

# A mark-specific quantile regression model

BY LIANQIANG QU

*School of Mathematics and Statistics, Central China Normal University,  
Wuhan, Hubei 430079, China*

LIUQUAN SUN

*Institute of Applied Mathematics, Academy of Mathematics and Systems Science,  
Chinese Academy of Sciences, Beijing, 100190, China  
slq@amt.ac.cn*

AND YANQING SUN

*Department of Mathematics and Statistics, University of North Carolina at Charlotte,  
Charlotte, U.S.A.*

## SUMMARY

Quantile regression has become a widely used tool for analyzing competing risks data. However, quantile regression for competing risks data with a continuous mark is still scarce. The mark variable is an extension of cause-of-failure in a classical competing risks model where cause of failure is replaced by a continuous mark only observed at uncensored failure times. An example of the continuous mark variable is the genetic distance that measures dissimilarity between the infecting virus and the virus contained in the vaccine construct. In this article, we propose a novel mark-specific quantile regression model. The proposed estimation method borrows strength from data in a neighborhood of a mark and is based on an induced smoothed estimation equation, which is very different from the existing methods for competing risks data with discrete causes. The asymptotic properties of the resulting estimators are established across mark and quantile continuums. In addition, a mark-specific quantile-type vaccine efficacy is proposed and its statistical inference procedures are developed. Simulation studies are conducted to evaluate the finite sample performances of the proposed estimation and hypothesis testing procedures. An application to the first HIV vaccine efficacy trial is provided.

*Some key words:* Competing risks; Continuous mark; Hypothesis testing; Mark-specific quantile regression; Survival data; Vaccine efficacy.

## 1. INTRODUCTION

### 1.1. Background

Quantile regression provides a comprehensive description on different parts of the conditional distribution of responses (Koenker & Bassett, 1996), and it has become a widely used tool in the survival analysis. For example, Powell (1984, 1986) modified the least absolute deviation procedure to analyze censored observations. Portnoy (2003) developed a recursively reweighted estimation procedure by using the principle of self-consistency for the Kaplan-Meier estimator. Peng & Huang (2008) proposed a recursive series of estimating equations for a sequence of quantiles based on the martingale feature associated with censored data. De Backer et al. (2019)

suggested an adaptive method to analyze survival data through modifying the so-called check function.

Competing risks data are common in survival analysis. When the competing causes of failures are finite, Peng & Fine (2007) proposed a nonparametric quantile inference method for cause-specific failure probabilities. Peng & Fine (2009) presented a competing risks quantile regression based on the cause-specific cumulative incidence function. Sun et al. (2012) developed a generalized linear quantile regression for competing risks data when the failure type may be missing. More related works are referred to Lee & Han (2016), Ahn & Kim (2018), Choi et al. (2018) and Farcomeni & Geraci (2020), among others. Competing risks models with continuous causes-of-failure (marks) are useful with many important applications (Sun et al., 2009, 2020). Our research is motivated by a data set from an HIV vaccine efficacy trial, in which the vaccine may only provide protection for HIV strains genetically similar to the HIV virus or viruses represented in the vaccine. The similarity between the infecting virus and the virus contained in the vaccine construct can be measured by the genetic divergence (or distance). Thus, the genetic divergence of infecting HIV viruses from the HIV strain represented in the vaccine needs to be taken into account to properly assess vaccine efficacy.

The mark variable is a measure of the genetic distance between two aligned HIV sequences, which is defined as the weighted percent mismatch of amino acids between the two HIV sequences. Since this distance may be unique for all infected subjects and the genetic diversity of HIV is extensive, it is natural to consider the mark as a continuous variable. Furthermore, during the observing period, the volunteers are potentially at risk of HIV infection from more than one mutually exclusive strains of viruses, and the mark is only observed when HIV infection occurs. If HIV infection does not occur, then the mark is undefined and is not meaningful. Thus, this situation can be considered as a competing risks setting, where causes of failure are replaced by a continuous mark only observed at uncensored failure times, and the mark is considered as continuous causes of failure (Sun et al., 2009). A preliminary analysis of the data is given in Fig. 1, which plots the curves of the mark-specific cumulative incidence functions separately for subjects stratified by treatment, age at the median values and behavioral risk score (Gilbert et al., 2008). Fig. 1 is quite suggestive of the effects of age and behavioral risk score on the mark-specific cumulative incidence functions.

When analyzing continuous mark data, the existing methods developed for discrete competing risks can no longer be applied. First, the mark is from a continuous distribution, and observations at a specified value of the mark are sparse. This feature of data is very different from that of discrete competing risks data. In addition, inspecting Fig. 1 reveals that the effects of covariates on the conditional quantiles of the failure time may vary nonlinearly with the mark. But the methods of Peng & Fine (2009), Sun et al. (2012), Ahn & Kim (2018), Choi et al. (2018) and Farcomeni & Geraci (2020) assume that the effects of covariates are constant at each given quantile level. Therefore, suitable methods are needed to analyze the varying effects of covariates with the mark. Moreover, although the marginal quantile regression methods such as in Portnoy (2003) and Peng & Huang (2008) can be used to assess the vaccine efficacy, it may fail to reveal the important relation between the vaccine efficacy and infecting viruses if the mark is ignored. A toy example is given at near the end of Section 1.3. The aim of this paper is to develop a quantile regression methodology for analyzing survival data with continuous marks and assessing the HIV vaccine efficacy. The proposed method allows the covariate effects to vary nonlinearly with the mark.

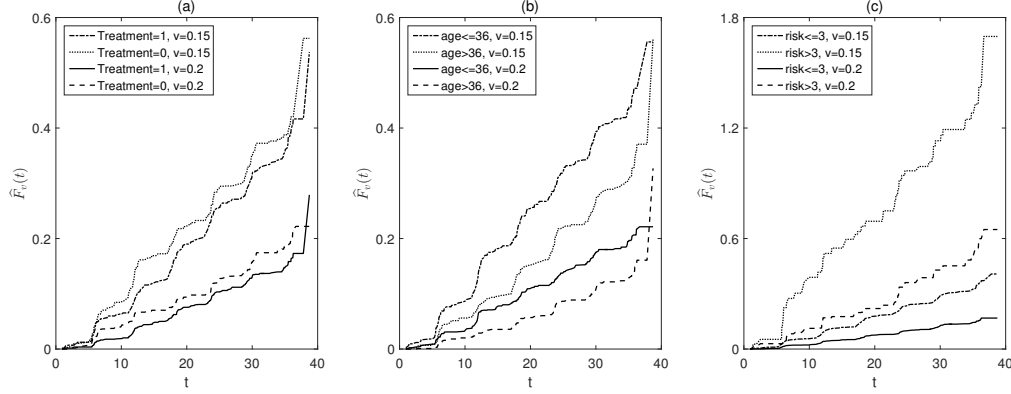


Fig. 1. The estimated mark-specific cumulative incidence functions for the vaccine trial data. The mark-specific cumulative incidence function (Gilbert et al., 2008) is defined as  $F_v(t) = \lim_{h \rightarrow 0} P(T \leq t, v \leq V < v + h)/h$ , where  $T$  is the time infected with HIV and  $V$  denotes the mark (i.e., weighted percent mismatch of amino acids).

### 1.2. Mark-specific quantile regression model

Let  $T$  be the failure time of interest and  $C$  be the censoring time. Let  $V$  denote a continuous mark variable, and  $Z = (1, \tilde{Z}^\top)^\top$ , where  $\tilde{Z}$  is a  $p$ -dimensional covariate vector. Assume that  $C$  is independent of  $(T, V)$  given  $Z$ . Further denote  $X = \min(T, C)$  as the observed time and  $\Delta = I(T \leq C)$  as the censoring indicator, where  $I(\cdot)$  denotes an indicator function. The mark is observed only when the corresponding failure time is uncensored. If  $\Delta = 0$ , then  $V$  is undefined and is not meaningful. The conditional mark-specific cumulative incidence function is defined as

$$F_v(t|Z) = \lim_{h \rightarrow 0} P(T \leq t, v \leq V < v + h|Z)/h.$$

The mark-specific cumulative incidence function is an extension of the cause-specific cumulative incidence function, where the cause of the failure time is replaced by a continuous mark (Gilbert et al., 2008). Suppose that  $F_v(t|Z) \leq \tilde{\tau}$  for some constant  $\tilde{\tau} > 0$ , and the support of the mark  $V$  is taken to be  $[0, 1]$ , rescaling  $V$  if necessary. For  $v \in [0, 1]$  and  $\tau \in (0, \tilde{\tau})$ , we define the  $\tau$ th conditional mark-specific quantile by

$$Q_v(\tau|Z) = \inf\{t : F_v(t|Z) \geq \tau\}.$$

Under the competing risks framework, the mark  $V$  is only meaningfully defined when failure occurs and it can not be treated as a covariate. The proposed conditional mark-specific quantile  $Q_v(\tau|Z)$  is not the conventional conditional quantile function  $\tilde{Q}(\tau|Z, v)$  obtained by treating  $V$  as a covariate, where  $\tilde{Q}(\tau|Z, v) = \inf\{t : F(t|Z, v) \geq \tau\}$  and  $F(t|Z, v) = P(T \leq t|Z, V = v)$  is the conditional distribution of  $T$  given  $Z$  and  $V = v$ . The conditional distribution function  $F(t|Z, v)$ , and therefore  $\tilde{Q}(\tau|Z, v)$ , is not identifiable under the competing risks setting.

