Statistica Sinica **28** (2018), 1333-1349
doi:https://doi.org/10.5705/ss.202016.0246

SEMIPARAMETRIC RANDOM-EFFECTS CONDITIONAL DENSITY MODELS FOR LONGITUDINAL ANALYSIS WITH CONCOMITANT INTERVENTION

Tianqing Liu, Colin O. Wu, Zhaohai Li and Yuanzhang Li

Jilin University, National Heart Lung and Blood Institute, The George Washington University and Walter Reed Army Institute of Research

Abstract: Longitudinal data in biomedical studies often involve concomitant interventions in addition to the pre-specified repeatedly measured outcome and covariate variables. Since a concomitant intervention is often initiated when a patient exhibits an undesirable health trend, adequate statistical methods should properly incorporate the starting time of a concomitant intervention in order to reduce the potential bias of the estimated intervention effects. We propose in this paper a class of semiparametric random-effects conditional density models for evaluating the distributions and concomitant intervention effects with longitudinal observations. These models simultaneously incorporate concomitant intervention effects and intra-subject longitudinal dependence structures, and quantify the change of the distribution functions through the ratio of two conditional density functions. The conditional density ratio is assumed to have a parametric form, while the baseline density function is nonparametric. We develop a likelihood-based method for estimating the parameters and a goodness-of-fit test for testing the validity of the models. Finite sample properties of our estimation and testing procedures are illustrated through a simulation study and an application to a longitudinal clinical trial in depression and heart disease.

Key words and phrases: Concomitant intervention, conditional density ratio, conditional likelihood, longitudinal data, random-effects conditional density model.

1. Introduction

In longitudinal clinical trials and epidemiological studies, patients or study participants are repeatedly observed over time, and concomitant interventions are often given to patients, due to ethical reasons, who exhibit undesirable trends of health status during the study period. A main objective in such longitudinal studies is to evaluate the temporal trends of some health outcomes and the effects of certain covariates of interest, such as the study subjects' baseline characteristics and some pre-specified treatments, on the distributions of these

health outcomes. Well-known regression methods for longitudinal analysis, including various types of parametric and nonparametric models proposed in the literature, for example,Verbeke and Molenberghs (2000), Diggle et al. (2002) and Fitzmaurice et al. (2009), have focused on the estimation and inferences of the effects of covariates that do not depend on the outcome variables. These regression methods, however, may lead to misspecified models and biased estimates of the concomitant intervention effects because the initiation of concomitant interventions could depend on the study subject's temporal trends of the response variable.

The Enhancing Recovery in Coronary Heart Disease (ENRICHD) study is a typical example of longitudinal clinical trials, which involves a concomitant intervention in addition to the randomly assigned treatment regimens and other covariates whose values do not depend on the outcome variables. The primary objective of this randomized clinical trial is to evaluate the efficacy of a 6-month cognitive behavior therapy (CBT) versus the usual cardiovascular care (UC) on overall mortality, cardiovascular events, and depression severity in patients with depression or low perceived social support after acute myocardial infarction; here depression severity was measured by the Hamilton Rating Scale for Depression (HRSD) and the Beck Depression Inventory (BDI) with higher HRSD and BDI scores indicating worsened depression. In addition to the randomized CBT treatment, patients who had high baseline depression scores or nondecreasing BDI trends were eligible for pharmacotherapy with antidepressants as a concomitant intervention, and antidepressants were also prescribed at the requests of the patients or their primary-care physicians. Major findings of the trial and justifications of its design can be found in ENRICHD (2001, 2003). Taylor et al. (2005) investigated the effects of pharmacotherapy on cardiovascular morbidity and mortality among 1834 depressed ENRICHD patients, and found that pharmacotherapy improved survival for this patient population.

To evaluate the effects of pharmacotherapy on depression severity measured by the BDI scores, Wu, Tian and Bang (2008) showed that the mixed-effects models, without taking the antidepressant starting time into account, led to biased estimates of pharmacotherapy effects, and they proposed a varying-coefficient model using the data from patients who started pharmacotherapy during the 6-month treatment period to show the beneficial effects of pharmacotherapy for lowering the patients' BDI scores. The varying-coefficient model of Wu, Tian and Bang (2008) does not use the data from patients who have already received pharmacotherapy at baseline or have not received pharmacotherapy during the

study. As an extension based on the framework of shared-parameter models (e.g., Follmann and Wu (1995)), Wu, Tian and Jiang (2011) proposed a change-point shared-parameter model for evaluating the concomitant intervention effects which was capable of incorporating the information from all the patients in the study. Their method is limited to modeling the conditional means of the outcome variables before and after the concomitant intervention through some known parametric distribution functions and estimating the parameters through a computationally intensive maximum likelihood procedure. Xing and Ying (2012) studied a semiparametric change-point regression model based on a counting process formulation; their regression model assumes environmental change-points with unknown number and locations, and differs from the setup of subject-specific concomitant interventions.

We develop a class of semiparametric random-effects conditional density (RECD) models for evaluating the conditional distributions of the outcome variable and the concomitant intervention effects in a longitudinal study. By quantifying the distribution functions of the outcome variable before and after the concomitant intervention through some random-effects, our models assume that the ratio of the conditional density functions of the subject's time-dependent outcome variable has a known form specified by some unknown parameters, while the underlying baseline density remains nonparametric. This modeling framework has the attractive feature that it simultaneously incorporates the intra-subject longitudinal dependence structure and the concomitant intervention effects. Our RECD models do not require the conditional distributions to be completely specified by a parametric family, they can be applied to studies with patients who may or may not receive concomitant interventions during the study. We develop a conditional likelihood-based estimation method for parameter estimation and inferences, and an information matrix-based goodness-of-fit test statistic for testing the validity of the models. Our application to the ENRICHD pharmacotherapy data and simulation results suggest that the proposed method leads to adequate parameter estimates with longitudinal data when a concomitant intervention is present.

2. Random-Effects Conditional Distribution Models

We introduce a general framework for longitudinal data with a concomitant intervention and a random-effects modeling approach which can simultaneously account for concomitant intervention effects and longitudinal dependence among

the observations within the same subject. Further discussions are given in Section S1 of Supplementary Materials.

