Published in final edited form as:

Biom J. 2011 July ; 53(4): 557–577. doi:10.1002/bimj.201000180.

Approximate Nonparametric Corrected-score Method for Joint Modeling of Survival and Longitudinal Data Measured with Error

Jean de Dieu Tapsoba^{*1}, Shen-Ming Lee¹, and C.Y. Wang²

¹Department of Statistics, Feng Chia University, Taichung, Taiwan 40724, R.O.C.

²Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, U.S.A.

Abstract

We consider the problem of jointly modeling survival time and longitudinal data subject to measurement error. The survival times are modeled through the proportional hazards model and a random effects model is assumed for the longitudinal covariate process. Under this framework, we propose an approximate nonparametric corrected-score estimator for the parameter, which describes the association between the time-to-event and the longitudinal covariate. The term nonparametric refers to the fact that assumptions regarding the distribution of the random effects and that of the measurement error are unnecessary. The finite sample size performance of the approximate nonparametric corrected-score estimator is examined through simulation studies and its asymptotic properties are also developed. Furthermore, the proposed estimator and some existing estimators are applied to real data from an AIDS clinical trial.

Keywords

Corrected score; Cumulant generating function; Measurement error; Proportional hazards; Random effects

1 Introduction

In many longitudinal studies researchers are interested in investigating the relationship between some longitudinal covariates and survival time. However, a routinely occurring complication is that the true longitudinal covariates are subject to biological fluctuation and difficult to measure accurately sometimes. A typical example is the CD4 cell count data arising from the AIDS clinical trial ACTG 175, a randomized study designed primarily to compare four treatment regimens in HIV patients (Hammer *et al.*, 1996). Along with some baseline characteristics and the time to AIDS diagnosis or death, the CD4 cell counts were collected with the purpose of examining the association between the CD4 cell count and the time to AIDS or death. It is well established in the literature that observations of the CD4 count are subject to measurement error (Degrootola and Tu, 1994). If not accounted for, measurement error in the covariates may introduce substantial bias in estimations based on routine regression procedures and lead to unreliable inference (Prentice, 1982; Huges, 1993).

^{*}Corresponding author: tapsoba1@yahoo.fr, Phone: +886 4 2451 7250, ext 4409, Fax: +886 4 2451 7092.

Conflict of Interest Statement

The authors have declared no conflict of interest.

There has been an increasing interest in the literature in developing methods for jointly modeling event time and longitudinal covariate measured with error. Most of these methods, in addition to assuming a model for the survival time, also postulate another model for the true longitudinal covariate process. Relevant references dealing with methods for joint modeling survival and longitudinal data measured with error include the followings. Wulfsohn and Tsiatis (1997) assumed a linear random effects model for the true longitudinal covariate process and the Cox proportional hazards model (Cox, 1972) for the survival time. Wang, Wang and Wang (2000) also considered a linear random effects model for the longitudinal data. They further assumed that the survival time data follow the Cox proportional hazards model, whose covariates are the random effects of the longitudinal data model. Henderson, Diggle and Dobson (2000) formulated a class of joint models for the survival and longitudinal covariate process. Jacqmin-Gadda, Commenges and Dartigues (2006) used a piecewise polynomial random effects model for the longitudinal covariate and a log-normal model for the survival time. More recently, Wu, Liu and Hu (2010) considered a nonparametric mixed-effects model for the longitudinal covariates, which are linked to the survival time through a shared random effects model. A common aspect of the estimation methods developed in the aforementioned papers is the need to specify the distributions of the covariate and measurement errors. Functional approaches, which are more flexible, have been proposed in an effort to relax the distributional assumptions. These approaches usually assume a linear random effects model for the longitudinal data and the proportional hazards model for the survival time data. They include the conditional score (Tsiatis and Davidian, 2001) and the corrected score (Wang, 2006; Song and Wang, 2008) methods, which yield consistent estimators without any specification of the underlying distribution function of the longitudinal covariate. A limitation of these methods is their reliance on the arbitrary assumption of normally distributed measurement errors. Such assumption may not hold in general situation and its inappropriateness may compromise the bias reduction ability of the bias-correction methods and lead to invalid results.

This paper aims to relax the distributional assumption on both the longitudinal covariate and the measurement errors in joint modeling of survival and longitudinal data measured with error. Assuming a random effects model for the true longitudinal covariate process and the proportional hazards model for the survival time, we develop an approximate nonparametric corrected score estimator for the parameter that characterizes the relationship between the survival and longitudinal data. The approach is free from assumptions regarding the distributions of the random effects and measurement errors. The estimation procedure uses cumulant expansion techniques (Kendall and Stuart, 1977, Chapter 3) and also involves estimation of the error moments for the correction of the bias of the naive partial likelihood score.

The remainder of this paper is structured as follows. We formulate the submodels for the longitudinal and survival data in Section 2. In Section 3, we first review briefly two existing functional methods, namely Wang (2006)'s corrected-score approach, which is here referred to as the parametric correction, and the conditional score method. We later present a new method, called the approximate nonparametric corrected score method in this section. In Section 4, the finite sample performance of the proposed estimator, which is not difficult to implement, is investigated via a simulation study. The application to the ACTG-175 data is presented in Section 5. Finally, we conclude with some remarks in Section 6.

2 Formulation of the joint model

Let n denote the number of subjects, which are followed up over a study period of positive and finite length L . For each subject i , let $\tilde{T}_i$ represent the unobserved survival time that is assumed to be right-censored, and C_i denote the censoring time, $i = 1, \dots, n$. The observed

survival data for subject i are $T_i = \min(\tilde{T}_i, C_i)$, the observed survival time, and Δ_i the failure time indicator, which takes on the value 1 if $\tilde{T}_i$ is an observed survival time and 0 otherwise, $i = 1, \dots, n$. Furthermore, let $Z_i(u)$ be the true underlying time-dependent covariate process that is unobservable and $\mathbf{Z}_{0i}(u) = \{Z_{01i}(u), \dots, Z_{0Ki}(u)\}'$ represents a K -dimensional covariate, which consists of both time-independent and time-varying covariates that can be measured accurately at any time u . Some longitudinal measurements of size m_i are taken on the covariate Z_i at ordered occasions t_{ij} , with $t_{im_i} = T_i$. The univariate assumption for the covariate $Z_i(u)$ is simply for convenience and the general multivariate case can be extended straightforwardly but with more complicated calculations.

2.1 Submodel for the longitudinal data

The true longitudinal covariate data, $Z_i(t_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, m_i$ are not observable in the situation under consideration. Letting W_{ij} denote the measurement of the error-prone covariate Z_i at time t_{ij} , we postulate the following additive measurement model for the longitudinal observations:

$$W_{ij} = f(t_{ij}, \alpha_i) + e_{ij},$$

where α_i is a p -dimensional vector of random effects, e_{ij} is the measurement error at time t_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m_i$ and $f(u, \alpha_i)$ denotes the true longitudinal covariate process $Z_i(u)$. Here, we assume that $p \geq 2$ and $f(u, \alpha_i)$ is a differentiable function of α_i . For simplicity in the following we take $f(u, \alpha_i)$ to be of the form $\mathbf{D}(u)\alpha_i$ where $\mathbf{D}(u)$ is the row-vector $(1, u, \dots, u^{p-1})$ for any u . Note that some linearization techniques can be used in cases where $f(u, \alpha_i)$ is nonlinear in some of the components of α_i . The vector of random effects, α_i is assumed to follow an unspecified distribution function with mean μ_α and variance Σ_α . The errors, e_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m_i$ are independent and identically distributed zero-mean random variables with variance σ^2 . Moreover, $\mathbf{e}_i = (e_{i1}, \dots, e_{im_i})'$ is assumed to be independent of α_i and $(T_i, C_i, \Delta_i, m_i, t_i, \mathbf{Z}_{0i})$, where $\mathbf{t}_i = \{t_{i1}, \dots, t_{im_i}\}$ denotes the set of measurement time points for subject i . In addition, it is also assumed that $(T_i, C_i, \Delta_i, m_i, \mathbf{t}_i, \alpha_i, \mathbf{e}_i, \mathbf{Z}_{0i})$, $i = 1, \dots, n$, are independent and identically distributed and the $\mathbf{t}_i$'s are possibly different among subjects. The notations $\mathbf{W}_i = (W_{i1}, \dots, W_{im_i})'$ and $\mathbf{W} = (\mathbf{W}'_1, \dots, \mathbf{W}'_n)'$ are used further.