The conditional mark-specific cumulative incidence function is an extension of the cause-specific cumulative incidence function in a competing risks setting, where the cause of the failure time is replaced by a continuous mark only observed at the failure time (Gilbert et al., 2008). Thus, the conditional mark-specific quantile function is an extension of the cause-specific quantile function in a competing risks setting for continuous cause of failure, and is defined analogously to that for the competing risks data with finitely many competing risks (Peng & Fine, 2009). The conditional mark-specific quantile function  $Q_v(\tau|Z)$  can be interpreted as the earliest time given covariate  $Z$  at which the proportion of subjects whose failures have occurred with mark  $V = v$  exceeds  $\tau$ . For the HIV vaccine efficacy trials,  $Q_v(\tau|Z)$  can be interpreted as the

first time given covariate  $Z$  at which the proportion of volunteers who have infected HIV with mark  $V = v$  exceeds  $\tau$ .

Because  $F_v(t|Z) = F(t|Z, v)f_V(v|Z)$ , where  $f_V(v|Z)$  is the conditional density function of  $V$  given  $Z$ , it follows that  $Q_v(\tau|Z) = \tilde{Q}(\tau^*(Z, v)|Z, v)$  with  $\tau^*(Z, v) = \tau/f_V(v|Z)$ , and the range of  $\tau$  is not necessarily bounded by 1. This is different from the competing risks quantile regression of Peng & Fine (2009) for discrete mark in which case  $f_V(v|Z)$  is the conditional probability mass function and  $\tau = \tau^*(Z, v)f_V(v|Z)$  is bounded by 1 due to  $0 \leq \tau^*(Z, v) \leq 1$ . In what follows, we consider  $\tau < \tilde{\tau} = \min_{v,z} \max_t F_v(t|Z = z)$  (see Section 2.3 for more details about the choice of the upper quantile).

We propose a novel mark-specified quantile regression model to capture the nonlinear interaction effects between the covariates and the mark on the failure. Specifically, for  $v \in [0, 1]$  and  $\tau \in (0, \tilde{\tau})$ , the model postulate that

$$Q_v(\tau|Z) = \exp\{Z^\top \beta_\tau^*(v)\}, \quad (1)$$

where  $\beta_\tau^*(v) = (\beta_{0\tau}^*(v), \beta_{1\tau}^*(v), \dots, \beta_{p\tau}^*(v))^\top$  is a  $(p+1)$ -dimensional vector of unknown continuous functions of  $v$  and  $\tau$ , and characterizes the varying effects of  $Z$  on the conditional mark-specific quantile of the failure time with respect to  $V$ . By setting the first component of  $Z$  as 1, model (1) has a nonparametric baseline function  $\exp\{\beta_{0\tau}^*(v)\}$ . Model (1) has a similar form of the varying-coefficient quantile regression model (Kim, 2007) in the absence of censored data:

$$\tilde{Q}(\tau|Z, v) = \exp\{Z^\top \tilde{\beta}_\tau^*(v)\}.$$

However, since  $Q_v(\tau|Z) = \tilde{Q}(\tau^*(Z, v)|Z, v)$ ,  $\tau^*(Z, v) = \tau/f_V(v|Z)$  depends on  $f_V(v|Z)$  and  $(Z, v)$ , and  $f_V(v|Z)$  is not identifiable under the competing risks setting, the method of Kim (2007) can not be directly applied to estimate  $\tilde{\beta}_{\tau^*(Z,v)}^*(v)$  or  $\beta_\tau^*(v)$  for model (1). In this paper, we develop an induced smoothing procedure (Brown & Wang, 2007) to estimate  $\beta_\tau^*(v)$  under model (1), which can be fast implemented using widely available numerical methods, such as the Newton-Raphson algorithm.

### 1.3. Quantile-type vaccine efficacy

The proposed model has applications in sieve analysis of vaccine efficacies. To evaluate the HIV vaccine efficacy, write the covariate as  $Z = (1, Z_1, Z_2^\top)^\top$ , where  $Z_1$  is the treatment (vaccine) group indicator and  $Z_2$  is a vector of other covariates. We define the mark-specific quantile-type vaccine efficacy as

$$\text{QVE}_\tau(v) = \frac{Q_v(\tau|1, Z_1 = 1, Z_2)}{Q_v(\tau|1, Z_1 = 0, Z_2)} - 1.$$

The function  $\text{QVE}_\tau(v)$  characterizes the nonlinear dependence on  $v$  of the ratio of the conditional mark-specific quantile of the failure time at level  $\tau$  under vaccine assignment ( $Z_1 = 1$ ) compared to under placebo assignment ( $Z_1 = 0$ ). A positive value of  $\text{QVE}_\tau(v)$  indicates the effectiveness of the vaccine that it takes longer time to reach same percentage ( $\tau$ ) of the mark-specific infections/diseases for the vaccine group as opposed to the placebo group. The larger the value of  $\text{QVE}_\tau(v)$  the more effective the vaccine is. It is close to zero if and only if the conditional mark-specific quantiles of the failure time have no clear differences between the vaccine and placebo groups. Here we focus on the log-linear model for  $Q_v(\tau|Z)$ . This is because under model (1),  $\text{QVE}_\tau(v) = \exp\{\beta_{1\tau}^*(v)\} - 1$  is free of  $Z_2$  but depends on the mark value  $v$ , which simplifies the inference procedure for  $\text{QVE}_\tau(v)$ .

The mark-specified quantile regression model (1) complements the modeling approaches based on the mark-specific hazard functions (Sun et al., 2009; Han et al., 2017; Sun et al.,

2020) by allowing covariate effects varying over  $\tau$ . It also complements the marginal quantile regression models (Portnoy, 2003; Peng & Huang, 2008) by providing additional insights on how the relation between the quantile and covariates changes with the mark. To illustrate the difference from the marginal quantile regression model, we consider a toy example. Let  $\beta_{0\tau}^*(v) = \tau(\gamma_0 + v)^2$  and  $\beta_{1\tau}^*(v) = \tau\gamma_1\{2(\gamma_0 + v) + \gamma_1\}$ . Assume that  $V$  is from  $U(0, 1)$ , and  $\log(T)$  from  $U(0, (V + \gamma_0 + \gamma_1 Z_1)^2)$  given  $V$  and  $Z_1$ , where  $U(0, c)$  denotes a uniform distribution on  $(0, c)$  ( $c > 0$ ). Thus, the marginal quantile of  $T$  given  $Z_1$  is  $Q(\tau|Z_1) = \exp\{\tau\gamma_0(1 + \gamma_0) + \tau\gamma_1(1 + 2\gamma_0 + \gamma_1)Z_1\}$ , where  $Q(\tau|Z) = \inf\{t : F(t|Z) \geq \tau\}$  and  $F(t|Z) = P(T \leq t|Z)$ . If  $\gamma_1 = -(2\gamma_0 + 1)$ , then  $\text{QVE}_\tau = Q(\tau|Z_1 = 1)/Q(\tau|Z_1 = 0) - 1 \equiv 0$ . But  $\text{QVE}_\tau(v) = \exp\{\tau\gamma_1[2(\gamma_0 + v) + \gamma_1]\} - 1$ , which is not zero if  $\gamma_1 \neq 0$ . That is, the important vaccine effects can be missed without the consideration of the mark under this case.

We also consider a cumulative version of  $\text{QVE}_\tau(v)$ , which is defined as  $\text{CQVE}_\tau(v) = \int_a^v \text{QVE}_\tau(u) du$  with  $0 < a < 1$ . The quantity can be used to assess the vaccine efficacy over a range of marks for  $v \in [a, b] \subset (0, 1)$  and quantile levels  $\tau \in [\tau_0, \tau_U]$ . We construct simultaneous confidence bands for  $\text{CQVE}_\tau(v)$ , and propose test statistics to evaluate the mark-specific vaccine efficacy based on the estimator of  $\text{CQVE}_\tau(v)$  for  $v \in [a, b] \subset (0, 1)$ .

## 2. ESTIMATION PROCEDURES

### 2.1. Induced smoothing estimators

Suppose that we observe  $n$  independent and identically distributed copies of  $(X, \Delta, \Delta V, Z)$ , denoted by  $(X_i, \Delta_i, \Delta_i V_i, Z_i)$  ( $i = 1, \dots, n$ ). In what follows, assume  $T_i$  are continuous random variables. Let  $N_i(t, v) = I(X_i \leq t, \Delta_i = 1, V_i \leq v)$  be the marked point counting process with a jump at an uncensored failure time  $X_i$  and the associated mark  $V_i$ . Define  $G(t|Z_i) = P(C_i \geq t|Z_i)$  as the survival function of the censoring time. Since  $C_i$  is independent of  $(T_i, V_i)$  given  $Z_i$ , it can be checked that

$$F_v(t|Z_i) = \lim_{h \rightarrow 0} \frac{1}{2h} E \left\{ \int_0^1 \int_0^t \frac{1}{G(s|Z_i)} I(|u - v| \leq h) N_i(ds, du) | Z_i \right\},$$

which suggests that we can borrow strength from data in a neighborhood of a mark. For each  $v \in (0, 1)$ , we propose the following mark-specific localized estimating equation to estimate  $\beta_\tau^*(v)$ :

$$U_n\{\beta_\tau(v)\} = 0,$$

where

$$U_n(\xi) = \frac{1}{n} \sum_{i=1}^n Z_i \left[ \int_0^1 \int_0^L \frac{1}{G(t|Z_i)} I\{\log t \leq Z_i^\top \xi\} K_h(u - v) N_i(dt, du) - \tau \right],$$

$L$  is the follow-up time satisfying  $P(X_i \geq L) > 0$ ,  $K_h(x) = K(x/h)/h$ ,  $K(x)$  is a kernel function with support on  $(-1, 1)$  and  $h$  is a bandwidth.

