

Research Article

A Semiparametric Marginalized Model for Longitudinal Data with Informative Dropout

Mengling Liu¹ and Wenbin Lu²

¹ Division of Biostatistics, Department of Environmental Medicine, New York University School of Medicine, New York, NY 10016, USA

² Department of Statistics, North Carolina State University, Raleigh, NC 27695, USA

Correspondence should be addressed to Mengling Liu, mengling.liu@nyu.edu

Received 1 April 2011; Revised 11 July 2011; Accepted 27 July 2011

Academic Editor: Yangxin Huang

Copyright © 2012 M. Liu and W. Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a marginalized joint-modeling approach for marginal inference on the association between longitudinal responses and covariates when longitudinal measurements are subject to informative dropouts. The proposed model is motivated by the idea of linking longitudinal responses and dropout times by latent variables while focusing on marginal inferences. We develop a simple inference procedure based on a series of estimating equations, and the resulting estimators are consistent and asymptotically normal with a sandwich-type covariance matrix ready to be estimated by the usual plug-in rule. The performance of our approach is evaluated through simulations and illustrated with a renal disease data application.

1. Introduction

Longitudinal studies often encounter data attrition because subjects drop out before the designated study end. Both statistical analysis and practical interpretation of longitudinal data can be complicated by dropouts. For example, in the Modification of Diet in Renal Disease (MDRD) study [1, 2], one main interest was to investigate the efficacy of interventions of blood pressure control and diet modification on patients with impaired renal functions. The primary outcome was glomerular filtration rate (GFR), which measured filtering capacity of kidneys, and was repeatedly measured over the study period. However, some patients could leave the study prematurely for kidney transplant or dialysis, which precluded further GFR measurements. This resulted in a dropout mechanism that could relate to patients' kidney function and correlate with their GFR values. Other patients were followed to the end of the study or dropped out due to independent reasons. Thus, statistical analysis of longitudinal

GFR needs to take into consideration the presence of mixed types of informative and independent dropouts.

Many statistical models and inference approaches have been proposed to accommodate the nonignorable missingness into modeling longitudinal data (see reviews [3–8]). According to the target of inference and the interpretation of model parameters, existing methods can be classified into three categories: subject-specific inference, event-conditioning inference, and marginal inference. First, a widely used modeling strategy for longitudinal data with informative dropouts is to specify their joint distribution via shared or correlated latent variables. Under such model assumptions, the longitudinal parameters have a conditional, subject-specific interpretation (e.g., [9–11]). But the interpretation of longitudinal parameters usually changes with the number and characteristics of latent variables assumed, for example, a single random intercept versus a random intercept plus a random slope.

Second, event-conditioning approaches have also been widely used when the target of inference is within subgroups of patients with particular dropout patterns or when the dropout can potentially change the material characteristic of the longitudinal process (e.g., death). The inference is usually conducted conditioning on the dropout pattern or on the occurrence of the dropout event. Thus, model parameters have an event-conditioning subpopulation-averaged interpretation, for example, pattern-mixture models for the group expectation of each dropout pattern [3, 12]; treatment effects among survivors [13]; gender and age effects in mortal cohort [14]. Because the interpretation of such models is made by conditioning on a future event, event-conditioning approaches may be natural in a retrospective setting but may not be directly useful for the evaluation of treatment efficacy prospectively.

Lastly, when the research objective is to study covariate effects at population level in a dropout-free situation, marginal models address this concern directly. When data are without missing or missing completely at random (using Rubin’s definition on missingness [15]), the estimation of model parameters can be carried out by the generalized estimating equation (GEE) approach assuming a “working” correlation matrix [16]. When dropouts are missing at random, the inverse probability-weighted GEE methods are commonly used [17, 18]. In the presence of informative dropouts, the class of selection models that were originally proposed to adjust selection bias in econometrics [19] have been widely used for the marginal analysis of longitudinal data [20–22]. Recently, the marginalized transition model [23] and marginalized pattern-mixture model [24] were proposed for binary longitudinal data with finite nonignorable nonresponse patterns. These marginalized approaches provide a powerful tool for studying the marginal association between longitudinal outcomes and covariates while incorporating nonignorable nonresponses.

In this paper, we shall adopt the idea of shared latent variables to account for the dependence between longitudinal responses and informative dropouts while focusing on marginal inference for the longitudinal responses. Here dropouts can occur on a continuous time scale. We develop an effective estimation procedure built on a series of asymptotically unbiased estimating equations with light computational burden. The resulting estimators for longitudinal parameters are shown to be consistent and asymptotically normal, with a sandwich-type variance-covariance matrix that can be estimated by the usual plug-in rule.

The remainder of the paper is organized as follows. In Section 2, we introduce the notation and the proposed semiparametric marginalized model. In Section 3, a simple estimating equation-based procedure is first proposed for the situation with pure informative dropouts and is extended to a more general situation where there is a mixture of random dropouts and informative dropouts. Asymptotic properties of resulting estimators are also studied.

Simulation studies and an application to a renal disease data set are given in Section 4. Some remarks are discussed in Section 5. All the technical details are provided in the appendix.

2. Notation and Model Specification

2.1. Data Notation

Consider that a longitudinal study follows n subjects over time period $[0, \tau]$. For the i th subject $i = 1, \dots, n$, the *complete-data* consist of $\{Y_{ij}, X_{ij}, t_{ij}, j = 1, \dots, n_i\}$, where Y_{ij} is the value of response at the j th observation time t_{ij} and X_{ij} is a $p \times 1$ vector of covariates associated with response Y_{ij} . Note that X_{ij} includes baseline covariates that are separately denoted by Z_i and potential time-dependent covariates. Let T_i denote the informative dropout time and C_i denote the random censoring time that is independent of (Y_{ij}, T_i) given the covariates. In practice, we observe (T_i^*, δ_i) , where $T_i^* = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$ taking the value of 1 if the informative dropout time is observed and 0 otherwise. Throughout the paper, let $I(\cdot)$ denote the indicator function. Due to the dropout, longitudinal responses and covariates can only be observed at $t_{ij} \leq T_i^*$. Hence, the *observed data* are $\{Y_{ij}, X_{ij}, t_{ij}, T_i^*, \delta_i, i = 1, \dots, n, j = 1, \dots, m_i\}$, where $m_i = \sum_{j=1}^{n_i} I(t_{ij} \leq T_i^*)$.

