}\tilde{W}_I(dx)+o_p(1)$. Similar arguments lead to the second and the third terms in (3.20) to be $o_p(1)$. Hence

$$\int_a^v n^{1/2}(\hat{\beta}(u)-\beta(u))\,du = -\int_a^v(\Sigma(x))^{-1}\tilde{W}_I(dx)+o_p(1) = -\tilde{W}_{\Sigma^{-1}}(v)+o_p(1),$$

which converges weakly to a p-dimensional mean-zero Gaussian martingale, $W_{\Sigma^{-1}}(v)$, with continuous paths. The covariance matrix of $W_{\Sigma^{-1}}(v)$ equals $\mathrm{Cov}(W_{\Sigma^{-1}}(v)) = \int_a^v \Sigma(u)^{-1}\Sigma(u)\,\Sigma(u)^{-1}\,du = \int_a^v \Sigma(u)^{-1}\,du$. *Q.E.D.*

Proof of Theorem 3.5. Note that:

$$\begin{aligned}
\hat{M}_{ki}(t,v) &= M_{ki}(t,v) - \int_0^t\int_a^v Y_{ki}(s)\exp((\hat{\beta}(u))^T Z_{ki})[\hat{\Lambda}_{0k}(ds,du)-\Lambda_{0k}(ds,du)]\\
&\quad -\int_0^t\int_a^v Y_{ki}(s)[\exp((\hat{\beta}(u))^T Z_{ki})-\exp((\beta(u))^T Z_{ki})]\Lambda_{0k}(ds,du).
\end{aligned} \qquad (3.21)$$

Consider the approximation decomposition

$$exp(\hat{\beta}(u)^T Z_{ki}) - exp(\beta(u)^T Z_{ki}) = (\hat{\beta}(u)-\beta(u))^T exp(\beta(u)^T Z_{ki}) + o(\|\hat{\beta}(u)-\beta(u)\|),$$

and

$$\begin{aligned}
\hat{\Lambda}_{0k}(t,v) - \Lambda_{0k}(t,v) &= \int_0^t \int_a^v \frac{dN_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,\hat{\beta}(u))} - \Lambda_{0k}(t,v) \\
&= \int_0^t \int_a^v \left[\frac{1}{n_k S_k^{(0)}(s,\hat{\beta}(u))} - \frac{1}{n_k S_k^{(0)}(s,\beta(u))}\right] N_{k\cdot}(ds,du) \\
&\quad + \int_0^t \int_a^v \frac{M_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,\beta(u))} + o_p(n_k^{-1/2}) \\
&= \int_0^t \int_a^v \frac{(S_k^{(1)}(s,\beta(u)))^T(\beta(u)-\hat{\beta}(u))}{n_k S_k^{(0)}(s,\hat{\beta}(u)) S_k^{(0)}(s,\beta(u))} N_{k\cdot}(ds,du) \\
&\quad + \int_0^t \int_a^v \frac{M_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,\beta(u))} + o_p(n_k^{-1/2}) \\
&= \int_0^t \int_a^v \frac{(S_k^{(1)}(s,\beta(u)))^T(\beta(u)-\hat{\beta}(u))\lambda_{0k}(s,u)}{S_k^{(0)}(s,\beta(u))}\, dsdu \\
&\quad + \int_0^t \int_a^v \frac{M_{k\cdot}(ds,du)}{n_k S_k^{(0)}(s,\beta(u))} + o_p(n_k^{-1/2}), \qquad (3.22)
\end{aligned}$$

where $M_{k\cdot}(t,v) = \sum_{i=1}^{n_k} M_{ki}(t,v)$.

From (3.21), we have

$$\begin{aligned}
W_k(t,v,z) &= n^{-1/2}\sum_{i=1}^{n_k} g_k(Z_{ki},z) M_{ki}(t,v) \\
&\quad + n_k n^{-1/2} \int_0^t \int_a^v \frac{S_{kg}^{(0)}(s,z,\beta(u))[(S_k^{(1)}(s,\beta(u)))^T(\hat{\beta}(u)-\beta(u))]}{S_k^{(0)}(s,\beta(u))} \lambda_{0k}(s,u)\, dsdu \\
&\quad - n^{-1/2} \int_0^t \int_a^v \frac{S_{kg}^{(0)}(s,z,\beta(u))}{S_k^{(0)}(s,\beta(u))} M_{k\cdot}(ds,du) \\
&\quad - n_k n^{-1/2} \int_0^t \int_a^v (\hat{\beta}(u)-\beta(u))^T S_{kg}^{(1)}(s,z,\beta(u)) \lambda_{0k}(s,u)\, dsdu + o_p(1).
\end{aligned} \qquad (3.23)$$