Since  $U_n(\xi)$  is only monotone but not continuous, an exact zero-crossing of  $U_n(\xi)$  may not exist. To be more specific, Fig. S1 in the Supplementary Material presents  $U_n\{\beta_\tau(v)\}$  as a function of  $\beta_{1\tau}^*(v)$  with  $\tau = 0.2$  and  $v = 0.5$  for model M3 studied in Section 4, that is,  $\beta_{0\tau}^*(v)$  and  $\beta_{2\tau}^*(v)$  are assumed to be known. The sample size is 1500 and the bandwidth is 0.2. It can be seen that  $U_n\{\beta_\tau(v)\}$  is very jagged and may flat at 0.43 (the value of  $\beta_{1\tau}^*(0.5)$ ), which results in numerical challenges in computing the solution of  $U_n(\xi)$ , particularly with multiple covariates.

In addition, since  $U_n(\xi)$  is nondifferentiable, the variance estimation of the resulting estimators can be very difficult.

To address these issues, we next propose an induced smoothing method to approximate  $U_n(\xi)$  using continuously differentiable functions (Brown & Wang, 2007). Specifically, let  $\Gamma_\tau(v)$  be a given  $(p+1) \times (p+1)$  matrix such that  $\|\Gamma_\tau(v)\|_F = O(1)$ , where  $\|D\|_F$  denotes the Frobenius norm of any matrix  $D$ . Similarly to Brown & Wang (2007), a smoothed version of  $U_n\{\beta(v)\}$  can be constructed by  $S_n\{\beta_\tau(v)\} = E_W[U_n\{\beta_\tau(v) + (nh)^{-1/2}\Gamma_\tau(v)W\}]$ , where  $W$  is a random vector from  $N(0, \mathbb{I}_{p+1})$  independent of  $(X_i, \Delta_i, \Delta_i V_i, Z_i)$ , and  $\mathbb{I}_{p+1}$  denotes the identity matrix of size  $p+1$ . A direct calculation shows that

$$S_n\{\beta_\tau(v)\} = \frac{1}{n} \sum_{i=1}^n Z_i \left[ \int_0^1 \int_0^L \frac{1}{G(t|Z_i)} \Phi \left\{ \frac{Z_i^\top \beta_\tau(v) - \log t}{(nh)^{-1/2} \gamma_i} \right\} K_h(u-v) N_i(dt, du) - \tau \right],$$

where  $\Phi(x)$  denotes the cumulative function of the standard normal distribution and  $\gamma_i = Z_i^\top \Gamma_\tau(v) \Gamma_\tau(v)^\top Z_i$ .

In practice, the survival function  $G(t|Z)$  is usually unknown, but can be estimated from the observed data. For instance, if the censoring is dictated by administrative decisions or appears to be independent of the covariates, we can use the Kaplan-Meier estimator for the censoring distribution. When the censoring depends on the covariates, we can estimate  $G(t|Z)$  by specifying a semiparametric regression model for the censoring time, such as the Cox model. Here for notational simplicity, we just consider that the censoring is independent of the covariates, and the estimation procedures are similar for the case when the censoring depends on the covariates. Let  $\hat{G}(t)$  be the survival function of the censoring time, and we focus on the Kaplan-Meier estimator of  $G(t)$ , denoted by  $\hat{G}(t)$ . In addition, in the numerical studies below, we take  $\Gamma_\tau(v) = \mathbb{I}_{p+1}$  for computational convenience.

Let  $\hat{S}_n(b)$  denote the estimating equation by replacing  $G(t|Z_i)$  with  $\hat{G}(t)$  in  $S_n(b)$ . We can show (as in the proof of Theorem 1 below) that  $\hat{S}_n\{\beta_\tau(v)\}$  is generally close to  $U_n\{\beta_\tau(v)\}$ . This fact can also be confirmed in the aforementioned Fig. S1, where  $\hat{S}_n\{\beta_\tau(v)\}$  approximates  $U_n\{\beta_\tau(v)\}$  well and has a unique solution. In what follows, we propose to estimate  $\beta_\tau^*(v)$  by the solution to  $\hat{S}_n\{\beta_\tau(v)\} = 0$ , denoted by  $\hat{\beta}_\tau(v)$ .

## 2.2. Asymptotic properties

We establish the uniform consistency and asymptotic normality of  $\hat{\beta}_\tau(v)$  for  $(v, \tau) \in \mathcal{B} = [a, b] \times [\tau_0, \tau_U]$ , where  $0 < a < b < 1$  and  $0 < \tau_0 < \tau_U < \tilde{\tau}$ . We assume the following conditions.

*Condition 1.*  $P(X \geq L) > 0$ , and  $Z$  is bounded almost surely.

*Condition 2.* Each component of  $\beta_\tau^*(v)$  is Lipschitz continuous with respect to  $(v, \tau) \in \mathcal{B}$ .

*Condition 3.* The conditional density function  $f(t, v|z)$  of  $(T, V)$  given  $Z = z$  is twice continuously differentiable with respect to  $t$  and  $v$ , and is bounded uniformly in  $(t, v, z) \in [0, L] \times [a, b] \times \mathcal{C}$ , where  $\mathcal{C}$  denotes the support of  $Z$ . In addition, the conditional density function  $g(v|z, t)$  of  $V$  given  $(Z, T) = (z, t)$  is bounded away from 0 and  $\infty$  uniformly in  $(t, v, z) \in [0, L] \times [a, b] \times \mathcal{C}$ .

*Condition 4.*  $\inf_{(v, \tau) \in \mathcal{B}} \lambda_{\min}(A_\tau(v)) > 0$ , where  $\lambda_{\min}(\cdot)$  denotes the minimum eigenvalue of a matrix, and

$$A_\tau(v) = E[Z Z^\top \exp(Z^\top \beta_\tau^*(v)) f(\exp(Z^\top \beta_\tau^*(v)), v|Z)].$$

*Condition 5.* The kernel function  $K(x)$  is symmetric with support  $[-1, 1]$ , and has bounded variation satisfying  $\int K(u)du = 1$ .

*Condition 6.* The bandwidth satisfies  $nh^2 \rightarrow \infty$  and  $nh^5 \rightarrow 0$ .

Conditions 1-4 are standard assumptions for quantile regression methods in the context of survival analysis, which are analogous to those in Peng & Huang (2008) and Qian & Peng (2010). Conditions 2 and 3 are needed for the uniform consistency and the asymptotic normality of  $\hat{\beta}_\tau(v)$ . Condition 4 is the assumption to ensure the identifiability of  $\beta_\tau^*(v)$ . Conditions 5 and 6 are standard assumption for kernel smoothing techniques. The uniform consistency and asymptotic normality of  $\hat{\beta}_\tau(v)$  are given in the following two theorems.

**THEOREM 1.** Suppose that Conditions 1-6 hold. Then  $\|\hat{\beta}_\tau(v) - \beta_\tau^*(v)\| = o_p(1)$  uniformly in  $(v, \tau) \in \mathcal{B}$ .

**THEOREM 2.** Suppose that Conditions 1-6 hold. Then  $(nh)^{1/2}\{\hat{\beta}_\tau(v) - \beta_\tau^*(v)\}$  converges in distribution to a zero-mean normal random vector with covariance matrix  $\Omega_\tau(v)$  for  $(v, \tau) \in \mathcal{B}$ , where  $\Omega_\tau(v) = A_\tau(v)^{-1}D(v, \tau)A_\tau(v)^{-1}$ ,

$$D(v, \tau) = \nu_0 E[Z Z^\top I\{\log T \leq Z^\top \beta_\tau^*(v)\} G(T)^{-1} g(v|Z, T)],$$

and  $\nu_0 = \int K(u)^2 du$ .

In order to estimate  $\Omega_\tau(v)$ , we need to estimate  $A_\tau(v)$  and  $D_\tau(v, \tau)$ . First, by checking the proofs of Theorems 1 and 2, we can consistently estimate  $A_\tau(v)$  by

$$\hat{A}_\tau(v) = n^{-1/2} h^{1/2} \sum_{i=1}^n \int_0^1 \int_0^L \frac{Z_i Z_i^\top}{\hat{G}(t) \gamma_i} \phi\left\{ \frac{Z_i^\top \hat{\beta}_\tau(v) - \log t}{(nh)^{-1/2} \gamma_i} \right\} K_h(u - v) N_i(dt, du),$$

where  $\phi(x)$  is the density function of the standard normal distribution. In addition, we can consistently estimate  $D(v, \tau)$  by

$$\hat{D}(v, \tau) = n^{-1} h \sum_{i=1}^n \hat{\eta}_i(v, \tau) \hat{\eta}_i(v, \tau)^\top,$$

where

$$\hat{\eta}_i(v, \tau) = Z_i \left[ \int_0^1 \int_0^L I\{\log t \leq Z_i^\top \hat{\beta}_\tau(v)\} K_h(u - v) \hat{G}(t)^{-1} N_i(dt, du) - \tau \right].$$

Thus,  $\Omega_\tau(v)$  can be consistently estimated by  $\hat{\Omega}_\tau(v)$ , where

$$\hat{\Omega}_\tau(v) = \hat{A}_\tau(v)^{-1} \hat{D}(v, \tau) \hat{A}_\tau(v)^{-1}. \quad (2)$$

*Remark 1.* Theorem 2 implies that the covariance matrix of  $\hat{\beta}_\tau(v)$  is not affected asymptotically by  $\hat{G}(t)$ . One intuitive explanation is that  $\hat{G}(t)$  is irrelevant to the mark in estimating  $\hat{\beta}_\tau(v)$ , and the former converges at a faster rate. Similar results have also been obtained by Zhang et al. (2022) for quantile regression models, where the inference on the parameters of interest is not affected asymptotically by the estimation of nuisance parameters.