2.1. Parametric random-effects conditional distribution models

Let N be the number of randomly selected subjects. The ith subject has $(n_i + 1)$ visits and the observation (T_{ij}, Y_{ij}, X_i) at the jth visit, $0 \le j \le n_i$, where T_{ij} is the study time defined as the time elapsed from the beginning of the study to the jth visit, X_i is a time-invariant covariate vector, and Y_{ij} is the real-valued or discrete outcome variable. For simplicity, we assume that the study involves only one concomitant intervention, and we denote by $S_i \in [0, +\infty)$ the ith subject's intervention starting time or change-point time and $\delta_{ij} = I(T_{ij} \ge S_i)$ the intervention indicator at the jth visit. In most clinical trials, the initial visit time is set to zero, $T_{i0} = 0$. In general, the choice of time depends on the study objectives, and the starting time may not be zero. We take $\widetilde{T}_{ij} = T_{ij} - T_{i0}$ to be the time elapsed from the baseline to the jth visit, and $R_{ij} = T_{ij} - S_i$ to be the time elapsed from the intervention starting time to the jth visit. A positive (or negative) value of R_{ij} suggests that the jth visit of the ith subject is after (or before) the intervention starting time S_i.

We assume throughout that no subject has taken the intervention before or at the start of the study, so that $S_i > T_{i0}$. To evaluate the designated treatment effects of a clinical trial, it is common to exclude subjects who have already taken an alternative intervention at baseline. Since not every subject changes from without intervention to intervention during the study, the ith subject's change-point time is observed if $T_{i0} < S_i \le T_{in_i}$. If $S_i > T_{in_i}$, the subject's change-point time is "right censored". The indicator variable for censoring κ_i is $\kappa_i = 0$ if $T_{i0} < S_i \le T_{in_i}$ and 1 if $S_i > T_{in_i}$. The observed change-point times are $\{\bar{S}_i^{(c)} = (S_i^{(c)}, \kappa_i) : i = 1, \cdots, N\}$, where $S_i^{(c)} = S_i$ if $\kappa_i = 0$ and $S_i^{(c)} = T_{in_i}$ if $\kappa_i = 1$. The observed data are $\{(T_{ij}, X_i, Y_{ij}, \bar{S}_i^{(c)}), 0 \le j \le n_i, 1 \le i \le N\}$. The set of visit times is $T_i = (T_{i0}, T_{i1}, \cdots, T_{in_i})^\mathsf{T}$, and the corresponding outcome values are $Y_i = (Y_{i0}, \cdots, Y_{in_i})^\mathsf{T}$.

Since the distribution of S_i may depend on $\{X_i, T_i\}$ as well as some unobserved variables, we assume that there is a latent random vector Δ_i, which depends on $\{X_i, T_i\}$ through the conditional density $f_{\Delta_i}(\cdot|\{X_i, T_i\})$, so that the conditional density of S_i given $\{\Delta_i, X_i, T_i\}$ is $f_{S_i}(\cdot|\{\Delta_i, X_i, T_i\})$. The conditional distribution of the outcome variable Y_i depends on the random variables $\{\Delta_i, X_i, T_i, S_i\}$ and can be constructed as follows. Let $f_{ij}(y|\Delta_i, X_i, T_i, S_i)$ be the conditional density of Y_{ij} given $\{\Delta_i, X_i, T_i, S_i\}$, which specifies a random-effects

model for Y_{ij}. For the sake of generality, $f_{ij}(\cdot|\cdot)$ may generally refer to the density function with respect to Lebesgue measure when Y_{ij} is a continuous random variable on the real line, or a probability function when Y_{ij} is a discrete or categorical random variable. Similar to the settings in the mixed-effects models (e.g., Verbeke and Molenberghs (2000)), we assume that, within each $i = 1, \ldots, N$, the outcome observations $\{Y_{i0},\ Y_{i1}, \cdots, Y_{in_i}\}$ are independent given $\{\Delta_i, X_i, T_i, S_i\}$.

Let β be the unknown parameter vector of interest that is used to characterize the time trends of Y_i and the concomitant intervention effects, and let ϕ, ψ, and φ be the unknown nuisance parameters, so that the conditional density functions are specified by $f_{S_i}(\cdot|\cdot;\beta,\psi)$, $f_{\Delta_i}(\cdot|\cdot;\varphi)$ and

$$f_{Y_i}(Y_i|\Delta_i, O_i, S_i;\beta,\phi) = \prod_{j=0}^{n_i} f_{ij}(Y_{ij}|\Delta_i, O_i, S_i;\beta,\phi),$$

where $O_i = \{X_i, T_i\}$. The joint likelihood of $\{Y_i, S_i, \Delta_i\}$ given O_i is

$$f_{Y_i}(Y_i|\Delta_i, O_i, S_i;\beta,\phi) f_{S_i}(S_i|\Delta_i, O_i;\beta,\psi) f_{\Delta_i}(\Delta_i|O_i;\varphi). \tag{2.1}$$

Since (2.1) belongs to a parametric family and the change-point time S_i may not be observed, the log-likelihood function $\ell_F(\theta)$ for the possibly censored observations $\{Y_i, \bar{S}_i^{(c)}\}$ conditioning on O_i, $i = 1, \cdots, N$, can be constructed based on (2.1) with unknown parameters $\theta = (\beta^\mathsf{T}, \phi^\mathsf{T}, \psi^\mathsf{T}, \varphi^\mathsf{T})^\mathsf{T}$. We can, in principle, estimate θ by maximizing the log-likelihood function $\ell_F(\theta)$. However, in applications, it may be difficult to correctly specify (2.1), so that more flexible statistical models and computationally feasible procedures are needed in practice.

2.2. Semiparametric random-effects conditional density models

Using the data structure of Section 2.1 and the assumption that, within the ith subject, $Y_{i0}, Y_{i1}, \cdots, Y_{in_i}$ are independent given $\{\Delta_i, X_i, T_i, S_i\}$, our semiparametric random-effects conditional density (RECD) models for $f_{Y_i}(\cdot|\cdot)$ do not require a fully parametric family and incorporate a simple structure to characterize the concomitant intervention effects. The conditional likelihood of $\{Y_i, S_i\}$ given O_i can then be constructed as in Section 2.1 by substituting $f_{Y_i}(\cdot|\cdot)$ of (2.1) with the conditional density given in the corresponding RECD model.

