Regression Analysis when Covariates are Regression Parameters of a Random Effects Model for Observed Longitudinal Measurements

C. Y. Wang[1], Naisyin Wang[2] and Suojin Wang[2]

[1] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109-1024, USA. *email:* cywang@fhcrc.org
[2] Department of Statistics, Texas A&M University, College Station, TX 77843-3143

September 27, 1999

SUMMARY

We consider regression analysis when covariate variables are the underlying regression coefficients of another linear mixed model. A naive approach is to use each subject's repeated measurements, which are assumed to follow a linear mixed model, and obtain subject-specific estimated coefficients to replace the covariate variables. However, directly replacing the unobserved covariates in the primary regression by these estimated coefficients may result in a significantly biased estimator. The aforementioned problem can be evaluated as a generalization of the classical additive error model where repeated measures are considered as replicates. To correct for these biases, we investigate a pseudo expected estimating equation (EEE) estimator, a regression calibration (RC) estimator and a refined version of the RC estimator. For linear regression, the first two estimators are identical under certain conditions. However, when the primary regression model is a nonlinear model, the RC estimator is usually biased. We thus consider a refined regression calibration estimator whose performance is close to that of the pseudo EEE estimator but does not require numerical integration. The RC estimator is also extended to the proportional hazards regression model. In addition to the distribution theory, we evaluate the methods through simulation studies. The methods are applied to analyze a real dataset from a child growth study.

Keywords: Censored regression; Estimating equation; Generalized linear model; Maximum likelihood; Measurement error; Random effects; Regression calibration.

Abbreviated title: Random-effect-coefficient Covariates Model

1 Introduction

The associations of longitudinal measurements of certain variables and the response variable in a primary regression model are often of interest. For example, longitudinal serum CA 125 levels are often used as a screening test for ovarian cancer in high risk women. Also, the survival of AIDS patients is associated with repeated $CD4^+$ lymphocyte counts. Using the subject-specific long-term averages of longitudinal measurements as covariates has been a common practice. However, in some applications, it is the rate of change over time rather than the average that is the major risk factor.

One particular motivating example for the current paper is a study on adult obesity (Whitaker et al., 1997). In routine health care for children, growth is monitored by recording heights and weights over time, among other measurements. A clinician is interested in the extent to which growth data from childhood can predict the likelihood of obesity in adulthood. One summary measurement for height and weight is body mass index (BMI). The BMI is defined as a subject's weight/height^2 (kg/m^2). Note that, in practice, the measurement times are usually unbalanced and missing data often occur. From the longitudinal BMI measurements over time, a clinician may wish to obtain some summary information which can be further used for the purpose of the intervention of adult obesity. Taking a long term average of BMI over a period of childhood was considered in Wang and Pepe (1999). In this paper, we consider a more general approach which allows such summary information to be regression coefficients based on the longitudinal BMI data. For example, assuming a child's BMI z-scores follow a linear trend between ages 3-5, researchers are interested in the effects of the initial BMI z-score at age 3 and the rate of change over time during this preschool age category (Whitaker et al.). The unobserved underlying parameters in the longitudinal linear model are the covariates of the primary model.

There are considerable number of publications where repeated measures are treated as replicates over time. That is, if we denote the measurements by $W_i(t_{ij}) \equiv W_{ij}, j = 1, \ldots, k_i$, then $W_{ij} = X_i + U_{ij}$, where each of the U_{ij} has mean 0. See Fuller (1987) and Carroll, et al. (1995). In the additive measurement error model, it is known that a significant bias problem may occur if one replaces the long-term average X_i by the observed average, $k_i^{-1} \sum_{i=1}^{n} W_{ij}$, of a subject for the primary regression model. However, to our knowledge, there is little work that has been done when the covariate variables of the primary regression model are random coefficients of another random effects model. A related problem in pharmacokinetics was investigated in Higgins et al. (1997). See also Zucker et al. (1995) for related work on inferences in random effects regression.

The structure of interest here can be described by two components: The first component contains repeated observed measurements which are assumed to follow a linear random effects model over possibly unequally-spaced times. The second component is the primary regression where the random coefficients of the random effects model are the covariate variables. A naive approach is to estimate the subject-specific parameters from say a least squares estimator and then use these estimates as replacements for the underlying covariates. The discussion in Section 2 indicates that this estimator is biased even for simple linear regression.

In this paper, we consider three approaches. We first apply the pseudo expected estimating equation (EEE) estimator (Wang and Pepe, 1999) to this particular data structure; see Section 3. The drawback of this approach is that except for a linear normal regression analysis, the estimating procedure usually involves numerical integration. Alternatively, one may consider a regression calibration (RC) estimator which replaces the underlying unobserved covariates by certain proper conditional expectations (Carroll, et al., 1995, Chapter 3). An interesting finding here is that, for normal random replicates, the conditional distributions of the unobserved covariates given the subject-specific least squares estimated coefficients are equivalent to those given all the random replicates. That is, to perform the RC (or later a refined RC) estimation, one only needs certain summary statistics from the secondary linear mixed model. Detailed studies about the RC estimator as well as the refined RC estimator in the generalized linear model setting are given in Section 4. In Section 5, we consider the case when the primary regression follows the Cox (1972) regression model and the outcome variable, failure time, may be censored. The RC estimator in this case is investigated. Simulation results on linear, logistic, Cox regression are presented in Section 6. In Section 7, we present analyses of data from the study of childhood predictors of adult obesity. Some possible extensions are discussed in Section 8.

