CONSULTANT'S FORUM

Semiparametric Modeling of Longitudinal Measurements and Time-to-Event Data—A Two-Stage Regression Calibration Approach

Wen Ye,[1,*] Xihong Lin,[2] and Jeremy M. G. Taylor[1]

[1]Department of Biostatistics, University of Michigan, Ann Arbor,
Michigan 48109, U.S.A.
[2]Department of Biostatistics, Harvard School of Public Health, Boston,
Massachusetts 02115, U.S.A.
*email: wye@umich.edu

SUMMARY. In this article we investigate regression calibration methods to jointly model longitudinal and survival data using a semiparametric longitudinal model and a proportional hazards model. In the longitudinal model, a biomarker is assumed to follow a semiparametric mixed model where covariate effects are modeled parametrically and subject-specific time profiles are modeled nonparametrially using a population smoothing spline and subject-specific random stochastic processes. The Cox model is assumed for survival data by including both the current measure and the rate of change of the underlying longitudinal trajectories as covariates, as motivated by a prostate cancer study application. We develop a two-stage semiparametric regression calibration (RC) method. Two variations of the RC method are considered, risk set regression calibration and a computationally simpler ordinary regression calibration. Simulation results show that the two-stage RC approach performs well in practice and effectively corrects the bias from the naive method. We apply the proposed methods to the analysis of a dataset for evaluating the effects of the longitudinal biomarker PSA on the recurrence of prostate cancer.

KEY WORDS: Joint modeling; Longitudinal data; Semiparametric mixed models; Smoothing splines; Survival analysis.

1. Introduction

It is common in medical research to collect both censored time-to-event data as well as longitudinal measurements of a biomarker. For example, in AIDS research, both time to death and repeated measures of CD4 counts are often collected (Tsiatis, DeGruttola, and Wulfsohn, 1995; Wang and Taylor, 2001), and in prostate cancer research, both time to cancer recurrence and repeated measures of prostate-specific antigen (PSA) are commonly observed (Law, Taylor, and Sandler, 2002). One is often interested in estimating the relative risk of the event of interest associated with the longitudinal trajectory of the biomarker. A challenge in analyzing such data is that longitudinal data are only measured intermittently and are typically subject to measurement error. Hence the true covariate values at each failure time are not available for many subjects. There is a large body of recent literature on joint modeling of longitudinal and survival data, see Tsiatis and Davidian (2004) and Yu et al. (2004) for reviews.However, the majority of the existing methods mainly focus on parametric models for longitudinal data using linear mixed models (Tsiatis et al., 1995; Faucett and Thomas, 1996; Wulfsohn

and Tsiatis, 1997; Xu and Zeger, 2001) or linear stochastic mixed models (Bycott and Taylor, 1998; Wang and Taylor, 2001).

Our model is motivated by a study of cancer recurrence in prostate cancer patients after radiation therapy. In prostate cancer patients, a significant elevation of the PSA level is known to precede clinical recurrence. It is of substantial interest to estimate the dependence of the risk of cancer recurrence on the PSA level and its rate of change (Sartor et al., 1997). The PSA level was only measured intermittently and is subject to measurement error. A challenge in jointly modeling cancer recurrence and PSA trajectories is the presence of complicated nonlinear longitudinal PSA trajectories (Figure 1). Specifically, following radiation therapy, PSA declines immediately, keeps steady during remission, and then rises exponentially before prostate cancer recurrence. In addition, the individual PSA trajectories vary substantially across subjects (Figure 2). These data show that the traditional Cox model/linear mixed model framework is not suitable for modeling the PSA data and a more flexible nonparametric model is desirable.
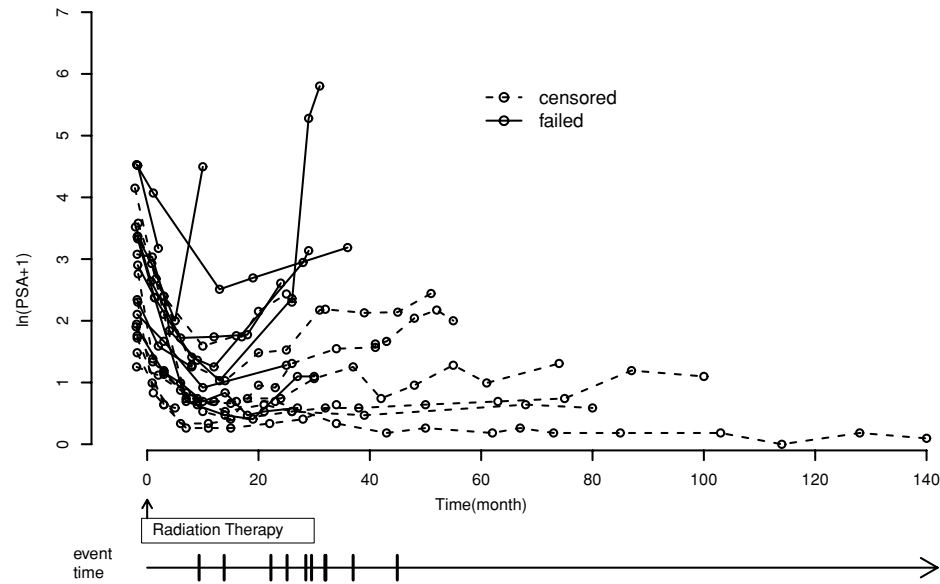
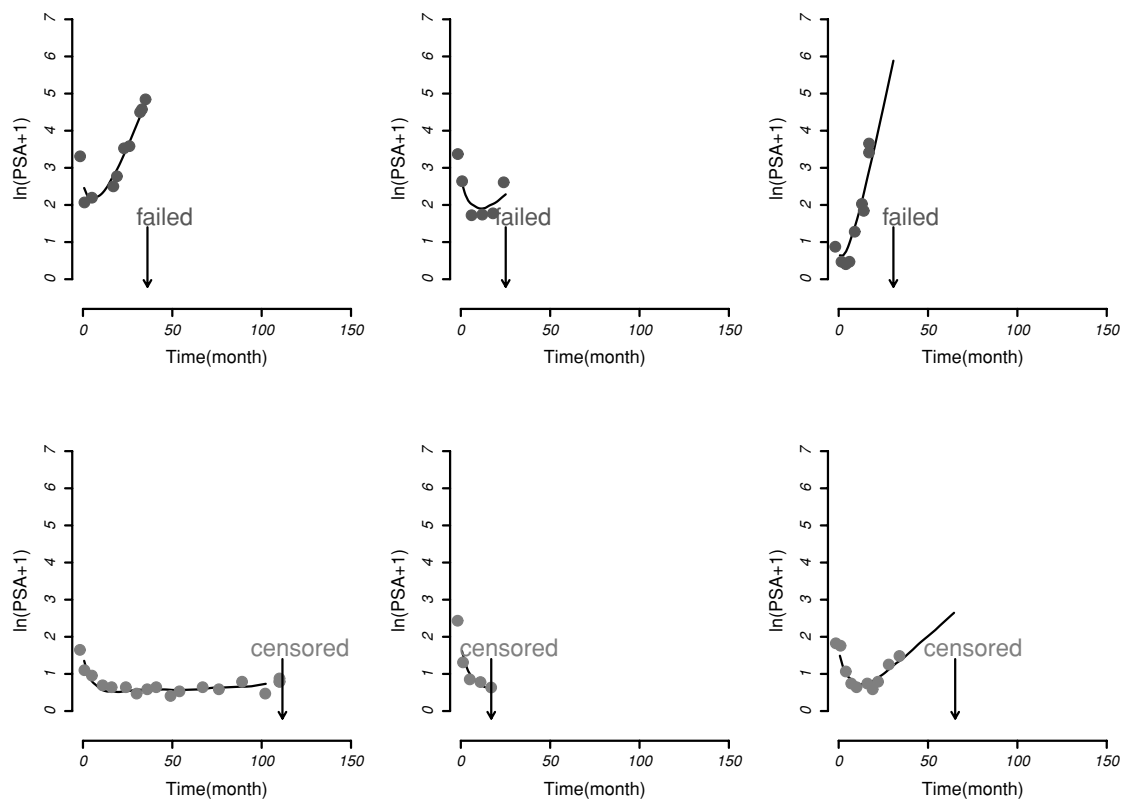**Figure 1.**   Longitudinal PSA profiles.