We consider here the case of continuous Y_{ij}'s. The discrete case can be found in Section S1 of Supplementary Materials. Let $h_i(y|\Delta_i, X_i, T_{i0}, S_i)$ be an unknown density function depending on $\{\Delta_i, X_i, T_{i0}, S_i\}$. Our RECD model has the form

$$\begin{cases} f_{i0}(y|\Delta_i, O_i, S_i) = h_i(y|\Delta_i, X_i, T_{i0}, S_i), \\ f_{ij}(y|\Delta_i, O_i, S_i) = g_i(y|Z_{ij}), \\ g_i(y|Z_{ij}) = \exp\{\alpha_{ij} + \gamma(y) Z_{ij}^{\mathsf{T}}\beta\} h_i(y|\Delta_i, X_i, T_{i0}, S_i), \end{cases} \tag{2.2}$$

where $\beta = (\beta_1, \cdots, \beta_p)^{\mathsf{T}}$ is the unknown parameter vector of interest, $\gamma(y)$ is a known function of y, $Z_{ij} = (Z_{ij1}, \cdots, Z_{ijp})^{\mathsf{T}} = \zeta(\kappa_i, X_i, \widetilde{T}_{ij}, \delta_{ij}, \delta_{ij}R_{ij})$ is a pre-specified p-dimensional function of $\{\kappa_i, X_i, \widetilde{T}_{ij}, \delta_{ij}, \delta_{ij}R_{ij}\}$, and the α_{ij}'s are normalizing constants such that $\int f_{ij}(y|\Delta_i, X_i, T_i, S_i)dy = 1$. With $\widetilde{T}_{ij} = T_{ij}$ and $R_{ij} = T_{ij} - S_i$, the effects of trial time and antidepressant use are determined by $Z_{ij}^{\mathsf{T}}\beta$, which can be specified by

$$Z_{ij}^{\mathsf{T}}\beta = \kappa_i\beta_{01} + T_{ij}\beta_1 + (1-\kappa_i)(\beta_{00} + \delta_{ij}\beta_2 + \delta_{ij}R_{ij}\beta_3), \tag{2.3}$$

where $\beta = (\beta_{00}, \beta_{01}, \beta_1, \beta_2, \beta_3)^{\mathsf{T}}$, $\kappa_i = 0$ if $0 \le S_i \le T_{in_i}$ and $\kappa_i = 1$ if $S_i > T_{in_i}$. The function (2.3) assumes that patients with different values of κ_i have possibly different intercepts β_{01} and β_{00}, and the antidepressant effects are described by the coefficients β_2 and β_3. When $\beta_{00} = \beta_{01} = \beta_0$ for some constant β_0, it follows from (2.3) that $Z_{ij}^{\mathsf{T}}\beta = \beta_0 + T_{ij}\beta_1 + (1-\kappa_i)(\delta_{ij}\beta_2 + \delta_{ij}R_{ij}\beta_3)$. Different forms of $\gamma(\cdot)$ correspond to various conventionally used densities in the literature. Commonly used forms of $\gamma(\cdot)$ include $\gamma(y) = y$ and $\gamma(y) = \log(y)$ (Anderson (1979); Kay and Little (1987); Qin et al. (2002)).

The log-density ratio of (2.2) for any Z_{ij}, y_1, and y_2 is

$$\log\frac{g_i(y_2|Z_{ij})}{g_i(y_1|Z_{ij})} = \log\frac{h_i(y_2|\Delta_i, X_i, T_{i0}, S_i)}{h_i(y_1|\Delta_i, X_i, T_{i0}, S_i)} + Z_{ij}^{\mathsf{T}}\beta\{\gamma(y_2) - \gamma(y_1)\}, \tag{2.4}$$

provided that $h_i(y_1|\Delta_i, X_i, T_{i0}, S_i)$ and $h_i(y_2|\Delta_i, X_i, T_{i0}, S_i)$ are both positive. It follows from (2.4) that the likelihood ratio between the points y_1 and y_2 is modulated by the covariate Z_{ij} through its linear combination $Z_{ij}^{\mathsf{T}}\beta$. In particular, if $\gamma(y) = y$, Z_{ij} is one-dimensional and the response Y_{ij} is binary, then

$$\beta = \log\frac{g_i(1|Z_{ij}+1)/g_i(0|Z_{ij}+1)}{g_i(1|Z_{ij})/g_i(0|Z_{ij})},$$

which is exactly the log odds ratio. This suggests that the β of (2.4) can be viewed as a generalized log-odds ratio for a given $\gamma(y)$ and an arbitrary response Y_{ij}. For vector-valued Z_{ij}, it follows from (2.4) that the lth component of β is the change of the log-density ratio associated with a unit increase of the lth component of Z_{ij}, when all other components of Z_{ij} are fixed.

3. Estimation and Inference Methods

We develop a conditional-likelihood method for the estimation of the pa-

rameters β and the construction of a goodness-of-fit statistic based on the corresponding information matrix to test the validity of the RECD model (2.2). Unlike the fully parametric likelihood approach, our estimation method treats the other parameters and the baseline density $h_i(\cdot|\cdot)$ as nuisance parameters. Existing methods and approaches for the estimation of $h_i(\cdot|\cdot)$ are discussed in Sections S2 and S3 of Supplementary Materials.

3.1. Conditional likelihood estimation method

Following Kalbfleisch (1978), Liang and Qin (2000) and Chan (2013), we consider the log-conditional likelihood function

$$\ell_C(\beta) = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i} \ell_{ijk}(\beta), \tag{3.1}$$

where $\ell_{ijk}(\beta) = -\log(1+\exp[\{\gamma(Y_{ik})-\gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\beta])$, $Z_{ijk} = Z_{ij} - Z_{ik}$, and $Z_{i0} = 0_{p\times 1}$. Here, (3.1) is an extension of the pairwise log-conditional likelihood (Liang and Qin (2000); Chan (2013)) to the RECD model (2.2). Let $\hat{\beta}$ be the maximum conditional likelihood estimator of β,

$$\hat{\beta} = \arg\max_{\beta} \ell_C(\beta). \tag{3.2}$$

Justification for (3.1) is given in Lemma 1 of Supplementary Materials.

Taking the partial derivatives of $\ell_C(\beta)$ with respect to the components of β, we have the estimating equations

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i} U_{ijk}(\beta) = 0, \tag{3.3}$$

where $U_{ijk}(\beta) = \partial\ell_{ijk}(\beta)/\partial\beta$. The global maximizer $\hat{\beta}$ of (3.2), if it exists, is a solution of (3.3). Let $\beta^* = (\beta_1^*, \cdots, \beta_p^*)^{\mathsf{T}}$ be the true value of β. It then follows from Lemma 1 in Section S7 of Supplementary Materials that

$$E\left\{\frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i} U_{ijk}(\beta^*)\right\} = 0.$$

Using the asymptotic derivations as in Crowder (1986) and Liang and Zeger (1986), we have the following.