2 Model and Naive Estimator

Except in Section 5 where Cox regression is considered, let Y_i be the outcome variable for the ith subject, $i = 1, \ldots, n$. Assume that repeated measurements $W_{ij} \equiv W_i(t_{ij}), j = 1, \ldots, k_i$, follow a random effects model

$$W_{ij} = \boldsymbol{D}'_{ij}\boldsymbol{X}_i + U_{ij},\ \ i = 1, \ldots, n, \tag{1}$$

where $\boldsymbol{D}_{ij} \equiv \boldsymbol{D}_i(t_{ij})$ are $q \times 1$ observed covariate vector, often with the first element being 1, and $\boldsymbol{X}_i = (X_{i1}, \cdots, X_{iq})'$ contains the ith set of random coefficients. For subject i, it is assumed that the number of measurements $k_i > q$. Consider $W_i(t)$ as an exposure measurement of subject i at time t. A typical example is that $\boldsymbol{D}_{ij} = (1, t_{ij})'$, with $q = 2$, and X_{i1}, X_{i2} correspond to the initial exposure level and the rate of change, respectively, of subject i. Note that time t_{ij} may be unequally-spaced. We further assume that $\boldsymbol{X}_i$ is multivariate-normal with mean $\boldsymbol{\mu}_x$, and variance-covariance matrix $\boldsymbol{\Sigma}_x$. Also let $\boldsymbol{U}_i \equiv (U_{i1}, \ldots, U_{ik_i})'$, which follows a multivariate normal distribution and satisfies that $\mathrm{E}(\boldsymbol{U}_i|\boldsymbol{X}_i) = \boldsymbol{0}$ and $\mathrm{var}(\boldsymbol{U}_i|\boldsymbol{X}_i) = \boldsymbol{\Sigma}_{ui}$. This includes correlated measurement errors (Wang et al., 1996).

Let $\boldsymbol{W}_i \equiv (W_{i1}, \ldots, W_{ik_i})'$. We assume that the primary regression model satisfies

$$\mathrm{E}(Y_i|\boldsymbol{W}_i, \boldsymbol{X}_i) = g(\boldsymbol{\mathcal{X}}'_i\boldsymbol{\beta}), \tag{2}$$

for some known function g, $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_q)'$ are the regression parameters of interest, and $\boldsymbol{\mathcal{X}}_i = (1, \boldsymbol{X}_i')'$. It is further stipulated that $\boldsymbol{W}_i$ is a surrogate for $\boldsymbol{X}_i$ such that the distribution of $(Y_i|\boldsymbol{W}_i, \boldsymbol{X}_i)$ is that of $(Y_i|\boldsymbol{X}_i)$, independent of $\boldsymbol{W}_i$. Note that (2) may be extended to include $\boldsymbol{\alpha}'\boldsymbol{Z}_i$ for other observable covariates $\boldsymbol{Z}_i$. When $\boldsymbol{D}_{ij} = (1, t_{ij})'$ with $q = 2$, we are concerned with the effects of the initial exposure level and the rate of change, on outcome Y.

One naive approach to this problem is to replace the unobserved $\boldsymbol{X}_i$ by the least squares (LS) estimator from model (1) for each subject. Let $\boldsymbol{D}_i$ be the ith design matrix where the jth row contains $\boldsymbol{D}_{ij}'$. Denote the LS estimates of $\boldsymbol{X}_i$ by $\boldsymbol{X}_i^{(ls)}$, *i.e.* $\boldsymbol{X}_i^{(ls)} = (\boldsymbol{D}_i'\boldsymbol{D}_i)^{-1}\boldsymbol{D}_i'\boldsymbol{W}_i$. To demonstrate the existence of bias, we consider linear regression $\mathrm{E}(Y_i|\boldsymbol{X}_i) = \boldsymbol{\mathcal{X}}_i'\boldsymbol{\beta}$. Write $\boldsymbol{X}_i^{(ls)} = \boldsymbol{X}_i + \boldsymbol{d}_i$, with $\mathrm{E}(\boldsymbol{d}_i) = 0$ and $\mathrm{cov}(\boldsymbol{X}_i^{(ls)}|\boldsymbol{X}_i) \equiv \boldsymbol{\Sigma}_{(wi)} = (\boldsymbol{D}_i'\boldsymbol{D}_i)^{-1}\boldsymbol{D}_i'\Sigma_{ui}\boldsymbol{D}_i(\boldsymbol{D}_i'\boldsymbol{D}_i)^{-1}$. Similar to Fuller (1987, Chapter 3), one can easily see that regressing Y_i on $\boldsymbol{X}_i^{(ls)}$ leads to biased estimates of $\boldsymbol{\beta}$ with an attenuation factor, $\boldsymbol{\Sigma}_x\left\{\boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_{(wi)}\right\}^{-1}$, on the vector of regression slopes. In general though, the individual slopes may not necessarily be underestimated.