Figure 2.   Observed PSA measurements and predicted longitudinal PSA curves using ORC under the spline SPMM for six randomly selected patients.

To deal with the nonlinear trajectory of longitudinal co-variates, Pauler and Finkelstein (2002) modeled the PSA profile with a two-piece linear spline. Yu et al. (2004) used a double exponential mixed-effect model. Brown, Ibrahim, and DeGruttola (2005) proposed a Bayesian regression spline method. However, these nonlinear models are mainly based on stringent parametric assumptions, and may not be flex-ible enough to capture complicated time trajectories, espe-cially the considerable variation of the shapes of individual trajectories.

We propose in this article a two-stage semiparametric joint model for longitudinal and survival data, where a

semiparametric stochastic mixed model (SPMM; Zhang et al., 1998) is assumed for the longitudinal data and a Cox model is assumed for the event time. Specifically, the longitudinal SPMM models covariate effects parameterically, the individual trajectories nonparametrically using a population smoothing spline and subject-specific random stochastic process. This model not only provides a flexible way to capture the curvy trajectory, but also provides the rate of change of the curve as a by-product. The Cox model for survival data assumes the hazard rate to depend on the current value of the underlying smooth subject-specific trajectory and its rate of change. We propose two types of regression calibration methods in our two-stage semiparametric joint model: risk set regression calibration (RRC; Tsiatis et al., 1995) and ordinary regression calibration (ORC).

The remainder of this article is organized as follows. Section 2 describes the model. Section 3 presents the two-stage estimating procedure. Section 4 describes the two regression calibration methods. Section 5 gives simulation results. Section 6 applies this method to the Michigan Prostate Cancer Study (MPCS), followed by discussion in Section 7.

2. The Semiparametric Joint Models for Longitudinal and Survival Data

2.1 *The Survival Model Associating the Event Time to the Biomarker*

Let T_i and C_i be the event and censoring times, respectively, for the ith subject ($i = 1, \ldots, m$). We only observe $Y_i = \min(T_i, C_i)$ along with an indicator Δ_i that is equal to 1 if the subject fails and equal to 0 if the subject is censored. Let $\mathbf{X}_i^*(t) = \{X_i^*(u), \, 0 < u \leq t\}$ be the true history of the biomarker up to time t. $\mathbf{X}_i^*(Y_i)$ is the hypothetical true history of a time-dependent covariate, which is not observed. Only intermittent measurements with measurement error $\mathbf{X}_i = (X_{i1}, \ldots, X_{in_i})^T$ are observed at time points $\mathbf{t}_i = (t_{i1}, \ldots, t_{in_i})$. We assume that the censoring time C_i is independent of T_i.

To associate the event time t with the true history of the biomarker up to time t, $\mathbf{X}_i^*(t)$, we assume the hazard is a function of the current biomarker value using a Cox model

$$\lambda\{t; \mathbf{X}_i^*(t), \mathbf{Z}_i\} = \lambda_0(t)\exp\{\alpha X_i^*(t) + \boldsymbol{\gamma}^T \mathbf{Z}_i\}, \quad (1)$$

where $\lambda_0(\cdot) \geq 0$ is the baseline hazard function and $\mathbf{Z}_i$ is a vector of other covariates.

In the prostate cancer study the event time is also allowed to depend on the rate of change of the biomarker trajectory. We hence also consider the following hazard model:

$$\lambda\{t; \mathbf{X}_i^*(t), \mathbf{Z}_i\} = \lambda_0(t)exp\{\alpha_1 X_i^*(t) + \alpha_2 X_i^{*\prime}(t) + \boldsymbol{\gamma}^T \mathbf{Z}_i\}, \quad (2)$$

where $X_i^{*\prime}(t)$ is the current value of the slope of the true trajectory $X_i^*(t)$ at time t. Our goal is to estimate the risk coefficients $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ and $\boldsymbol{\gamma}$.

2.2 *The Semiparametric Stochastic Mixed Model for the Longitudinal Biomarker Trajectory*

We assume the observed intermittent longitudinal measurements X_{ij} measure the unknown true values X_{ij}^* with additive errors as

$$X_{ij} = X_i^*(t_{ij}) + \varepsilon_{ij}, \quad i = 1, \ldots, m \quad j = 1, \ldots, n_i, \quad (3)$$

where $\varepsilon_{ij} \sim \text{iid } N(0, \sigma_\varepsilon^2)$. Motivated by the MPCS data, we assume that the true unobserved subject-specific trajectory $X_i^*(t)$ follows an SPMM (Zhang et al., 1998):

$$X_i^*(t) = \mathbf{Z}_i^T \boldsymbol{\beta} + \psi(t) + \mathbf{U}_i(t)^T \mathbf{b}_i + W_i(t), \quad (4)$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients associated with the covariates $\mathbf{Z}_i$, $\psi(\cdot)$ is an unknown smooth function of the population trajectory that is estimated using a smoothing spline, $\mathbf{b_i}$ is a vector of subject-specific random effects that are associated with the covariates $\mathbf{U}_i(t)$ and is assumed to follow $N\{0, \mathbf{D}(\boldsymbol{\phi})\}$, and $W_i(t_{ij})$ is a mean 0 integrated Wiener (IW) stochastic process, whose covariance takes the form

$$\gamma(s, t) = \text{cov}\{W_i(s), W_i(t)\}$$
$$= \frac{1}{6}[\xi \min^2(t, s)\{3 \times \max(t, s) - \min(t, s)\}]. \quad (5)$$

We also assume ε_{ij}, $\mathbf{b}_i$, and $W_i(t_{ij})$ to be mutually independent.