2.2. Semiparametric Marginalized Latent Variable Model

We first introduce the composition of our proposed model and then discuss the model motivation and interpretation. The first component is a marginal generalized linear model for longitudinal responses Y_{ij} 's:

$$g\{E(Y_{ij} | X_{ij})\} \triangleq g\{\mu_{ij}\} = \beta'_M X_{ij}, \quad (2.1)$$

where $g(\cdot)$ is a known link function and μ_{ij} denotes the marginal expectation. The second component is a linear transformation model for the informative dropout time T_i :

$$H(T_i) = -\theta' Z_i + \eta_i, \quad (2.2)$$

where $H(\cdot)$ is an unspecified monotone transformation function and η_i is assumed to follow a known continuous distribution $F(\cdot)$ that is independent of Z_i . The last component is a conditional mean model characterizing the dependence between longitudinal responses and informative dropouts:

$$g\{E(Y_{ij} | X_{ij}, \eta_i)\} \triangleq g\{\mu_{ij}(\eta_i)\} = \Delta_{ij} + \alpha' b_{ij}(\eta_i), \quad (2.3)$$

where the latent random effects $\mathbf{b}_i(\eta_i) = \{b'_{i1}(\eta_i), \dots, b'_{im_i}(\eta_i)\}$ are investigator-specified functions of η_i and covariates, and Δ_{ij} is an implicit parameter whose value is determined by the integral equation matching the conditional mean model (2.3) with the corresponding marginal model (2.1), that is,

$$g^{-1}(\beta'_M X_{ij}) = E(Y_{ij} | X_{ij}) = E\{E(Y_{ij} | X_{ij}, \eta_i)\} = \int g^{-1}\{\Delta_{ij} + \alpha' b_{ij}(\eta)\} dF(\eta). \quad (2.4)$$

The marginal mean model (2.1) directly specifies the marginal relationship between the responses and covariates, and β_M is the $p \times 1$ marginal regression parameters of main interest. Next, the semiparametric linear transformation model (2.2) is chosen to provide a flexible survival model for the informative dropout time while it still can be easily incorporated into model (2.3) for the dependence of the longitudinal responses and informative dropouts. Model (2.2) includes the proportional hazards model [25], the proportional odds model [26], and the Box-Cox transformation model as special cases and has been studied intensively in survival analysis literature [27–29]. In addition, as we present in Section 3, the explicit assumption on the error distribution in (2.2) can facilitate the “marginalization” procedure for parameter estimation.

The conditional mean model (2.3) is motivated by the construction of the marginalized random-effects model [30, 31]. As a motivating example, we consider a continuous Gaussian process following a simple random-effects model, $Y_{ij} = \beta'X_{ij} + b_{0i} + b_{1i}t_{ij} + \varepsilon_{ij}$, where (b_{i0}, b_{i1}) are the random intercept and slope, and error terms ε_{ij} , $j = 1, \dots, n_i$ are assumed to follow $N(0, \Sigma_i)$ but independent of (b_{i0}, b_{i1}) or η_i . Note that ε_{ij} 's can still exhibit temporal dependence in addition to what has been accounted by the random effects, that is, Σ_i with AR(1) covariance structure. Furthermore, as in joint modeling approaches via latent variables, the joint distribution of (b_{0i}, b_{1i}, η_i) is assumed to be $N\left(\mathbf{0}, \begin{pmatrix} \Sigma_{b,C} \\ C', 1 \end{pmatrix}\right)$. It is easy to see that the conditional mean has the expression as model (2.3),

$$E(Y_{ij} | X_{ij}, \eta_i) = \Delta_{ij} + C_1\eta_i + C_2\eta_it_{ij}. \quad (2.5)$$

We use model (2.3) primarily as a parsimonious model for the dependence structure between the longitudinal responses and informative dropout times. However, note that although model (2.3) takes a similar form as the marginalized random-effects model, it does not intend to fully specify the joint distribution of the repeated measurements since model (2.3) only specifies the conditional mean function and there is no conditional independence assumed.

Note that Δ_{ij} is the solution of (2.4), and thus its value implicitly depends on β_M , α , the formulation of $b_{ij}(\eta_i)$ and the distribution of η_i . The specification of $b_{ij}(\eta_i)$ reflects investigator's assumptions on the dependence structure among the longitudinal responses and their association with the dropout times. It is well known that the dependence assumptions between longitudinal measurements and informative dropouts are usually unverifiable from the observed data, but it intrinsically affects the inference about β_M . Thus, a sensitivity analysis under various assumptions is always warranted. It is clear that the sensitivity analysis can be easily conducted within the framework of model (2.3). For example, the analysis may start with a large model in the specification for $\alpha'b_{ij}(\eta_i)$, for example, $\alpha_1\eta_i + \alpha_2\eta_it_{ij} + \alpha_3\eta_iX_{ij}$, and then examine the statistical significance of estimates for α 's to further simplify the model. As shown in the next subsection, complex structure can be imposed on $b_{ij}(\eta_i)$ without introducing much extra computation. Lastly, we note that the marginal interpretation of longitudinal parameters in model (2.1) is invariant under different specifications of the conditional mean model (2.3).

3. Estimation and Asymptotic Properties

3.1. Conditional Generalized Estimating Equation

First, assume that η_i is known. We construct a “conditional” generalized estimating equation for $\mathcal{B} = (\beta'_M, \alpha)'$. More specifically, the estimating function $U(\mathcal{B})$ is specified as

$$\sum_{i=1}^n \mathbf{A}'_i(\eta_i) \mathbf{W}_i [\mathbf{Y}_i - \boldsymbol{\mu}_i(\eta_i)] = \sum_{i=1}^n \mathbf{A}'_i(\eta_i) \mathbf{W}_i \left[\mathbf{Y}_i - g^{-1} \{ \boldsymbol{\Delta}_i + \mathbf{b}_i(\eta_i) \alpha \} \right], \quad (3.1)$$

where \mathbf{Y}_i denotes the $m_i \times 1$ vector of observed responses of subject i ; $\boldsymbol{\Delta}_i = (\Delta_{i1}, \dots, \Delta_{im_i})'$; $\mathbf{b}_i(\eta_i) = \{b_{i1}(\eta_i), \dots, b_{im_i}(\eta_i)\}'$; $\boldsymbol{\mu}_i(\eta_i) = \{\mu_{i1}(\eta_i), \dots, \mu_{im_i}(\eta_i)\}'$; $\mathbf{A}_i(\eta_i) = [\partial \boldsymbol{\mu}_i(\eta_i) / \partial \beta'_M, \partial \boldsymbol{\mu}_i(\eta_i) / \partial \alpha]'$; \mathbf{W}_i is a $m_i \times m_i$ weight matrix.