3 Pseudo Expected Estimating Equation Estimator

Wang and Pepe (1999) proposed an expected estimating equation approach for measurement error problems. In our problem, let $\boldsymbol{\psi}(Y_i, \boldsymbol{X}_i, \boldsymbol{\beta})$ be the estimating score when data are complete. If the conditional distribution of $\boldsymbol{X}_i$ given $(Y_i, \boldsymbol{W}_i)$ is known, then the EEE estimator solves $\sum_{i=1}^n \mathrm{E}\{\boldsymbol{\psi}(Y_i, \boldsymbol{X}_i, \boldsymbol{\beta})|Y_i, \boldsymbol{W}_i\} = 0$. Note that $P(Y_i|\boldsymbol{X}_i, \boldsymbol{W}_i)P(\boldsymbol{W}_i|\boldsymbol{X}_i)P(\boldsymbol{X}_i) = P(Y_i|\boldsymbol{X}_i)P(\boldsymbol{W}_i|\boldsymbol{X}_i)P(\boldsymbol{X}_i)$. Hence, the EEE may be written as

$$\sum_{i=1}^{n} \frac{\int \boldsymbol{\psi}(Y_i, \boldsymbol{x}, \boldsymbol{\beta})P(Y_i|\boldsymbol{x})P(\boldsymbol{W}_i|\boldsymbol{x})P(\boldsymbol{x})d\boldsymbol{x}}{\int P(Y_i|\boldsymbol{x})P(\boldsymbol{W}_i|\boldsymbol{x})P(\boldsymbol{x})d\boldsymbol{x}} = 0. \tag{3}$$

Calculating (3) generally requires numerical integration in the case of nonlinear regression. Moreover, the nuisance parameters in the $(\boldsymbol{X}, \boldsymbol{W})$ model, including $\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \boldsymbol{\Sigma}_{ui}$, need to be estimated. Let $\boldsymbol{\eta}$ be the vector of nuisance parameters. The EEEs for $\boldsymbol{\eta}$ can also be obtained similarly by taking conditional expectations; see Wang and Pepe for details. In this paper, we consider writing a method of moments estimate of $\boldsymbol{\eta}$, $\widehat{\boldsymbol{\eta}}$, as a solution to a class of equations $\sum_{i=1}^{n} \boldsymbol{\Psi}_N(\boldsymbol{W}_i, \boldsymbol{\eta}) = 0$, where $\boldsymbol{\Psi}_N(\boldsymbol{W}_i, \boldsymbol{\eta})$ satisfies $\mathrm{E}\{\sum_{i=1}^{n} \boldsymbol{\Psi}_N(\boldsymbol{W}_i, \boldsymbol{\eta})\} = 0$. Note that the left hand side of (3) is a function of both $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$. It is convenient to write (3) as $\sum_{i=1}^{n} \boldsymbol{\Psi}_P(Y_i, \boldsymbol{W}_i, \boldsymbol{\beta}, \boldsymbol{\eta}) = 0$, and the pseudo EEE estimator of $\boldsymbol{\beta}$ is then obtained for given $\boldsymbol{\eta} = \widehat{\boldsymbol{\eta}}$. The subscript "$P$" and "$N$" in $\boldsymbol{\Psi}_P$ and $\boldsymbol{\Psi}_N$ indicate their association with the "primary" and "nuisance" parameters, respectively.

In general, estimation of the nuisance parameters involved in the distribution of $\boldsymbol{W}_i$ can be obtained by the restricted maximum likelihood estimator or the method of moments. When Y given $\boldsymbol{X}$ is linear, (3) has an explicit form and hence numerical integration is not needed. In addition to the nuisance parameters associated with $\boldsymbol{W}_i$, the nuisance vector $\boldsymbol{\eta}$ contains an additional parameter σ_e, the standard deviation of the $(Y, \boldsymbol{X})$ regression errors. By calculating $\mathrm{var}(Y_i|\boldsymbol{W}_i)$, a method of moments estimator for σ_e^2 can be obtained by solving

$$\sum_{i=1}^{n} \left[\{Y_i - \mathrm{E}(\boldsymbol{\mathcal{X}}_i'|\boldsymbol{W}_i)\boldsymbol{\beta}\}^2 - \sigma_e^2 - \boldsymbol{\beta}(\boldsymbol{\Sigma}_x \boldsymbol{D}_i')(\boldsymbol{D}_i \boldsymbol{\Sigma}_x \boldsymbol{D}_i' + \boldsymbol{\Sigma}_{ui})^{-1} \boldsymbol{D}_i \boldsymbol{\Sigma}_x \boldsymbol{\beta}' \right] = 0. \quad (4)$$

In this case, $\boldsymbol{\Psi}_N(\boldsymbol{W}_i, \boldsymbol{\delta})$ should be written as $\boldsymbol{\Psi}_N(Y_i, \boldsymbol{W}_i, \boldsymbol{\beta}, \boldsymbol{\eta})$. We now summarize a result regarding the EEE and the likelihood-based score.