Theorem 1. *If the model* (2.2) *and the conditions* C1-C5 *in Section* S7 *of Supplementary Materials are satisfied, then* $\hat{\beta}$ *is a consistent estimator of* β *and* $\sqrt{N}(\hat{\beta}-\beta^*) \xrightarrow{d} N(0,\Sigma)$ *as* $N\to\infty$, *where* $\Sigma = D^{-1}VD^{-\mathsf{T}}$,

$$D = \lim_{N\to\infty} -N^{-1}\sum_{i=1}^{N}\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i}\frac{\partial U_{ijk}(\beta^*)}{\partial \beta^{\mathsf{T}}},$$

$$V = \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\left\{\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i} U_{ijk}(\beta^*)\right\}^{\otimes 2},$$

with $e^{\otimes 2}$ *representing the matrix* ee^{T}.

The proof of Theorem 1 is given in Section S7 of Supplementary Materials. The results of Theorem 1 can be used to make asymptotically approximate inferences for β when (2.2) holds.

3.2. Goodness-of-fit tests based on information matrix

Using the well-known information matrix equality, White (1982) proposed an information matrix test (IMT) for detecting parametric model mis-specifications under the ordinary likelihood situation. In the current context, given that (2.2) may not hold for some longitudinal datasets, we propose an IMT for testing its validity and derive the asymptotic distributions of the IMT test statistic under the null hypothesis. We define, from (3.1),

$$H_N(\beta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i}\frac{\partial \ell_{ijk}^2(\beta)}{\partial\beta\partial\beta^{\mathsf{T}}},$$

$$K_N(\beta) = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=0}^{n_i}\sum_{j=k+1}^{n_i}\left\{\frac{\partial \ell_{ijk}(\beta)}{\partial\beta}\right\}^{\otimes 2},$$

whose expressions are given in Section S4 of Supplementary Materials. If (2.2) holds then, by Lemma 2 in Section S7 of Supplementary Materials, $E\{H_N(\beta^*)\} = E\{K_N(\beta^*)\}$. We can use the fact that $E\{H_N(\beta^*) - K_N(\beta^*)\} = 0$ under (2.2) to construct an information matrix based goodness-of-fit statistic for testing its validity.

For $1 \le b \le a \le p$ and $1 \le l = a + (b-1)p - (b-1)b/2 \le p(p+1)/2$, let

$$\begin{aligned} w_l^{(ijk)}(\beta) &= \left\{\frac{\partial \ell_{ijk}(\beta)}{\partial\beta_a}\right\}\left\{\frac{\partial \ell_{ijk}(\beta)}{\partial\beta_b}\right\} + \frac{\partial \ell_{ijk}^2(\beta)}{\partial\beta_a\partial\beta_b} \\ &= \frac{\exp[2\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\beta]\{\gamma(Y_{ij}) - \gamma(Y_{ik})\}^2 Z_{ijka}Z_{ijkb}}{\{1+\exp[\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\beta]\}^2} \\ &\quad - \frac{\exp[\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\beta]\{\gamma(Y_{ij}) - \gamma(Y_{ik})\}^2 Z_{ijka}Z_{ijkb}}{\{1+\exp[\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\beta]\}^2}. \end{aligned}$$

Since the $p \times p$ symmetric matrix $K_N(\beta^*) - H_N(\beta^*)$ can be estimated by the $\sqrt{N}$-consistent estimator $K_N(\hat{\beta}) - H_N(\hat{\beta})$, the test statistic can be formed using the "indicators" $1/N \sum_{i=1}^{N} \sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} w_l^{(ijk)}(\hat{\beta})$, $l = 1, \cdots, p(p+1)/2$, which are the lower triangular elements of $K_N(\hat{\beta}) - H_N(\hat{\beta})$. Since, White (1982), it is often inappropriate to base the test on all $p(p+1)/2$ indicators, we construct the test statistics using a subset of them, the "diagonal indicators", which are the diagonal elements of the matrix $K_N(\hat{\beta}) - H_N(\hat{\beta})$.

Define, for $1 \le d \le p(p+1)/2$, $w^{(ijk)}(\beta) = (w_1^{(ijk)}(\beta), \cdots, w_d^{(ijk)}(\beta))^{\mathsf{T}}$,

$$W_N(\beta) = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} w^{(ijk)}(\beta) \quad \text{and} \quad \widetilde{W}_N = W_N(\hat{\beta}),$$

where $l = 1, \cdots, d$ have been reassigned appropriately.

Theorem 2. *Let* $G = \lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} \partial w^{(ijk)}(\beta^*)/\partial \beta^{\mathsf{T}}$. *If* (2.2) *and the conditions* C1-C8 *in Section* S7 *of Supplementary Materials are satisfied, then* $\sqrt{N}\widetilde{W}_N \xrightarrow{d} N(0, \Psi)$ *as* $N \to \infty$, *where*

$$\Psi = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} \left\{ w^{(ijk)}(\beta^*) + GD^{-1}U_{ijk}(\beta^*) \right\} \right]^{\otimes 2}.$$

The proof of Theorem 2 is given in Section S7 of Supplementary Materials. Since Ψ is unknown, it has to be estimated in practice. To do so, we first estimate G by

$$\begin{aligned} \hat{G} &= N^{-1} \sum_{i=1}^{N} \sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} \frac{\partial w^{(ijk)}(\hat{\beta})}{\partial \beta^{\mathsf{T}}} \\ &= N^{-1} \sum_{i=1}^{N} \sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} \left(\frac{\partial w_1^{(ijk)}(\hat{\beta})}{\partial \beta}, \cdots, \frac{\partial w_d^{(ijk)}(\hat{\beta})}{\partial \beta} \right)^{\mathsf{T}}, \end{aligned}$$

where $\partial w_l^{(ijk)}(\hat{\beta})/\partial \beta = A_{ab}^{(ijk)} Z_{ijk}$ and

$$\begin{aligned} A_{ab}^{(ijk)} = {} & \frac{3\exp[2\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\hat{\beta}] - \exp[\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\hat{\beta}]}{\{1 + \exp[\{\gamma(Y_{ik}) - \gamma(Y_{ij})\}Z_{ijk}^{\mathsf{T}}\hat{\beta}]\}^3} \\ & \times \{\gamma(Y_{ik}) - \gamma(Y_{ij})\}^3 Z_{ijka} Z_{ijkb}. \end{aligned}$$