It is easy to see that $U(\mathcal{B})$ has mean zero at the true parameter values \mathcal{B}_0 under model (2.3). Note that the vector of marginal parameters β_M is implicitly present in $U(\mathcal{B})$ with $\boldsymbol{\Delta}_i$ through the constrain equation (2.4). Thus, the Jacobian matrix $\mathbf{A}_i(\eta_i)$ needs to be derived using both the constrain (2.4) and models (2.1) and (2.3), which is different from the ordinary GEE. More specifically, entries of the Jacobian matrix $\mathbf{A}_i(\eta_i)$ are given by

$$\begin{aligned} \frac{\partial \mu_{ij}(\eta_i)}{\partial \beta_M} &= X_{ij} Y_{ij} [E\{Y_{ij}(\eta_i)\}]^{-1} Y_{ij}(\eta_i), \\ \frac{\partial \mu_{ij}(\eta_i)}{\partial \alpha} &= \left\{ b_{ij}(\eta_i) - E\{b_{ij}(\eta_i) Y_{ij}(\eta_i)\} [E\{Y_{ij}(\eta_i)\}]^{-1} \right\} Y_{ij}(\eta_i), \end{aligned} \quad (3.2)$$

where $Y_{ij} = 1/\dot{g}\{g^{-1}(\beta'_M X_{ij})\}$, $Y_{ij}(\eta_i) = 1/\dot{g}\{g^{-1}(\boldsymbol{\Delta}_{ij} + \alpha' \mathbf{b}_{ij}(\eta_i))\}$, and we use $\dot{a}(x)$ to denote the derivative of a function $a(x)$ throughout this paper. In particular, we have $Y_{ij} = \mu_{ij}(1 - \mu_{ij})$ and $Y_{ij}(\eta_i) = \mu_{ij}(\eta_i)\{1 - \mu_{ij}(\eta_i)\}$ under the logit-link function for binary data; $Y_{ij} = \mu_{ij}$ and $Y_{ij}(\eta_i) = \mu_{ij}(\eta_i)$ under the log-link function for count data. Thus, under these canonical link functions, $Y_{ij} = \text{Var}(Y_{ij} | X_{ij})$ and $Y_{ij}(\eta_i) = \text{Var}(Y_{ij} | X_{ij}, \eta_i)$ are the marginal variance and conditional variance of the responses, respectively. In addition, these formulations also facilitate our selection of the weight matrix \mathbf{W}_i . For example, for binary longitudinal data with logit-link function, we can choose a weight matrix as $\mathbf{W}_i^{-1} = \text{diag}\{\text{Var}(Y_{ij} | X_{ij}, \eta_i), j = 1, \dots, m_i\}$.

It is clear that the implementation of the estimating function (3.1) requires the knowledge of η_i , which is an unknown quantity and has to be estimated first. The estimation of the semiparametric linear transformation model (2.2) has been studied by many authors [27–29]. In particular, Chen et al. [27] proposed a class of martingale-based estimating equations,

$$\begin{aligned} \sum_{i=1}^n \int_0^\infty Z_i [dN_i(t) - I(T_i^* \geq t) d\Lambda\{\theta' Z_i + H(t)\}] &= 0, \\ \sum_{i=1}^n [dN_i(t) - I(T_i^* \geq t) d\Lambda\{\theta' Z_i + H(t)\}] &= 0, \quad \forall t \geq 0, \end{aligned} \quad (3.3)$$

where $N_i(t) = I(T_i^* \leq t, \delta_i = 1)$. Then an iterative algorithm can be carried out to solve θ and H simultaneously. We estimate (θ, H) using the approach of Chen et al. [27] and shall denote the estimates as $\hat{\Theta} \triangleq (\hat{\theta}, \hat{H})$.

3.2. Estimation Procedure for Pure Informative Dropouts

We first consider the situation of pure informative dropouts, that is, $\delta_i \equiv 1$. Define $\hat{\eta}_i = \hat{H}(T_i) + \hat{\theta}'Z_i$ and replace η_i 's in (3.1) with their estimated counterparts $\hat{\eta}_i$'s. Denote the resulting estimating function by $U(\mathcal{B}; \hat{\Theta})$ and define the estimator of \mathcal{B} as the solution to $U(\mathcal{B}; \hat{\Theta}) = 0$. The estimation of \mathcal{B} entails an iteration between solving nonlinear equations for Δ_{ij} and updating a Newton-Raphson equation for \mathcal{B} . More specifically, given the current estimated value of $\mathcal{B}^{(j)}$ at the j th step, we first estimate $\Delta_{ij}^{(j)}$ from

$$g^{-1}\{\beta_M^{(j)'} X_{ij}\} = \int g^{-1}\{\Delta_{ij} + \alpha^{(j)'} b_{ij}(\eta)\} dF(\eta), \quad (3.4)$$

and then update the parameters \mathcal{B} by

$$\mathcal{B}^{(j+1)} = \mathcal{B}^{(j)} + \left\{ \sum_{i=1}^n \mathbf{A}_i^{(j)'}(\hat{\eta}_i) \mathbf{W}_i^{(j)} \mathbf{A}_i^{(j)}(\hat{\eta}_i) \right\}^{-1} \left\{ \sum_{i=1}^n \mathbf{A}_i^{(j)'}(\hat{\eta}_i) \mathbf{W}_i^{(j)} [\mathbf{Y}_i - \boldsymbol{\mu}_i^{(j)}(\hat{\eta}_i)] \right\}, \quad (3.5)$$

where $\mathbf{A}_i^{(j)}(\hat{\eta}_i)$, $\mathbf{W}_i^{(j)}$, and $\boldsymbol{\mu}_i^{(j)}(\hat{\eta}_i)$ are evaluated at the current parameter values $\mathcal{B}^{(j)}$ and $\Delta_{ij}^{(j)}$. The algorithm is iterated until it converges. Because η_i is assumed to follow an explicit parametric distribution F , it greatly simplifies the marginalization procedure (3.4). We propose to use the Gaussian-quadrature approach [32] to numerically evaluate (3.4) and $\mathbf{A}_i^{(j)}(\hat{\eta}_i)$. Since the integrand of (3.4) is monotonic in $\Delta_{ij}^{(j)}$ and so is the whole integral, it is easy to calculate a large number of $\Delta_{ij}^{(j)}$, $i = 1, \dots, n$, $j = 1, \dots, m_i$, in all iterative steps. Moreover, the numerical integration is only upon the one-dimensional space of η_i and requires light computation even with complex structure assumed on $b_{ij}(\eta_i)$. The proposed iterative algorithm has been implemented using "R" codes, which are available from the authors upon request.