*Remark 2.* As suggested in Brown & Wang (2007), we can also apply an iterative procedure to solve  $\hat{S}_n\{\beta_\tau(v)\} = 0$ , which can simultaneously estimate  $\beta_\tau^*(v)$  and the asymptotic covariance matrix. Specifically, define  $\beta_\tau^{(k)}(v)$  as the estimate of  $\beta_\tau^*(v)$  at the  $k$ th iteration.

*Step 0.* Choose an initial estimate  $\beta_\tau^{(0)}(v)$  by using the solution to  $\hat{S}_n\{\beta_\tau(v)\} = 0$  with  $\Gamma_\tau(v) = \mathbb{I}_{p+1}$ , and let  $k = 0$ .

*Step 1.* Update  $\Gamma_\tau(v) = [\hat{\Omega}_\tau^{(k)}(v)]^{1/2}$ , where  $\hat{\Omega}_\tau^{(k)}(v)$  is obtained by replacing  $\hat{\beta}_\tau(v)$  with  $\beta_\tau^{(k)}(v)$  in  $\hat{\Omega}_\tau(v)$  defined by (2)

260 *Step 2.* Obtain  $\beta_\tau^{(k+1)}(v)$  by solving  $\hat{S}_n\{\beta_\tau(v)\} = 0$  with  $\Gamma_\tau(v) = [\hat{\Omega}_\tau^{(k)}(v)]^{1/2}$ .

*Step 3.* Set  $k = k + 1$ . Repeat Steps 1 and 2 until convergence of both  $\beta_\tau^{(k)}(v)$  and  $\hat{\Omega}_\tau^{(k)}(v)$  is achieved to a specified tolerance.

However, this procedure is very time-consuming for estimating  $\beta_\tau^*(v)$  with respect to  $v$  and  $\tau$ .

*Remark 3.* Our proposed method can be extended to other link functions for  $Q_v(\tau|Z)$ . A more general approach is to generalize model (1) to

$$Q_v(\tau|Z) = H\{Z^\top \beta_\tau^*(v)\},$$

where  $H(\cdot) > 0$  is a known monotone link function (Sun et al., 2012; Peng & Fine, 2009). Under this setting, the estimating equation takes the form

$$\frac{1}{n} \sum_{i=1}^n Z_i \left[ \int_0^1 \int_0^L \frac{1}{\hat{G}(t)} \Phi \left\{ \frac{Z_i^\top \beta_\tau(v) - H^{-1}(t)}{(nh)^{-1/2} \gamma_i} \right\} K_h(u - v) N_i(dt, du) - \tau \right] = 0.$$

Let  $A_\tau^*(v) = E[Z Z^\top \dot{H}(Z^\top \beta_\tau^*(v)) f(H(Z^\top \beta_\tau^*(v)), v|Z)]$ , and

$$D^*(v, \tau) = \nu_0 E[Z Z^\top I\{H^{-1}(T) \leq Z^\top \beta_\tau^*(v)\} G(T)^{-1} g(v|Z, T)],$$

where  $\dot{H}(x)$  denotes the first derivative of  $H(x)$ . Then Theorems 1 and 2 still hold by replacing  $A_\tau(v)$  and  $D(v, \tau)$  with  $A_\tau^*(v)$  and  $D^*(v, \tau)$ , respectively.

Let  $B_\tau^*(v) = \int_a^v \beta_\tau^*(u) du$  denote the cumulative regression coefficient, which can be estimated by  $\hat{B}_\tau(v) = \int_a^v \hat{\beta}_\tau(u) du$ . Define  $\omega_\tau(s, u) = E[Z I(s \leq T \leq \exp\{Z^\top \beta_\tau^*(u)\}) g(u|Z, T)]$ ,  $y(t) = P(X \geq t)$ , and  $M^C(t) = N^C(t) - \int_0^t Y(s) d\Lambda^C(s)$ , where  $N^C(t) = I(X \leq t, \Delta = 0)$ ,  $Y(t) = I(X \geq t)$ , and  $\Lambda^C(t)$  is the cumulative hazards function of the censoring time. The following theorem establishes the weak convergence of  $\hat{B}_\tau(v)$ .

275 **THEOREM 3.** Suppose that Conditions 1-6 hold. Then  $n^{1/2}\{\hat{B}_\tau(v) - B_\tau^*(v)\}$  converges weakly on  $[a, b] \times [\tau_0, \tau_U]$  to a  $(p+1)$ -dimensional Gaussian process with mean zero and covariance matrix  $\Psi(v_1, v_2, \tau_1, \tau_2) = E\{\varphi_1(v_1, \tau_1)\varphi_1(v_2, \tau_2)^\top + \varphi_2(v_1, \tau_1)\varphi_2(v_2, \tau_2)^\top\}$  at  $(v_1, \tau_1)$  and  $(v_2, \tau_2)$ , where

$$\begin{aligned} \varphi_1(v, \tau) &= \left[ \int_a^v \int_0^L A_\tau(u)^{-1} I\{\log t \leq Z^\top \beta_\tau^*(u)\} \frac{N(dt, du)}{G(t)} - \tau \int_a^v A_\tau(u)^{-1} du \right] Z, \\ \varphi_2(v, \tau) &= \int_a^v \int_0^L A_\tau(u)^{-1} \omega_\tau(s, u) y(s)^{-1} dM^C(s) du, \end{aligned}$$

280 and  $\varphi_1(v, \tau)$  and  $\varphi_2(v, \tau)$  are independent to each other.

The covariance function  $\Psi(v_1, v_2, \tau_1, \tau_2)$  can be consistently estimated by

$$\hat{\Psi}(v_1, v_2, \tau_1, \tau_2) = \frac{1}{n} \sum_{i=1}^n \left[ \hat{\varphi}_{1i}(v_1, \tau_1) \hat{\varphi}_{1i}(v_2, \tau_2)^\top + \hat{\varphi}_{2i}(v_1, \tau_1) \hat{\varphi}_{2i}(v_2, \tau_2)^\top \right],$$



where

$$\begin{aligned}\hat{\varphi}_{1i}(v, \tau) &= \left[ \int_a^v \int_0^L \hat{A}_\tau(u)^{-1} I\{\log t \leq Z_i^\top \hat{\beta}_\tau(u)\} \frac{N_i(dt, du)}{\hat{G}(t)} - \tau \int_a^v \hat{A}_\tau(u)^{-1} du \right] Z_i, \\ \hat{\varphi}_{2i}(v, \tau) &= \int_a^v \int_0^L \hat{A}_\tau(u)^{-1} \hat{\omega}_\tau(s, u) \bar{Y}(s)^{-1} d\hat{M}_i^C(s) du, \\ \hat{\omega}_\tau(s, v) &= \frac{1}{n} \sum_{i=1}^n \int_a^b \int_0^L \left[ Z_i I(s \leq t \leq \exp\{Z_i^\top \hat{\beta}_\tau(v)\}) K_h(u - v) \frac{N_i(dt, du)}{\hat{G}(t)} \right],\end{aligned}$$

$\bar{Y}(s) = n^{-1} \sum_{i=1}^n I(X_i \geq s)$ ,  $\hat{M}_i^C(s) = I(X_i \leq s, \Delta_i = 0) - \int_0^s Y_i(u) d\hat{\Lambda}^C(u)$  and  $\hat{\Lambda}^C(s)$  is the Nelson-Aalen estimator of  $\Lambda^C(s)$ . 285

### 2.3. Tuning parameters selection

We provide some practical guidance on how to select the tuning parameters, including the bandwidth parameter used in induced smoothing, the range of quantile levels and the range of  $V$ . Bandwidth selection is often a critical part of nonparametric regression. Here we use an  $M$ -fold cross-validation method to choose the bandwidth parameter (Tian et al., 2005). Specifically, we randomly divide the data into  $M$  roughly equal-sized groups. Let  $D_k$  denote the  $k$ th subgroup of data, and  $\hat{\beta}_\tau^{(-k)}(v)$  be an estimate of  $\beta_\tau^*(v)$  using the data from all subgroups other than  $D_k$ . Under model (1), we have 290

$$E \left[ \int_0^v \int_0^{\exp\{Z_i^\top \beta_\tau^*(u)\}} \frac{N_i(ds, du)}{G(s|Z_i)} \middle| Z_i \right] = \int_0^v F_u(\exp\{Z_i^\top \beta_\tau^*(u)\} | Z_i) du = \tau v. \quad \text{295}$$

Thus, the  $k$ th prediction error can be given by

$$\text{PE}_k(h) = \sum_{i \in D_k} \int_{\tau_0}^{\tau_U} \int_a^b \left[ \int_0^v \int_0^L \frac{I\{\log s \leq Z_i^\top \hat{\beta}_\tau^{(-k)}(u)\}}{\hat{G}(s)} N_i(ds, du) - \tau v \right]^2 dv d\tau.$$