A natural cubic smoothing spline $\psi(t)$ can be viewed as generated from a partial IW stochastic process (Wahba, 1978).Suppose in the SPMM model (4), $\mathbf{U}_i(t) = (1, t)^T$ and $\psi_i(t) = \psi(t) + \mathbf{U}_i(t)^T \mathbf{b}_i + W_i(t)$ denotes the individual trajectory of subject i. Because the population time trajectory $\psi(t)$ is estimated as a smoothing spline, and $W_i(t)$ is assumed to be a mean zero IW process, the best linear unbiased prediction (BLUP) estimator of the $\psi_i(t)$ is a smoothing spline. A key advantage of using this method is that we allow for both the population longitudinal trajectory to be modeled nonparametrically as a smoothing spline and subject-specific curves to be estimated nonparametrically as smoothing splines.

3. Two-Stage Estimation in the Semiparametric Joint Model

3.1 *The Approximate Survival Model Conditional on the Observed Longitudinal Measures*

As the Cox model (1) is a special case of (2), we focus our discussions below on model (2). The Cox model (2) specifies the dependence of the hazard rate on the true biomarker trajectory $X_i^*(t)$ and its derivative $X_i^{*\prime}(t)$. Because $X_i^*(t)$ is unknown, one cannot estimate $(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ by directly fitting the Cox model (2). Following Prentice (1982) and Dafni and Tsiatis (1998), we use the survival model $\lambda\{t; \mathbf{X}_i^*(t), \mathbf{Z}_i\}$ in (2) and the measurement error model $X_i(t) \,|\, X_i^*(t)$ in (3) to derive the hazard function conditional on the error-prone longitudinal trajectory $\lambda\{t; \mathbf{X}_i(t), \mathbf{Z}_i\}$.

Assume measurement errors are nondifferential (Carroll et al., 2006), that is, $\lambda\{t \,|\, \mathbf{X}_i^*(t), \mathbf{X}_i(t), \mathbf{Z}_i\} = \lambda\{t \,|\, \mathbf{X}_i^*(t), \mathbf{Z}_i(t)\}$. Some calculations show that the induced hazard function $\lambda\{t; \mathbf{X}_i(t), \mathbf{Z}_i\}$ can be written as

$$\lambda\{t; \mathbf{X}_i(t), \mathbf{Z}_i\}$$
$$= \lambda_0(t)e^{\boldsymbol{\gamma}^T \mathbf{Z}_i} E\left(e^{\alpha_1 X_i^*(t) + \alpha_2 X_i^{*\prime}(t)} \,\middle|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t\right). \quad (6)$$

One can see that the conditional expectation $E(e^{\alpha_1 X_i^*(t) + \alpha_2 X_i^{*\prime}(t)} \,|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ is a complicated function of the unknown baseline hazard $\lambda_0(t)$ and the conditional distributions of the covariate $X_i^*(t)$ given $\mathbf{X}_i(t)$ and $\mathbf{Z}_i$, and is thus not tractable. Similar to Dafni and Tsiatis (1998), we propose to approximate this expectation by a first-order

approximation, which is often called a regression calibration approximation, as

$$E\left(e^{\alpha_1 X_i^*(t) + \alpha_2 X_i^{*\prime}(t)} \,\big|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t\right)$$
$$\approx e^{\alpha_1 E\left(e^{X_i^*(t)} \,\big|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t\right) + \alpha_2 E\left(e^{X_i^{*\prime}(t)} \,\big|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t\right)}. \quad (7)$$

Let $\hat{X}_i^*(t) = E(X_i^*(t) \,|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$, and $\hat{X}_i^{*\prime}(t) = E(X_i^{*\prime}(t) \,|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$, then the corresponding induced partial likelihood for model (2) is

$$L(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = \prod_{i=1}^{k} \frac{\displaystyle\prod_{l \in F(t_i)} e^{\boldsymbol{\gamma}^T \mathbf{Z}_l} e^{\alpha_1 \hat{X}_l^*(t) + \alpha_2 \hat{X}_l^{*\prime}(t)}}{\left(\displaystyle\sum_{l \in R(t_i)} e^{\boldsymbol{\gamma}^T Z_l} e^{\alpha_1 \hat{X}_l^*(t) + \alpha_2 \hat{X}_l^{*\prime}(t)}\right)^{J_i}}, \quad (8)$$

where $t_1 < t_2 < \cdots < t_k$ are the k observed unique ordered failure times, J_i is the number of failures at t_i, $R(t_i)$ is the risk set just prior to t_i, and $F(t_i)$ is the set of subjects failing at t_i. The risk coefficients are estimated by maximizing this induced partial likelihood.

3.2 *Smoothing Spline Estimation in the Semiparametric Mixed Model*

We describe in this section smoothing spline estimation of the population curve $\psi(t)$ and the subject-specific curves $\psi_i(t)$ in the semiparametric mixed model (4). Let $\mathbf{t}^0 = (t_1^0, \ldots, t_p^0)$ be a vector of ordered distinct values of the time points $t_{ij} (i = 1, \ldots, m, \ j = 1, \ldots, n_i)$, which are the knots for smoothing spline. Let $\boldsymbol{\Psi} = \{\psi(t_1^0), \ldots \psi(t_p^0)\}^T$ and $\mathbf{N}_i$ be the incidence matrix for the ith subject connecting $\mathbf{t}_i = (t_{i1}, \ldots, t_{in_i})^T$ and $\mathbf{t}_0$ such that the (j, l)th elements of $\mathbf{N}_i$ is 1 if $t_{ij} = t_l^0$ and 0 otherwise $(j = 1, \ldots, n_i, \ l = 1, \ldots, p)$. Let $\mathcal{U}_i = \{\mathbf{U}_i^T(t_{i1}), \mathbf{U}_i^T(t_{i2}), \ldots, \mathbf{U}_i^T(t_{in_i})\}^T$ and $\mathbf{W}_i = \{W_i(t_{i1}), W_i(t_{i2}), \ldots, W_i(t_{in_i})\}^T$. Denote $\boldsymbol{\Gamma}_i$ the covariance matrix of $\mathbf{W}_i$, where the (j, l)th element of $\boldsymbol{\Gamma}_i$ is $\gamma(t_{ij}, t_{il})$.