3.3. Estimation Procedure for Mixed Types of Dropouts

We generalize the proposed estimation function (3.1) to accommodate the situation where there are mixed informative dropouts and random censoring. More specifically, the modified estimating equation is given by

$$U^*(\mathcal{B}; \Theta) = \sum_{i=1}^n \mathbf{A}_i^{*'}(\eta_i^*, \delta_i) \mathbf{W}_i^* [\mathbf{Y}_i - \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i)] = 0, \quad (3.6)$$

where $\eta_i^* = H(T_i^*) + \theta'Z_i$; the j th component of $\boldsymbol{\mu}_i^*(\eta_i^*, \delta_i)$ is

$$\mu_{ij}^*(\eta_i^*, \delta_i) = \delta_i g^{-1}\{\Delta_{ij} + \alpha' b_{ij}(\eta_i^*)\} + (1 - \delta_i) E[g^{-1}\{\Delta_{ij} + \alpha' b_{ij}(\eta)\} \mid \eta \geq \eta_i^*], \quad (3.7)$$

and the Jacobian matrix $\mathbf{A}_i^*(\eta_i^*, \delta_i) = [\partial \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i) / \partial \beta_M', \partial \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i) / \partial \alpha']$. When $\delta_i = 1$, the i th component of $U^*(\mathcal{B}; \Theta)$ is the same as the one in (3.1). For $\delta_i = 0$, the entries of $\mathbf{A}_i^*(\eta_i^*, 0)$ are given by

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}_{ij}^*(\eta_i^*, 0)}{\partial \beta_M} &= X_{ij} \mathbf{Y}_{ij} [\mathbf{E}\{\mathbf{Y}_{ij}(\eta_i)\}]^{-1} \mathbf{E}\{\mathbf{Y}_{ij}(\eta) \mid \eta \geq \eta_i^*\}, \\ \frac{\partial \boldsymbol{\mu}_{ij}^*(\eta_i^*, 0)}{\partial \alpha} &= \mathbf{E}\{b_{ij}(\eta) \mathbf{Y}_{ij}(\eta) \mid \eta \geq \eta_i^*\} - \mathbf{E}\{b_{ij}(\eta_i) \mathbf{Y}_{ij}(\eta_i)\} [\mathbf{E}\{\mathbf{Y}_{ij}(\eta_i)\}]^{-1} \mathbf{E}\{\mathbf{Y}_{ij}(\eta) \mid \eta \geq \eta_i^*\}. \end{aligned} \quad (3.8)$$

In addition, the entries of the weight matrix can be changed to $\text{Var}\{\mathbf{Y}_{ij} \mid X_{ij}, \eta_i^*, \delta_i\}$ accordingly. Conditional expectations of various functions given $\eta \geq \eta_i^*$ are computed using the Gaussian-quadrature method. Let $\hat{\eta}_i^* = \widehat{H}(T_i^*) + \hat{\theta}' Z_i$ and replace η_i^* 's in (3.6) with their estimated counterparts $\hat{\eta}_i^*$'s. Denote the resulting estimating function by $U^*(\mathcal{B}; \hat{\Theta})$. Then the estimator $\hat{\mathcal{B}}$ of \mathcal{B} can be obtained from the equation $U^*(\mathcal{B}; \hat{\Theta}) = 0$ using the same iterative algorithm described in the previous subsection.

3.4. Asymptotic Properties of $\hat{\mathcal{B}}$

In this subsection, we establish the asymptotic properties of $\hat{\mathcal{B}}$. Towards this end, we need the following assumptions.

- (C1) The covariates X_{ij} 's are bounded with probability 1.
- (C2) The true parameter values \mathcal{B}_0 and θ_0 belong to the interior of a known compact set, and the true transformation function H_0 has a continuous and positive derivative.
- (C3) Let $\Lambda(\cdot)$ denote the cumulative hazard function of η_i . Define $\lambda(t) = \dot{\Lambda}(t)$ and $\psi(t) = \dot{\lambda}(t)/\lambda(t)$. Then $\lambda(\cdot)$ is positive, $\psi(\cdot)$ is continuous, and $\lim_{t \rightarrow -\infty} \lambda(t) = 0 = \lim_{t \rightarrow -\infty} \psi(t)$.
- (C4) τ is finite and satisfies $P(T > \tau) > 0$ and $P(C = \tau) > 0$.
- (C5) The matrix $\Omega \equiv \mathbf{E}\{\mathbf{A}_1^*(\eta_1^*, \delta_1) \mathbf{W}_1 \mathbf{A}_1^*(\eta_1^*, \delta_1)\}$ is positive finite, and the number of repeated measurements $m_i \ll N$.

The regularity conditions (C1)–(C4) are also used by Chen et al. [27] to derive the consistency and asymptotic normality of the estimators $\hat{\Theta}$. Condition (C5) is needed to establish the consistency and asymptotic normality of $\hat{\mathcal{B}}$, which is given in the following theorem.

Theorem 3.1. *Under conditions (C1)–(C5), with probability 1, $|\hat{\mathcal{B}} - \mathcal{B}_0| \rightarrow 0$. In addition, one has, as $n \rightarrow \infty$,*

$$\sqrt{n}(\hat{\mathcal{B}} - \mathcal{B}_0) \rightarrow_D N(0, \Omega^{-1} V \Omega^{-1}). \quad (3.9)$$

The definition of V and a sketch of the proof for Theorem 3.1 are given in the appendix. The asymptotic variance-covariance matrix can be consistently estimated by its empirical counterpart $\hat{\Omega}^{-1} \hat{V} \hat{\Omega}^{-1}$, which can be easily obtained using the usual plug-in rule.

4. Numerical Studies

4.1. Simulations

We conducted a series of simulation studies to evaluate the finite-sample performance of our proposed approach. Consider a binary longitudinal process with the marginal probability of success as $g^{-1}(\beta_0 + \beta_1 t + \beta_2 Z)$, where g was the logit-link function; observations occurred at $\mathbf{t}_i = \{t_{ij} = j, j = 0, 1, \dots, 5\}$; Z was generated from a Bernoulli distribution with the success probability of 0.5, and $(\beta_0, \beta_1, \beta_2) = (-1.5, 0.3, 1)$. The informative dropout time T_i was generated from a linear transformation model $H(T_i) = -\theta Z_i + \eta_i$, where $\theta = -0.5$. We considered three distributions for η_i : the standard normal distribution (N.), the extreme value distribution (E.), and the logistic distribution (L.), corresponding to the normal transformation model, the Cox proportional hazard model, and the proportional odds model, respectively, for the informative dropout time. We then generated the binary response Y_{ij} independently from a Bernoulli distribution with the success probability of $g^{-1}\{\Delta_{ij} + \alpha b_{ij}(\eta_i)\}$, where Δ_{ij} was calculated to match the marginal mean value as in (2.4) and α indicated the level of dependence. We considered several combinations to specify the dependence between longitudinal outcomes and informative dropout times. More specifically, when $\alpha = 0$, there was no informative dropouts; when $\alpha = 0.5$ (or 0.25) and $b_{ij}(\eta_i) = \eta_i$, the dependence existed and was linear in the latent variable η_i ; when $\alpha = 0.25$ (or 0.5) and $b_{ij}(\eta_i) = \eta_i t_{ij}$, the dependence was present through an interaction between the latent variable and the observation time.