The optimal bandwidth can be obtained by

$$h_{\text{opt}} = \arg \min_h \sum_{k=1}^M \text{PE}_k(h).$$

In practice, however, the cross-validation method may be time-consuming. Alternatively, we could choose the bandwidth  $h = \varpi \hat{\sigma}_V n_0^{-1/4}$  by using the rule of thumb bandwidth, where  $\hat{\sigma}_V$  is the estimated standard error of the observed marks,  $n_0$  is the number of observed failure times, and  $\varpi > 0$  is a prespecified constant (Sun et al., 2009; Han et al., 2017). Simulation results presented in Section 4 show that the bandwidth  $h_{\text{opt}}$  leads to smaller estimated standard deviation in estimation and higher power for the tests. 300

For the range of quantile levels, the lower and upper quantiles,  $\tau_0$  and  $\tau_U$ , are required to satisfy the conditions that there is no censoring below the  $\tau_0$ th quantile, and  $\tau_U < \min_{v,z} \max_t F_v(t|Z = z)$ . In practice,  $\tau_0$  can be chosen to be close to 0 if censoring occurs at early stages (He et al., 2022). The upper quantile  $\tau_U$  can be chosen such that  $\tau_U < \min_{v,z} \max_t \hat{F}_v(t|Z = z)$ , where  $\hat{F}_v(t|Z = z)$  is an estimate of  $F_v(t|Z = z)$ . For a discrete covariate  $Z$  or a stratified version based on  $Z$ ,  $F_v(t|Z = z)$  can be estimated by 305

$$\hat{F}_v(t|Z = z) = \frac{1}{n} \sum_{i=1}^n \int_0^1 \int_0^t \frac{I(Z_i = z)}{\hat{G}(s)} K_h(u - v) N_i(ds, du).$$

By plotting the mark-specific cumulative incidence function estimates stratified on covariates as in Fig. 1, for example, we can choose  $\tau_U$  such that all estimated curves exceed it in the right tails.

For the range of  $V$ , as discussed in Sun et al. (2009), we assume that the mark variable  $V$  has a known and bounded support, and that without loss of generality, this support is taken to be  $[0, 1]$ , rescaling  $V$  if necessary (Sun et al., 2009). In practice, the range of  $V$  can be taken to be  $(V_{\min}, V_{\max})$  and  $[a, b] \subset (V_{\min}, V_{\max})$ , where  $V_{\min} = \min_{i:\delta_i=1} V_i$  and  $V_{\max} = \max_{i:\delta_i=1} V_i$ .

### 3. INFERENCE FOR VACCINE EFFICACY

The confidence bands for the regression coefficients and the mark-specific quantile-type vaccine efficacy are provided in the Supplementary Material. Here we consider to test the mark-specific quantile-type vaccine efficacy. If the vaccine has no efficacy, then the vaccine will provide no protection against any infecting strain of virus. As a result, the  $\tau$ th mark-specific quantiles should have no significant difference between the vaccine and placebo groups for all  $v \in [a, b]$  and  $\tau \in [\tau_0, \tau_U]$ . That is,  $\text{QVE}_\tau(v) \equiv 0$  for all  $(v, \tau) \in \mathcal{B}$ . Thus, it is interested in testing the efficacy over a range of  $v$  and  $\tau$  to assess the overall clinical/public health benefit of the vaccine. The first set of hypotheses is

$$\begin{aligned} H_{10} : \text{QVE}_\tau(v) &\equiv 0 \text{ for all } (v, \tau) \in \mathcal{B} \\ \text{versus } H_{11} : \text{QVE}_\tau(v) &\neq 0 \text{ for some } (v, \tau) \in \mathcal{B} \\ \text{or } H_{12} : \text{for each given } \tau, &\text{QVE}_\tau(v) \geq 0 \text{ with strict inequality for at least some } v. \end{aligned}$$

To test  $H_{10}$ , we consider the following statistics:

$$\begin{aligned} \mathcal{T}_{11} &= \sup_{\tau \in [\tau_0, \tau_U]} \int_a^b n \{ \widehat{\text{CQVE}}_\tau(v) \}^2 / \hat{\zeta}_{0\tau}(v) dv \text{ and} \\ \mathcal{T}_{12}(\tau) &= n^{1/2} \widehat{\text{CQVE}}_\tau(b) / \hat{\zeta}_{0\tau}^{1/2}(b), \end{aligned}$$

where  $\hat{\zeta}_{0\tau}(v) = n^{-1} \sum_{i=1}^n \hat{\vartheta}_{0i}^2(\tau, v)$  with

$$\hat{\vartheta}_{0i}(\tau, v) = - \int_a^v \exp\{\hat{\beta}_{1\tau}(u)\} e_1^\top \hat{A}_\tau(u)^{-1} [\hat{\varphi}_{1i}(v, \tau) + \hat{\varphi}_{2i}(v, \tau)] du \quad (3)$$

and  $e_1 = (0, 1, 0, \dots, 0)^\top \in \mathbb{R}^{p+1}$ . The test statistic  $\mathcal{T}_{11}$  captures general departures  $H_{11}$ , while  $\mathcal{T}_{12}(\tau)$  is sensitive to the alternative  $H_{12}$ , which is likely to be positive when  $H_{12}$  holds. The statistics  $\mathcal{T}_{11}$  and  $\mathcal{T}_{12}$  are close to zero when  $\text{QVE}_\tau(v) \equiv 0$ , and hence we shall reject  $H_{10}$  if  $\mathcal{T}_{11} > c_{11}(\alpha)$  and  $\mathcal{T}_{12}(\tau) > c_{12}(\alpha)$ , where  $c_{11}(\alpha)$  and  $c_{12}(\alpha)$  are the critical values. By some arguments similar to Section 2.2, we have that under  $H_{10}$ ,  $n^{1/2} \widehat{\text{CQVE}}_\tau(v)$  is asymptotically equivalent to  $n^{-1/2} \sum_{i=1}^n \hat{\vartheta}_{0i}(\tau, v)$ , where  $\hat{\vartheta}_{0i}(\tau, v)$  is defined in (3). Thus,  $c_{12}(\alpha)$  can be taken as the upper  $\alpha$ -quantile  $z_\alpha$  of the standard normal distribution. To obtain the critical value  $c_{11}(\alpha)$ , we consider a resampling technique (Lin et al., 1993). Let

$$\mathcal{T}_{11}^* = \sup_{\tau \in [\tau_0, \tau_U]} \int_a^b \left\{ n^{-1/2} \sum_{i=1}^n W_i \hat{\vartheta}_{0i}(\tau, v) \right\}^2 / \hat{\zeta}_{0\tau}(v) dv,$$

where  $W_i$  ( $i = 1, \dots, n$ ) are independent standard normal variables and are independent of the observed data. According to the arguments of Lin et al. (1993), the null distribution of  $\mathcal{T}_{11}$  can be approximated by the conditional distribution of  $\mathcal{T}_{11}^*$  given the observed data, which can be obtained by repeatedly generating the normal random sample  $W_i$  ( $i = 1, \dots, n$ ) while fixing

the observed data. Thus, the critical values  $c_{11}(\alpha)$  can be taken as the  $(1 - \alpha)$ -percentile of the conditional distributions of  $\mathcal{T}_{11}^*$ . 345

When the null  $H_{10}$  is rejected, one may wish to test whether the quantile-type vaccine efficacy varies with respect to  $v$  and  $\tau$ . In addition, if the vaccine is effective, then the vaccine will afford protection against the infecting strain of virus, and the vaccine efficacy will decrease as  $v$  increases (Sun et al., 2009). This leads to  $\text{QVE}_\tau(v)$  decreasing with  $v$  for each given  $\tau$ . Thus, we consider the following set of hypotheses: 350

$$\begin{aligned} H_{20} : \text{QVE}_\tau(v) &= \psi \text{ for all } (v, \tau) \in \mathcal{B} \\ \text{versus } H_{21} : \text{QVE}_\tau(v) &\neq \psi \text{ for some } (v, \tau) \in \mathcal{B} \\ \text{or } H_{22} : &\text{for each given } \tau, \text{QVE}_\tau(v) \text{ decreases as } v \text{ increases,} \end{aligned}$$

where  $\psi$  is some unspecified constant. 355

Let  $a_1$  and  $\tau_0^*$  be two specified constants such that  $a < a_1 < b$  and  $\tau_0 < \tau_0^* < \tau_U$ . To test  $H_{20}$ , we propose the following test statistics:

$$\mathcal{T}_{21}(\tau, v) = n^{1/2} \left\{ \frac{\widehat{\text{CCQVE}}_\tau(v)}{(v - a)(\tau - \tau_0)} - \frac{\widehat{\text{CCQVE}}_{\tau_U}(b)}{(b - a)(\tau_U - \tau_0)} \right\} / \hat{\zeta}_{1\tau}^{1/2}(v),$$

and

$$\mathcal{T}_{22}(\tau) = n^{1/2} \int_{a_1}^b \left\{ \frac{\widehat{\text{CQVE}}_\tau(v)}{v - a} - \frac{\widehat{\text{CQVE}}_\tau(b)}{b - a} \right\} / \hat{\zeta}_{2\tau}^{1/2}(v) dv, \quad \text{360}$$

where  $\widehat{\text{CCQVE}}_\tau(v) = \int_{\tau_0}^\tau \widehat{\text{CQVE}}_s(v) ds$ ,  $\hat{\zeta}_{1\tau}(v) = n^{-1} \sum_{i=1}^n \hat{\vartheta}_{1i}^2(\tau, v)$  and  $\hat{\zeta}_{2\tau}(v) = n^{-1} \sum_{i=1}^n \hat{\vartheta}_{2i}^2(\tau)$ . Here,  $\hat{\vartheta}_{1i}(\tau, v)$  and  $\hat{\vartheta}_{2i}(\tau)$  are defined as

$$\hat{\vartheta}_{1i}(\tau, v) = \frac{\int_{\tau_0}^\tau \hat{\vartheta}_{0i}(s, v) ds}{(v - a)(\tau - \tau_0)} - \frac{\int_{\tau_0}^{\tau_U} \hat{\vartheta}_{0i}(s, b) ds}{(b - a)(\tau_U - \tau_0)},$$

and

$$\hat{\vartheta}_{2i}(\tau) = \int_{a_1}^b \left\{ \frac{\hat{\vartheta}_{0i}(\tau, v)}{v - a} - \frac{\hat{\vartheta}_{0i}(\tau, b)}{b - a} \right\} dv. \quad \text{365}$$