Proposition 1: For any $g(\boldsymbol{\mathcal{X}}_i'\boldsymbol{\beta})$, if $\boldsymbol{\psi}(Y_i, \boldsymbol{X}_i, \boldsymbol{\beta}) = (\partial/\partial\boldsymbol{\beta})\ln\{P(Y_i|\boldsymbol{X}_i)\}$, then the EEE for $\boldsymbol{\beta}$, $\mathrm{E}\{\boldsymbol{\psi}(Y_,\boldsymbol{X}_i, \boldsymbol{\beta})|Y_i, \boldsymbol{W}_i\}$, is identical to $(\partial/\partial\boldsymbol{\beta})\ln\{P(Y_i|\boldsymbol{W}_i)\}$. Furthermore, under the linear case, $Y_i = \boldsymbol{\mathcal{X}}_i'\boldsymbol{\beta} + e_i$, where e_i's are independent of $\boldsymbol{X}_i$'s, if $(\boldsymbol{X}_i', \boldsymbol{U}_i', e_i)'$ is multivariate normally distributed, $\boldsymbol{D}_i = \boldsymbol{D}_1$ for $i = 1, \ldots, n$, and σ_e^2 is estimated by solving (4), then solving $\sum_{i=1}^{n} \mathrm{E}\{\boldsymbol{\psi}(Y_i, \boldsymbol{X}_i, \boldsymbol{\beta})|Y_i, \boldsymbol{W}_i\} = 0$ is equivalent to solving $\sum_{i=1}^{n} \boldsymbol{\psi}\{Y_i, \mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i), \boldsymbol{\beta}\} = 0$.

The proof of Proposition 1 is given in the Appendix. As a result, in linear regression under the conditions described above, the EEE is equivalent to the complete data estimating equation by replacing the unobserved $\boldsymbol{X}_i$ by $\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i)$. In other words, in this case, the pseudo EEE estimator leads to the regression calibration estimator.

4 Regression Calibration

4.1 *The method*

Proposition 1 provides a situation, namely the linear case, in which the RC estimator is consistent and is in fact a pseudo likelihood estimator. We now discuss the use of the RC estimator for a more general setting, the generalized linear model in which given $\boldsymbol{X}_i$, the distribution of Y_i belongs to a natural exponential family. That is, $\mathrm{E}(Y_i|\boldsymbol{X}_i) = g(\boldsymbol{\mathcal{X}}'_i\boldsymbol{\beta})$; $\mathrm{var}(Y_i|\boldsymbol{X}_i) = \gamma\kappa_i^{-1}v\{g(\boldsymbol{\mathcal{X}}'_i\boldsymbol{\beta})\}$, where γ, κ_i and $v(\cdot)$ are respectively a scale parameter, a prior weight, and a variance function, and $g(\cdot)$ is a given function. The inverse function, $g^{-1}(\cdot)$, is usually referred to as the link function. For example, in the aforementioned linear case, the link is the identity function.

The RC approach may be implemented by replacing the unobserved $\boldsymbol{X}_i$ with its conditional expectation given the observed longitudinal measurements, $\boldsymbol{W}_i$. In the classical measurement error model such that $W_{ij} = X_i + U_{ij}$, U_{ij} is independent of $U_{ij'}$ for $j \neq j'$, $\mathrm{E}(X_i|\boldsymbol{W}_i)$ can be easily shown to be the same as $\mathrm{E}(X_i|\overline{W}_i)$, where $\overline{W}_i = \sum_{j=1}^{k_i} W_{ij}/k_i$. Thus, the simple summary statistic $\overline{W}_i$ is sufficient to serve as a surrogate for X_i. In our problem, one obvious question is whether a subject's LS estimated coefficient estimator, $\boldsymbol{X}_i^{(ls)}$, is sufficient to serve as a surrogate for $\boldsymbol{X}_i$. That is, instead of replacing $\boldsymbol{X}_i$ by $\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i)$, is it equally efficient if one replaces it by $\mathrm{E}\{\boldsymbol{X}_i|\boldsymbol{X}_i^{(ls)}\}$?

Under a Gaussian model described in Section 2, if $\boldsymbol{\Sigma}_{ui} = \sigma_u^2\mathbf{I}_{k_i}$, then it can be readily seen that $\boldsymbol{X}_i^{(ls)}$ is a sufficient statistic for $\boldsymbol{X}_i$, and thus $(\boldsymbol{X}_i|\boldsymbol{W}_i)$ and $(\boldsymbol{X}_i|\boldsymbol{X}_i^{(ls)})$ have the same distribution. Therefore, $\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i) = \mathrm{E}\{\boldsymbol{X}_i|\boldsymbol{X}_i^{(ls)}\}$, $\mathrm{var}(\boldsymbol{X}_i|\boldsymbol{W}_i) = \mathrm{var}\{\boldsymbol{X}_i|\boldsymbol{X}_i^{(ls)}\}$.