For each scenario, we considered samples of size 100 and 200 and conducted 500 runs of simulations. The Gaussian-quadrature approximation was calculated using 50 grid points. We first considered the situation of pure informative dropouts and generated the dropout time T_i from the transformation model with $H_0(t) = 2\{\arctan(t) + \pi/2\}$. Under the assumptions of η_i following the normal, the extreme value, and the logistic distributions, the average numbers of repeated measurements were 3.91, 3.37, and 3.94, respectively. The estimation results on β_1 and α are summarized in Table 1. The proposed estimators are unbiased under all simulated scenarios, and the Wald-type 95% confidence intervals all have reasonable empirical coverage probabilities. The performance of the proposed method is consistent with different distributional assumptions of η_i and different specifications of the dependence structure, and the results improve as the sample size increases.

Next, we consider the situation where there are mixed informative dropouts and random censoring. For simplicity, let C_i be an administrative censoring at the end of the study, that is, $\tau = 6$. The informative dropout time T_i was generated from the transformation model with $H(t) = \log(t) - 1$ and η_i followed the standard normal distribution. This yielded the proportion of informative dropouts of 69.6% and the average number of repeated measurements about 4. Other settings were kept the same as in the previous simulations. The simulation results are presented in Table 2. Again, the proposed approach gives unbiased parameter estimates and reasonable coverage probabilities under all the scenarios. For comparison, we also implemented the ordinary GEE method [16]. When the informative dropout is absent, that is, $\alpha = 0$, the GEE method yields consistent parameter estimates of β_1 as expected. But when there is informative dropout ($\alpha \neq 0$), the performance of the GEE method deteriorates quickly as the magnitude of the dependence between the longitudinal data and informative dropout increases.

Last, we conducted sensitivity analysis for the proposed approach and our simulations consisted of two parts. First, as discussed in Section 2, to better characterize the dependence

Table 1: Simulation results for pure informative dropouts.

F	α	$b_{ij}(\eta_i)$	N	β_1			α				
				Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
N.	0	η_i	100	0.001	0.084	0.082	0.948	0.004	0.130	0.127	0.954
			200	0.006	0.057	0.058	0.952	-0.002	0.090	0.091	0.942
	0.5	η_i	100	0.008	0.083	0.082	0.952	0.008	0.148	0.142	0.936
			200	0.005	0.056	0.058	0.954	0.001	0.102	0.101	0.944
	0.25	$\eta_i t_{ij}$	100	0.004	0.102	0.101	0.940	0.006	0.083	0.081	0.938
			200	0.009	0.071	0.071	0.948	-0.001	0.060	0.058	0.940
E.	0	η_i	100	0.007	0.107	0.097	0.920	0.001	0.150	0.140	0.942
			200	0.002	0.069	0.069	0.950	-0.003	0.010	0.097	0.958
	0.5	η_i	100	0.010	0.097	0.097	0.948	0.025	0.161	0.161	0.962
			200	0.009	0.071	0.069	0.932	0.002	0.111	0.113	0.950
	0.25	$\eta_i t_{ij}$	100	0.026	0.138	0.135	0.936	0.001	0.124	0.114	0.926
			200	0.009	0.089	0.095	0.954	0.001	0.079	0.082	0.952
L.	0	η_i	100	0.004	0.079	0.077	0.932	0.001	0.076	0.077	0.950
			200	0.002	0.057	0.055	0.952	-0.001	0.058	0.054	0.938
	0.5	η_i	100	0.009	0.079	0.076	0.944	0.007	0.114	0.105	0.928
			200	0.002	0.053	0.054	0.964	0.006	0.074	0.075	0.956
	0.25	$\eta_i t_{ij}$	100	0.003	0.096	0.096	0.954	0.008	0.069	0.067	0.942
			200	0.002	0.067	0.068	0.964	0.004	0.047	0.047	0.958

In Tables 1–3: F : error distribution of the semiparametric transformation model; N : sample size; SSE: sample standard deviations of estimates; SEE: mean of estimates standard errors; CP: 95% coverage probability of Wald-type confidence interval.

Table 2: Simulation results for mixed types of dropouts.

α	$b_{ij}(\eta_i)$	N	Proposed β_1			Proposed α			GEE β_1		
			Bias	SSE	CP	Bias	SSE	CP	Bias	SSE	CP
0	η_i	100	0.003	0.075	0.958	-0.007	0.158	0.956	0.002	0.068	0.946
		200	0.005	0.056	0.954	-0.003	0.114	0.944	0.005	0.049	0.956
0.25	η_i	100	0.006	0.073	0.948	0.009	0.156	0.964	0.061	0.068	0.878
		200	0.004	0.055	0.950	-0.003	0.116	0.936	0.057	0.050	0.768
0.50	η_i	100	0.008	0.076	0.948	0.010	0.174	0.952	0.116	0.070	0.660
		200	0.006	0.051	0.960	-0.008	0.121	0.956	0.113	0.049	0.380
0.25	$\eta_i t_{ij}$	100	0.005	0.106	0.948	0.005	0.106	0.960	0.233	0.074	0.110
		200	0.009	0.078	0.936	-0.007	0.077	0.940	0.233	0.054	0.000
0.50	$\eta_i t_{ij}$	100	0.007	0.097	0.954	0.010	0.141	0.950	0.397	0.081	0.000
		200	0.008	0.071	0.940	-0.008	0.104	0.936	0.395	0.059	0.000

structure between longitudinal responses and informative dropouts, we would suggest to start with a large model in the specification for $\alpha' b_{ij}(\eta_i)$ and then examine the statistical significance of estimates for α 's to further simplify the model. We simulated data from a simple model with either $\alpha' b_{ij}(\eta_i) = \alpha_1 \eta_i$ or $\alpha' b_{ij}(\eta_i) = \alpha_2 \eta_i t_{ij}$, and then applied the proposed approach by assuming a bigger model (2.3) as $\Delta_{ij} + \alpha_1 \eta_i + \alpha_2 \eta_i t_{ij}$. The simulation results are summarized in the top panel in Table 3. The proposed method can reasonably well estimate all the parameters, and in particular, could correctly indicate the unnecessary zero term. Second, we simulated data from $\Delta_{ij} + \alpha_1 \eta_i + \alpha_2 \eta_i t_{ij}$ but fitted misspecified models that omitted

Table 3: Sensitivity analysis for misspecified models under mixed types of dropouts.

True	Fitted	N	Proposed β_1		Proposed α_1		Proposed α_2	
			Bias	CP	Bias	CP	Bias	CP
0.25 η_i	$\alpha_1\eta_i + \alpha_2\eta_it_{ij}$	100	0.009	0.946	0.014	0.956	0.002	0.950
		200	0.009	0.964	-0.002	0.948	0.002	0.968
0.25 η_it_{ij}	$\alpha_1\eta_i + \alpha_2\eta_it_{ij}$	100	0.005	0.946	-0.004	0.954	0.007	0.956
		200	0.009	0.936	0.001	0.956	-0.007	0.952
0.5 $\eta_i + 0.1\eta_it_{ij}$	$\alpha_1\eta_i$	100	0.071	0.860	0.119	0.918		
		200	0.069	0.788	0.103	0.896		
0.2 $\eta_i + 0.25\eta_it_{ij}$	$\alpha_2\eta_it_{ij}$	100	-0.017	0.924			0.086	0.850
		200	-0.017	0.902			0.078	0.692

some terms. The results are summarized in the lower panel of Table 3. It is evident that misspecified models lead to biased estimates for both longitudinal regression coefficients and dependence parameters.