2.2 Submodel for the survival times

The survival times are linked to the longitudinal covariates via the proportional hazards function. Assuming a non-informative censoring, as well as the independence of survival times to measurement times and errors, the hazards for subject i is defined:

$$\lambda_i(u|Z_i, \mathbf{Z}_{0i}) = \lim_{du \rightarrow 0} \{\text{pr}(u \leq \tilde{T}_i < u + du | \tilde{T}_i \geq u, \alpha_i, \mathbf{Z}_{0i}) / du\} = \lambda_0(u) \exp\{\beta' \mathbf{X}_i(u)\},$$

where $\mathbf{X}_i(u) = \{Z_i(u), \mathbf{Z}'_{0i}(u)\}'$, $\beta = (\beta_1, \beta_2)'$ and $\lambda_0(\cdot)$ is an unspecified baseline hazards function. The main challenge is that $Z_i(u)$ is not observed. Hence, inference of the parameter estimation will be based on the finite number of longitudinal data $\mathbf{W}$.

3 Estimation

We are mainly concerned with the characterization of the relationship between the survival and the longitudinal covariate data. Hence, we aim to estimate the parameter β in the survival time submodel. As the underlying longitudinal covariate process is unobserved, a

naive approach is to replace subject-specific random effects with their least squares estimators. Next, treating α_j as a vector of fixed parameters, we proceed as follows for its estimation at any given time. Let $m_j(u)$ denote the number of longitudinal observations for subject i by time u . Also, let $\mathbf{W}_j(u) = \{W_{j1}, \dots, W_{jm_j(u)}\}'$, and $\mathbf{e}_j(u) = \{e_{j1}, \dots, e_{jm_j(u)}\}'$. At any time u such that $m_j(u) \geq p$, the least square estimator of α_j is

$\hat{\alpha}_j(u) = \{\mathbf{D}_j'(u)\mathbf{D}_j(u)\}^{-1}\mathbf{D}_j'(u)\mathbf{W}_j(u)$, where $\mathbf{D}_j(u)$ is the $m_j(u) \times p$ -matrix whose j th row is $\mathbf{D}(t_{ij})$.

Note that $\hat{\alpha}_j(u) = \alpha_j + \hat{\mathbf{e}}_j(u)$, where $\hat{\mathbf{e}}_j(u) = \{\mathbf{D}_j'(u)\mathbf{D}_j(u)\}^{-1}\mathbf{D}_j'(u)\mathbf{e}_j(u)$. Therefore, given α_j , $\hat{\alpha}_j(u)$ has mean α_j . Now, define $\hat{\mathbf{Z}}_j(u) \equiv \mathbf{D}(u)\hat{\alpha}_j(u)$,

$\hat{\mathbf{X}}_j(u) \equiv \{\hat{\mathbf{Z}}_j(u), \mathbf{Z}_{0j}'(u)\}'$, and $\mathbf{R}_j(u) \equiv \{\mathbf{D}(u)\hat{\mathbf{e}}_j(u), \mathbf{0}_k'\}'$. It is clear that $E\{\hat{\mathbf{Z}}_j(u)|\mathbf{Z}_j(u)\} = \mathbf{Z}_j(u)$ and $\hat{\mathbf{X}}_j(u) = \mathbf{X}_j(u) + \mathbf{R}_j(u)$. Furthermore, let $Y_j(u) \equiv I(T_j \geq u, t_{jp} \leq u)$ denote the at-risk process and define the counting process increment $dN_j(u) \equiv I(\Delta_j = 1, u \leq T_j < u + du, t_{jp} \leq u)$, where $I(\cdot)$ is an indicator function. The "ideal" regression corresponds to the case when the true longitudinal covariate process is observed. In this case, the parameter of interest, β , is estimated based on the usual Cox partial likelihood score function, which is defined as follows.

$$U_n^{id}(\beta, \mathbf{X}) \equiv \frac{1}{n} \sum_{i=1}^n \int_0^L \left\{ \mathbf{X}_i(u) - \frac{S_1(u, \beta, \mathbf{X})}{S_0(u, \beta, \mathbf{X})} \right\} dN_i(u),$$

where $\mathbf{X} \equiv (\mathbf{X}_1', \dots, \mathbf{X}_n')'$ and $S_k(u, \beta, \mathbf{X}) \equiv n^{-1} \sum_{i=1}^n Y_i(u) \mathbf{X}_i^k(u) \exp\{\beta' \mathbf{X}_i(u)\}$, $k = 0, 1$.

The "ideal" estimator of β , denoted by $\hat{\beta}_{id}$ is the solution to $U_n^{id}(\beta, \mathbf{X}) = 0$. Since $\mathbf{X}_j(u)$ is not observable in practice, the "naive" regression consists of simply substituting $\hat{\mathbf{X}}_j(u)$ for $\mathbf{X}_j(u)$ in the Cox partial likelihood score at any time u . The resulting "naive" estimator, denoted by $\hat{\beta}_{na}$, satisfies $U_n^{id}(\beta, \hat{\mathbf{X}}) = 0$, where $\hat{\mathbf{X}} \equiv (\hat{\mathbf{X}}_1', \dots, \hat{\mathbf{X}}_n')'$. However, as pointed out by Prentice (1982) and Wang, Wang and Wang (2000), $\hat{\beta}_{na}$ is biased for β and the naive approach may lead to erroneous conclusions. In the following, we present some functional methods, which adjust for the measurement errors without the need to specify the distribution function of the random effects.

3.1 Parametric corrected-score approach

The corrected score approach is a very attractive functional method that was first introduced by Nakamura (1992) for the Cox regression with covariates subject to measurement error. It accounts for the errors in the covariates by removing the error-induced bias of the "naive" score function, which is obtained by simply substituting the observed error-prone covariates for the unobserved ones in the usual Cox partial likelihood score function. Although the correction method does not require any specification of the distribution function of the covariates, it is a routine to make a distributional assumption on the measurement error. Nakamura (1992)'s corrected score estimator assumes normally distributed measurement error in time-independent covariates. Under the same assumption, Wang (2008) proposed a nonparametric maximum likelihood approach for Cox regression with subject-specific measurement error. In jointly modeling survival time and longitudinal covariates measured with error, Wang (2006) and Song and Wang (2008) developed corrected score estimators, which also assume normal errors. In this subsection we adopt the idea of the corrected score approach of Wang (2006) and present an estimator, which we term as the parametric corrected score estimator. Define

$$V_{ki}(u, \beta) \equiv E[\mathbf{R}_i^k(u) \exp\{\beta' \mathbf{R}_i(u)\}] \text{ and } V_k(u, \beta) \equiv n^{-1} \sum_{i=1}^n V_{ki}(u, \beta) \text{ for } k = 0, 1 \text{ and } i = 1, \dots, n.$$

With the assumption that $\mathbf{e}_j$ and α_j are independent, and some simple calculations, it can be established that

$$\frac{E\{S_1(u, \beta, \widehat{\mathbf{X}})\}}{E\{S_0(u, \beta, \widehat{\mathbf{X}})\}} = \frac{s_1(u, \beta)}{s_0(u, \beta)} + \frac{v_1(u, \beta)}{v_0(u, \beta)},$$

where $s_k(u, \beta)$ and $v_k(u, \beta)$ are the limits of $S_k(u, \beta, \mathbf{X})$ and $V_k(u, \beta)$, respectively, for $k = 0, 1$, and fixed u and β . Hence, it is clear that the expectation of $U_n^{na}(\beta_0)$ is not zero, where β_0 denotes the true value of β . Let $\Omega_i(u) = \sigma^2 \mathbf{D}(u) \{\mathbf{D}'_i(u) \mathbf{D}_i(u)\}^{-1} \mathbf{D}'(u)$ and $\Sigma_i(u) \equiv \text{diag} [\Omega_i(u), \mathbf{0}_{K \times K}]$ for any u and $i, i = 1, \dots, n$. For normally distributed measurement errors, $E[\exp\{\beta' \mathbf{R}_i(u)\}] = \tilde{V}_0(u, \beta)$ and $E[\mathbf{R}_i(u) \exp\{\beta' \mathbf{R}_i(u)\}] = \tilde{V}_1(u, \beta)$, where for any u and β , $\tilde{V}_0(u, \beta) \equiv \exp\{\beta' \Sigma_i(u) \beta / 2\}$ and $\tilde{V}_1(u, \beta) \equiv \Sigma_i(u) \beta \exp\{\beta' \Sigma_i(u) \beta / 2\}$, $i = 1, \dots, n$. A corrected score estimating function of the Wang (2006) type is defined as follows.

$$U_n^{pc}(\beta) \equiv \frac{1}{n} \sum_{i=1}^n \int_0^L \left\{ \widehat{\mathbf{X}}_i(u) - \frac{S_1(u, \beta, \widehat{\mathbf{X}})}{S_0(u, \beta, \widehat{\mathbf{X}})} + \frac{\tilde{V}_1(u, \beta)}{\tilde{V}_0(u, \beta)} \right\} dN_i(u), \quad (1)$$

where $\tilde{V}_k(u, \beta) = n^{-1} \sum_{i=1}^n \tilde{V}_{ki}(u, \beta)$, $k = 0, 1$. The parametric corrected score estimator for β , denoted by $\hat{\beta}_{pc}$, solves $U_n^{pc}(\beta) = 0$. The large sample properties of this estimator were derived in Wang (2006).