We estimate the nonparametric function $\psi(t)$ in (4) using a smoothing spline. Assuming that the smoothing parameter λ and the variance component $\{\sigma_\varepsilon^2, \boldsymbol{\phi}^T, \xi\}$ are known, the fixed effect $\boldsymbol{\beta}$ and the smoothing spline estimator of the population mean curve $\psi(t)$ can be estimated by maximizing the penalized likelihood

$$-\sum_{i=1}^{m} \left(\mathbf{X_i} - \mathbf{Z_i^T}\boldsymbol{\beta} - \mathbf{N_i}\boldsymbol{\Psi}\right)^T \mathbf{V_i}^{-1}\left(\mathbf{X_i} - \mathbf{Z_i^T}\boldsymbol{\beta} - \mathbf{N_i}\boldsymbol{\Psi}\right)$$
$$-\sum_{i=1}^{m} \frac{1}{2}\log|\mathbf{V_i}| - \frac{\lambda}{2}\boldsymbol{\Psi}^T\mathbf{K}\boldsymbol{\Psi}, \quad (9)$$

where $\mathbf{V}_i = \mathcal{U}_i\mathbf{D}(\boldsymbol{\phi})\mathcal{U}_i^T + \boldsymbol{\Gamma}_i + \mathbf{I}_i\sigma_\varepsilon^2$, $\mathbf{I}_i$ is an $n_i \times n_i$ identity matrix, and $\mathbf{K}$ is smoothing spline smoothing matrix (Green and Silverman, 1994). Following Zhang et al. (1998), we can formulate the smoothing spline semiparametric mixed model (4) as a working linear mixed model by writing the smoothing spline estimator of $\boldsymbol{\Psi}$ as the BLUP estimator of a linear combination of fixed effects and random effects. Specifically, write $\boldsymbol{\Psi} = \mathbf{T}\boldsymbol{\delta} + \mathbf{B}\mathbf{a}$, where $\mathbf{T} = (\mathbf{1}, \mathbf{t}^0)^T$, $\mathbf{B} = \mathbf{L}(\mathbf{L}^T\mathbf{L})^{-1}$, $\mathbf{K} = \mathbf{L}\mathbf{L}^T$, $\mathbf{L}$ is a $p \times (p - 2)$ full rank matrix. Then smoothing spline estimation in (4) can proceed by fitting the linear mixed

model

$$\mathbf{X}_i = \mathbf{N}_i\mathbf{T}\boldsymbol{\delta} + \mathbf{Z}_i\boldsymbol{\beta} + \mathbf{N}_i\mathbf{B}\mathbf{a} + \mathcal{U}_i\mathbf{b}_i + \mathbf{W}_i + \boldsymbol{\epsilon}_i, \quad (10)$$

where $\boldsymbol{\delta}$ and $\boldsymbol{\beta}$ are regression coefficients, $\mathbf{a} \sim N(0, \tau\mathbf{I})$, $\mathbf{b}_i \sim N\{0, \mathbf{D}(\boldsymbol{\phi})\}$, $\mathbf{W}_i \sim N\{0, \boldsymbol{\Gamma}(\xi)\}$ are random effects, and $\tau = 1/\lambda$. Estimation of the smoothing parameter λ and the variance component $\{\sigma_\varepsilon^2, \boldsymbol{\phi}^T, \xi\}$ can proceed using restricted maximum likelihood (REML) in (10). The smoothing spline estimator of $\boldsymbol{\Psi}$ can be obtained from fitting model (10) using the BLUP as $\hat{\boldsymbol{\Psi}} = \mathbf{T}\hat{\boldsymbol{\delta}} + \mathbf{B}\hat{\mathbf{a}}$. For details, see Zhang et al. (1998).

We focus here on estimating the true continuous subject-specific trajectoies $\mathbf{X}_i^*(t)$ and its derivative $\mathbf{X}_i^{*\prime}(t)$. Specifically, for subject i, the true marker value $X_i^*(t)$ at time t and a subset of the complete history of the observed longitudinal biomarker measurements $\underline{\mathbf{X}}_i = \{X_i(t_{i1}), \ldots, X_i(t_{ij})\}^T \subseteq \mathbf{X}_i$ have a joint multivariate normal distribution:

$$\begin{bmatrix} X_i^*(t) \\ X_i(t_{i1}) \\ X_i(t_{i2}) \\ \vdots \\ X_i(t_{ij}) \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{Z}_i^T\boldsymbol{\beta} + \psi(t) \\ \mathbf{Z}_i^T\boldsymbol{\beta} + \psi(t_{i1}) \\ \mathbf{Z}_i^T\boldsymbol{\beta} + \psi(t_{i2}) \\ \vdots \\ \mathbf{Z}_i^T\boldsymbol{\beta} + \psi(t_{ij}) \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}\right).$$

Let $\mathcal{U}_i = \{\mathbf{U}_i^T(t_{i1}), \mathbf{U}_i^T(t_{i2}), \ldots, \mathbf{U}_i^T(t_{ij})\}^T$ and $\underline{\mathbf{t}}_i = (t_{i1}, \ldots, t_{ij})^T$ be the corresponding subset of the random effect matrix and the time vector, thus we have

$$\boldsymbol{\Sigma}_{11} = \mathbf{U}_i(t)\mathbf{D}(\boldsymbol{\phi})\mathbf{U}_i(t)^T + \frac{1}{3}\xi t^3$$
$$\boldsymbol{\Sigma}_{22} = \mathcal{U}_i\mathbf{D}(\boldsymbol{\phi})\mathcal{U}_i^T + \boldsymbol{\Gamma}_i(\underline{\mathbf{t}}_i, \underline{\mathbf{t}}_i; \xi) + \mathbf{I}_{\mathbf{j}\times\mathbf{j}} \times \sigma_\varepsilon^2$$
$$\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^T = \mathbf{U}_i(t)\mathbf{D}(\boldsymbol{\phi})\mathcal{U}_i^T + \boldsymbol{\Gamma}(t, \underline{\mathbf{t}}_i),$$

where $\boldsymbol{\Gamma}(t, \underline{\mathbf{t}}_i)$ is the covariance matrix of $\mathbf{W}_i$, the (l, l')th element of $\boldsymbol{\Gamma}(\underline{\mathbf{t}}_i, \underline{\mathbf{t}}_i; \xi)$ is $\gamma(t_{il}, t_{il'}; \xi)$, and the lth element of $\boldsymbol{\Gamma}(t, \underline{\mathbf{t}}_i; \xi)$ is $\gamma(t, t_{il}; \xi)$.