4.2. Application to Renal Disease Data from MDRD Study

Here we considered a subgroup of 129 patients with low-protein diet in MDRD study B, among whom, 62 patients were randomized to the group of normal-blood-pressure control and 67 patients were randomized to the group of low-blood-pressure control. Besides the randomized intervention, other covariates included time in study (`time`), baseline disease progression status (`Prog`), baseline blood pressure (`Bp`), and log-transformed baseline urine protein level (`log.Pt`). There were 52 (40.3%) patients left the study prematurely for kidney transplant or dialysis and were treated as informative dropouts.

We applied the proposed approach to estimate the marginal effects of covariates on GFR values. To account for the possible informative dropouts, we assumed that the dependence term $a'b_{ij}(\eta_i)$ had a form of $\alpha_1\eta_i + \alpha_2\eta_it_{ij}$, analogous to the joint modeling approach with latent random intercept and random slope used in Schluchter et al. [33]. We considered the situations of η_i following the standard normal, the extreme value, or the standard logistic distributions. Because the outcome was a continuous variable, we used the identity link function.

Our results are presented in Table 4 and compared with the results from the ordinary GEE [16] with an independent working correlation matrix. More specifically, the slope estimates from the proposed approach indicate a much faster decreasing rate of GFR (e.g., `time Est` = -0.27, `SE` = 0.03, under the normality assumption for η_i) than the result from the ordinary GEE method (`time Est` = -0.14, `SE` = 0.03). A possible explanation is that those patients remaining under observation usually have better kidney functions and thus higher GFR values. The ordinary GEE approach that treats the observed patients as random representatives of the population tends to underestimate the degressive trend of GFR.

The estimates for the intervention on blood pressure control show positive effect of the low-blood-pressure control on the longitudinal GFR development. Although the results are not statistically significant, the estimates from the proposed method (e.g., `Intervention Est` = 0.82, `SE` = 1.07, under the normality assumption for η_i) are about twice large of the values from the ordinary GEE method (`Intervention Est` = 0.35, `SE` = 0.90). Moreover, for the proposed approach, the results under different distributional assumptions for η_i are quite similar. The estimates of the dependency parameters (α_1, α_2) are positive and statistically significant.

Table 4: Estimates of regression coefficients for the MDRD study.

Variable	Proposed						GEE	
	Normal		EV		Logistic		Est	SE
	Est	SE	Est	SE	Est	SE		
Intercept	18.54	0.96	18.57	0.91	18.58	1.11	18.57	0.78
Time	-0.27	0.03	-0.29	0.04	-0.28	0.03	-0.14	0.03
Intervention	0.82	1.06	0.74	1.01	0.71	1.17	0.35	0.90
Prog	-0.14	1.07	-0.08	1.02	-0.14	1.19	-0.14	0.91
Bp	-0.15	1.38	-0.20	1.34	-0.07	1.48	-0.36	0.49
log.Pt	-1.09	0.39	-1.09	0.37	-1.12	0.42	-0.61	0.38
η_i	1.91	0.50	1.38	0.38	1.11	0.28		
$\eta_i t_{ij}$	0.14	0.04	0.14	0.04	0.08	0.03		

Intervention: blood pressure control (1: low and 0: normal); Prog: baseline disease progression status (1: yes and 0: no); Bp: baseline blood pressure; log.Pt: baseline log-transformed urine protein level.

This indicates that higher GFR values are positively associated with longer dropout times in the study. In addition, our proposed approach shows that the baseline urine protein level is significantly associated with the longitudinal GFR development, but the ordinary GEE method does not show such significance. The results obtained using our proposed method are also consistent with those reported in Schluchter et al. [33].

5. Discussion

In this paper, we propose a semiparametric marginalized model for marginal inference of the relationship between longitudinal responses and covariates in the presence of informative dropouts. The regression parameters represent the covariate effects on the population level. The proposed estimators are expected to be insensitive to misspecification of the latent variable distribution [31], which is desirable pertaining to the sensitivity analysis on unverifiable assumptions for the informative dropouts. In practice, the choice between marginal models and other types of joint modeling approaches should be determined by study objective.

To estimate the regression parameters in the proposed marginalized model, we proposed a class of simple conditional generalized estimating equations and demonstrated its computational convenience. In general, a likelihood-based approach can be used to achieve more efficient inference and is also of great interest. For example, a marginalized random effects model [30, 31] can be used for the longitudinal process and a frailty model [34] can be used for the dropout time. Furthermore, latent variables (b_{ij}, η_i) can be modeled by a copula distribution or non-Gaussian distributions [35]. The likelihood-based methods enjoy the high efficiency and facilitate the implementation of classical model selection procedures, such as AIC or BIC; however, intensive computations are often involved when the dimension of random effects is high.

Appendix

A. Proof of Theorem 3.1

Under the conditions (C1)–(C4), Chen et al. [27] established the consistency of $\hat{\theta}$ and the uniform consistency of a transformed $\widehat{H}(\cdot)$, that is, $\sup_{t \in [0, \tau]} |\exp\{\widehat{H}(t)\} - \exp\{H_0(t)\}| = o_p(1)$.

We first derive the consistency of the proposed estimator $\widehat{\mathcal{B}}$, which is the solution of the equation $U^*(\mathcal{B}; \widehat{\Theta}) = 0$. Note that $-(1/n)(\partial U^*(\mathcal{B}; \widehat{\Theta})/\partial \mathcal{B}) = (1/n) \sum_n \mathbf{A}_i^{*'}(\widehat{\eta}_i^*, \delta_i) \mathbf{W}_i \mathbf{A}_i^*(\widehat{\eta}_i^*, \delta_i)$ is a positive definite matrix, and by the law of large numbers and the consistency of $\widehat{\theta}$ and \widehat{H} , it converges uniformly to a deterministic positive definite matrix $\Omega(\mathcal{B})$ over a compact set of \mathcal{B} . In addition, we have that $(1/n)U^*(\mathcal{B}; \widehat{\Theta})$ converges uniformly to a deterministic function $u^*(\mathcal{B}; \Theta_0)$ satisfying $u^*(\mathcal{B}_0; \Theta_0) = 0$ and $-(\partial u^*(\mathcal{B}; \Theta_0)/\partial \mathcal{B})|_{\mathcal{B}=\mathcal{B}_0} = \Omega(\mathcal{B}_0) = \Omega$. Thus, the estimating equation $U^*(\mathcal{B}; \widehat{\Theta}) = 0$ exists a unique solution $\widehat{\mathcal{B}}$. Since \mathcal{B}_0 is the unique solution of $u^*(\mathcal{B}_0; \Theta_0) = 0$, the consistency of $\widehat{\mathcal{B}}$ easily follows.