Then, a consistent estimator of Ψ under the model (2.2) is

$$\widetilde{\Psi} = \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} \left\{ w^{(ijk)}(\hat{\beta}) + \hat{G}\hat{D}^{-1}U_{ijk}(\hat{\beta}) \right\} \right]^{\otimes 2}.$$

Let $\widetilde{W}_{Nl} = 1/N \sum_{i=1}^{N} \sum_{k=0}^{n_i} \sum_{j=k+1}^{n_i} w_l^{(ijk)}(\hat{\beta})$, $l = 1, \cdots, d$. By Theorem 2, we can test the validity of (2.2) using the goodness-of-fit statistic

$$Q_{\max} = \max_{l \in \{1 \cdots, d\}} |\widetilde{W}_{Nl}|, \tag{3.4}$$

where large values of $Q_{\max}$ indicate evidence of its invalidity. The critical values of the test statistic $Q_{\max}$ are computed using

$$\widetilde{W}_N = (\widetilde{W}_{N1}, \cdots, \widetilde{W}_{Nd})^{\mathsf{T}} \xrightarrow{d} N(0, \Psi) \; as \; N \to \infty,$$
$$P_{H_0}(Q_{\max} > t) = 1 - P_{H_0}(|\widetilde{W}_{N1}| \leq t, \cdots, |\widetilde{W}_{Nd}| \leq t)$$

and multivariate integration.

4. Simulation Study

We considered three simulation settings to investigate the finite sample properties of the estimation and inference procedures for the RECD model (2.2). Each simulated sample contained either $N = 200$ or $N = 500$ subjects. Following the simulation design considered by Wu, Tian and Jiang (2011), we assigned, for each of the N subjects, 30 "scheduled visits" at time points $\{0, 0.2 + e_1, \cdots, 5.8 + e_{29}\}$ where e_l, $l = 1, \ldots, 29$, were independently generated from the uniform $U(-0.1, 0.1)$ distribution. We allowed each "scheduled visit" in $\{0, 0.2 + e_1, \cdots, 5.8 + e_{29}\}$ to have 40% probability of being missing. The resulting time points lead to an unbalanced longitudinal sample that, to some degree, resembles the time design points of the NGHS data. The observed visiting times are $\{T_{ij} : i = 1, \ldots, N; j = 0, \ldots, n_i\}$ with possibly unequal n_i for different subjects. In all simulation settings, we performed the IMTs on the diagonal indicators. We present the results for the conditional normal density model, where the variances of outcome variables are unknown constants. Additional results for the remaining two simulation settings are given in Section S5 of Supplementary Materials.

For the ith subject, we generated the unobserved random variable Δ_i and S_i from (a) $\Delta_i \sim N(0, 1)$, $S_i \sim N(2 + 0.1\Delta, 0.25)$ and (b) $\Delta_i \sim Exp(1)$, $S_i \sim Exp(1 + 3\Delta_i^2)$, where $Exp(\lambda)$ denotes the exponential distribution with mean λ. The outcome variables of the first model were generated form

$$Y_{ij} | \{\Delta_i, T_i, S_i\} \sim N(\Delta_i, \sigma_Y^2 Z_{ij}^{\mathsf{T}} \beta, \sigma_Y^2), \quad \text{for} \quad j = 0, \cdots, n_i, \tag{4.1}$$

where $Z_{i0} = 0_{4\times 1}$, $Z_{ij} = (1, \widetilde{T}_{ij}, \delta_{ij}, \delta_{ij} R_{ij})^{\mathsf{T}}$, $\widetilde{T}_{ij} = T_{ij} - T_{i0}$, $R_{ij} = T_{ij} - S_i$, $\sigma_Y = 1$, and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^{\mathsf{T}} = (0, 0.2, 0.2, -0.3)^{\mathsf{T}}$. The effects of the concomitant intervention were determined by β_2 and β_3, where $\beta_2 = 0.2$ was the initial increase

Table 1. Estimation and testing results of the 1,000 simulated samples from (4.1), $Y_{ij}|\{\Delta_i, T_i, S_i\} \sim N(\Delta_i, \sigma_Y^2 Z_{ij}^{\mathsf{T}}\beta, \sigma_Y^2)$. Columns 1 and 2: Sample size and true parameter values. Columns 3 to 6: Biases and root mean square errors (RMSEs) of the maximum conditional likelihood estimator $\hat{\beta}$ and the maximum likelihood estimator $\tilde{\beta}$. Column 7: Empirical coverage probabilities (ECPs) of the 95% confidence intervals (CIs) for $\hat{\beta}$. Column 8: Empirical levels (ELs) of the information matrix test (IMT).

N		Bias		RMSE		ECP	EL of IMT
		$\hat{\beta}$	$\tilde{\beta}$	$\hat{\beta}$	$\tilde{\beta}$		
(a) $\Delta_i \sim N(0,1)$, $S_i \sim N(2+0.1\Delta_i, 0.25)$							
200	$\beta_0 = 0$	0.0037	0.0033	0.0911	0.0873	0.958	0.053
	$\beta_1 = 0.2$	0.0002	−0.0005	0.0457	0.0443	0.956	
	$\beta_2 = 0.2$	0.0022	0.0028	0.0687	0.0673	0.945	
	$\beta_3 = -0.3$	−0.0005	0.0002	0.0490	0.0477	0.954	
500	$\beta_0 = 0$	−0.0022	−0.0030	0.0568	0.0552	0.955	0.049
	$\beta_1 = 0.2$	0.0019	0.0020	0.0304	0.0295	0.947	
	$\beta_2 = 0.2$	−0.0027	−0.0026	0.0447	0.0435	0.954	
	$\beta_3 = -0.3$	−0.0020	−0.0020	0.0320	0.0312	0.948	
(b) $\Delta_i \sim Exp(1)$, $S_i \sim Exp(1+3\Delta_i^2)$							
200	$\beta_0 = 0$	0.0010	0.0095	0.0855	0.0829	0.950	0.056
	$\beta_1 = 0.2$	0.0009	0.0325	0.0200	0.0373	0.946	
	$\beta_2 = 0.2$	−0.0016	−0.1086	0.0637	0.1252	0.951	
	$\beta_3 = -0.3$	−0.0015	−0.0365	0.0259	0.0441	0.936	
500	$\beta_0 = 0$	0.0016	0.0120	0.0534	0.0522	0.943	0.053
	$\beta_1 = 0.2$	−0.0006	0.0326	0.0124	0.0347	0.950	
	$\beta_2 = 0.2$	−0.0008	−0.1115	0.0409	0.1187	0.956	
	$\beta_3 = -0.3$	0.0007	−0.0361	0.0157	0.0391	0.953	

of the mean outcome values at the start of the intervention and $\beta_3 = -0.3$ was the rate of decrease in the mean outcome values as the intervention duration time increases.