Replacing the parameters in model (4) with their estimates, the conditional expectation of $X_i^*(t)$ given $\underline{\mathbf{X}}_i$ can be estimated as

$$\hat{E}_{\underline{\mathbf{X}}_i, \mathbf{Z}_i}\left[X_i^*(t)\right] = \hat{\psi}(t) + \mathbf{Z}_i^T\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}\left(\underline{\mathbf{X}}_i - \mathbf{Z}_i^T\hat{\boldsymbol{\beta}} - \mathbf{N}_i\hat{\boldsymbol{\Psi}}_i\right), \quad (11)$$

where $\hat{\boldsymbol{\Psi}}_i = \{\hat{\psi}(t_{i1}), \ldots \hat{\psi}(t_{ij})\}^T$. It follows from (11) that the conditional expectation of the derivative process $X_i^{*\prime}(t)$ can be estimated as

$$\hat{E}_{\underline{\mathbf{X}}_i, \mathbf{Z}_i}\left[X_i^{*\prime}(t)\right] = \hat{\psi}'(t) + \frac{\partial\hat{\boldsymbol{\Sigma}}_{12}}{\partial t}\hat{\boldsymbol{\Sigma}}_{22}^{-1}\left(\underline{\mathbf{X}}_i - \mathbf{Z}_i^T\hat{\boldsymbol{\beta}} - \mathbf{N}_i\hat{\boldsymbol{\Psi}}_i\right). \quad (12)$$

To make inference for the risk coefficient estimators, we use the standard errors calculated based on the induced partial likelihood as if all the true covariate values were known. This method does not take into account the uncertainty of these estimated time-varying covariates, $\hat{E}(X_i^*(t) \,|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ and $\hat{E}(X_i^{*\prime}(t) \,|\, \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$. Therefore, the estimated standard errors for the risk coefficients are likely to be biased and tend to be smaller than the true variance of these risk coefficient estimates. The extent of this bias will be investigated in simulation studies.

4. Regression Calibration Methods for Estimation in the Semiparametric Joint Model

We discuss two regression calibration methods in this section. They differ in calculations of the conditional expectations $E(X_i^*(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ and $E(X_i^{*\prime}(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ at time t. The first method, RRC, estimates the true subject-specific longitudinal values $X_i^*(t)$ at any time t by fitting the SPMM (4) using subjects in each risk set. The second less computationally intensive method, ORC, estimates $X_i^*(t)$ at any time t by fitting the SPMM only once using all available longitudinal measures.

4.1 *The Risk Set Regression Calibration (RRC) Method*

Suppose that at any time t, the history of the observed longitudinal biomarker values up to time t, that is, $\mathbf{X}_i(t)$, for all the subjects still at risk at time t follow the SPMM (4). Let $\mathbf{X}_i = \{X(t_{i1}), \ldots, X(t_{ij})\}^T$, where t_{ij} is the time point at which the last measurement before t is measured. At the first stage, we fit the SPMM (4) using the data $\mathbf{X}_i$ and calculate $\hat{E}(X_i^*(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ and $\hat{E}(X_i^{*\prime}(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ using (11) and (12). At the second stage, one replaces the conditional expectation in the partial likelihood (8) at time t by $\hat{E}(X_i^*(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ and $\hat{E}(X_i^{*\prime}(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ for each subject i who has survived up to time t.

Note that in this two-stage method, three approximations are used. First, the conditional expected relative risk is approximated by the first-order approximation (7). This approximation works well when the conditional distribution is concentrated. Second, within each risk set, the joint distribution of the true marker value at time t, $X_i^*(t)$, and the history of the observed longitudinal measurements, $\mathbf{X}_i(t)$, are generally not normal. Even if this normality assumption holds at $t = 0$, because withdrawal from each risk set is related to the biomarker, as more people experience the failure event, the distribution of the subject-specific random effects deviates from the normal distribution. Therefore, the normality assumption for the risk set at each event time is only an approximation. Third, $E_{\mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t} X_i^*(t)$ and $E_{\mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t} X_i^{*\prime}(t)$ are approximated with their BLUP estimators, which is the Laplace approximation of these expectations and the Laplace approximation often works well for continuous cases (Breslow and Lin, 1995).

4.2 *The Ordinary Regression Calibration (ORC) Method*

Attenuation in regression coefficients in proportional hazards models arising from covariate measurement error in the case of time-independent covariates has been extensively investigated (Prentice, 1982; Abelson and Prentice, 1997). Xie, Wang, and Prentice (2001) proposed an RRC estimator for this problem and showed that under extreme dependence of either failure or censoring time on the covariate, their RRC estimator eliminates much of the bias that presents in an ORC estimator. However, under most of their simulation conditions, both RRC and ORC estimators perform similarly well. Because the RRC estimation procedure recalibrates at each risk set, it is computationally expensive. In view of the results in Xie et al. (2001), we also develop an ORC estimator for the semiparametric joint model, which is much less time consuming than the RRC estimator.

In the first stage of the ORC estimation procedure, parameters in model (4) are estimated from a single fit to all

the available longitudinal data for all subjects. Let $\underline{\mathbf{X}}_i = \{X(t_{i1}), \ldots, X(t_{in_i})\}^T$ denote all the observed longitudinal measures for subject i. One then fits the SPMM to the longitudinal data $\underline{\mathbf{X}}_i$. The BLUP estimates of $X_i^*(t)$ at each failure time are then calculated using (11) and (12) for all the subjects at each failure time. In the second stage, these BLUP estimates are substituted to the partial likelihood (8) for estimating the regression coefficients in the Cox model (2). Note that in this procedure, in contrast to the RRC procedure, future data are also used to calculate the BLUP estimates $\hat{X}_i^*(t)$ for each t.

4.3 *Comparison of Risk Set Regression Calibration with Ordinary Regression Calibration*

In both the RRC and ORC procedures, the true longitudinal biomarker values in the partial likelihood function are substituted by some predicted values. How well these procedures perform is directly related to the bias and variability of these predicted biomarker values. In general, we expect the BLUP subject-specific curve estimates using the RRC procedure to be less biased but more variant than those using the ORC procedure.

Because withdrawal from the risk set is related to the biomarker, subjects at higher risk tend to have less biomarker measurements. Therefore, if the failure of the event strongly depends on the longitudinal covariate, the parameters in model (4) from a single fit to the complete data using the ORC can be biased due to the informative dropout, and the corresponding $\hat{E}(X_i^*(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ and $\hat{E}(X_i^{*\prime}(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ are biased. In contrast, the RRC estimator is developed to accommodate this change of risk set over time, and BLUP estimates using the RRC procedure tend to have less bias than those in the ORC procedure. However, in a finite sample situation, the RRC procedure, which recalibrates $\hat{X}_i^*(t)$ at each risk set, is subject to a decreasing number of subjects over time. In addition, in the RRC procedure, $\hat{X}_i^*(t)$ are obtained through extrapolation in time in each risk set, whereas in the ORC procedure interpolation is used to predict these covariate values for most of the time points. Therefore, the BLUP estimates using the RRC procedure are often more variant.