We further define

$$\mathcal{T}_{21} = \sup_{v \in [a_1, b]} \sup_{\tau \in [\tau_0^*, \tau_U]} |\mathcal{T}_{21}(\tau, v)|.$$

When  $\text{QVE}_\tau(v)$  is a constant function,  $\mathcal{T}_{21}$  and  $\mathcal{T}_{22}(\tau)$  are likely to be zero. The statistic  $\mathcal{T}_{21}$  captures general departure  $H_{21}$ , while  $\mathcal{T}_{22}(\tau)$  is sensitive to monotone alternative  $H_{22}$ , which is likely to be positive when  $H_{22}$  holds. Hence, we will reject  $H_{20}$  if  $\mathcal{T}_{21} > c_{21}(\alpha)$  and  $\mathcal{T}_{22}(\tau) > c_{22}(\alpha)$ , where  $c_{21}(\alpha)$  and  $c_{22}(\alpha)$  are the critical values. Under  $H_{20}$ ,  $\mathcal{T}_{22}(\tau)$  is asymptotically standard normal, and  $c_{22}(\alpha)$  can be taken as the upper  $\alpha$ -quantile  $z_\alpha$  of the standard normal distribution. The critical value  $c_{21}(\alpha)$  can be obtained through the following resampling technique. Let 370

$$\mathcal{T}_{21}^* = \sup_{v \in [a_1, b]} \sup_{\tau \in [\tau_0^*, \tau_U]} \left| n^{-1/2} \sum_{i=1}^n W_i \hat{\vartheta}_{1i}(\tau, v) / \hat{\zeta}_{1\tau}^{1/2}(v) \right|. \quad \text{375}$$

Then the null distribution of  $\mathcal{T}_{21}$  can be approximated by the conditional distribution of  $\mathcal{T}_{21}^*$  given the observed data, and the critical value  $c_{21}(\alpha)$  can be taken as the  $(1 - \alpha)$ -percentile of the conditional distribution of  $\mathcal{T}_{21}^*$ .

#### 4. SIMULATION STUDIES

In this section, we conduct simulation studies to evaluate the finite sample performance of the proposed method. We first generate  $Z^* = (Z_1^*, Z_2^*)$  from a multivariate normal distribution with mean 0, variance 1 and correlation 0.5. Then let  $\tilde{Z}_1 = I(Z_1^* > 0)$  and  $\tilde{Z}_2 = \Phi(Z_2^*)$ . Under these settings,  $\tilde{Z}_1$  follows a Bernoulli distribution with success probability 0.5,  $\tilde{Z}_2$  follows a uniform distribution  $U(0, 1)$ , and  $\tilde{Z}_1$  is related to  $\tilde{Z}_2$ . The conditional density of  $V$  equals 1 if  $\mu\tilde{Z}_1 = 0$  and is  $2v$  otherwise, where  $\mu$  is a constant given below. The failure time  $T$  satisfies  $P(T \leq t | V, \tilde{Z}) = \Phi(\log t - \gamma(V)^\top \tilde{Z})$ , where  $\gamma(v) = (\gamma_1(v), \gamma_2(v))^\top$ . Let  $Z = (1, \tilde{Z}^\top)^\top$  with  $\tilde{Z} = (\tilde{Z}_1, \tilde{Z}_2)^\top$ . Under the preceding settings, the underlying quantile regression model takes the form

$$\log\{Q_v(\tau | Z)\} = \Phi^{-1}(\tau) + \left[ \gamma_1(v) + \mu \left\{ \Phi^{-1}\left(\frac{\tau}{2v}\right) - \Phi^{-1}(\tau) \right\} \right] \tilde{Z}_1 + \gamma_2(v) \tilde{Z}_2.$$

Here, we set  $\gamma_1(v) = \gamma_{11} + \gamma_{12}v$  and  $\gamma_2(v) = 0.5(1 + v^2)$ . In the study, four models are considered:

$$\begin{aligned} \text{(M1)} : (\mu, \gamma_{11}, \gamma_{12}) &= (0, 0, 0); & \text{(M2)} : (\mu, \gamma_{11}, \gamma_{12}) &= (0, 0.4, 0); \\ \text{(M3)} : (\mu, \gamma_{11}, \gamma_{12}) &= (1, 0.43, 0); & \text{(M4)} : (\mu, \gamma_{11}, \gamma_{12}) &= (1, 0.9, -0.6). \end{aligned}$$

Model (M1) is considered for the null hypothesis  $H_{10}$  of no vaccine efficacy, and models (M2)-(M4) are considered for the alternative hypotheses  $H_{11}$  and  $H_{12}$ . The departure from  $H_{10}$  increases as the model moves from (M2) to (M4). Model (M2) is also considered for the null hypothesis  $H_{20}$  of constant vaccine efficacy, while (M3)-(M4) are considered for the alternatives  $H_{21}$  and  $H_{22}$ . The departure from  $H_{20}$  increases as the model moves from (M3) to (M4). The censoring time  $C$  is generated from an exponential distribution with mean  $c$ , where  $c$  is chosen to give censoring rate of about 40% under models (M1)-(M4).

Set  $[a, b] = [0.3, 0.8]$  and  $[\tau_0, \tau_U] = [0.1, 0.4]$ . For the calculations of  $\widehat{\text{CQVE}}_\tau(v)$  and the tests studied in Section 3, we take the grid of 50 evenly spaced points in  $[a, b]$  and  $[\tau_0, \tau_U]$ . In the test statistics  $\mathcal{T}_{21}$  and  $\mathcal{T}_{22}(\tau)$ , the parameters  $a_1$  and  $\tau_0^*$  are taken as 0.35 and 0.15, respectively. The kernel function is set to be the Epanechnikov kernel function, that is,  $K(x) = 0.75(1 - x^2)I(|x| < 1)$ . The bandwidth  $h$  is chosen using 5-fold cross-validation method, and the optimal bandwidth is denoted by  $h_{\text{opt}}$ . For a sensitivity analysis, we also consider  $h=0.15$  or  $0.2$  by using the rule of thumb bandwidth. The critical values are calculated using the resampling method with 1000 simulated realizations. The results presented below are based on 1000 replications with sample sizes  $n = 1000$  and  $1500$ .

Tabs. 1 and S1-S3 in the Supplementary Material report the empirical biases (Bias), the empirical standard deviations (SD), the average of the estimated standard deviations (ESD) of  $\widehat{\text{QVE}}_\tau(v)$  and  $\widehat{\text{CQVE}}_\tau(v)$ , and the coverage probabilities (CP) of the pointwise 95% confidence bands for  $\text{QVE}_\tau(v)$  and  $\text{CQVE}_\tau(v)$ . The results suggest that the proposed estimators perform reasonably well. Specifically, the proposed estimators are nearly unbiased, the estimated standard deviations agree well with the empirical standard deviations, and the coverage probabilities of the pointwise 95% confidence bands are close to the nominal level. The performance of the proposed estimators becomes better when the sample size increases from 1000 to 1500. In addition, the optimal

bandwidth  $h_{\text{opt}}$  leads to smaller estimated standard deviations for  $\widehat{\text{CQVE}}_{\tau}(v)$  and  $\widehat{\text{QVE}}_{\tau}(v)$  than  $h = 0.15$  and  $0.2$ .

Figs. S2-S5 in the Supplementary Material further display the estimated curves of  $\text{QVE}_{\tau}(v)$  and  $\text{CQVE}_{\tau}(v)$  with the pointwise and simultaneous confidence bands under models M1-M4 for  $h_{\text{opt}}$  and  $n = 1000$ . It can be seen that the estimated curves are close to their true curves and the confidence bands cover the entire true curves. The results for other settings are similar to those in Figs. S2-S5. In addition, Figs. S6-S9 in the Supplementary Material also display the estimated curves of  $\beta_{1\tau}^*(v)$  and  $\beta_{2\tau}^*(v)$  with the pointwise confidence intervals under models M1-M4 for  $h_{\text{opt}}$  and  $n = 1000$ . The results suggest that the estimators  $\hat{\beta}_{1\tau}(v)$  and  $\hat{\beta}_{2\tau}(v)$  are nearly unbiased, and the confidence intervals also cover the entire true curves.