Following (3.17) and (3.19) in the proof of Theorem 3.3, the second term of (3.23) equals

to

$$
\begin{aligned}
&(n_k/n)\Bigg\{\int_0^t\int_a^v \frac{[S_k^{(1)}(s,\beta(u))\otimes S_{kg}^{(0)}(s,z,\beta(u))]^T\lambda_{0k}(s,u)}{S_k^{(0)}(s,\beta(u))}(\Sigma(u))^{-1}\\
&\quad\times\Bigg[\int_0^1 K_h(x-u)\tilde{W}_I(dx)\Bigg]\,dsdu\Bigg\}^T+o_p(1)\\
&=(n_k/n)\Bigg\{\int_0^t\int_a^v \frac{[S_k^{(1)}(s,\beta(u))\otimes S_{kg}^{(0)}(s,z,\beta(u))]^T\lambda_{0k}(s,u)}{S_k^{(0)}(s,\beta(u))}(\Sigma(u))^{-1}\,ds\tilde{W}_I(du)\Bigg\}^T+o_p(1)\\
&=(n_k/n)\Bigg\{\int_a^v\Bigg[\int_0^t \frac{[S_k^{(1)}(s,\beta(u))\otimes S_{kg}^{(0)}(s,z,\beta(u))]^T\lambda_{0k}(s,u)}{S_k^{(0)}(s,\beta(u))}\,ds\Bigg](\Sigma(u))^{-1}\tilde{W}_I(du)\Bigg\}^T\\
&+o_p(1)
\end{aligned}
\tag{3.24}
$$

Similarly, the fourth term of (3.23) is equal to

$$
\Bigg\{\int_a^v\Bigg[\int_0^t (n_k/n)(S_{kg}^{(1)}(s,z,\beta(u)))^T\lambda_{0k}(s,u)\,ds\Bigg](\Sigma(u))^{-1}\tilde{W}_I(du)\Bigg\}^T+o_p(1).
\tag{3.25}
$$

Combining (3.23), (3.24) and (3.25), we have

$$
\begin{aligned}
&W_k(t,v,z)\\
&=n^{-1/2}\sum_{i=1}^{n_k}\int_0^t\int_a^v\left[g_k(Z_{ki},z)-\frac{S_{kg}^{(0)}(s,z,\beta(u))}{S_k^{(0)}(s,\beta(u))}\right]M_{ki}(ds,du)\\
&+(n_k/n)\Bigg\{\int_a^v\Bigg[\int_0^t\left(\frac{S_k^{(1)}(s,\beta(u))\otimes S_{kg}^{(0)}(s,z,\beta(u))}{S_k^{(0)}(s,\beta(u))}-S_{kg}^{(1)}(s,z,\beta(u))\right)^T\lambda_{0k}(s,u)\,ds\Bigg]\\
&\times(\Sigma(u))^{-1}\tilde{W}_I(du)\Bigg\}^T+o_p(n_k^{1/2})\\
&=n^{-1/2}\sum_{i=1}^{n_k}\int_0^t\int_a^v\left[g_k(Z_{ki},z)-\frac{S_{kg}^{(0)}(s,z,\beta(u))}{S_k^{(0)}(s,\beta(u))}\right]M_{ki}(ds,du)\\
&+n^{-1/2}\sum_{l=1}^{K}\sum_{i=1}^{n_l}\Bigg\{\int_0^\tau\int_a^v(R_k(t,u,z))^T(\Sigma(u))^{-1}\left[Z_{li}(s)-\frac{S_l^{(1)}(s,\beta(u))}{S_l^{(0)}(s,\beta(u))}\right]M_{li}(ds,du)\Bigg\}^T\\
&+o_p(1)
\end{aligned}
$$

$$= n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \int_0^\tau \int_a^v I(l=k) I(s \le t) \left[g_l(Z_{li}, z) - \frac{S_{lg}^{(0)}(s, z, \beta(u))}{S_l^{(0)}(s, \beta(u))} \right] M_{li}(ds, du)$$
$$+ n^{-1/2} \sum_{l=1}^{K} \sum_{i=1}^{n_l} \left\{ \int_0^\tau \int_a^v (R_k(t, u, z))^T (\Sigma(u))^{-1} \left[Z_{li}(s) - \frac{S_l^{(1)}(s, \beta(u))}{S_l^{(0)}(s, \beta(u))} \right] M_{li}(ds, du) \right\}^T$$
$$+ o_p(1). \tag{3.26}$$