If Y_i given $\boldsymbol{X}_i$ is linear, then Y_i given $\boldsymbol{X}_i^{(ls)}$ still has the linear form with *covariates* $\mathrm{E}\{\boldsymbol{X}_i|\boldsymbol{X}_i^{(ls)}\}$ and the same parameter $\boldsymbol{\beta}$. However, this property does not generally apply to nonlinear link functions. To further reduce the bias for cases with a nonlinear link, a refined RC approach is proposed in the next subsection.

4.2 *Refined RC Approach*

In logistic regression with covariate measurement error, the RC estimator may suffer from the bias problem particularly when the magnitude of the relative risk is large, although large $\boldsymbol{\Sigma}_{x|w}$ and incidence rate will also amplify the bias. To reduce the bias, we investigate a refined approach which aims at approximating the pseudo EEE while avoiding a numerical integration for the following models. First, we consider the probit regression where Y_i has a binary outcomes, 0, 1 and $\text{pr}(Y_i = 1|\boldsymbol{X}_i) = \Phi(\boldsymbol{\mathcal{X}}_i'\boldsymbol{\beta})$, with $\Phi(\cdot)$ being a standard normal distribution function. Similar to Carroll et al. (1984),

$$\text{pr}(Y_i = 1|\boldsymbol{W}_i) = \Phi\left\{\frac{\beta_0 + \text{E}(\boldsymbol{X}_i|\boldsymbol{W}_i)\boldsymbol{\beta}_{[0]}}{(1+\boldsymbol{\beta}_{[0]}'\boldsymbol{\Sigma}_{x|w}\boldsymbol{\beta}_{[0]})^{1/2}}\right\}, \tag{5}$$

where $\text{E}(\boldsymbol{X}_i|\boldsymbol{W}_i) = \boldsymbol{\mu}_x + \boldsymbol{\Sigma}_x\boldsymbol{D}_i'(\boldsymbol{D}_i\boldsymbol{\Sigma}_x\boldsymbol{D}_i' + \boldsymbol{\Sigma}_{ui})^{-1}\{\boldsymbol{W}_i - \boldsymbol{D}_i\boldsymbol{\mu}_x\}$, $\boldsymbol{\Sigma}_{x|w} = \text{var}(\boldsymbol{X}_i|\boldsymbol{W}_i) = \boldsymbol{\Sigma}_x - \boldsymbol{\Sigma}_x\boldsymbol{D}_i'(\boldsymbol{D}_i\boldsymbol{\Sigma}_x\boldsymbol{D}_i' + \boldsymbol{\Sigma}_{ui})^{-1}\boldsymbol{D}_i\boldsymbol{\Sigma}_x$. Therefore, for probit regression, a refined analysis may be made by utilizing a set of expected estimating equation based on (5).

For logistic regression, $\text{pr}(Y_i = 1|\boldsymbol{X}_i) = H(\boldsymbol{\mathcal{X}}_i'\boldsymbol{\beta})$, where $H(u) = \{1+\exp(-u)\}^{-1}$. As in Liang and Liu (1991) and Carroll et al. (1995, page 65), applying the probit-to-logistic approximation ($\Phi(v) \approx H(cv)$ with $c = 15\pi/(16\sqrt{3}) \approx 1.7$), we have

$$\text{pr}(Y_i = 1|\boldsymbol{W}_i) \approx H\left\{\frac{\beta_0 + \text{E}(\boldsymbol{X}_i|\boldsymbol{W}_i)\boldsymbol{\beta}_{[0]}}{(1+\boldsymbol{\beta}_{[0]}'\boldsymbol{\Sigma}_{x|w}\boldsymbol{\beta}_{[0]}/c^2)^{1/2}}\right\}.$$

The refined RC for this case can be obtained by solving the score equations based on the approximate likelihood above with $\text{E}(\boldsymbol{X}_i|\boldsymbol{W}_i)$ replaced by its estimate.

5 Proportional Hazards Regression as the Primary Model

We now extend the primary model (2) to the situation when the outcome variable may be censored. Let $(T_i, \delta_i), i = 1, \ldots, n$, be the observed times and noncensoring indicators for the

ith study subject, where T_i is the minimum of the failure time and the potential censoring time. We consider the Cox (1972) regression model $\lambda(t; \boldsymbol{X}_i) = \lambda_0(t)\exp(\boldsymbol{X}_i'\boldsymbol{\beta})$, where $\lambda_0(\cdot) \geq 0$ is an unspecified baseline hazard function and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)$. Again, covariate vector $\boldsymbol{X}_i$ consists of parameters from a random effects model from longitudinal measurements, and hence they are time-independent. We may extend the hazard function above to have $\boldsymbol{\alpha}'\boldsymbol{Z}_i$ for other observable time-dependent covariate $\boldsymbol{Z}_i$. The key point about the modeling of this problem is that the relative risk at time t of the disease onset is determined by some unobservable quantities, such as the initial level and the rate of change, of certain longitudinal measurements, but not solely determined by a single measurement.