To investigate the performance of the test procedures, we also provide the empirical sizes and powers of the test statistics  $\mathcal{T}_{11}$ ,  $\mathcal{T}_{12}(\tau)$ ,  $\mathcal{T}_{21}$  and  $\mathcal{T}_{22}(\tau)$  at the significance level of 0.05. The results are reported in Tab. 2. We see that the empirical sizes of all tests are close to the nominal level 5%, and all tests have reasonable powers to detect deviations from the null hypothesis. The powers of all tests increase as the simulation model moves from (M2) to (M4). In addition, the tests with the optimal bandwidth achieve higher powers than those with  $h = 0.15$  and  $0.2$ . Furthermore, we observe that the ESD of  $\widehat{\text{CQVE}}_{\tau}(v)$  with  $h = 0.2$  is smaller than that with  $h = 0.15$ , which suggests that the empirical powers of the tests increase as the bandwidth varies from 0.15 to 0.2. This is because that a larger bandwidth usually leads to a smaller ESD of  $\widehat{\text{CQVE}}_{\tau}(v)$ , but the biases retain approximately the same, resulting in increased power for the larger bandwidth. Such phenomenon was also observed by Sun et al. (2009) under the mark-specific proportional hazards model, which may be associated with the convergence rate of the normalized  $\widehat{\text{CQVE}}_{\tau}(v)$  to a Gaussian process.

## 5. REAL DATA ANALYSIS

For illustration purposes, we applied the proposed method to a dataset from the first HIV vaccine efficacy trial. The trial was carried out in North America and The Netherlands, and 5403 HIV-negative volunteers at risk were enrolled for acquiring HIV infection (Flynn et al., 2005). Volunteers were randomly assigned to receive either a recombinant glycoprotein 120 vaccine (AIDSVAX) or placebo in a 2:1 ratio, and were monitored for HIV infection at semi-annual HIV testing visits for 36 months. The vaccine may only provide protection for HIV strains genetically similar to the HIV virus or viruses represented in the vaccine. The similarity between the infecting virus and the virus contained in the vaccine construct was measured by the genetic distance, which was defined as the percent mismatch of amino acids between two aligned HIV sequences. Since the HIV-gp120 region contained neutralizing epitopes that potentially induced anti-HIV antibody responses that prevented HIV infection (Wyatt et al., 1998), we defined the mark  $V$  as the percent mismatch of amino acids in the whole gp120 region (581 amino acids long), where all possible mismatches of particular pairs of amino acids (e.g., A versus C) were weighted by the estimated probability of interchange (Nickle et al., 2005). During the trial, 368 individuals were infected with HIV, but 32 individuals had missing marks. Since our proposed method cannot be directly applied to handle missing data, we removed the 32 individuals with missing mark, and focused on the analysis of the remaining 336 samples, as in Sun et al. (2009) and Han et al. (2017). Each of the remaining 336 samples (217 vaccine and 119 placebo) had a unique mark, and the mark  $V$  ranged from 0.059 to 0.261.

Following Sun et al. (2009) and Han et al. (2017), we considered three covariates: treatment indicator ( $Z_1$ , taking value 1 if the volunteer was in the vaccine group and 0 otherwise), age

at the enrollment ( $Z_2$ , ranging 18-62 years with median of 36), and behavioral risk score ( $Z_3$ , taking values 0-7) as defined in Flynn et al. (2005). According to the preliminary analysis as in Fig. 1, we assumed that the data can be described by model (1) on  $[\tau_0, \tau_U] = [0.05, 0.15]$  and  $[a, b] = [0.1, 0.2]$ . We used the Epanechnikov kernel, and the grid points were taken the same as those in the simulation studies. We used the 5-fold cross-validation method to select the optimal bandwidth, and the optimal bandwidth  $h_{\text{opt}} = 0.04$  as shown in Fig. S10 of the Supplementary Material.

The estimated curves for  $\beta_{1\tau}^*(v)$ ,  $\beta_{2\tau}^*(v)$  and  $\beta_{3\tau}^*(v)$  with  $\tau = 0.05, 0.1$  and  $0.15$  and their pointwise confidence bands are given in Figs. S11-S13 of the Supplementary Material. We found that age and behavior risk score had significant effects on the risk of infection, but the vaccine did not seem to be significantly related to the risk of HIV infection. For examples, age had a significant negative effect for the mark larger than 0.14 at  $\tau = 0.05$ , but was nonsignificant over the whole mark interval for  $\tau = 0.1$  and  $0.15$ . Behavior risk score had a negative decreasing effects over the whole mark interval at  $\tau = 0.05$ , while it showed a negative inverted U-shaped pattern for  $\tau = 0.1$  and  $0.15$ . For comparison, we also analyzed the data with the method of Peng & Huang (2008) that does not consider the mark, and the comparison results are given in Figure S14 of the Supplementary Material. It can be seen that the curves estimated by Peng and Huang's method were very different from ours. In particular, Peng and Huang's method showed that the covariate effects did not change with the quantiles (for example, the estimated values for  $\beta_{1\tau}^*(v)$  were between 0.01991462 and 0.01991494), while our method indicated that the covariate effects were varying with the quantiles.

The test statistic  $\mathcal{T}_{11}$  for  $H_{10}$  versus  $H_{11}$  yielded a p-value of 0.193. The p-value of  $\mathcal{T}_{12}(\tau)$  was larger than 0.798 for testing  $H_{10}$  against  $H_{12}$ . These results suggested that the vaccine had no significant efficacy against HIV; see Fig. 2 for the plots of  $\widehat{\text{QVE}}_{\tau}(v)$  and  $\widehat{\text{CQVE}}_{\tau}(v)$ . In addition, we conducted the tests to evaluate whether the vaccine efficacy varied with the mark. The p-value for testing  $H_{20}$  versus  $H_{21}$  was 0.728 for  $\mathcal{T}_{21}$ , while the p-value for testing against  $H_{21}$  was larger than 0.793 for  $\mathcal{T}_{22}(\tau)$ . This indicated that the vaccine efficacy had no varying tendency on the considered mark interval. These results were consistent with those obtained by Han et al. (2017).

#### ACKNOWLEDGEMENT

The authors thank the Editor, Professor Paul Fearnhead, an Associate Editor and two reviewers for their constructive and insightful comments and suggestions that greatly improve the article. The authors gratefully acknowledge Global Solutions for Infectious Diseases (GSID), Dr. Faruk Sinangil and Dr. Peter Gilbert for providing data from the phase 3 HIV vaccine trial VAX004. Lianqiang Qu's research was partly supported by the National Natural Science Foundation of China (No. 12001219). Liuquan Sun's research was partly supported by the National Natural Science Foundation of China (No. 12171463). Yanqing Sun's research was partly supported by the NIAID NIH award number R37AI054165 and the National Science Foundation grant DMS1915829.

#### SUPPLEMENTARY MATERIAL

Supplementary material available at *Biometrika* online includes the proofs of Theorems 1-3, confidence bands for vaccine efficacy, a discussion on multivariate continuous marks, and additional simulation and application results.

## REFERENCES

- AHN, K. W. & KIM, S. (2018). Variable selection with group structure in competing risks quantile regression. *Statist. Med.*, **37**, 1577–1586.
- BROWN, B. M. & WANG, Y.-G. (2007). Induced smoothing for rank regression with censored survival times. *Statist. Med.*, **26**, 828–836.
- CHOI, S., KANG, S. & HUANG, X. (2018). Smoothed quantile regression analysis of competing risks. *Biom. J.* **60**, 934–946.
- DE BACKER, M., EL GHOUGH, A. & VAN KEILEGOM, I. (2019). An adapted loss function for censored quantile regression. *J. Amer. Statist. Assoc.* **114**, 1126–1137.
- FARCOMENI, A. & GERACI, M. (2020). Multistate quantile regression models. *Statist. Med.* **39**, 45–56.
- FLYNN, N. M., FORTHAL, D. N., HARRO, C. D., JUDSON, F. N., MAYER, K. H., PARA, M. F. & THE RGP120 HIV VACCINE STUDY GROUP (2005). Placebocontrolled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. *J. Infect. Dis.* **191**, 654–665.
- GILBERT, P. B., MCKEAGUE, I. W. & SUN, Y. (2008). The 2-sample problem for failure rates depending on a continuous mark: an application to vaccine efficacy. *Biostatistics* **9**, 263–276.
- HAN, D., SUN, L., SUN, Y. & QI, L. (2017). Mark-specific additive hazards regression with continuous marks. *Lifetime Data Anal.* **23**, 467–494.
- HE, X., PAN X., TAN, K. M. & ZHOU, W.-X. (2022). Scalable estimation and inference for censored quantile regression process. *Ann. Statist.* **50**, 2899–2924.
- KIM, M.-O. (2007). Quantile regression with varying coefficients. *Ann. Statist.* **35**, 92–108.
- KOENKER, R. & BASSETT, G. (1978). Regression quantiles. *Econometrica* **46**, 33–50.
- LEE, M. & HAN, J. (2016). Covariate-adjusted quantile inference with competing risks. *Comput. Stat. Data Anal.* **101**, 57–63.
- LIN, D. Y., WEI, L. J., & YING, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80**, 557–572.
- 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, Univ. Washington.
- PENG, L., & FINE, J. P. (2007). Nonparametric quantile inference with competing risks. *Biometrika* **94**, 735–744.
- PENG, L., & FINE, J. P. (2009). Competing risks quantile regression. *J. Amer. Statist. Assoc.* **104**, 1440–1453.
- PENG, L., & HUANG, Y. (2008). Survival analysis based on quantile regression models. *J. Amer. Statist. Assoc.* **103**, 637–649.
- PORTNOY, S. (2003). Censored regression quantiles. *J. Amer. Statist. Assoc.* **98**, 1001–1012.
- POWELL, J. (1984). Least absolute deviations estimation for the censored regression model. *J. Econometrics* **25**, 303–325.
- POWELL, J. (1986). Censored regression quantiles. *J. Econometrics* **32**, 143–155.
- QIAN, J. & PENG, L. (2010). Censored quantile regression with partially functional effects. *Biometrika* **97**, 839–850.
- SUN, Y., GILBERT, P. B. & MCKEAGUE, I. W. (2009). Proportional hazards models with continuous marks. *Ann. Statist.* **37**, 394–426.
- SUN, Y., LI, M. & GILBERT, P. B. (2013). Mark-specific proportional hazards model with multivariate continuous marks and its application to HIV vaccine efficacy trials. *Biostatistics* **14**, 60–74.
- SUN, Y., QI, L., HENG, F. & GILBERT, P. B. (2020). A hybrid approach for the stratified mark-specific proportional hazards models with missing covariates and missing marks, with applications to dengue vaccine efficacy trials. *J. R. Stat. Soc. Ser. C* **69**, 791–814.
- SUN, Y., WANG, H. & GILBERT, P. B. (2012). Quantile regression for competing risks data with missing cause of failure. *Statist. Sinica* **22**, 703–728.
- TIAN, L., ZUCKER, D & WEI, L. J. (2005). On the Cox model with time-varying regression coefficients. *J. Amer. Statist. Assoc.* **100**, 172–183.
- WYATT, R., KWONG, P. D., DESJARDINS, E., SWEET, R. W., ROBINSON, J., HENDRICKSON, W. A. & SO-DROSKI, J. G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**, 705–711.
- ZHANG, Y., WANG, H. J. & ZHU, Z. (2022). Single-index thresholding in quantile regression. *J. Amer. Statist. Assoc.* **117**, 2222–2237.