We considered two estimators of β, the maximum conditional likelihood estimator $\hat{\beta}$ and the maximum likelihood estimator $\tilde{\beta}$. Here, $\tilde{\beta}$ was computed by maximizing the likelihood function of the fully parametric model of (4.1), with $\Delta_i \sim N(\mu, \sigma_\Delta^2)$ and $S_i \sim N(\alpha_0 + \alpha_1\Delta_i, \sigma_S^2)$. Here μ, σ_Δ^2, α_0, α_1 and σ_S^2 are nuisance parameters in this fully parametric model. In case (a), $\Delta_i \sim N(0,1)$ and $S_i \sim N(2+0.1\Delta, 0.25)$, so that the parametric model was correctly specified. In case (b), the data were generated from (4.1) with $\Delta_i \sim Exp(1)$ and $S_i \sim Exp(1+3\Delta^2)$, so that the parametric model was misspecified.

The simulation was repeated 1,000 times. Table 1 summarizes the biases and root mean square errors (RMSE) of the estimators, the empirical coverage probabilities (ECP) of the 95% confidence intervals (CI) for the maximum conditional

likelihood estimator $\hat{\beta}$, and the empirical levels (EL) of the IMT for the model (4.1). When the fully parametric model was correctly specified, the RMSEs of $\tilde{\beta}$ are slightly smaller than that of $\hat{\beta}$. When the fully parametric model was misspecified, $\hat{\beta}$ has significantly smaller RMSEs than that of $\tilde{\beta}$. Since it is often difficult in applications to correctly specify the distributions of S_i and Δ_i, the RMSE results from Table 1 suggest that $\hat{\beta}$ is more desirable than $\tilde{\beta}$ in practice. The ECPs of the normal approximate CIs for $\hat{\beta}$ are close to the nominal level of 95%. The ELs of the IMT, given by the P-values of IMT statistic $Q_{\max}$, get closer to the nominal significance level of 0.05 under H_0 when N increases from 200 to 500. These results of the IMT empirical levels are then consistent with the asymptotic results of Theorem 2.

5. Application to ENRICHD Pharmacotherapy Data

Our objective for the ENRICHD Pharmacotherapy data is to evaluate the additional effects of using antidepressants on the trends of depression severity measured by the BDI scores for the patients who received antidepressants during the six-month psychosocial treatment period. Since patients in the UC arm did not have repeatedly measured BDI scores during the 6-month treatment period, we applied our model and method to the sub-sample of the ENRICHD patients in the CBT treatment arm, which has been studied by Wu, Tian and Jiang (2011). Within our sample, 95 started antidepressant during the treatment period and 486 did not use antidepressants before the end of the treatment period. Our sample has 36 more patients than the sample of Wu, Tian and Jiang (2011), which included only those who attended 5 or more CBT sessions.

Let Y_{ij}, T_{ij}, S_i, and $R_{ij} = T_{ij} - S_i$ be the ith patient's BDI score, trial time (days/100), starting time (days/100) of antidepressant use, and antidepressant duration time (days/100), respectively, at the jth visit. For all $1 \le i \le N$, $T_{i0} = 0$, T_{in_i} is the trial time at the last visit, and $\delta_{ij} = I(T_{ij} \ge S_i)$ is the indicator of antidepressant use at T_{ij}. Setting $\gamma(y)$ of (2.2) to $\gamma(y) = \log(1+y)$, our semiparametric conditional density model for this ENRICHD sample is

$$\begin{cases} f_{i0}(y|\Delta_i, T_i, S_i) = h_i(y|\Delta_i, S_i), \\ f_{ij}(y|\Delta_i, T_i, S_i) = \exp(\alpha_{ij} + \log(1+y) Z_{ij}^{\mathsf{T}} \beta) h_i(y|\Delta_i, S_i), \end{cases} \tag{5.1}$$

where $h_i(y|\Delta_i, S_i)$ is an unknown density function depending on $\{\Delta_i, S_i\}$, and the α_{ij}'s are normalizing constants such that $\int f_{ij}(y|\Delta_i, T_i, S_i)dy = 1$. With $\widetilde{T}_{ij} = T_{ij}$ and $R_{ij} = T_{ij} - S_i$, the effects of trial time and antidepressant use are

determined by $Z_{ij}^{\mathsf{T}}\beta$, specified as

$$\begin{aligned}
(a)\quad & Z_{ij}^{\mathsf{T}}\beta = \beta_0 + \log(1+T_{ij})\beta_1 + \delta_{ij}\beta_2 + \log(1+\delta_{ij}R_{ij})\beta_3,\\
& \beta = (\beta_0,\beta_1,\beta_2,\beta_3)^{\mathsf{T}};\\
(b)\quad & Z_{ij}^{\mathsf{T}}\beta = \kappa_i\beta_{01} + \log(1+T_{ij})\beta_1\\
& \qquad +(1-\kappa_i)\{\beta_{00} + \delta_{ij}\beta_2 + \log(1+\delta_{ij}R_{ij})\beta_3\},\\
& \beta = (\beta_{00},\beta_{01},\beta_1,\beta_2,\beta_3)^{\mathsf{T}},
\end{aligned}$$

where $\kappa_i = 0$ if $0 \le S_i \le T_{in_i}$ and $\kappa_i = 1$ if $S_i > T_{in_i}$. Here (b) is a more general model than (a), because it assumes that intercepts are possibly different for patients with different values of κ_i. For both (a) and (b), the antidepressant effects are described by β_2 and β_3.