The balance between bias and variation of these BLUP estimates determines the relative performance of the RRC and ORC procedures. For instance, when the failure rate is small and events are rare, the change of the distribution of the risk set over time might be ignorable. In this case, the nonparametric BLUP estimates of the individual longitudinal trajectories $\hat{X}_i^*(t)$ and their derivatives $\hat{X}_i^{*\prime}(t)$ can approximate $E(X_i^*(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ and $E(X_i^{*\prime}(t) \mid \mathbf{X}_i(t), \mathbf{Z}_i, Y_i \geq t)$ quite well, and the ORC procedure might perform as well as or even better than the RRC procedure when considering the bias/variance tradeoff. Because the ORC procedure is much less time consuming, it is worth studying and comparing the properties of the relative risk regression coefficient estimates using these two procedures.

5. Simulation Study

We perform extensive simulation studies to compare the finite sample size performance of the RRC with the ORC methods under different scenarios. We considered the effect of (1) the magnitude of measurement error σ_ε^2, (2) the magnitude of the

risk coefficient α, (3) the censoring rate, and (4) the sampling frequency of longitudinal measurements. Each of these four simulation studies involves 200 Monte Carlo datasets of sample size $m = 300$ each.

Let $\mathbf{t}_{\text{knot}}$ be a vector of 19 equally spaced time points in $[0, 144]$. For each subject i, the true longitudinal trajectory $X_i^*(t)$ was generated from the following model:

$$X_i^*(t) = \psi(t) + g_i(t), \tag{13}$$

where $\psi(t) = \log(5.7 \times e^{-0.19 \times t} + 0.244 \times e^{0.05 \times t} + 1)$ is the mean population curve, and $g_i(t)$ is a subject-specific cubic smoothing spline with 19 distinct knots $\mathbf{t}_{\text{knot}}$, characterizing the subject-specific deviation from $\psi(t)$. The values of $g_i(t)$ at the distinct knots $\mathbf{t}_{\text{knot}}$ are given by $\mathbf{1}b_{0i} + \mathbf{t}_{\text{knot}} b_{1i} + \mathbf{B}\mathbf{a}_i$, where the 19×17 matrix $\mathbf{B}$ is defined in Section 3.2., $b_{0i} \sim N(0, 0.1^2)$, $b_{1i} \sim N(0, 0.02^2)$, and $\mathbf{a}_i \sim N(0, 0.0025^2\mathbf{I}_{17 \times 17})$.

For simplicity, to study the influence of the four aforementioned factors on the properties of $\hat{\alpha}_{ORC}$ and $\hat{\alpha}_{RRC}$, we considered the survival model (1), in which the risk depends on the current value of the longitudinal biomarker. For each simulation study, we varied one factor. Besides the ORC and the RRC methods, the "last value carried forward (LVCF)" method was also used as a comparison. This method simply pulls forward the nearest preceding value of the marker and treats it as if it were the current value of the marker at the failure time. We also considered a fourth method, which uses the true unobserved $X_i^*(t)$ at each event time and serves as a benchmark, against which the ORC, RRC, and LVCF can be compared. The properties examined include the bias or the relative bias, the empirical standard error (EMP_SE), the estimated standard error (EST_SE), the mean squared error (MSE), and the coverage rate of the nominal 95% confidence interval (CI).

We first describe the basic setting of the model considered in our simulation, and then explain the variation of the relevant factors in each of these simulation studies. In the basic model, measurements of $X_i^*(t)$ were taken to follow a nominal time schedule of $t_{ij} = (0, 1, 2, \ldots, 144)$, that is, every one time unit from 0 to 144. The missing rate of measurements was 60% for time <12, and increased at a rate of 5% every 12 time units, and for time ≥ 60, the missing rate was always 90%. The observed measurements were simulated with a normal measurement error, $\epsilon_{ij} \sim N(0, 0.3^2)$. For each subject i, the failure time T_i was generated independently with the first event from a nonhomogeneous Poisson process with the intensity function

$$\lambda_i(t) = \lambda_0(t)e^{\alpha X_i^*(t)}, \tag{14}$$

where $\lambda_0(t) \equiv 0.001$ and $\alpha = 2.0$. The censoring time C_i was generated independently from a Weibull distribution with the hazard function $\lambda_C(t) = \lambda_c \gamma_c(\lambda_c t)^{\gamma_c - 1}$, where $\lambda_c = 0.008$ and $\gamma_c = 1.614$. If T_i and C_i are both >144, subject i is set to be censored at $t = 144$. The marginal censoring probability is approximately 30%.

In the first simulation study, we investigated the impact of the measurement error σ_ε^2 in (3) by comparing the above basic setting to two other settings, one with a smaller measurement error $\sigma_\varepsilon^2 = 0.1^2$ and the other with a larger measurement error $\sigma_\varepsilon^2 = 0.6^2$. As shown in Table 1, all three estimates, LVCF, ORC, and RRC, are attenuated and the bias becomes more se-

Table 1
Simulation study to examine the impact of the measurement error on the performance of the risk coefficient estimates using the four methods: the true longitudinal biomarker (TRUE), last value carried forward (LVCF), ordinary regression calibration (ORC), and risk set regression calibration (RRC)

Method	Average $\hat{\alpha}$	Bias	Empirical SE	Estimated SE	MSE	95% CI coverage
TRUE	2.016	0.016	0.142	0.134	0.020	0.935
$\sigma_\varepsilon = 0.1$						
LVCF	1.840	−0.160	0.123	0.126	0.041	0.740
ORC	1.972	−0.026	0.139	0.131	0.020	0.930
RRC	1.960	−0.040	0.143	0.132	0.022	0.905
$\sigma_\varepsilon = 0.3$						
LVCF	1.491	−0.509	0.105	0.109	0.270	0.005
ORC	1.840	−0.160	0.135	0.126	0.044	0.710
RRC	1.828	−0.172	0.142	0.128	0.050	0.670
$\sigma_\varepsilon = 0.6$						
LVCF	0.937	−1.063	0.079	0.082	1.136	0.000
ORC	1.638	−0.362	0.131	0.122	0.149	0.165
RRC	1.644	−0.356	0.141	0.123	0.146	0.235

vere as the measurement error increases. In general, the ORC estimator and the RRC estimator perform similarly, and both correct a lot of the bias in the naive method LVCF. At $\sigma_\varepsilon^2 = 0.1^2$ and $\sigma_\varepsilon^2 = 0.3^2$, the two regression calibration estimators perform fairly well, and the ORC estimator is slightly better than the RRC estimator. At $\sigma_\varepsilon^2 = 0.6^2$, all these three estimators have more bias, but the RRC estimator is slightly less biased and has a slightly higher coverage rate than the ORC estimator. For both regression calibration estimators, the estimated standard errors of the risk coefficient estimates are in general slightly smaller than their empirical counterparts. This is because by using the standard error formula for the Cox proportional hazards model in the second stage, we do not take account of the variability of the plugged in BLUP estimates of the time-varying covariates. Our results show that the difference is small. Table 1 shows the ratio of the average estimated SE and the empirical SE ranges from 0.87 to 0.92 for RRC and from 0.93 to 0.94 for ORC. The naive LVCF method tends to have smaller empirical SEs and estimated SEs than those of the two regression calibration estimators, but ORC and RRC have much better performance than LVCF in terms of MSEs. We also investigated bootstrap method for estimating SEs of α in the basic setting for the ORC method. One hundred bootstrap samples were used for each simulated dataset. The results show that the bootstrap SE is very close to the empirical SE for the ORC methods (bootstrap SE = 0.131, empirical SE = 0.133, average estimated SE = 0.126).