3.2 Conditional score method

The conditional score method is a functional approach that accounts for the bias by conditioning on a sufficient statistic for $\mathbf{X}_i(u)$. Assuming that the measurement errors are normally distributed and following Tsiatis and Davidian (2001), it can be shown that $\chi_i(u, \beta) = \widehat{\mathbf{X}}_i(u) + \Sigma_i(u) \beta dN_i(u)$ is a sufficient statistic for $\mathbf{X}_i(u)$. Moreover, the conditional score estimating function is specified as follows:

$$U_n^{cs}(\beta) \equiv \frac{1}{n} \sum_{i=1}^n \int_0^L \left\{ \chi_i(u, \beta) - \frac{G_1(u, \beta)}{G_0(u, \beta)} \right\} dN_i(u),$$

where $G_k(u, \beta) \equiv n^{-1} \sum_{i=1}^n Y_i(u) \chi_i^k(u, \beta) \exp\{\beta' \chi_i(u, \beta) - \beta' \Sigma_i(u) \beta / 2\}$, $k = 0, 1$. The conditional score estimator, denoted by $\hat{\beta}_{cs}$, solves $U_n^{cs}(\beta) = 0$. Under the normal error assumption, $\hat{\beta}_{cs}$ is consistent. Note that this assumption may not always be valid in practice. In the following we consider modifying the estimating function in (1) to adjust for situations, where the distribution function of the measurement errors may deviate from normality.

3.3 Approximate nonparametric corrected-score approach

Relaxing the assumption regarding the distribution function of the measurements errors has triggered considerable research, which has focused so far on the proportional hazards model with time-independent covariates measured with errors. Nonparametric corrected-score estimators, which do not need distributional assumptions on the covariates and errors, were proposed by Huang and Wang (2000), Gorfine, Hsu and Prentice (2004), and Song and Huang (2005) when replicated data of the time-independent covariates are available. Note that the problems of these papers were not on joint modeling of longitudinal and survival data. Here the error-prone covariate is measured longitudinally and we propose a correction approach, which is free from distributional assumptions on the longitudinal covariate and the measurement errors. To avoid the need to specify the error distribution, the method assumes that the moment generating function of the errors exists, and uses the residuals from the least

square estimation of subject-specific random coefficients to estimate $E[\exp\{\beta' \mathbf{R}_i(u)\}]$ and $E[\mathbf{R}_i(u) \exp\{\beta' \mathbf{R}_i(u)\}]$ nonparametrically, $i = 1, \dots, n$. For notational convenience, let $d_i(u)$ denote the $m_i(u)$ -dimensional vector $\mathbf{D}(u)\{\mathbf{D}'_i(u)\mathbf{D}_i(u)\}^{-1}\mathbf{D}'_i(u)$. With the assumption that $e_{ij}, j = 1, \dots, m_i, i = 1, \dots, n$, are independent and identically distributed, the application of the cumulant expansion (Kendall and Stuart, 1977, Chapter 3) yields

$$E[\exp\{\beta' \mathbf{R}_i(u)\}] = \exp\left\{\sum_{r=1}^{\infty} \sum_{j=1}^{m_i(u)} \beta_j^r d_{ij}^r(u) \kappa_r / r!\right\}, \quad (2)$$

where κ_r denotes the r th cumulant of the error e_{11} . Also, $E[\mathbf{R}_i(u) \exp\{\beta' \mathbf{R}_i(u)\}]$ can be obtained from (2) by noting that this quantity is $(\beta' / \beta) E[\exp\{\beta' \mathbf{R}_i(u)\}]$, $i = 1, \dots, n$. Note that κ_r can be expressed in terms of the first r moments, $\nu_1, \dots, \nu_r$, of the errors for any positive r using the following recursive formula (Smith, 1995). $\kappa_1 = \nu_1$ and for $r \geq 2$,

$$\kappa_r = \nu_r - \sum_{k=1}^{r-1} \frac{(r-1)!}{k!(r-k-1)!} \kappa_{r-k} \nu_k. \quad (3)$$

For instance, $\kappa_2 = \nu_2 - \nu_1^2$, $\kappa_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3$ and $\kappa_4 = \nu_4 - 4\nu_1\nu_3 - 3\nu_2^2 + 12\nu_1\nu_2\nu_1 - 6\nu_1^4$. There may be situations where the moment $\nu_k, k \geq 1$, is not available and then has to be estimated from the data. For any positive k and u such that $m_i(u) \geq p$, a method of moments estimator

for ν_k can be obtained at time u as follows. Let $\mathbf{P}_i(u) = \mathbf{I}_{m_i(u)} - \mathbf{D}_i(u)\{\mathbf{D}'_i(u)\mathbf{D}_i(u)\}^{-1}\mathbf{D}'_i(u)$, $\mathbf{K}_i(u) = \text{diag}[P_{i11}(u), \dots, P_{im_i(u)m_i(u)}(u)]$ and $\tilde{\mathbf{e}}_i(u) = \{\mathbf{K}_i(u)\}^{-1/2}\hat{\mathbf{e}}_i(u)$, where $\mathbf{I}_{m_i(u)}$ is the identity matrix of size $m_i(u)$, and $P_{ijl}(u), l = 1, \dots, m_i(u)$, are the entries on the main diagonal of $\mathbf{P}_i(u)$, $i = 1, \dots, n$. Letting $\tilde{e}_{ij}(u)$ be the j th element of $\tilde{\mathbf{e}}_i(u)$, it is not difficult to

see that $\tilde{e}_{ij}(u) = \sum_{l=1}^{m_i(u)} H_{ijl}(u) e_{il}(u)$, where $H_{ijl}(u)$ is the (j, l) entry of the matrix $\mathbf{H}_i(u) = \{\mathbf{K}_i(u)\}^{-1/2}\mathbf{P}_i(u)$. Moreover, let $\tilde{\nu}_{ijk}(u)$ denote the moment of order k of $\tilde{e}_{ij}(u)$. It can be easily shown that

$$\tilde{\nu}_{ijk}(u) = \sum_{k_1 + \dots + k_{m_i(u)} = k} \frac{k!}{k_1! \dots k_{m_i(u)}!} \nu_{k_1} H_{ij1}^{k_1}(u) \dots \nu_{k_{m_i(u)}} H_{ijm_i(u)}^{k_{m_i(u)}}(u),$$

where $\nu_0 = 1$ and $\nu_1 = 0$. For illustrative purpose, a few moments are given as follows. We

$$\tilde{\nu}_{ij1}(u)$$

$$= \nu_1 \sum_{l=1}^{m_i(u)} H_{ijl}(u), \tilde{\nu}_{ij2}(u)$$

$$= \nu_2 \sum_{l=1}^{m_i(u)} H_{ijl}^2(u), \tilde{\nu}_{ij3}(u) = \nu_3 \sum_{l=1}^{m_i(u)} H_{ijl}^3(u), \text{ and } \tilde{\nu}_{ij4}(u)$$

$$= \nu_4 \sum_{l=1}^{m_i(u)} H_{ijl}^4(u)$$

have that $+6\nu_2^2 \sum_{l < l'} H_{ijl}^2(u) H_{ijl'}^2(u)$. Therefore, method of moment estimators for ν_1, ν_2, ν_3 and ν_4 are given respectively as

$$\begin{aligned}\widehat{v}_1(u) &= \sum_{i=1}^n \sum_{j=1}^{m_i(u)} I(m_i(u) > p) \tilde{e}_{ij}(u) / M_1(u), \\ \widehat{v}_2(u) &= \sum_{i=1}^n \sum_{j=1}^{m_i(u)} I(m_i(u) > p) \tilde{e}_{ij}^2(u) / M_2(u), \\ \widehat{v}_3(u) &= \sum_{i=1}^n \sum_{j=1}^{m_i(u)} I(m_i(u) > p) \tilde{e}_{ij}^3(u) / M_3(u), \\ \widehat{v}_4(u) &= \sum_{i=1}^n I(m_i(u) > p) \left\{ \sum_{j=1}^{m_i(u)} \tilde{e}_{ij}^4(u) - 6\widehat{v}_2^2(u) \sum_{j=1}^{m_i(u)} \sum_{l < l'} H_{ijl}^2(u) H_{ijl'}^2(u) \right\} / M_4(u),\end{aligned}$$

where $M_k(u) = \sum_{i=1}^n \sum_{j=1}^{m_i(u)} \sum_{l=1}^{m_i(u)} I(m_i(u) > p) H_{ijl}^k(u)$, $k = 1, 2, 3, 4$. For moments of higher orders, their estimators can be obtained similarly. Let $\widehat{v}_k(u)$ denote the estimator for v_k at time u , $k = 1$. An estimator of the cumulant κ_r , denoted by $\widehat{\kappa}_r(u)$, can be deduced by substituting $\widehat{v}_k(u)$ for v_k , $k = 1, \dots, r$, in (3) at time u . A practical inconvenience associated with the cumulant expansion in (3.2) is the infinite number of terms involved in the expansion. In practice, we suggest approximating $E[\exp\{\boldsymbol{\beta}' \mathbf{R}_A(u)\}]$ by truncating the infinite series at some finite order. Let R denote the truncation order and