When $h_i(\cdot|\cdot)$ is the density of a log-normal distribution, (5.1) is equivalent to

$$\begin{cases}
\log(1+Y_{i0})|\{\Delta_i,T_i,S_i\} \sim N(\Delta_{1i},\Delta_{2i}),\\
\log(1+Y_{ij})|\{\Delta_i,T_i,S_i\} \sim N(\Delta_{1i}+\Delta_{2i}Z_{ij}^{\mathsf{T}}\beta,\Delta_{2i}), \quad j=1,\cdots,n_i,
\end{cases} \tag{5.2}$$

where $\Delta_i = (\Delta_{1i},\Delta_{2i})^{\mathsf{T}}$ and $Z_{ij}^{\mathsf{T}}\beta$ are specified either in (a) or (b). Under (5.2), the covariate effects β can be interpreted as a scale-invariant regression coefficient vector for the conditional mean regression model. A positive (negative) value of β_2 represents the mean increase (decrease) of BDI scores at the time of antidepressant use. A positive (negative) value of β_3 represents the increase (decrease) of BDI scores following a unit increase of the antidepressant duration time. Negative values of β_2 and β_3 represent the benefit of antidepressant use for lowering BDI scores. Since the distribution of Δ_i is unknown, the conditional likelihood method of Section 3.1 has to be used to estimate β under (5.2), which leads to the same estimators as (5.1).

Table 2 shows the maximum conditional likelihood estimators of β under (5.1) with the covariate effects (a) and (b), the corresponding 95% CI's for β_j, $j = 0,1,2,3$, and the p-values of the Wald-type test for testing $\beta_j = 0$, $j = 0,1,2,3$. To evaluate the validity of the model (5.1), we performed the IMT on the diagonal indicators. The p-values of the IMT were 0.1545 and 0.1578 for covariate settings (a) and (b), respectively, which suggest that (5.1) provides an acceptable fit to the data. Since (5.1) under (b) is a more general model than (5.1) under (a), the 95% CI's and p-values of β_{00} and β_{01} suggest that (5.1) under (b) is a more appropriate model than (5.1) under (a). Because $\delta_{ij} = 0$ when $\kappa_i = 1$, the estimates for β_3 under (a) and (b) of (5.1) use almost the same information, hence, are almost the same. Because of the addition of the intercept term β_{01}

Table 2. The maximum conditional likelihood (CL) estimators of the parameters under (5.1) with covariate effects (a) and (b), their corresponding 95% confidence intervals (CI) for β_j, the p-values of the Wald-type test for testing $\beta_j = 0$, and the p-values of the IMT test statistics for the validity of (5.1) based on the ENRICHD Pharmacotherapy Data.

Covariate Effects	Parameter	Maximum CL Estimator	95% CI	P-value of Wald-type test	P-value of IMT for Model (5.1)
(a)	β_0	-2.3583	$(-2.8801, -1.8366)$	< 0.0001	0.1545
	β_1	-5.1193	$(-5.9348, -4.3038)$	< 0.0001	
	β_2	0.2188	$(-1.4293, 1.8668)$	0.7947	
	β_3	-0.6251	$(-3.2808, 2.0306)$	0.6446	
(b)	β_{00}	-0.5043	$(-1.5106, 0.5020)$	0.3260	0.1578
	β_{01}	-3.0954	$(-3.7325, -2.4583)$	< 0.0001	
	β_1	-5.0800	$(-5.8942, -4.2657)$	< 0.0001	
	β_2	-0.0101	$(-1.7331, 1.7128)$	0.9908	
	β_3	-0.6779	$(-3.3316, 1.9758)$	0.6166	

for $\kappa_i = 1$ in (5.1) under (b), the estimates of β_2 differ between (a) and (b). The negative estimates of β_2 and β_3 from (5.1) under (b) suggest that the use of antidepressant may have some beneficial effect of lowering BDI scores over time for this patient population, where the large p-values for β_2 and β_3 suggest that the evidence of beneficial effect for antidepressant use is at most weak under the current semiparametric model and moderate sample size. Wu, Tian and Jiang (2011) detected some significant statistical evidence for the beneficial effect of antidepressant use, their conclusions obtained based on the stronger assumption of parametric models.

Under (5.2) with the covariate effects (b), the predictor $\hat{\Delta}_i = (\hat{\Delta}_{1i}, \hat{\Delta}_{2i})^{\mathsf{T}}$ of Δ_i can be obtained as the solution to the equations $(n_i+1)\Delta_{1i} = \sum_{j=0}^{n_i} \log(1 + Y_{ij}) - \Delta_{2i} \sum_{j=0}^{n_i} Z_{ij}^{\mathsf{T}}\hat{\beta}$ and $\Delta_{2i} = (n_i+1)^{-1} \sum_{j=0}^{n_i} \{\log(1+Y_{ij}) - \Delta_{1i} - \Delta_{2i} Z_{ij}^{\mathsf{T}}\hat{\beta}\}^2$. The overall trend can be estimated by $\hat{Y}_{i0} = \hat{\Delta}_{1i}$ and $\hat{Y}_{ij} = \hat{\Delta}_{1i} + \hat{\Delta}_{2i} Z_{ij}^{\mathsf{T}}\hat{\beta}$, $i = 1, \cdots, N$, $j = 1, \cdots, n_i$. In Figure 1, the estimated overall trends of BDI scores for six patients with concomitant intervention are plotted based on points $(\log(1 + T_{ij}), \log(1 + \hat{Y}_{ij}))$, $j = 0, \cdots, n_i$. The overall decreasing trend of BDI scores represents the benefit of antidepressant use for these patients.

6. Conclusions and Discussions

We developed in this paper a class of RECD models for evaluating the distributions and concomitant intervention effects with longitudinal data. Under these models, the conditional density ratio is assumed to have a parametric form, while