To prove the asymptotic normality, by the Taylor expansion, we have

$$\sqrt{n}(\widehat{\mathcal{B}} - \mathcal{B}_0) = \left\{ -\frac{1}{n} \frac{\partial U^*(\mathcal{B}^*; \widehat{\Theta})}{\partial \mathcal{B}} \right\}^{-1} \frac{1}{\sqrt{n}} U^*(\mathcal{B}_0; \widehat{\Theta}), \quad (\text{A.1})$$

where \mathcal{B}^* lies between $\widehat{\mathcal{B}}$ and \mathcal{B}_0 . Since $\widehat{\mathcal{B}}$ is consistent and $-(1/n)(\partial U^*(\mathcal{B}; \widehat{\Theta})/\partial \mathcal{B})$ converges uniformly to $\Omega(\mathcal{B})$, we have that $-(1/n)(\partial U^*(\mathcal{B}^*; \widehat{\Theta})/\partial \mathcal{B})$ converges to Ω . Furthermore, the Taylor expansion of $U^*(\mathcal{B}_0; \widehat{\Theta})$ around Θ_0 gives

$$\begin{aligned} \frac{1}{\sqrt{n}} U^*(\mathcal{B}_0; \widehat{\Theta}) &= \frac{1}{\sqrt{n}} \left[\sum_{i=1}^n \mathbf{A}_i^{*'}(\eta_i^*, \delta_i) \mathbf{W}_i \{ \mathbf{Y}_i - \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i) \} \right. \\ &\quad \left. - \sum_{i=1}^n \mathbf{A}_i^{*'}(\eta_i^*, \delta_i) \mathbf{W}_i \{ \boldsymbol{\mu}_i^*(\widehat{\eta}_i^*, \delta_i) - \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i) \} \right] \\ &= \frac{1}{\sqrt{n}} \left[U^*(\mathcal{B}_0; \Theta_0) - \sum_{i=1}^n \mathbf{A}_i^{*'}(\eta_i^*, \delta_i) \mathbf{W}_i \left\{ \frac{\partial \boldsymbol{\mu}_i^*(\eta, \cdot)}{\partial \eta} \bigg|_{(\eta_i^*, \delta_i)} \right\} \right. \\ &\quad \left. \times \left\{ (\widehat{\theta} - \theta_0)' (Z_i + O_i) + \widehat{H}(T_i^*; \theta_0) - H_0(T_i^*) \right\} \right] + o_p(1), \end{aligned} \quad (\text{A.2})$$

where $O_i = \partial \widehat{H}(T_i^*; \theta) / \partial \theta|_{\theta=\theta_0}$ and its asymptotic representation can be found in the appendix of Chen et al. [27]. The asymptotic representations of $\widehat{\theta}$ and \widehat{H} have also been derived by Chen et al. [27],

$$\begin{aligned} \sqrt{n}(\widehat{\theta} - \theta) &= \Sigma_*^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\tau \{ Z_i - \mu_b(t) \} dM_i(t) + o_p(1), \\ \Lambda^* \{ \widehat{H}(t; \theta_0) \} - \Lambda^* \{ H_0(t) \} &= n^{-1} \sum_{i=1}^n \int_0^t \frac{\lambda^* \{ H_0(s) \}}{B_2(s)} dM_i(s) + o_p(n^{-1/2}), \end{aligned} \quad (\text{A.3})$$

where $N_i(t) = \delta_i I(T_i^* \leq t)$ and $\xi_i(t) = I(T_i^* \geq t)$ are the counting and at-risk processes, respectively, $M_i(t) = N_i(t) - \int_0^t \xi_i(s) d\Lambda \{ \theta_0' Z_i + H_0(s) \}$ is the mean-zero martingale process, and

the definitions of Σ_* and the functions $\Lambda^*(\cdot)$, $\lambda^*(\cdot)$, $B_2(\cdot)$, and $\mu_b(\cdot)$ are given in the appendix of Chen et al. [27].

Plugging these terms back to the expansion of $(1/\sqrt{n})U^*(\mathcal{B}_0; \hat{\Theta})$ specified in (A.2), some rearrangement yields that it is equal to

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\mathbf{A}_{i \cdot}'(\eta_i^*, \delta_i) \mathbf{W}_i \{ \mathbf{Y}_i - \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i) \} - \int_0^\tau [Q_1 \{ Z_i - \mu_b(t) \} + Q_2(t)] dM_i(t) \right) + o_p(1), \quad (\text{A.4})$$

where

$$Q_1 = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \left[\mathbf{A}_{i \cdot}'(\eta_i^*, \delta_i) \mathbf{W}_i \left\{ \frac{\partial \boldsymbol{\mu}_i^*(\eta, \cdot)}{\partial \eta} \bigg|_{(\eta_i^*, \delta_i)} \right\} \right] (Z_i + O_i)' \Sigma_*^{-1}, \quad (\text{A.5})$$

$$Q_2(t) = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \left[\mathbf{A}_{i \cdot}'(\eta_i^*, \delta_i) \mathbf{W}_i \left\{ \frac{\partial \boldsymbol{\mu}_i^*(\eta, \cdot)}{\partial \eta} \bigg|_{(\eta_i^*, \delta_i)} \right\} \right] \frac{B(t, T_i^*) \xi_i(t)}{B_2(t)}.$$

The definition of $B(t, s)$ can be found in Chen et al. [27].

Hence, $(1/\sqrt{n})U^*(\mathcal{B}_0; \hat{\Theta})$ has been written as a standardized summation of independent terms with mean zero. By the central limit theorem, it is asymptotically equivalent to a multivariate Gaussian variable with zero mean and covariance matrix V , which is the limit of

$$\frac{1}{n} \sum_{i=1}^n \left(\mathbf{A}_{i \cdot}'(\eta_i^*, \delta_i) \mathbf{W}_i \{ \mathbf{Y}_i - \boldsymbol{\mu}_i^*(\eta_i^*, \delta_i) \} - \int_0^\tau [Q_1 \{ Z_i - \mu_b(t) \} + Q_2(t)] dM_i(t) \right)^{\otimes 2}. \quad (\text{A.6})$$

From (A.1), it is easy to see that the estimator $\hat{\mathcal{B}}$ is asymptotically normal with mean zero and the variance-covariance matrix $\Omega^{-1} V \Omega^{-1}$, which can be consistently estimated by its empirical counterpart $\hat{\Omega}^{-1} \hat{V} \hat{\Omega}^{-1}$ using the usual plug-in rule.