$\widehat{V}_0(U, \boldsymbol{\beta}) \equiv n^{-1} \sum_{i=1}^n \exp\{\sum_{r=1}^R \sum_{j=1}^{m_i(u)} \beta_1^r d_{ij}^r(u) \widehat{\kappa}_r(u) / r!\}$ and $\widehat{V}_1(u, \boldsymbol{\beta}) \equiv (\boldsymbol{\beta} / \boldsymbol{\beta})' \widehat{V}_0(u, \boldsymbol{\beta})$. A modification of (3.1) is given by the following estimating function.

$$U_n^{nc}(\boldsymbol{\beta}) \equiv \frac{1}{n} \sum_{i=1}^n \int_0^L \left\{ \widehat{\mathbf{X}}_i(u) - \frac{S_1(u, \boldsymbol{\beta}, \widehat{\mathbf{X}})}{S_0(u, \boldsymbol{\beta}, \widehat{\mathbf{X}})} + \frac{\widehat{V}_1(u, \boldsymbol{\beta})}{\widehat{V}_0(u, \boldsymbol{\beta})} \right\} dN_i(u).$$

The proposed approximate nonparametric corrected-score estimator for the primary parameter $\boldsymbol{\beta}$, denoted by $\widehat{\boldsymbol{\beta}}_{nc}$, satisfies $U_n^{nc}(\boldsymbol{\beta}) = 0$. This equation can be solved using the standard Newton Raphson iterative procedure. Define $\boldsymbol{\Gamma}_n^{nc}(\boldsymbol{\beta}) \equiv -(\partial / \partial \boldsymbol{\beta}) U_n^{nc}(\boldsymbol{\beta})$ and let $\tilde{s}_k(u, \boldsymbol{\beta}_0)$ denote the limit of $S_k(u, \boldsymbol{\beta}_0, \widehat{\mathbf{X}})$ as n goes to infinity, for $k = 0, 1$. In addition, let $dN(u)$ be the limit of $n^{-1} \sum_{i=1}^n dN_i(u)$ and define

$$\boldsymbol{\Psi}_i(\boldsymbol{\beta}_0) \equiv \int_0^L \left\{ \widehat{\mathbf{X}}_i(u) - \frac{\tilde{s}_1(u, \boldsymbol{\beta}_0)}{\tilde{s}_0(u, \boldsymbol{\beta}_0)} + \frac{v_1(u, \boldsymbol{\beta}_0)}{v_0(u, \boldsymbol{\beta}_0)} \right\} dN_i(u) + \left\{ \frac{\widehat{V}_1(u, \boldsymbol{\beta}_0)}{v_0(u, \boldsymbol{\beta}_0)} - \frac{S_{1i}(u, \boldsymbol{\beta}_0, \widehat{\mathbf{X}})}{\tilde{s}_0(u, \boldsymbol{\beta}_0)} + \frac{S_{0i}(u, \boldsymbol{\beta}_0, \widehat{\mathbf{X}}) \tilde{s}_1(u, \boldsymbol{\beta}_0)}{\tilde{s}_0(u, \boldsymbol{\beta}_0)^2} - \frac{\widehat{V}_{0i}(u, \boldsymbol{\beta}_0) v_1(u, \boldsymbol{\beta}_0)}{v_0(u, \boldsymbol{\beta}_0)^2} \right\} dN(u).$$

The asymptotic properties of $\widehat{\boldsymbol{\beta}}_{nc}$ are given in the following results.

Proposition. Under conditions (A1)-(A5) given in the Appendix of the Supplementary Materials, $\widehat{\boldsymbol{\beta}}_{nc}$ is a consistent estimator for $\boldsymbol{\beta}_0$. Moreover, $\sqrt{n}(\widehat{\boldsymbol{\beta}}_{nc} - \boldsymbol{\beta}_0)$ asymptotically follows a zero-mean normal distribution with variance $\{\boldsymbol{\Gamma}^{-1}(\boldsymbol{\beta}_0)\}' \boldsymbol{\Psi}(\boldsymbol{\beta}_0) \{\boldsymbol{\Gamma}^{-1}(\boldsymbol{\beta}_0)\}$, where $\boldsymbol{\Gamma}(\boldsymbol{\beta}_0)$ and $\boldsymbol{\Psi}(\boldsymbol{\beta}_0)$ are the limits of $\boldsymbol{\Gamma}_n^{nc}(\boldsymbol{\beta}_0)$ and $n^{-1} \sum_{i=1}^n \boldsymbol{\Psi}_i(\boldsymbol{\beta}_0) \boldsymbol{\Psi}_i'(\boldsymbol{\beta}_0)$, respectively, as n goes to infinity.

The proof of the proposition is provided in the Appendix of the Supplementary Materials. The estimation of the nuisance parameter $\eta = \{\sigma^2, \mu'_\alpha, \text{vec}(\boldsymbol{\Sigma}_\alpha)\}'$ may be of interest in some applications. Based on the method of moments, an estimating equation for $\boldsymbol{\eta}$ is

$$\sum_{i=1}^n I(m_i > p) \Phi_i(\boldsymbol{\eta}) = 0, \text{ where } \Phi_i(\boldsymbol{\eta}) = [\phi_{i1}(\boldsymbol{\eta}), \phi'_{i2}(\boldsymbol{\eta}), \text{vec}\{\phi_{i3}(\boldsymbol{\eta})\}]', \phi_{i1}(\boldsymbol{\eta}) = (\mathbf{W}_i - \mathbf{D}_i \hat{\boldsymbol{\alpha}}_i)' (\mathbf{W}_i -$$

$\mathbf{D}_i(\hat{\boldsymbol{\alpha}}_i) - \sigma^2(m_i - p)$, $\boldsymbol{\phi}_2(\boldsymbol{\eta}) = \hat{\boldsymbol{\alpha}}_i - \boldsymbol{\mu}_\alpha$ and $\boldsymbol{\phi}_3(\boldsymbol{\eta}) = \hat{\boldsymbol{\alpha}}_i \hat{\boldsymbol{\alpha}}_i' - \boldsymbol{\mu}_\alpha \boldsymbol{\mu}_\alpha' - \boldsymbol{\Sigma}_\alpha - \sigma^2(\mathbf{D}_i' \mathbf{D}_i)^{-1}$. Here $\boldsymbol{\alpha}_i = \hat{\boldsymbol{\alpha}}_i(t_{im_i})$ and $\mathbf{D}_i = \mathbf{D}_i(t_{im_i})$. Furthermore, let $\hat{\boldsymbol{\eta}}$ denote the estimator of $\boldsymbol{\eta}$ which solves this estimating equation and $\boldsymbol{\Theta} = (\boldsymbol{\beta}', \boldsymbol{\eta}')'$, $\hat{\boldsymbol{\Theta}} = (\hat{\boldsymbol{\beta}}_{nc}', \hat{\boldsymbol{\eta}}_0')'$ and $\boldsymbol{\Theta}_0 = (\boldsymbol{\beta}_0', \boldsymbol{\eta}_0')'$, where $\boldsymbol{\eta}_0$ is the vector of true values of the nuisance parameters. The asymptotic variance covariance matrix of $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)$ can be estimated by the robust sandwich estimator, $\omega(\hat{\boldsymbol{\Theta}}) = \{\boldsymbol{\vartheta}^{-1}(\hat{\boldsymbol{\Theta}})\} \boldsymbol{\varphi}(\hat{\boldsymbol{\Theta}}) \{\boldsymbol{\vartheta}^{-1}(\hat{\boldsymbol{\Theta}})\}'$, where

$$\boldsymbol{\vartheta}(\boldsymbol{\Theta}) \equiv n^{-1} \sum_{i=1}^n (\partial/\partial \boldsymbol{\Theta}) \boldsymbol{\varphi}_i(\boldsymbol{\Theta}), \boldsymbol{\varphi}(\boldsymbol{\Theta}) \equiv n^{-1} \sum_{i=1}^n \boldsymbol{\varphi}_i(\boldsymbol{\Theta}) \boldsymbol{\varphi}_i'(\boldsymbol{\Theta}) \text{ and } \boldsymbol{\varphi}_i(\boldsymbol{\Theta}) \equiv \{\Psi_i'(\boldsymbol{\beta}), I(m_i > p) \Phi_i'(\boldsymbol{\eta})\}'$$