Prentice (1982) proposed the idea of RC when covariates are mismeasured, and Wang, et al. (1997) studied this method under a missing data framework, where only part of the X_i (a validation set) were available. See also Hu, et al. (1998). In our problem, the main difference compared to the measurement error model literature in Cox regression is that in addition to unobservable intercept X_{i1}, random slopes $(X_{i2}, \ldots, X_{iq})'$ in (1) are unobservable covariates also. Wulfsohn and Tsiatis (1997) investigated a related problem.

Let $r(\boldsymbol{X}_i, \boldsymbol{\beta}) = \exp(\boldsymbol{X}_i'\boldsymbol{\beta})$ be the relative risk function given covariates $\boldsymbol{X}_i$. The RC estimator of $\boldsymbol{\beta}$ solves

$$\sum_{i=1}^{n} \delta_i \left\{ \mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i) - \frac{\boldsymbol{S}^{(1)}(\boldsymbol{\beta}, T_i)}{S^{(0)}(\boldsymbol{\beta}, T_i)} \right\} = 0, \tag{6}$$

where $\boldsymbol{S}^{(m)}(\boldsymbol{\beta}, T_i) = n^{-1}\sum_{j=1}^{n} Y_j(T_i)\{\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i)\}^m r\{\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i), \boldsymbol{\beta}\}$, $m = 0, 1$, and $Y(t) = I[T \geq t]$ is an at risk indicator. Similar to the argument in Wang, et al. (1997), the RC estimator converges to some $\boldsymbol{\beta}_*$ which is a root of $\mathrm{E}\{\sum_{i=1}^{n} \boldsymbol{\Psi}_P(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\beta}, \boldsymbol{\eta})\} = 0$, where $\boldsymbol{\Psi}_P(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\beta}, \boldsymbol{\eta})$ is the score for subject i in (6). Similar to the standard regression problems, the asymptotic bias, $\boldsymbol{\beta}_* - \boldsymbol{\beta}$, of the RC estimator in the Cox regression could be enhanced by a large magnitude of the relative risk, and further amplified by large $\boldsymbol{\Sigma}_{x|w}$ and low censorship. Note that inference based on RC is usually valid when the relative risk of interest is not large, especially in the case of testing a null hypothesis of the existence of an exposure effect. However, one should be cautious under some extreme relative risk cases.

Let $N_i(t) = I[T_i \le t, \delta_i = 1]$ be the failure time counting process. Similar to Wang, et al., under some regular conditions $M_i(t) = N_i(t) - \int_0^t Y_i(u)\lambda_0(u) r\{\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i), \boldsymbol{\beta}\}du$, $i = 1, \ldots, n$, can be shown to be independent local square integrable martingales on [0,1] (Andersen and Gill, 1982). Let $s^{(0)}(\boldsymbol{\beta}, t) = \mathrm{E}\{S^{(0)}(\boldsymbol{\beta}, t)\}$ and $\boldsymbol{s}^{(1)}(\boldsymbol{\beta}, t) = \mathrm{E}\{\boldsymbol{S}^{(1)}(\boldsymbol{\beta}, t)\}$. Let $\boldsymbol{\Theta} = (\boldsymbol{\beta}', \boldsymbol{\eta}')'$. To write the estimating scores as an iid sum, it can be shown that

$$\boldsymbol{\Psi}_P(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\Theta}) = \int \left\{ \mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i) - \frac{\boldsymbol{s}^{(1)}(\boldsymbol{\beta}, u)}{s^{(0)}(\boldsymbol{\beta}, u)} \right\} dM_i(u) + o_p(n^{-1/2}).$$

Now consider covariance estimation. Let $\boldsymbol{\Psi}_N(\boldsymbol{W}_i, \boldsymbol{\eta})$ be an unbiased score for $\boldsymbol{\eta}$. Let $\boldsymbol{\Theta}_* = (\boldsymbol{\beta}_*', \boldsymbol{\eta}')'$, $a^{\otimes 2} = aa'$ for any vector a, $\boldsymbol{\Psi}(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\Theta}) = [\{\boldsymbol{\Psi}_P(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\Theta})\}', \{\boldsymbol{\Psi}_N(\boldsymbol{W}_i, \boldsymbol{\eta})\}']'$, $\boldsymbol{G}_n(\boldsymbol{\Theta}) = n^{-1}\sum_{i=1}^n (\partial/\partial\boldsymbol{\Theta})\boldsymbol{\Psi}(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\Theta})$. Then $n^{1/2}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_*)$ is asymptotically normally distributed with mean 0 and variance $\boldsymbol{G}_n^{-1}(\boldsymbol{\Theta}_*)V(\boldsymbol{\Theta}_*)\{\boldsymbol{G}_n^{-1}(\boldsymbol{\Theta}_*)\}'$, where $V(\boldsymbol{\Theta}) = n^{-1}\sum_{i=1}^n \boldsymbol{\Psi}(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\Theta})^{\otimes 2}$. Let $\widehat{\boldsymbol{\Psi}}$ be $\boldsymbol{\Psi}$ but with $\boldsymbol{\Psi}_P$ replaced by