Table 2 examines the effect of the size of the risk coefficient α on the performance of the proposed methods. When $\alpha = 1$ both the two regression calibration estimators behave well. The ORC estimator is slightly less biased and has a slightly higher coverage rate than the RRC estimator. As α increases both regression calibration estimators become more biased, although the biases are fairly small and are about 7–11%. However, the RRC estimator seems to be able to correct the bias better than the ORC estimator when α is large. When $\alpha = 3$, the bias of the RRC estimator is 2/3 of the bias of the

Table 2
Simulation study to examine the impact of magnitude of the risk coefficient α on the performance of the risk coefficient estimates using the four methods: the true longitudinal biomarker (TRUE), last value carried forward (LVCF), ordinary regression calibration (ORC), and risk set regression calibration (RRC)

Method	$\hat{\alpha}$	Average Relative Bias	Empirical SE	Estimated SE	MSE	95% CI coverage
$\alpha = 1$						
TRUE	1.012	0.012	0.066	0.070	0.004	0.960
LVCF	0.899	−0.101	0.056	0.066	0.013	0.680
ORC	0.976	−0.024	0.064	0.068	0.005	0.930
RRC	0.957	−0.033	0.063	0.068	0.006	0.890
$\alpha = 2$						
TRUE	2.009	0.0045	0.130	0.133	0.017	0.940
LVCF	1.498	−0.251	0.098	0.110	0.262	0.005
ORC	1.844	−0.078	0.130	0.127	0.041	0.745
RRC	1.828	−0.086	0.135	0.128	0.048	0.705
$\alpha = 3$						
TRUE	2.980	−0.0067	0.188	0.201	0.036	0.965
LVCF	1.681	−0.440	0.130	0.137	1.757	0.000
ORC	2.637	−0.121	0.200	0.198	0.172	0.480
RRC	2.763	−0.079	0.214	0.209	0.102	0.750

ORC estimator. The RRC estimator also has a much higher coverage rate, 75% versus 49% of the ORC estimator.

We also studied the effect of censoring rate and frequency of repeated measurements. As the censoring rate increases from 8% to 60%, the bias in the naive LVCF estimator becomes slightly worse, and the bias in the two regression calibration estimators improves slightly (ORC: 9% to 7%, RRC: 10% to 5%). When the censoring rate is 8%, the ORC estimator is slightly better than the RRC estimator. As the censoring rate increases, the difference becomes smaller. When censoring rate is 60%, the RRC estimator becomes slightly better than the ORC estimator. As the median number of covariate measurements per person increases from 8 to 32, as expected, the bias in all three estimators decreases (ORC: 12.5% to 5%, RRC: 12.5% to 5.5%). The two regression calibration estimators behave similarly. Although increasing the sampling frequency decreases bias, in practice one must also consider the cost of data collection of extra longitudinal measurements.

6. Application to the Prostate Cancer Study

We applied the two proposed regression calibration procedures to the MPCS, a study of cancer recurrence in prostate cancer patients after radiation therapy (Taylor, Yu, and Sandler, 2005). A major interest of this study is to estimate the dependence of the risk of cancer recurrence on the PSA level and the rate of change of the PSA level. A total of 934 patients were enrolled from July 1987 to February 2000.

Since the start of radiation therapy, posttreatment longitudinal PSA values were measured about every 6 months to the time of clinical recurrence of cancer, loss of follow-up, death from nonprostate cancer disease, or censoring at the end of study. The median follow-up from entry of the study was 44.5 months (range 0.6–144.5 months). Clinical recurrence (local recurrence, distant metastasis, or regional recurrence) was observed in 140 (15.0%) of 934 patients. All other patients

are treated as censored. During the study follow-up, a total of 59 patients received salvage hormone therapy (HT), among whom 15 (25.4%) experienced clinical failure. PSA measurements after HT in these patients were not included in the data analysis. The average number of PSA measurements is 6.6 ± 4.1 (range 1–29). Tumor stage III and above are used as the reference group against tumor stages I and II. The baseline covariates include patient age at the entry of the study (age), baseline PSA (B_PSA), tumor stage (T-stage), Gleason score (GS), and total dose of radiation therapy (dose). The PSA values measured at the baseline and follow-up were log transformed as LNPSA = ln (PSA+1), B_LNPSA = ln (B_PSA+1), to make the normality assumption more plausible.

In the first stage of the analysis we fit the SPMM (4) to the longitudinal PSA data. The fixed effect covariates vector Z_i include GS, T-stage, and age. The random effects $\mathbf{U}_i = (1, t_{ij})^T$, and the random stochastic process W_i is the mean zero IW process. The population longitudinal trajectory $\psi(t)$ was estimated using a smoothing spline. The two-stage RRC and ORC methods were used to estimate the subject-specific curves $\hat{X}_i^*(t)$ and the regression coefficients in the Cox model. The unit for the rate of change is LNPSA unit per month. A few randomly selected fitted subject-specific LNPSA curves using the ORC procedure under the SPMM (4) are shown in Figure 2. These plots demonstrate that the individual smoothing spline estimates using (4) capture individual profiles very well.

The Cox proportional hazard models were used to estimate the dependence of the risk of cancer recurrence on the longitudinal PSA profile and other covariates. We started with model (2) for the survival data, in which $X_i^*(t)$=LNPSA, $X_i^*(t) = $ LNPSA' (t), and the time-independent covariates included GS, T-stage, age, dose, and B_LNPSA. HT was included in the model as a time-dependent covariate, where HT(t) equals 1 if a subject has had the HT at time t, and 0 otherwise. To check whether the assumed linear relation between the hazard and the rate of change of LNPSA is correct, we did some exploratory analysis fitting a penalized spline to LNPSA. The results suggested (not shown here) that log-hazard rate is linear in the square root of LNPSA'(t). We hence used the square root of LNPSA'(t) instead of LNPSA'(t) in our final model. Furthermore, to obtain better standard error estimates for all the parameters, we used the bootstrap method with 800 bootstrap samples for the ORC analysis. We did not do bootstrap for the RRC analysis because the computational time would be excessive.