4 Simulation study

This section presents a simulation study to investigate the performance of the approximate nonparametric corrected score estimator. We considered a single error-prone longitudinal covariate Z_i and modeled its process using the linear random effects model $Z_i(u) = \alpha_{1i} + \alpha_{2i}u$. The observations of Z_i were generated according to the model $W_{ij} = Z_i(t_{ij}) + e_{ij}$, where e_{ij} is the error and $t_{ij} = 0, 4, 8, 10, 16, 24$. The survival times were generated according to the hazards function $\lambda_i(u) = \lambda_0(u) \exp\{\beta_1 Z_i(u) + \beta_2 Z_{0i}\}$, where Z_{0i} was simulated from a Bernoulli trial with probability of success $p = 0.5$, $\boldsymbol{\beta} = \{-\ln(2), \ln(3)\}'$ and $\lambda_0 = 1$. A common censoring time was assumed for all individuals with a censoring rate of 50%. The available longitudinal observations were obtained prior to the observed event times.

Furthermore, α_{1i} was simulated from the normal distribution $N(\mu_{\alpha,1}, \sigma_{\alpha,1,1}^2)$ with $\mu_{\alpha,1} = 2$ and $\sigma_{\alpha,1,1}^2 = 1$. Meanwhile, α_{2i} was generated independently of α_{1i} and followed the uniform distribution $U(2\mu_{\alpha,2}, 0)$, where $\mu_{\alpha,2} = -0.1$. To explore the robustness of the methods with respect to the distribution function of the measurement errors, three distinct scenarios were considered. Under the first scenario, the e_{ij} 's followed the normal distribution $N(0, \sigma^2)$. The other two scenarios correspond to situations, where the errors distribution deviates from normality. Hence, under the second scenario, a symmetric non-normal distribution was assumed for the e_{ij} 's, which were generated from the uniform distribution $\sigma U(-\sqrt{3}, \sqrt{3})$. In the third situation, the e_{ij} 's followed a skewed distribution, namely $(\chi_v^2 - v)\sigma/\sqrt{2v}$, where $v = 2$. In all three cases the value of σ^2 was set as 0.1 and 0.3 to depict different magnitudes of the measurement error. Moreover, the value of the truncation order was $R = 4$.

A total of 200 Monte Carlo samples were generated with a sample size $n = 200$ and 600 to examine the finite sample performance of the estimators. The parameter $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ was estimated using (i) the "ideal" regression assuming a known longitudinal covariate process, (ii) the "naive" estimation which ignores the measurement error, (iii) the parametric corrected score approach which depends on the normality assumption about the errors, (iv) the conditional score method which also needs the normal error assumption and (v) the approximate nonparametric correction method which does not place any distributional assumption on the errors. The results of the simulations are presented in Tables 1–3. In these tables, "Bias" denotes the average of the 200 estimates $\hat{\beta}_k$ of β_k minus β_k , "ASE" stands for the average of 200 estimates of the standard deviations of the estimate $\hat{\beta}_k$, "SDE" means the square root of the sample variance of 200 estimates of β_k , $k = 1, 2$, and "CP" is the coverage probability of the 95% Wald confidence interval.

Tables 1 shows the results of the simulations under scenario 1. It appears that for β_1 , the ideal estimator outperforms all the other estimators in terms of biases. It has smaller biases and exhibits coverage probabilities close to the nominal 95% level. As expected, all the bias-correction estimators perform better than the naive estimator with respect to the biases and coverage probabilities. However, their standard deviations are generally larger

than those of the naive estimator. Hence, the corrected scores and conditional score methods tend to inflate the variance while reducing the bias. It can also be seen that the parametric and approximate nonparametric corrected score approaches yield very close results. Furthermore, the conditional score estimator displays smaller biases and standard deviations than the parametric and approximate nonparametric corrected score estimators to some extent when $n = 200$ and $\sigma^2 = 0.3$. Hence, the conditional score estimator has slightly better small sample size performance compared with the corrected score estimators under the normal error assumption. This is consistent with findings in Wang (2006) and is understandable since the parametric and approximate nonparametric corrected score estimators are based on asymptotic correction of the naive score. The advantage of the conditional score becomes less apparent when $n = 600$. Regarding β_2 , it can be observed that the measurement errors affect the estimation, although, the effect is less pronounced than that on β_1 .

The results of the simulations under scenario 2 are reported in Table 2. For β_1 , the ideal estimator performs similarly as under the first scenario. The bias-correction estimators provide smaller biases and larger coverage probabilities than the naive estimator. Moreover, the parametric and approximate nonparametric approaches yield close results when $\sigma^2 = 0.1$ or $n = 600$. The nonparametric estimator seems to hold a small advantage over the parametric estimator in terms of biases and standard deviations when $n = 200$ and the error variance is substantial ($\sigma^2 = 0.3$). In this situation, the conditional score estimator shows smaller bias and standard deviation compared with the corrected-score estimators.

Table 3 displays the results of the simulations under scenario 3. It can be observed that, for β_1 , the ideal estimator outperforms all the other estimators. The naive estimator performs the worst with respect to the biases and coverage probabilities. A striking difference between these results and those presented in the previous tables is the notable advantage of the approximate nonparametric corrected score estimator over the parametric corrected score and conditional score estimators with respect to the biases and standard deviations. To illustrate the fact that the estimators requiring the normal error assumption may not be consistent, we additionally perform some simulations under the setting of scenario 3 with an increased sample size $n = 2000$. The results in Table 4 indicate that, for β_1 , the approximate nonparametric estimator performs satisfactorily in terms of biases, standard deviations and coverage probabilities. On the other hand, the parametric corrected score and conditional score estimators show coverage probabilities lower than 90% and large biases when $\sigma^2 = 0.3$.

Overall, for β_1 , the ideal estimator performs better than all the other estimators across all scenarios in terms of biases. Note that this estimator is not useful in real applications since the true process of the longitudinal covariate is not observable. The naive estimator displays greater biases and lower coverage probabilities compared with the bias-correction estimators in all cases. The parametric and approximate nonparametric corrected-score estimators have very close performances when the measurement errors are normally distributed, in which case they are outperformed by the conditional score estimator. They also have similar performances when the errors follow a symmetric non-normal distribution and σ^2 is small. The approximate nonparametric corrected score estimator notably outperforms the parametric corrected score and conditional score estimators with regard to the biases and standard deviations when the distribution of the errors is skewed. Furthermore, the biases and standard deviations of the bias-reduction estimators are increasing functions of σ^2 and decreasing functions of the sample size n . Another observation is that, the parametric corrected score and conditional score methods also work well when σ^2 is small, regardless of whether or not the normal error assumption holds. However, they may lead to inconsistent estimators when the measurement errors are skewed with a large variance. For

β_2 , the estimators have similar performances as for β_1 , although, the impact of the measurement errors on the estimation is not as severe as on the estimation of β_1 . The estimation of the vector of nuisance parameters, $\eta = (\sigma^2, \mu_{\alpha,1}, \mu_{\alpha,2}, \sigma_{\alpha,1,1}^2, \sigma_{\alpha,2,2}^2)$, is influenced by the magnitude of the measurement errors as well. The biases and standard deviations of each of the components of η increase as the measurement errors become substantial.

5 The ACTG-175 data example

5.1 Data

We illustrate the approximate nonparametric correction method using data from the ACTG 175 study, which is a double-blind randomized trial comparing the efficacy of four different antiretroviral treatments: zidovudine (*zdv*) monotherapy, the combination of didanosine with zidovudine (*zdv/ddl*), zidovudine plus zalcitabine (*zdv/ddc*) and didanosine (*ddl*) alone. The study involved 2467 HIV-infected patients who were enrolled in the clinical trial units between December 1991 and October 1992 and followed longitudinally for almost three years (until November 1994). The time-dependent covariate was the CD4 cell count which was measured intermittently and approximately at 0, 8, 20, 32, ..., 168 weeks after randomization. The data on the CD4 count were collected for a secondary objective which was the investigation of a possible link between this covariate and the time to AIDS disease development or death. More details on the clinical trial can be found in Hammer *et al.* (1996). The time was rescaled with origin at the time of randomization because the subjects entered the study at different times. In the following analysis, we include individuals with more than 4 longitudinal observations, resulting in a subset of $n = 2094$ subjects, of whom 264 were diagnosed with AIDS or died in the study period. The censoring was due to administrative censoring in most cases.