$$\begin{aligned}\widehat{\boldsymbol{\Psi}}_P(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\beta}, \boldsymbol{\eta}) &= \delta_i \left\{ \mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i) - \frac{\boldsymbol{S}^{(1)}(\boldsymbol{\beta}, T_i)}{S^{(0)}(\boldsymbol{\beta}, T_i)} \right\} \\ &\quad - n^{-1}\sum_{j=1}^n \frac{\delta_j Y_i(T_j) r\{\mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i), \boldsymbol{\beta}\}}{S^{(0)}(\boldsymbol{\beta}, T_j)} \left\{ \mathrm{E}(\boldsymbol{X}_i|\boldsymbol{W}_i) - \frac{\boldsymbol{S}^{(1)}(\boldsymbol{\beta}, T_j)}{S^{(0)}(\boldsymbol{\beta}, T_j)} \right\} + o_p(n^{-1/2}).\end{aligned}$$

Similar to Wang (1999), it can be shown that $V(\boldsymbol{\Theta}) = \widehat{V}(\boldsymbol{\Theta}) + o_p(1)$, where $\widehat{V}(\boldsymbol{\Theta}) = n^{-1}\sum_{i=1}^n \widehat{\boldsymbol{\Psi}}(\boldsymbol{W}_i, T_i, \delta_i, \boldsymbol{\Theta})^{\otimes 2}$. The covariance matrix for $n^{1/2}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_*)$ can be consistently estimated by $\boldsymbol{G}_n^{-1}(\widehat{\boldsymbol{\Theta}})\widehat{V}(\widehat{\boldsymbol{\Theta}})\{\boldsymbol{G}_n^{-1}(\widehat{\boldsymbol{\Theta}})\}'$.

6 Simulation Studies

In this section we report some results of Monte-Carlo simulation studies on finite sample performance of the proposed estimators. In Table 1, Y_i given $\boldsymbol{X}_i$ is logistic. In Table 2, (T_i, δ_i) given $\boldsymbol{X}_i$ follows Cox regression. We first generated $\boldsymbol{X} = (X_1, X_2)'$ from a bivariate normal distribution with mean $\boldsymbol{\mu}_x = (0,0)'$, $\text{var}(X_1) = 1$, $\text{var}(X_2) = 0.25$, $\text{cor}(X_1, X_2) = -0.1$. Four replicates of W_{ij} were simulated from the model $W_{ij} = X_{i1} + X_{i2}t_{ij} + U_{ij}$, where t_{ij} were normally distributed with mean $j-1$, standard deviation 0.1. The choice of t_{ij} is analogous to the situation when a subject has 4 routine annual clinic visits, and each visit is allowed in a time window. We considered a random error process with $\boldsymbol{\Sigma}_{ui} = \sigma_u^2 \mathbf{I}_4$, and applied the method of moments to estimate $\boldsymbol{\eta}$. The true parameters and sample sizes n used in each simulation study are shown in the corresponding tables. In the tables, the "bias" is the average of $(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$'s from 500 runs, "s.d." denotes the sample standard deviation of the estimators, "mean(s.e.)" denotes the average of the estimated standard errors of the estimators. The 95% confidence interval coverage probabilities are included also.

Table 1 presents various estimates for (β_1, β_2) when Y_i given $\boldsymbol{X}_i$ is logistic. Biases from the naive estimator for β_1 are quite large. As described earlier, when the relative risk is large the RC estimator may have a bias problem and the refined RC estimator provides a better solution. When the relative risk and measurement error are moderate ($\beta_1 = \ln(2)$, $\beta_2 = \ln(1.5)$, $\sigma_u = 1$), it is seen that the RC estimator has limited bias. Larger measurement error ($\sigma_u = 1.414$) increases the bias. For the case with a large relative risk and $n = 500$, the RC estimator has fairly substantial biases, and subsequently lower coverage probabilities. The refined RC estimator performs similarly to the pseudo EEE estimator. Note that if we adjust β_0 such that the disease rate is high (say $\beta_0 = \ln(4)$), then the bias of the RC estimator will increase.

Table 2 studies the Cox regression model with the baseline hazard $\lambda(t; \boldsymbol{X}_i) = \exp(\boldsymbol{X}_i'\boldsymbol{\beta})$. The underlying covariates $\boldsymbol{X}_i$ and longitudinal measurements W_{ij} were generated similarly to those in Table 1. We used a common censoring time, such that the censoring rate was .5 and .1, respectively. It is seen that the naive estimator has a bias problem, especially for β_1. The RC estimator performs well when the relative risk is moderate and the censoring rate is

0.5, although the bias for β_2 is larger when $\sigma_u = 1.414$. Under moderate risk parameters such as $\boldsymbol{\beta}' = \{\ln(2), \ln(2)\}$, the RC biases were the largest when $\sigma_u = 1.414$ and the censoring rate was 0.1. Nevertheless, the most sensitive factor to the RC bias is the larger relative risk under $\boldsymbol{\beta}' = \{\ln(5), \ln(5)\}$.