The results are shown in Table 3. The regression coefficient estimates using the two regression calibration methods ORC and RRC are similar. The LNPSA level and the square root of the rate of change of LNPSA level are significant predictors of cancer recurrence. If a patient's PSA level doubles, the risk of cancer recurrence increases by 63% and 48% based on the ORC and RRC estimates, respectively. If the square root of a patient's rate of change of LNPSA is increased by 0.1, the chance that he experiences cancer recurrence is increased by 33% and 70% according to the ORC and RRC model, respectively. If the baseline PSA level doubles, the risk is increased by 22% (ORC) or 32% (RRC). Salvage HT has been shown to be able to effectively decrease the risk of cancer recurrence, at least in the short term. Even with the current PSA level and the rate of change being adjusted, the baseline PSA level

Table 3
Regression coefficient estimates in the Cox proportional hazard model for the prostate cancer data

Parameters	ORC				RRC		
	Estimate	Standard error	Hazard ratio	Bootstrap SE	Estimate	Standard error	Hazard ratio
LNPSA(t)	0.705	0.090	2.024	0.113	0.568	0.096	1.766
$\sqrt{\text{LNPSA}}'(t)$	2.835	0.787	17.03	0.913	5.335	0.920	207.5
I(T stage = 1)	−1.186	0.333	0.305	0.390	−1.381	0.332	0.251
I(T stage = 2)	−0.244	0.226	0.783	0.307	−0.390	0.227	0.677
GS	0.443	0.083	1.557	0.094	0.497	0.084	1.644
Dose	0.0756	2.99	1.08	3.33	−0.280	3.025	0.755
Age(10 years)	−0.547	0.141	0.579	0.158	−0.522	0.143	0.594
BLNPSA	0.290	0.106	1.336	0.146	0.397	0.106	1.488
HT(t)	−1.597	0.424	0.203	0.512	−1.339	0.394	0.262

is still shown to be an important positive predictor. Cancer T-stage, GS, and age are all found to be significant predictors for the risk of cancer recurrence. The bootstrap estimated SE is in general slightly larger than the SE obtained from the Cox proportional hazard model, the ratios range from 138% to 111%.

7. Discussions

We have proposed in this article two regression calibration procedures for jointly modeling survival and longitudinal data, allowing longitudinal data to have nonlinear population and subject-specific trajectories. A SPMM is used to estimate both the population and nonlinear subject-specific trajectories nonparametrically using smoothing splines. Our extensive simulation results show that the ORC estimator and the RRC estimator are attenuated but generally perform well. Both methods are able to correct much of the bias present in the naive LVCF method. Under most circumstances, the two regression calibration estimators behave very similarly. However, the ORC estimators tend to have relatively less bias in the situation of small measurement error and small risk coefficient. Under the opposite scenarios the RRC estimator behaves relatively better. The ORC estimators are computationally much faster compared to the RRC estimators, especially for large datasets.

An advantage of the two-stage regression calibration approach is that they can be easily implemented using existing software for the semiparametric stochastic mixed model (SPMM SAS macro) and the proportional hazards model (PROC PHREG). As shown in our simulation studies, the relative risk estimates $\hat{\alpha}$ still have downward biases of 7–10%, as a consequence of the bias and variation in the predicted marker values $\hat{X}_i^*(t)$. To achieve better regression coefficient estimators of α, one needs to seek an alternative approach that takes account of the informative drop-out in the longitudinal model and incorporates the uncertainty of measurement error into the relative risk model. This can be achieved by jointly maximizing a penalized likelihood constructed using both the longitudinal data process and time-to-event data, in which the survival model can be thought of as the model for dropout process. Computation of such a joint penalized likelihood approach is likely to be much more computationally

intensive compared with the proposed two-stage regression calibration methods. Future research is needed.

References

Abelson, R. P. and Prentice, D. A. (1997). Contrast tests of interaction hypotheses. *Psychological Methods* **2,** 315–328.

Breslow, N. E. and Lin, X. (1995). Bias correction in generalised linear mixed models with a single component of dispersion. *Biometrika* **82,** 81–91.

Brown, E. R., Ibrahim, J. G., and DeGruttola, V. (2005). A flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics* **61,** 64–73.

Bycott, P. and Taylor, J. (1998). A comparison of smoothing techniques for CD4 data measured with error in a time-dependent Cox proportional hazards model. *Statistics in Medicine* **17,** 2061–2077.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd edition. London: Chapman & Hall/CRC Press.

Dafni, U. G. and Tsiatis, A. A. (1998). Evaluating surrogate markers of clinical outcome when measured with error. *Biometrics* **54,** 1445–1462.

Faucett, C. L. and Thomas, D. C. (1996). Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Statistics in Medicine* **15,** 1663–1685.

Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. London: Chapman & Hall Ltd.

Law, N. J., Taylor, J. M. G., and Sandler, H. (2002). The joint modeling of a longitudinal disease progression marker

and the failure time process in the presence of cure. *Biostatistics (Oxford)* **3,** 547–563.

Pauler, D. K. and Finkelstein, D. M. (2002). Predicting time to prostate cancer recurrence based on joint models for non-linear longitudinal biomarkers and event time outcomes. *Statistics in Medicine* **21,** 3897–3911.

Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model (Corr: V71 p219). *Biometrika* **69,** 331–342.

Sartor, C. I., Strawderman, M. H., Lin, X.-H., Kish, K. E., McLaughlin, P. W., and Sandler, H. M. (1997). Rate of PSA rise predicts metastatic versus local recurrence after definitive radiotherapy. *International Journal of Radiology Oncology, Biology, Physics* **38,** 941–947.

Taylor, J. M., Yu, M., and Sandler, H. M. (2005). Individualized predictions of disease progression following radiation therapy for prostate cancer. *Journal of Clinical Oncology* **23,** 816–825.

Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14,** 809–834.

Tsiatis, A. A., DeGruttola, V., and Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association* **90,** 27–37.

Wahba, G. (1978). Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society, Series B: Methodological* **40,** 364–372.

Wang, Y. and Taylor, J. M. G. (2001). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* **96,** 895–905.

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53,** 330–339.

Xie, S. X., Wang, C. Y., and Prentice, R. L. (2001). A risk set calibration method for failure time regression by using a covariate reliability sample. *Journal of the Royal Statistical Society, Series B: Statistical Methodology* **63,** 855–870.

Xu, J. and Zeger, S. L. (2001). The evaluation of multiple surrogate endpoints. *Biometrics* **57,** 81–87.

Yu, M., Law, N. J., Taylor, J. M. G., and Sandler, H. M. (2004). Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica* **14,** 835–862.

Zhang, D., Lin, X., Raz, J., and Sowers, M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association* **93,** 710–719.