5.2 Analysis models and results

The $\log_{10}$ -transformation of CD4 counts was used to stabilize the variance and normalize the within-subject measurement error for the CD4 counts in the analysis and we assumed that the unobserved process of $\log_{10}\text{CD4}$ is $Z_i(u) = \alpha_{1i} + \alpha_{2i}u$ and $X_{ij} = Z_i(t_{ij}) + e_{ij}$, where X_{ij} and e_{ij} are the $\log_{10}\text{CD4}$ and the measurement error, respectively, for subject i at time t_{ij} . Figure 1 displays the $\log_{10}\text{CD4}$ profiles for ten subjects from the analyzed data set.

The time to AIDS disease development or death was linked to the $\log_{10}\text{CD4}$ and other baseline covariates through the hazards function, $\lambda_i(t) = \lambda_0(t) \exp\{\beta_1 Z_i(t) + \beta_2' \mathbf{Z}_{0i}\}$, where $\mathbf{Z}_0 = (Z_{01}, Z_{02}, Z_{03})'$ denotes the vector of treatments indicators, which are defined as follows.

$Z_{01} = 1$ if treatment = *zdv/ddl* and 0 otherwise,

$Z_{02} = 1$ if treatment = *zdv/ddc* and 0 otherwise,

$Z_{03} = 1$ if treatment = *ddl* and 0 otherwise.

The *zdv* treatment is considered as reference with which, all the other treatments are compared. We estimated the parameter $\beta = (\beta_1, \beta_2)'$ based on the naive regression, the parametric-corrected score method and the nonparametric corrected score approach using $R = 4$ as the value of the truncation order. The results are reported in Table 5. In addition to estimating the parameter β , we also estimated the vector of nuisance parameter $\eta = \{\sigma^2, \mu_{\alpha}', \text{vech}(\Sigma_{\alpha})'\}$, where σ^2 is the error variance, μ_{α} and Σ_{α} are the mean and variance covariance matrix of the vector of random effects $\alpha_i = (\alpha_{1i}, \alpha_{2i})'$, respectively. The results of this estimation are also presented in Table 5.

It appears that the estimate of σ^2 is significant (p -value is smaller than 0.0001) at the significance level of 5%. Hence, measurement error is present in the CD4 counts indeed. It can also be noted that the estimate of β_1 obtained from the naive regression, the conditional score, the parametric and approximately nonparametric correction methods are all significant (p -value is smaller than 0.0001). Therefore, there is an association between $\log_{10}\text{CD4}$ and the time to the clinical disease progression or death. The negative sign of the estimates of β_1 implies that the hazards function is a decreasing function of the $\log_{10}\text{CD4}$. Moreover, all the methods lead to insignificant estimates of each of the components of β_2 (the p -values are greater than 0.05), indicating that treatment is not associated with time to AIDS progression or death when $\log_{10}\text{CD4}$ is included in the proportional hazards model. The same observation was noted by Song, Davidian and Tsiatis (2002), although, a previous analysis considering only the treatment effects on hazards revealed that zdv is inferior when compared to the other treatments (Hammer *et al.*, 1996). Note that the conditional score, the parametric and approximate nonparametric correction approaches yield close results. This may be explained by the fact σ^2 is small, and is consistent with the results of the simulation studies when σ^2 is small. In addition, it can also be observed that the estimates of the components of the nuisance parameter, η , are all significant.

6 Conclusion

In this paper, we have presented a new corrected score approach for joint modeling of survival time and longitudinal covariate data measured with errors, considering a random effects model for the longitudinal covariate process and the Cox proportional hazards model for the survival time data. The parameter that characterizes the relationship between the survival time and the longitudinal covariate is treated as primary parameter for which, we have provided a new estimator, termed as the approximate nonparametric corrected-score estimator. This estimator performs well in many situations and its asymptotic properties have been established. It also has the noteworthy feature of being able to account for the measurement errors without distributional assumptions on the longitudinal covariate and the measurement errors. Moreover, when the distribution of the measurement error is skewed, the proposed approximate nonparametric corrected-score estimator has better finite sample performance than the corrected-score estimator of Wang (2006) and the conditional score estimator, which may not be consistent without the normal error assumption.

In some instances, nonlinear random effects models may be more appropriate for the longitudinal covariate process. One possibility is to use linearization techniques and apply the proposed approximate nonparametric correction method. An alternative approach is to model the process of the longitudinal covariate nonparametrically using the local linear regression or polynomial spline methods. We are currently considering these extensions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the National Institute of Health grants P01CA53996, R01ES017030 (C.Y. Wang) and the National Science Foundation grant 99-2118-M035-003-MY2 (S.M. Lee).

References

1. Anderson PK, Gill RD. Cox's regression model for counting process: a large sample study. *Annals of Statistics*. 1982; 10:1100–1120.

2. Cox DR. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society Series B*. 1972; 34:187–220.
3. Deguttola V, Tu XM. Modeling progression of CD4-lymphocyte count and its relationship to survival time. *Biometrics*. 1994; 50:1003–1014. [PubMed: 7786983]
4. Gorfine M, Hsu L, Prentice RL. Nonparametric correction for covariate measurement error in a stratified Cox model. *Biostatistics*. 2004; 5:75–87. [PubMed: 14744829]
5. Hammer SM, Katzenstein DA, Hughes MD, Gundacker H, Schooley RT, Haubrich MR, Henry WK, Lederman MM, Phair JP, Niu M, Hirsch MS, Merigan TC. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *The New England Journal of Medicine*. 1996; 335:1081–1090.
6. Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics*. 2000; 1:465–480. [PubMed: 12933568]
7. Huang Y, Wang CY. Cox regression with covariates unascertainable: a nonparametric-correction approach. *Journal of the American Statistical Association*. 2000; 95:1209–1219.
8. Hughes MD. Regression dilution in the proportional hazards model. *Biometrics*. 1993; 49:1056–1066. [PubMed: 8117900]
9. Jacqmin-Gadda H, Commenges D, Dartigues JF. Random changepoint model for joint modeling of cognitive decline and dementia. *Biometrics*. 2006; 62:254–260. [PubMed: 16542253]
10. Kendall, MG.; Stuart, A. The advanced theory of statistics. 4th edition. Vol. Volume 1. New York: Hafner; 1977.
11. Nakamura T. Proportional hazards model with covariates subject to measurement error. *Biometrics*. 1992; 48:829–838. [PubMed: 1420844]
12. Prentice R. Covariate measurement errors and parameter estimates in failure time regression. *Biometrika*. 1982; 69:331–342.
13. Rudin, W. Principle of mathematical analysis. McGraw-Hill; 1964.
14. Smith PJ. A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *The American Statistician*. 1995; 49:217–218.
15. Song X, Davidian M, Tsiatis AA. An estimator for the proportional hazards model with multiple longitudinal covariates measured with error. *Biostatistics*. 2002; 3:511–528. [PubMed: 12933595]
16. Song X, Huang Y. On corrected score approach for proportional hazards model with covariate measurement error. *Biometrics*. 2005; 61:702–714. [PubMed: 16135021]
17. Song X, Wang CY. Semiparametric approaches for joint modeling of survival time and longitudinal data with time-varying coefficients. *Biometrics*. 2008; 64:557–566. [PubMed: 17725812]
18. Tsiatis AA, Davidian M. A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika*. 2001; 88:447–458.
19. Wang CY. Corrected score estimator for joint modeling of longitudinal and failure time data. *Statistica Sinica*. 2006; 16:235–253.
20. Wang CY. Nonparametric maximum likelihood estimator for Cox regression with subject-specific measurement error. *Scandinavian Journal of Statistics*. 2008; 35:613–628.
21. Wang CY, Wang N, Wang S. Regression analysis when covariates are regression parameters of a random effects model for observed longitudinal measurements. *Biometrics*. 2000; 56:487–495. [PubMed: 10877308]
22. Wu L, Liu W, Hu XJ. Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors. *Biometrics*. 2010; 66:327–335. [PubMed: 19673859]
23. Wulfsohn M, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997; 53:330–339. [PubMed: 9147598]

Figure 1.
 $\log_{10}CD4$ profiles of ten individuals from the analyzed data set.

Table 1

Simulation results for Scenario 1 corresponding to normal measurement error; Ideal, "ideal" regression; Naive, "naive" estimation; CDS, conditional score method; PCR, parametric corrected score method; NCR, approximate nonparametric corrected score approach; Nuis, vector of nuisance parameters.