Further, we note that $S_{lg}^{(0)}(s, z, \beta(u))$, $S_{lg}^{(1)}(s, z, \beta(u))$, $S_l^{(0)}(s, \beta(u))$ and $S_l^{(1)}(s, z, \beta(u))$ in (3.26) can be replaced with their expectations $s_{lg}^{(0)}(s, z, \beta(u))$, $s_{lg}^{(1)}(s, z, \beta(u))$, $s_l^{(0)}(s, \beta(u))$ and $s_l^{(1)}(s, z, \beta(u))$ respectively. The resulted process is the sum of iid terms involving the integrations with respect to $M_{li}(s, u)$, which is equivalent to the one before the change by Lemma 2 of Gilbert, Mckeague and Sun (2006). Then it follows from the limit theorems of empirical process theory that $W(t, v, z)$ converges weakly to a mean zero Gaussian process. *Q.E.D.*

Table 3.1: Empirical size of goodness of fit test at the significant level 0.05

Model	(α, β, γ)	n	h	size
M_1	$(0., 0., .3)$	200	0.20	6.2
			0.25	6.6
			0.30	6.4
		350	0.20	6.3
			0.25	6.4
			0.30	5.9
		500	0.20	6.4
			0.25	6.3
			0.30	6.1
M_2	$(-.6, .6, .3)$	200	0.20	5.7
			0.25	5.2
			0.30	5.4
		350	0.20	4.9
			0.25	5.4
			0.30	5.4
		500	0.20	5.3
			0.25	5.3
			0.30	5.3
M_3	$(-.69, 0, .3)$	200	0.20	5.5
			0.25	5.9
			0.30	5.9
		350	0.20	6.7
			0.25	6.8
			0.30	6.6
		500	0.20	6.4
			0.25	6.3
			0.30	6.4
M_4	$(-1.5, 1.5, .3)$	200	0.20	4.0
			0.25	4.0
			0.30	4.2
		350	0.20	4.2
			0.25	3.9
			0.30	4.1
		500	0.20	4.4
			0.25	4.6
			0.30	5.0

Table 3.2: The power of goodness of fit test at significant level 0.05

Model	c	n	h	power
H_1	0.8	300	0.20	49.4
			0.25	49.3
			0.30	49.2
		500	0.20	71.3
			0.25	71.1
			0.30	70.6
		800	0.20	88.9
			0.25	88.4
			0.30	87.9
H_2	1.0	300	0.20	59.6
			0.25	60.3
			0.30	58.4
		500	0.20	81.3
			0.25	80.5
			0.30	80.8
		800	0.20	95.5
			0.25	95.3
			0.30	95.1
H_3	1.2	300	0.20	65.8
			0.25	64.8
			0.30	64.2
		500	0.20	87.3
			0.25	86.9
			0.30	86.4
		800	0.20	98.1
			0.25	97.7
			0.30	97.4
H_4	1.4	300	0.20	70.2
			0.25	69.4
			0.30	68.5
		500	0.20	90.5
			0.25	90.7
			0.30	90.4
		800	0.20	98.6
			0.25	98.3
			0.30	98.2

REFERENCES

[1] Aly, E. E., Kochar, S. C. and Mckeague, I. W. (1994). Some tests for comparing cumulative incidence functions and cause-specific hazard rates. *Journal of the American Statistical Association*, **89**, 994-999.

[2] Anderson, P. K. and Gill, R. (1982). Cox's regression model for counting process: a large sample study. *Ann. Statist.* , **10**, 1100-20.

[3] Buchbinder, S. P., Mehrotra, D. V., and Duerr, A. et al. (2008). Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): A double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet* , **372**, 1881–1893.

[4] Buus, S., Lauemoller, S. L., Worning, P. et al. (2003). Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. *Tissue Antigens*, **62**, 378–384.

[5] Flynn, N. M., Forthal, D. N., Harro, C. D., Judson, F. N., Mayer, K. H., Para, M. F., and the rgp120 HIV Vaccine Study Group. (2005). Placebo-controlled trial of a recombinant glycoprotein 120 vaccine to prevent HIV infection. *J Infect Dis*, **191**, 654–665.

[6] Gichangi, A., Vach, W. (2006). The analysis of competing risks data: a guided tour. *Statistics in Medicine (Forthcoming)*.

[7] Gilbert, P. B. (2000). Large sample theory of maximum likelihood estimates in semiparametric biased sampling models. *Annals of Statistics*, **28**, 151-194.