Acknowledgments

The authors would like to thank the editor and the referees for their instructive comments. This research was supported partially by NIH grants RO1 CA-140632 and RO3 CA-153083.

References

- [1] G. J. Beck, R. L. Berg, C. H. Coggins et al., "Design and statistical issues of the modification of diet in renal disease trial. The modification of diet in renal disease study group," *Controlled Clinical Trials*, vol. 12, pp. 566–586, 1991.
- [2] S. Klahr, A. S. Levey, G. J. Beck et al., "The effects of dietary protein restriction and blood-pressure control on the progression of chronic renal disease," *The New England Journal of Medicine*, vol. 330, no. 13, pp. 877–884, 1994.
- [3] J. W. Hogan and N. M. Laird, "Mixture models for the joint distribution of repeated measures and event times," *Statistics in Medicine*, vol. 16, no. 1–3, pp. 239–257, 1997.

- [4] J. W. Hogan and N. M. Laird, "Model-based approaches to analysing incomplete longitudinal and failure time data," *Statistics in Medicine*, vol. 16, no. 1–3, pp. 259–272, 1997.
- [5] J. W. Hogan, J. Roy, and C. Korkontzelou, "Tutorial in biostatistics. Handling drop-out in longitudinal studies," *Statistics in Medicine*, vol. 23, no. 9, pp. 1455–1497, 2004.
- [6] J. G. Ibrahim and G. Molenberghs, "Missing data methods in longitudinal studies: a review," *Test*, vol. 18, no. 1, pp. 1–43, 2009.
- [7] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*, John Wiley & Sons, New York, NY, USA, 1987.
- [8] A. A. Tsiatis and M. Davidian, "Joint modeling of longitudinal and time-to-event data: an overview," *Statistica Sinica*, vol. 14, no. 3, pp. 809–834, 2004.
- [9] V. De Gruttola and X. M. Tu, "Modelling progression of CD4-lymphocyte count and its relationship to survival time," *Biometrics*, vol. 50, no. 4, pp. 1003–1014, 1994.
- [10] M. D. Schluchter, "Methods for the analysis of informatively censored longitudinal data," *Statistics in Medicine*, vol. 11, no. 14–15, pp. 1861–1870, 1992.
- [11] M. C. Wu and R. J. Carroll, "Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process," *Biometrics*, vol. 44, no. 1, pp. 175–188, 1988.
- [12] R. J. A. Little, "Pattern-mixture models for multivariate incomplete data," *Journal of the American Statistical Association*, vol. 88, pp. 125–134, 1993.
- [13] B. F. Kurland and P. J. Heagerty, "Directly parameterized regression conditioning on being alive: analysis of longitudinal data truncated by deaths," *Biostatistics*, vol. 6, no. 2, pp. 241–258, 2005.
- [14] C. Dufouil, C. Brayne, and D. Clayton, "Analysis of longitudinal studies with death and drop-out: a case study," *Statistics in Medicine*, vol. 23, no. 14, pp. 2215–2226, 2004.
- [15] D. B. Rubin, "Inference and missing data," *Biometrika*, vol. 63, no. 3, pp. 581–592, 1976.
- [16] K. Y. Liang and S. L. Zeger, "Longitudinal data analysis using generalized linear models," *Biometrika*, vol. 73, no. 1, pp. 13–22, 1986.
- [17] J. M. Robins, A. Rotnitzky, and L. P. Zhao, "Analysis of semiparametric regression-models for repeated outcomes in the presence of missing data," *Journal of the American Statistical Association*, vol. 90, pp. 106–121, 1995.
- [18] G. Y. Yi and R. J. Cook, "Marginal methods for incomplete longitudinal data arising in clusters," *Journal of the American Statistical Association*, vol. 97, no. 460, pp. 1071–1080, 2002.
- [19] J. J. Heckman, "Sample selection bias as a specification error," *Econometrica*, vol. 47, pp. 153–162, 1979.
- [20] G. M. Fitzmaurice, G. Molenberghs, and S. R. Lipsitz, "Regression-models for longitudinal binary responses with informative drop-outs," *Journal of the Royal Statistical Society Series B*, vol. 57, pp. 691–704, 1995.
- [21] M. G. Kenward, "Selection models for repeated measurements with non-random dropout: an illustration of sensitivity," *Statistics in Medicine*, vol. 17, no. 23, pp. 2723–2732, 1998.
- [22] G. Molenberghs, M. G. Kenward, and E. Lesaffre, "The analysis of longitudinal ordinal data with non-random drop-out," *Biometrika*, vol. 84, no. 1, pp. 33–44, 1997.
- [23] B. F. Kurland and P. J. Heagerty, "Marginalized transition models for longitudinal binary data with ignorable and non-ignorable drop-out," *Statistics in Medicine*, vol. 23, no. 17, pp. 2673–2695, 2004.
- [24] K. J. Wilkins and G. M. Fitzmaurice, "A marginalized pattern-mixture model for longitudinal binary data when nonresponse depends on unobserved responses," *Biostatistics*, vol. 8, no. 2, pp. 297–305, 2007.
- [25] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society Series B*, vol. 34, pp. 187–220, 1972.
- [26] S. Bennett, "Analysis of survival data by the proportional odds model," *Statistics in Medicine*, vol. 2, no. 2, pp. 273–277, 1983.
- [27] K. Chen, Z. Jin, and Z. Ying, "Semiparametric analysis of transformation models with censored data," *Biometrika*, vol. 89, no. 3, pp. 659–668, 2002.
- [28] S. C. Cheng, L. J. Wei, and Z. Ying, "Analysis of transformation models with censored data," *Biometrika*, vol. 82, no. 4, pp. 835–845, 1995.
- [29] D. Clayton and J. Cuzick, "Multivariate generalizations of the proportional hazards model," *Journal of the Royal Statistical Society Series A*, vol. 148, pp. 82–117, 1985.
- [30] P. J. Heagerty, "Marginally specified logistic-normal models for longitudinal binary data," *Biometrics*, vol. 55, no. 3, pp. 688–698, 1999.

- [31] P. J. Heagerty and S. L. Zeger, "Marginalized multilevel models and likelihood inference," *Statistical Science*, vol. 15, no. 1, pp. 1–26, 2000.
- [32] W. H. Press, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, NY, USA, 2nd edition, 1992.
- [33] M. D. Schluchter, T. Greene, and G. J. Beck, "Analysis of change in the presence of informative censoring: application to a longitudinal clinical trial of progressive renal disease," *Statistics in Medicine*, vol. 20, no. 7, pp. 989–1007, 2001.
- [34] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding, *Statistical Models Based on Counting Processes*, Springer Series in Statistics, Springer, New York, NY, USA, 1993.
- [35] R. Tsonaka, G. Verbeke, and E. Lesaffre, "A semi-parametric shared parameter model to handle non-monotone nonignorable missingness," *Biometrics*, vol. 65, no. 1, pp. 81–87, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