		$\sigma^2 = 0.10$				$\sigma^2 = 0.30$				
		Bias	ASE	SDE	CP	Bias	ASE	SDE	CP	
$n = 200$	Ideal	β_1	-0.0063	0.0761	0.0732	0.9500	-0.0063	0.0761	0.0732	0.9500
		β_2	0.0300	0.2189	0.2097	0.9500	0.0300	0.2189	0.2097	0.9500
	Naive	β_1	0.0159	0.0751	0.0715	0.9300	0.0564	0.0731	0.0686	0.8600
		β_2	0.0172	0.2181	0.2132	0.9450	-0.0058	0.2180	0.2151	0.9450
	CDS	β_1	-0.0089	0.0801	0.0765	0.9450	-0.0139	0.0891	0.0843	0.9500
		β_2	0.0312	0.2219	0.2172	0.9400	0.0335	0.2301	0.2275	0.9350
	PCR	β_1	-0.0098	0.0808	0.0768	0.9450	-0.0168	0.0918	0.0854	0.9500
		β_2	0.0320	0.2224	0.2174	0.9400	0.0362	0.2321	0.2284	0.9350
	NCR	β_1	-0.0097	0.0808	0.0767	0.9450	-0.0166	0.0918	0.0853	0.9500
		β_2	0.0320	0.2224	0.2174	0.9400	0.0361	0.2321	0.2284	0.9350
	Nuis	σ^2	0.0003	0.0049	0.0047	0.9650	0.0008	0.0146	0.0141	0.9650
		$\mu_{a,1}$	0.0013	0.0714	0.0772	0.9100	0.0016	0.0746	0.0815	0.9150
		$\mu_{a,2}$	-0.0001	0.0042	0.0040	0.9350	-0.0001	0.0045	0.0043	0.9350
		$\sigma_{\alpha,1,1}^2$	-0.0106	0.1026	0.1076	0.9150	-0.0127	0.1125	0.1153	0.9050
	$\sigma_{\alpha,2,2}^2$	0.0000	0.0002	0.0003	0.9250	0.0000	0.0003	0.0003	0.9250	
$n = 600$	Ideal	β_1	-0.0014	0.0432	0.0447	0.9300	-0.0014	0.0432	0.0447	0.9300
		β_2	-0.0000	0.1232	0.1137	0.9300	-0.0000	0.1232	0.1137	0.9300
	Naive	β_1	0.0231	0.0427	0.0448	0.8700	0.0655	0.0414	0.0440	0.6050
		β_2	-0.0163	0.1233	0.1230	0.9300	-0.0413	0.1233	0.1226	0.9300
	CDS	β_1	-0.0009	0.0455	0.0478	0.9200	-0.0010	0.0501	0.0534	0.9200
		β_2	-0.0027	0.1254	0.1254	0.9300	-0.0039	0.1300	0.1301	0.9300
	PCR	β_1	-0.0011	0.0456	0.0479	0.9200	-0.0018	0.0505	0.0537	0.9300

		$\sigma^2 = 0.10$				$\sigma^2 = 0.30$			
		Bias	ASE	SDE	CP	Bias	ASE	SDE	CP
NCR	β_2	-0.0024	0.1255	0.1255	0.9300	-0.0031	0.1303	0.1303	0.9300
	β_1	-0.0011	0.0456	0.0478	0.9200	-0.0018	0.0505	0.0536	0.9300
	β_2	-0.0024	0.1255	0.1254	0.9300	-0.0031	0.1303	0.1303	0.9300
Nuis	σ^2	0.0000	0.0029	0.0029	0.9250	0.0000	0.0087	0.0088	0.9250
	$\mu_{\alpha,1}$	-0.0019	0.0410	0.0438	0.9200	-0.0018	0.0428	0.0453	0.9200
	$\mu_{\alpha,2}$	0.0000	0.0024	0.0025	0.9300	0.0000	0.0026	0.0027	0.9450
	$\sigma^2_{\alpha,1,1}$	-0.0040	0.0598	0.0599	0.9200	-0.0063	0.0650	0.0670	0.9300
	$\sigma^2_{\alpha,2,2}$	0.0000	0.0001	0.0001	0.9600	0.0000	0.0002	0.0002	0.9600

Table 2

Simulation results for scenario 2 when the measurement error follows $\sigma U(-\sqrt{3}, \sqrt{3})$; Ideal, "ideal" regression; Naive, "naive" estimation; CDS, conditional score method; PCR, parametric corrected score method; NCR, approximate nonparametric corrected score approach; Nuis, vector of nuisance parameters.

		$\sigma^2 = 0.10$				$\sigma^2 = 0.30$				
		Bias	ASE	SDE	CP	Bias	ASE	SDE	CP	
$n = 200$	Ideal	β_1	-0.0079	0.0762	0.0707	0.9600	-0.0079	0.0762	0.0707	0.9600
		β_2	-0.0097	0.2205	0.1952	0.9700	-0.0097	0.2205	0.1952	0.9700
	Naive	β_1	0.0141	0.0752	0.0715	0.9600	0.0552	0.0727	0.0703	0.8300
		β_2	-0.0242	0.2208	0.1996	0.9650	-0.0481	0.2208	0.2047	0.9700
	CDS	β_1	-0.0111	0.0805	0.0767	0.9700	-0.0165	0.0893	0.0871	0.9700
		β_2	-0.0099	0.2248	0.2026	0.9700	-0.0080	0.2338	0.2152	0.9700
	PCR	β_1	-0.0120	0.0812	0.0770	0.9700	-0.0194	0.0923	0.0884	0.9750
		β_2	-0.0091	0.2254	0.2029	0.9700	-0.0053	0.2361	0.2161	0.9700
	NCR	β_1	-0.0119	0.0811	0.0770	0.9700	-0.0186	0.0917	0.0881	0.9750
		β_2	-0.0091	0.2254	0.2029	0.9700	-0.0058	0.2358	0.2160	0.9700
	Nuis	σ^2	0.0004	0.0037	0.0037	0.9500	0.0011	0.0111	0.0110	0.9500
		$\mu_{a,1}$	-0.0012	0.0717	0.0692	0.9400	-0.0010	0.0747	0.0734	0.9450
$n = 600$	Ideal	β_1	-0.0052	0.0422	0.0438	0.9250	-0.0052	0.0422	0.0438	0.9250
		β_2	0.0034	0.1221	0.1241	0.9450	0.0034	0.1221	0.1241	0.9450
	Naive	β_1	0.0183	0.0415	0.0412	0.9150	0.0605	0.0402	0.0388	0.6650
		β_2	-0.0096	0.1223	0.1241	0.9400	-0.0331	0.1225	0.1235	0.9200
	CDS	β_1	-0.0060	0.0442	0.0441	0.9350	-0.0073	0.0485	0.0469	0.9400
		β_2	0.0042	0.1242	0.1263	0.9400	0.0054	0.1286	0.1300	0.9500
		$\sigma_{\alpha,1,1}^2$	-0.0101	0.1023	0.0979	0.9550	-0.0122	0.1114	0.1073	0.9500
		$\sigma_{\alpha,2,2}^2$	0.0000	0.0002	0.0002	0.9500	0.0000	0.0003	0.0003	0.9450
		$\mu_{a,2}$	0.0003	0.0042	0.0041	0.9200	0.0003	0.0045	0.0044	0.9250
		σ^2	0.0004	0.0037	0.0037	0.9500	0.0011	0.0111	0.0110	0.9500

		$\sigma^2 = 0.10$				$\sigma^2 = 0.30$			
		Bias	ASE	SDE	CP	Bias	ASE	SDE	CP
PCR	β_1	-0.0062	0.0443	0.0441	0.9350	-0.0082	0.0489	0.0471	0.9400
	β_2	0.0045	0.1243	0.1263	0.9400	0.0063	0.1289	0.1302	0.9500
NCR	β_1	-0.0062	0.0443	0.0441	0.9350	-0.0075	0.0487	0.0469	0.9350
	β_2	0.0045	0.1243	0.1263	0.9400	0.0059	0.1288	0.1235	0.9500
Nuis	σ^2	-0.0001	0.0021	0.0022	0.9350	-0.0004	0.0064	0.0065	0.9350
	$\mu_{\alpha,1}$	0.0025	0.0412	0.0392	0.9550	0.0028	0.0429	0.0412	0.9550
	$\mu_{\alpha,2}$	-0.0001	0.0024	0.0025	0.9200	-0.0001	0.0026	0.0027	0.9250
	$\sigma_{\alpha,1,1}^2$	0.0007	0.0595	0.0629	0.9000	0.0022	0.0651	0.0702	0.9100
	$\sigma_{\alpha,2,2}^2$	0.0000	0.0001	0.0002	0.9200	0.0000	0.0002	0.0002	0.9150

Table 3

Simulation results for scenario 3 when the measurement error follows $(\chi^2_2 - 2)\sigma/2$; Ideal, "ideal" regression; Naive, "naive" estimation; CDS, conditional score method; PCR, parametric corrected score method; NCR, approximate nonparametric corrected score approach; Nuis, vector of nuisance parameters.