[8] Gilbert,P. B., DeGruttola,V., Hammer,S. M., and Kuritzkes,D. R. (2001). Virological and regimen termination surrogate endpoints in AIDS clinical trials. *Journal of the American Medical Association*, **8**, 775–782.

[9] Gilbert, P. B., McKeague, I. W., Sun, Y. (2004). Tests for comparing mark-specific hazards and cumulative incidence functions. *Lifetime data analysis*, **10**, 5-28.

[10] Gilbert, P. B., Sun, Y. (2005). Failure time analysis of HIV vaccine effects on viral load and antiretroviral therapy initiation. *Biostatistics*, **6**, 374-94.

[11] Gilbert, P. B., McKeague, I. W. and Sun, Y. (2008). The 2-sample problem for failure rates depending on a continuous mark: an application to vaccine efficacy. *Biostatistics*, **9**, 2, 263-276.

[12] Gray, G. E., Bekker, L., Churchyard, G. J., et al. (2009). Did unblinding affect HIV risk behaviour and risk perception in the HVTN503/Phambili study? *Retrovirology*, **6**, (Suppl 3), 209.

[13] Heckerman, D., Kadie, C., and Listgarten, J. (2007). Leveraging information across HLA alleles/supertypes improves epitope prediction. *J Comput Biol*, **14**, 736-46.

[14] Hirsch, M. S., Frangoise, B., D'Aquila, R. T., et al. (2000). Antiretroviral Drug Resistance Testing in Adult HIV-1 Infection. *The Journel of American Medical Associaton*, **283**, 2417–2426.

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

[16] Huang, Y. and Louis, T. A. (1998). Nonparametric estimation of the joint distribution of survival time and mark variables. *Biometrika*, **85**, 785-798.

[17] Kalbfeisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, New York, Wiley.

[18] Kuk, A. Y. C. (1992). A semiparametric mixture model for the analysis of competing risks data. *Australian Journal of Statistics*, **34**, 169–180 .

[19] 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.

[20] Lee, T. E. and Wang, J. (2003). *Statistical Methods for Survival Data Analysis*, New York, Wiley.

[21] Lin, D. Y., Wei, L. J, and Ying, Z. (1993) Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. *Biometrika*, **80**, 557-572.

[22] Lunn, M. and McNeil, D. (1995). Applying Cox regression to competing risks. *Biometrics*, **51**, 524–532.

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

[24] Olschewski, M. and Schumacher, M. (1990). Statistical analysis of quality of life data in cancer clinical trials. *Statistics in Medicine*, **9**, 749–763.

[25] Pintilie, M. (2006). *Competing Risks: A Practical Perspective*, New York, Wiley.

[26] Pitisuttithum, P., Gilbert, P., Gurwith, M. et al. (2006). Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. *J Infect Dis*, **194**, 1661–1671.

[27] Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Fluornoy, N., Farewell, V. T. and Breslow, N.E. (1978). The analysis of failure time in the presence of competing risks. *Biometrics*, **34**, 541–554.

[28] Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S. et al. (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med*, **361**, 2209–2220.

[29] Scheike, T., Zhang, M. and Gerds, T. A. (2008). Predicting cumulative incidence probability by direct binomial regression.*Biometrika*, **95**, 205–220.

[30] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York, Wiley.

[31] Spiekerman, C. F and Lin, D. Y. (1996). Checking the marginal Cox model for correlated failure time data. *Biometrika*, **83**, 143-156.

[32] Sun, Y. (2001). Generalized nonparametric test procedures for comparing multiple cause-specific hazard rates. *Journal of Nonparametric Statistics*, **13**, 171–207.

[33] Sun, Y. and Wu, H. (2005). Semiparametric time-varying coefficients regression model for longitudinal data. *Scandinavian Journal of Statistics*, **32**, 21–47.

[34] Sun, Y. and Tiwari, R. C. (1995). Comparing cause-specific hazard rates of a competing risks model with censored data. *Institute of Mathematical Statistics*, **27**, 255–270.

[35] Sun, Y., Gilbert, P. B. and McKeague, I. W. (2009). Proportional Hazards Models With Continuous Marks. *The Annals of Statistics*, **37**, 394–426.

[36] UNAIDS, Report from a Meeting of the WHO-UNAIDS Vaccine Advisory Committee. (2001). Approaches to the development of broadly protective HIV vaccines: Challenges posed by genetic, biological and antigenic variability of HIV-1. *AIDS*, **15**, W1–W251.

[37] Van der Vaart, A. W. (1998). *Asymptotic Statistics*, Cambridge, Cambridge University Press.

[38] Wang,L. (2010). Introduction and Examples of Competing Risks in R.