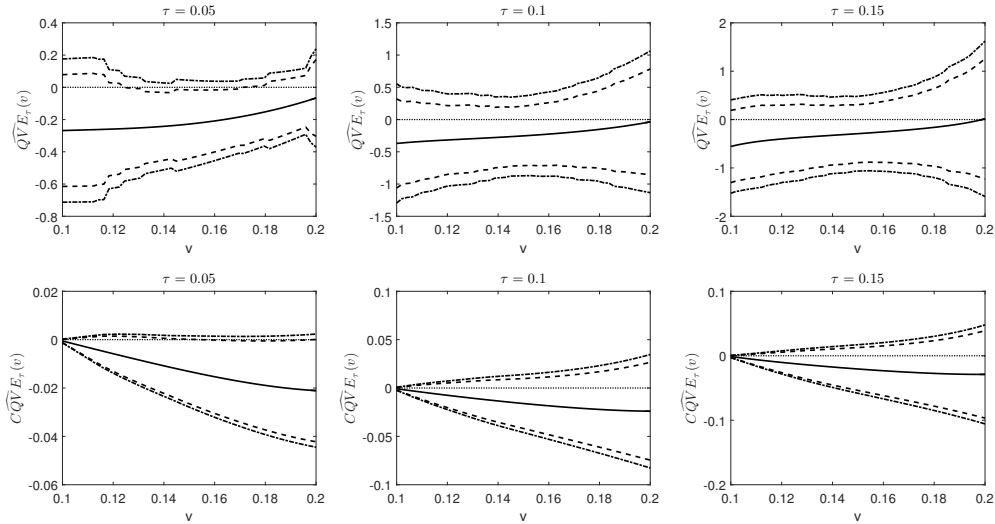
Table 1. *Simulation results for  $\widehat{QVE}_\tau(v)$  and  $\widehat{CQVE}_\tau(v)$  under model M1*

$\tau$	$v$	$h$	$n$	$\widehat{QVE}_\tau(v)$				$\widehat{CQVE}_\tau(v)$			
				Bias	ESD	SD	CP	Bias	ESD	SD	CP
0.1	0.6	0.15	1000	0.031	0.274	0.259	0.926	0.007	0.062	0.058	0.942
			1500	0.020	0.220	0.213	0.932	0.006	0.051	0.049	0.935
		0.2	1000	0.021	0.231	0.218	0.930	0.006	0.058	0.055	0.939
			1500	0.003	0.185	0.176	0.925	0.002	0.047	0.046	0.932
		$h_{\text{opt}}$	1000	0.014	0.198	0.188	0.932	0.004	0.053	0.051	0.938
			1500	0.004	0.175	0.153	0.952	0.003	0.046	0.042	0.946
	0.8	0.15	1000	0.015	0.270	0.249	0.937	0.012	0.085	0.079	0.958
			1500	0.019	0.220	0.211	0.930	0.010	0.069	0.067	0.946
		0.2	1000	0.012	0.236	0.212	0.934	0.009	0.081	0.076	0.958
			1500	0.012	0.189	0.178	0.916	0.004	0.066	0.063	0.942
		$h_{\text{opt}}$	1000	0.016	0.212	0.195	0.922	0.006	0.077	0.073	0.952
			1500	0.018	0.175	0.165	0.926	0.005	0.065	0.058	0.954
0.3	0.6	0.15	1000	0.022	0.238	0.235	0.938	0.005	0.055	0.051	0.945
			1500	0.014	0.194	0.185	0.930	0.004	0.044	0.045	0.938
		0.2	1000	0.012	0.200	0.198	0.935	0.004	0.051	0.048	0.952
			1500	0.014	0.165	0.154	0.941	0.004	0.042	0.041	0.951
		$h_{\text{opt}}$	1000	0.007	0.171	0.167	0.942	0.003	0.047	0.043	0.962
			1500	0.001	0.149	0.140	0.946	0.001	0.039	0.037	0.942
	0.8	0.15	1000	0.005	0.238	0.240	0.927	0.008	0.073	0.069	0.959
			1500	0.021	0.199	0.188	0.939	0.007	0.060	0.059	0.940
		0.2	1000	0.002	0.202	0.198	0.935	0.005	0.069	0.066	0.958
			1500	0.014	0.166	0.162	0.948	0.006	0.057	0.055	0.948
		$h_{\text{opt}}$	1000	0.002	0.177	0.176	0.924	0.003	0.066	0.063	0.950
			1500	0.008	0.154	0.146	0.956	0.002	0.055	0.051	0.962



Table 2. Empirical sizes and powers of the tests  $\mathcal{T}_{11}$ ,  $\mathcal{T}_{12}(\tau)$ ,  $\mathcal{T}_{21}$  and  $\mathcal{T}_{22}(\tau)$  at the nominal level 0.05. The reported values are in percentages.

Model	$h$	$n$	$\mathcal{T}_{11}$	$\mathcal{T}_{12}(0.1)$	$\mathcal{T}_{12}(0.2)$	$\mathcal{T}_{12}(0.3)$	$\mathcal{T}_{21}$	$\mathcal{T}_{22}(0.1)$	$\mathcal{T}_{22}(0.2)$	$\mathcal{T}_{22}(0.3)$
M1	0.15	1000	8.0	2.7	2.7	3.4	3.9	4.3	3.4	3.2
		1500	7.6	4.0	3.2	3.6	4.2	3.7	3.5	3.5
	0.2	1000	8.3	3.0	2.7	3.3	4.4	4.1	4.0	3.3
		1500	7.4	3.5	4.2	4.2	5.1	5.2	3.7	3.6
	$h_{\text{opt}}$	1000	7.0	3.8	2.6	3.4	5.8	3.6	4.0	3.6
		1500	7.2	2.2	3.4	2.2	3.8	4.2	4.6	5.0
M2	0.15	1000	53.2	73.1	82.0	82.2	3.1	5.6	4.4	4.1
		1500	76.1	89.4	94.5	95.1	2.9	4.4	5.5	5.4
	0.2	1000	61.7	74.0	84.8	85.9	2.9	4.4	4.2	4.9
		1500	86.1	90.9	96.6	96.9	4.7	6.0	5.7	5.1
	$h_{\text{opt}}$	1000	73.6	76.4	85.8	90.0	3.0	5.2	4.4	5.0
		1500	88.4	92.4	97.6	97.0	3.6	5.4	6.4	6.0
M3	0.15	1000	75.1	74.3	82.4	82.1	13.4	26.9	41.2	47.8
		1500	95.1	90.6	95.4	96.5	31.5	39.2	57.1	68.0
	0.2	1000	89.0	77.6	89.9	91.5	31.9	32.6	55.0	65.4
		1500	99.1	92.2	97.9	98.4	57.1	46.6	72.5	82.9
	$h_{\text{opt}}$	1000	96.2	82.6	91.4	94.0	53.0	42.2	67.2	77.8
		1500	99.4	95.0	99.2	98.5	69.2	56.8	80.2	88.6
M4	0.15	1000	98.1	94.0	98.7	97.1	31.5	49.8	69.0	72.6
		1500	100	98.5	100	99.9	61.7	68.3	85.6	91.0
	0.2	1000	99.6	95.6	99.1	98.7	57.1	58.6	78.8	84.9
		1500	100	98.7	99.9	99.9	82.8	79.7	93.1	97.1
	$h_{\text{opt}}$	1000	99.8	96.8	99.3	99.8	69.2	73.5	90.3	94.7
		1500	100	99.2	100	100	91.0	85.4	96.2	98.6

Fig. 2. Vaccine trial data analysis: The estimated curves of  $QVE_{\tau}(v)$  and  $CQVE_{\tau}(v)$  with  $\tau = 0.05, 0.1$  and  $0.15$ . The solid lines are the estimated functions, the dashed lines are the pointwise 95% confidence intervals, and the dash-dotted lines are the simultaneous 95% confidence bands.