		$\sigma^2 = 0.10$					$\sigma^2 = 0.30$				
		Bias		SDE		CP	Bias		SDE		CP
		Bias	ASE	SDE	CP		Bias	ASE	SDE	CP	
$n = 200$	Ideal	β_1	-0.0059	0.0763	0.0751	0.9450	-0.0059	0.0763	0.0751	0.9450	
		β_2	0.0177	0.2168	0.2187	0.9500	0.0177	0.2168	0.2187	0.9500	
	Naive	β_1	0.0155	0.0749	0.0727	0.9150	0.0517	0.0726	0.0685	0.8600	
		β_2	0.0031	0.2162	0.2203	0.9400	0.0180	0.2155	0.2205	0.9350	
	CDS	β_1	-0.0095	0.0801	0.0782	0.9400	-0.0205	0.0890	0.0848	0.9550	
		β_2	0.0170	0.2201	0.2238	0.9450	0.0220	0.2279	0.2318	0.9400	
	PCR	β_1	-0.0103	0.0808	0.0785	0.9450	-0.0236	0.0920	0.0862	0.9600	
		β_2	0.0178	0.2207	0.2241	0.9450	0.0248	0.2303	0.2328	0.9400	
	NCR	β_1	-0.0077	0.0798	0.0776	0.9350	-0.0108	0.0868	0.0817	0.9500	
		β_2	0.0163	0.2201	0.2237	0.9450	0.0176	0.2270	0.2305	0.9400	
	Nuis	σ^2	0.0014	0.0088	0.0090	0.9450	-0.0041	0.0264	0.0270	0.9450	
		$\mu_{a,1}$	0.0008	0.0712	0.0738	0.9500	0.0023	0.0746	0.0755	0.9500	
		$\mu_{a,2}$	-0.0001	0.0043	0.0042	0.9400	-0.0002	0.0046	0.0044	0.9500	
		$\sigma^2_{\alpha,1,1}$	-0.0086	0.1043	0.1011	0.9150	-0.0067	0.1142	0.1091	0.9450	
$n = 600$	Ideal	$\sigma^2_{\alpha,2,2}$	0.0000	0.0003	0.0002	0.9700	0.0000	0.0003	0.0003	0.9550	
		β_1	-0.0044	0.0422	0.0398	0.9400	-0.0044	0.0422	0.0759	0.9500	
	Naive	β_2	-0.0015	0.1226	0.1132	0.9600	-0.0015	0.1226	0.1132	0.9600	
		β_1	0.0166	0.0415	0.0382	0.9300	0.0523	0.0404	0.0373	0.7600	
	CDS	β_2	-0.0124	0.1225	0.1144	0.9400	-0.0320	0.1224	0.1153	0.9350	
		β_1	-0.0080	0.0442	0.0409	0.9350	-0.0188	0.0489	0.0455	0.9150	
	PCR	β_2	0.0019	0.1246	0.1163	0.9500	0.0093	0.1288	0.1213	0.9400	
		β_1	-0.0083	0.0443	0.0409	0.9350	-0.0197	0.0494	0.0457	0.9150	

		$\sigma^2 = 0.10$				$\sigma^2 = 0.30$			
		Bias	ASE	SDE	CP	Bias	ASE	SDE	CP
NCR	β_2	0.0022	0.1246	0.1163	0.9500	0.0103	0.1292	0.1214	0.9400
	β_1	-0.0059	0.0439	0.0405	0.9550	-0.0083	0.0472	0.0437	0.9450
	β_2	0.0008	0.1244	0.1161	0.9450	0.0034	0.1278	0.1202	0.9350
Nuis	σ^2	0.0006	0.0051	0.0052	0.9250	0.0017	0.0152	0.0157	0.9250
	$\mu_{\alpha,1}$	-0.0007	0.0412	0.0398	0.9650	-0.0002	0.0429	0.0421	0.9600
	$\mu_{\alpha,2}$	0.0002	0.0024	0.0022	0.9450	0.0002	0.0025	0.0024	0.9700
$\sigma^2_{\alpha,1,1}$		0.0008	0.0603	0.0576	0.9400	-0.0008	0.0660	0.0638	0.9400
$\sigma^2_{\alpha,2,2}$		0.0000	0.0001	0.0001	0.9500	0.0000	0.0002	0.0002	0.9400

Table 4

Simulation results for scenario 3 when the measurement error follows $(\chi^2_2 - 2)\sigma/2$; Ideal, "ideal" regression; Naive, "naive" estimation; CDS, conditional score method; PCR, parametric corrected score method; NCR, approximate nonparametric corrected score approach; Nuis, vector of nuisance parameters.

$n = 2000$		$\sigma^2 = 0.10$				$\sigma^2 = 0.30$			
		Bias	ASE	SDE	CP	Bias	ASE	SDE	CP
Ideal	β_1	-0.0008	0.0231	0.0240	0.9350	-0.0008	0.0231	0.0240	0.9350
	β_2	0.0076	0.0667	0.0687	0.9150	0.0076	0.0667	0.0687	0.9150
Naive	β_1	0.0208	0.0226	0.0240	0.7950	0.0568	0.0220	0.0236	0.2750
	β_2	-0.0050	0.0668	0.0691	0.9200	-0.0258	0.0667	0.0690	0.9100
CDS	β_1	-0.0032	0.0241	0.0255	0.9350	-0.0126	0.0267	0.0286	0.8800
	β_2	0.0090	0.0679	0.0703	0.9200	0.0142	0.0701	0.0726	0.9200
PCR	β_1	-0.0033	0.0241	0.0255	0.9350	-0.0129	0.0267	0.0286	0.8800
	β_2	0.0091	0.0679	0.0703	0.9200	0.0145	0.0702	0.0726	0.9200
NCR	β_1	-0.0010	0.0239	0.0253	0.9400	-0.0019	0.0256	0.0274	0.9450
	β_2	0.0078	0.0678	0.0701	0.9200	0.0081	0.0695	0.0719	0.9300
Nuis	σ^2	0.0030	0.0027	0.0029	0.9300	0.0008	0.0082	0.0088	0.9300
	$\mu_{\beta,1}$	-0.0033	0.0229	0.0233	0.9400	-0.0032	0.0239	0.0247	0.9500
	$\mu_{\beta,2}$	-0.0001	0.0013	0.0014	0.9500	-0.0001	0.0014	0.0015	0.9400
	$\sigma^2_{\alpha,1,1}$	0.0024	0.0330	0.0294	0.9550	0.0016	0.0361	0.0329	0.9550
	$\sigma^2_{\alpha,2,2}$	0.0000	0.0001	0.0001	0.9550	0.0000	0.0001	0.0001	0.9500

Table 5

Results of the analysis of the ACTG-175 data; Naive, "naive" regression; CDS, conditional score method; PCR, parametric corrected score method; NCR, approximate nonparametric corrected score approach; Est, estimate; SE, standard error.

Parameter	Naive			CDS			
	Est	SE	p-value	Est	SE	p-value	
β	β_1	-1.8881	0.1256	< 0.0001	-1.9144	0.1307	< 0.0001
	β_{21}	0.0042	0.1815	0.9816	0.0186	0.1884	0.9215
	β_{22}	0.1663	0.2401	0.4884	0.1875	0.2482	0.4501
	β_{23}	0.2832	0.2106	0.1786	0.3085	0.2144	0.1501
PCR							
NCR							
	β_1	-1.9150	0.1312	< 0.0001	-1.9178	0.1343	< 0.0001
	β_{21}	0.0189	0.1886	0.9202	0.0197	0.1882	0.9167
	β_{22}	0.1880	0.2485	0.4494	0.1829	0.2492	0.4630
	β_{23}	0.3089	0.2146	0.1501	0.3078	0.2147	0.1517
Nuisance parameter							
η	σ^2	0.014133	0.000715	< 0.000001			
	$\mu_{a,1}$	2.571319	0.003504	< 0.000001			
	$\mu_{a,2}$	-0.002156	0.000090	< 0.000001			
	$\sigma^2_{\alpha,1,1}$	0.016791	0.000849	< 0.000001			
	$\sigma_{a,1,2}$	0.000031	0.000016	0.049600			
	$\sigma^2_{\alpha,2,2}$	0.000016	0.000001	< 0.000001			

$\beta=(\beta_1, \beta_2)'$ is the vector of primary parameters which characterizes the relationship between the time to AIDS or death and the covariates; α_1 and α_2 are the coefficients of log10CD4 count and the vector of dummy variables for treatment, respectively; η is the nuisance parameter, whose components are the variance of the errors, the mean of the random effects and the non-redundant entries of the covariance matrix of the random effects.