[39] Yeni, P. G, Hammer, S. M., Carpenter, C. C., et al. (2002). Antiretroviral Treatment for Adult HIV Infection in 2002. *The Journel of American Medical Associaton*, **288**, 222–235.

[40] Aalen, O. O. and Johansen, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations, *Scand. J. Statist.* , **5**, 141–150.

[41] Cai, Z. and Sun, Y. (2003). Local linear estimation for time-dependent coefficients in Cox's regression models. *Scand. J. Statist.*, **30**, 93–111.

[42] Flynn, N. M., Forthal, D. N., Harro, C. D., Judson, F. N., Mayer, K. H., Para, M. F., Gilbert, P. B. and The rgp120 HIV Vaccine Study Group (2005). Placebo-controlled phase 3 trial of recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. *Journal of Infectious Diseases*, **191** 654–665.

[43] Gilbert, P.B., McKeague, I.W., Sun, Y. (2004). Tests for comparing mark-specfic hazards and cumulative incidence functions. *Lifetime Data Analysis*, **10**, 5–28.

[44] Gilbert, P.B., McKeague, I.W., Sun, Y. (2007). The two-sample problem for failure rates depending on a continuous mark: an application to vaccine efficacy. *Biostatistics*, **9**,263–276.

[45] Graham, B.S. (2002). Clinical trials of HIV vaccines. *Annual Review of Medicine*, **53**, 207–21.

[46] Huang, Y. and Louis, T.A. (1998). Nonparametric estimation of the joint distribution of survival time and mark variables. *Biometrika*, **85**, 785–798.

[47] Kalbfleisch J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, New York, Wiley.

[48] Lagakos, S. W. and Schoenfeld, D. A. (1984). Properties of proportional-hazards score tests under misspecified regression models. *Biometrics*, **40**, 1037-1048

[49] Lagakos, S. W. (1988b). Effects of mismodeling and mismeasuring explanatory variables on tests of association with a response variable. *Statistics in Medicine*, **7**, 257-274.

[50] Lin, D. Y., and Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association*, **84**, 1074–1078.

[51] Lin,D.Y., Wei,L.J, and Z.Ying. (1993) Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. *Biometrika*, **80**, 557-572.

[52] Nabel, G. J. (2001). Challenges and opportunities for development of an AIDS vaccine. *Nature*, **410**, 1002–7.

[53] Nickle, D.C., Heath, L., Jensen, M.A., Gilbert, P.B., Kosakovsky Pond, S.L.K., Mullins, J.I. (2005). Amino acid substitution matrices for HIV-1 subtype B. *Technical Report, University of Washington.*

[54] Olschewski, M. and Schumacher, M. (1990). Statistical analysis of quality of life in cancer clinical trials. *Statistics in Medicine*, **9**, 749–763.

[55] Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Fluornoy, N., Farewell, V. T. and Breslow, N. E. (1978). The analysis of failure time in the presence of competing risks. *Biometrics*, **34**, 541–554.

[56] Robins, J. M., Hernan, M. A., Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, **11**, 550–560.

[57] Schumacher, M. (1984). Two-sample tests of Cramér-von Mises and Kolmogorov-Smirnov type for randomly censored data. *Internat. Statist. Rev.*, **52**, 263–281.

[58] Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, New York, Wiley.

[59] Spiekerman, C.F and Lin, D.Y. (1996). Checking the marginal Cox model for correlated failure time data.*Biometrika*, **83**, 143-156.

[60] Struthers, C. and Kalbfleisch, J. D. (1986). Misspecified Proportional Hazards Model. *Biometrika*, **74**, 363-369.

[61] Sun, Y., Gilbert, P. B. and McKeague, I. W. (2009). Proportional Hazards Models With Continuous Marks. *The Annals of Statistics*, **37**, 394–426.

[62] Sun, Y. and Gilbert, B. P. (2011). Estimation of stratified mark-specific proportional hazards models with missing marks. To appear in Scandinavian Journal of Statistics.

[63] Sun, Y. and Wu, H. (2005). Semiparametric time-varying coefficients regression model for longitudinal data. *Scandinavian Journal of Statistics*, **32**, 21–47.

[64] Tsiatis, A. A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences USA*, **72**, 20–22.

[65] Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

[66] UNAIDS (2004). Joint United Nations Programme for HIV/AIDS. AIDS Epidemic Update.

[67] Wyatt, R., Kwong, P. D., Desjardins, E., Sweet, R. W., Robinson, J., Hendrickson, W. A., Sodroski, J. G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature*, **393**, 705–711.