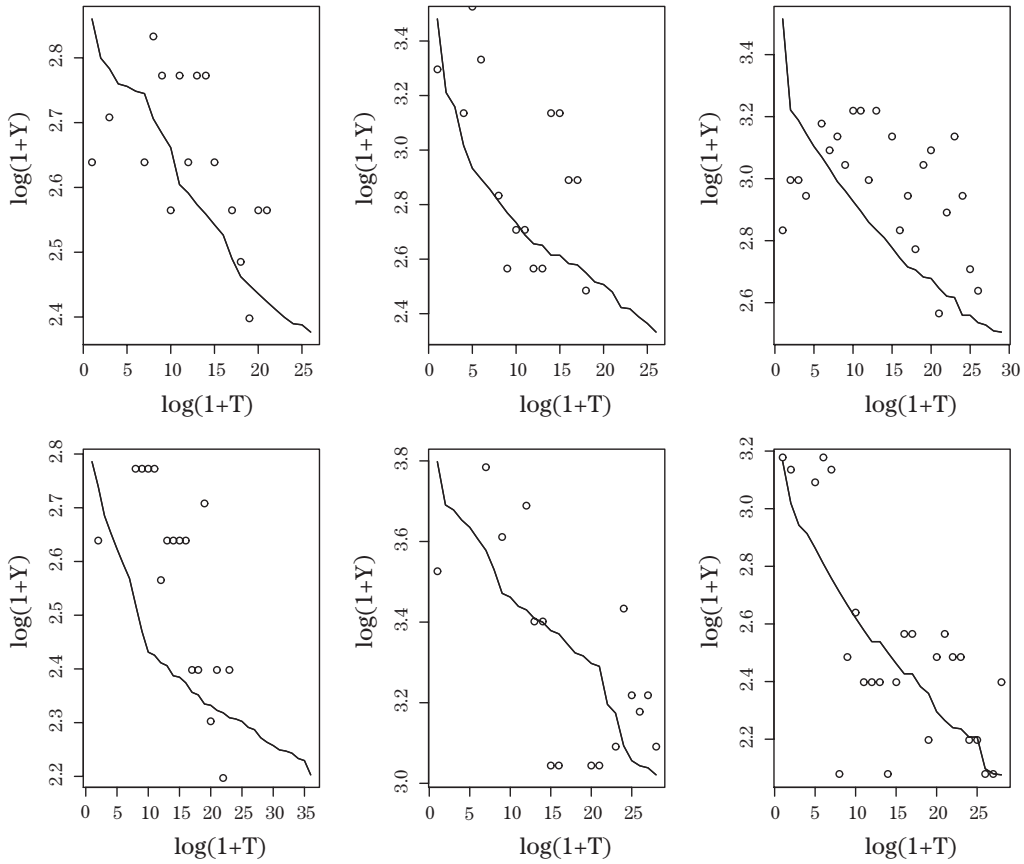

Figure 1. The estimated overall trend of BDI scores for six patients with concomitant intervention based on the ENRICHD Pharmacotherapy Data.

the baseline density function is nonparametric. We further proposed a likelihood-based method for estimating the parameters and a goodness-of-fit test for testing the validity of the models, and derived the consistency and asymptotic normality of the conditional likelihood estimators. We illustrated the practical values of the RECD models through a simulation study and an application to a longitudinal clinical trial for depression.

Supplementary Materials

The supplementary materials contain further discussions of the models, derivations of the likelihood functions under the RECD models, additional simulation results, and proofs of Theorems 1 and 2.

Acknowledgment

Research of the first, third and fourth authors was partly supported by the ILIR (Independent Laboratory In-House Research), Department of the Army, W81XWH-11-C-0082. The views expressed are those of the authors and should not be construed to represent the positions of the Department of the Army or Department of Defense. The ENRICHD study was supported by contracts NO1-HC-55140 through NO1-HC-55148 from the National Heart, Lung and Blood Institute, National Institutes of Health. Tianqing Liu was partly supported by the NSFC (No.11201174) and the Natural Science Foundation for Young Scientists of Jilin Province, China (No.20150520054JH).

References

Anderson, J. A. (1979). Multivariate logistic compounds. *Biometrika* **66**, 17-26.

Chan, K. C. G. (2013). Nuisance parameter elimination for proportional likelihood ratio models with nonignorable missingness and random truncation. *Biometrika* **100**, 269–276.

Crowder, M. (1986). On consistency and inconsistency of estimating equations. *Econometric Theory* **2**, 305–330.

Diggle, P. J., Heagerty, P., Liang, K. Y. and Zeger, S. L. (2002). *Analysis of Longitudinal Data* (2nd edition). Oxford: Oxford University Press.

ENRICHD Investigators (2001). Enhancing recovery in coronary heart disease patients (ENRICHD): study intervention rationale and design. *Psychosomatic Medicine* **63**, 747–755.

ENRICHD Investigators (2003). Enhancing recovery in coronary heart disease patients (ENRICHD): the effects of treating depression and low perceived social support on clinical events after myocardial infarction. *J. Am. Med. Assoc.* **289**, 3106–3116.

Fitzmaurice, G., Davidian, M., Verbeke, G. and Molenberghs, G. (2009). *Longitudinal Data Analysis*, Boca Raton: Chapman & Hall/CRC.

Follmann, D. and Wu, M. C. (1995). An approximate generalized linear model with random effects for informative missing data. *Biometrics* **51**, 151–168.

Kalbfleisch, J. D. (1978). Likelihood methods and nonparametric tests. *J. Am. Statist. Assoc.* **73**, 167–170.

Kay, R. and Little, S. (1987). Transformations of the explanatory variables in the logistic regression model for binary data. *Biometrika* **74**, 495–501.

Liang, K. Y. and Qin, J. (2000). Regression analysis under non-standard situations: a pairwise pseudolikelihood approach. *J. R. Statist. Soc. B* **62**, 773–786.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Qin, J., Berwick, M., Ashbolt, R. and Dwyer, T. (2002). Quantifying the change of melanoma incidence by Breslow thickness. *Biometrics* **58**, 665–670.

Taylor, C. B., Youngblood, M. E., Catellier, D., Veith, R. C., Carney, R. M., Burg, M. M., Kaufmann, P., Shuster, J., Mellman, T., Blumenthal. and others (2005). Effects of an-

tidepressant medication on morbidity and mortality in depressed patients after myocardial infarction. *Archives of General Psychiatry* **62**, 792–798.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*, New York: Springer.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* **50**, 1–25.

Wu, C. O., Tian, X. and Bang, H. (2008). A varying-coefficient model for the evaluation of time-varying concomitant intervention effects in longitudinal studies. *Statistics in Medicine* **27**, 3042–3056.

Wu, C. O., Tian, X. and Jiang, W. (2011). A shared parameter model for the estimation of longitudinal concomitant intervention effects. *Biostatistics* **12**, 737–749.

Xing, H. and Ying, Z. (2012). A semiparametric change-point regression model for longitudinal observations. *J. Am. Statist. Assoc.* **107**, 1625–37.

School of Mathematics, Jilin University, Changchun 130012, China.

E-mail: tianqingliu@outlook.com

Office of Biostatistics Research, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892-7913, U.S.A.

E-mail: wuc@nhlbi.nih.gov

Department of Statistics, The George Washington University, Rome Hall, 7th Floor, 801, 22nd Street, NW Washington, D.C. 20052, USA.

E-mail: zli@gwu.edu

Division of Preventive Medicine, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA.

E-mail: Yuanzhang.li2.civ@mail.mil

(Received May 2016; accepted February 2017)