7 Data Analysis

We illustrate the methods using data from the longitudinal study of childhood growth, described in the introduction. There were 327 subjects included in the analysis, in which each subject had at least three measurements of BMI z-score (BMI-z) between ages 2.75 and 5.25 years. The age interval (2.75, 5.25) was used to include the routine health visits for ages 3 and 5. We consider logistic regression and the outcome variable of interest is adult obesity. Subjects were categorized as obese or non-obese as adults using their average BMI between 21 and 29 years of age. The objective of the analysis here is to determine the extent to which the initial BMI-z value at age 3 and the rate of change of the BMI-z values from ages 3-5 are predictive of adult obesity. Fifty two of these 327 subjects have been classified as obese.

For subject i, let W_{ij} denote the BMI-z measurements which are available at times $\{t_{ij}, j = 1, \cdots, k_i\}$ in the interval 3 to 5 years of age. Figure 1 displays the longitudinal BMI-z values from a simple random sample of 30 subjects. Least squares lines were also given. Overall, it seems reasonable to assume approximate linear trend of $W_i(t)$ in t for these subjects.

We shifted the time axis so that time 0 denotes age 3, and thus the intercept is the mean effect at age 3. We assume that the data follow a logistic regression model with covariates $\boldsymbol{X} = (X_1, X_2)'$, where X_1 is the underlying initial BMI-z value at age 3 and X_2 is the rate of change of a subject's BMI-z values over time. The objective of our study differs from that of a classical measurement error analysis mainly in that we are interested in the effects of the initial BMI-z value at age 3 as well as the rate of change of the longitudinal BMI-z values over a subject's childhood, rather than the effect of the overall average BMI-z, on adult obesity.

The naive estimates for $(\beta_0, \beta_1, \beta_2)$ were (-2.012, 1.064, 2.722) with s.e.'s (0.190, 0.209, 0.527). The RC estimates were (-2.084, 1.124, 5.012) with s.e.'s (0.202, 0.216, 1.007). The refined RC estimates were (-2.271, 1.223, 5.465) with s.e.'s (0.281, 0.248, 1.338). The pseudo-EEE estimates were (-2.266, 1.190, 5.186) with s.e.'s (0.282, 0.253, 1.533). The nuisance parameter estimates solving $\boldsymbol{\Psi}_N(\boldsymbol{W}, \boldsymbol{\eta}) = \mathbf{0}$ were $\widehat{\boldsymbol{\mu}}_x = (-0.035, 0.031)'$, $\widehat{\text{var}}(X_1) = 0.73$, $\widehat{\text{var}}(X_2) = 0.05$, $\widehat{\text{cor}}(X_1, X_2) = -0.23$ and $\widehat{\sigma}_u^2 = 0.12$. From the naive estimates, a child's BMI-z value at age 3 had an odds ratio (e^{β_1}) of approximately 2.9; the rate of change had an odds ratio (e^{β_2}) of approximately 15. From the RC, the refined RC and the pseudo-EEE, the BMI-z value at age 3 had an odds ratio of approximately 3.1 to 3.4; the rate of change had an odds ratio of approximately 150-236. In summary, both the initial BMI-z value at age 3 and the rate of change appear to have significant effects on adult obesity.

8 Extensions

In this paper, we have developed new approaches for modeling covariates from longitudinal measurements, in relation to an outcome variable. To our knowledge, this model is novel and applicable to medical and epidemiological studies. The pseudo EEE, RC and refined RC estimators are easy to implement. These methods, however, depend on the multivariate-normal assumptions on covariates and longitudinal measurements. The use of functional methods (Carroll et al., 1995, Chapter 6) to this problem may avoid the normality assumption.

When the primary model is Cox regression, estimation methods other than RC may be technically complicated. However, a slightly improved method may be obtained by adopting the idea of a risk set RC of Clayton (1991) to this more general set up. In a classical measurement error problem, Xie, et al. (1999) showed that this type of risk set RC estimator is promising when $|\beta_1|$, the coefficient of a mismeasured covariate X, is not larger than say $\ln(4)$ while all covariates are standardized. In our more complicated model, some preliminary results show similar conclusions, although the covariance estimation in this particular setup is somewhat difficult. Further development of other approaches, such as the nonparametric maximum likelihood estimation, is also a challenging topic for future research.

Acknowledgments

This research was supported by US National Institutes of Health grants: AG 15026, CA 53996, HL 57288 (C.Y. Wang), CA 74552 (N. Wang), CA 57030 (S. Wang), P30-E509106 to TAMU Center for Environmental and Rural Health (N. Wang, S. Wang); and by the TAMU IRI Program (S. Wang). The authors are thankful to M. Pepe and B. Whitaker for the data on child growth.

References

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The Annals of Statistics* **10**, 1100-1120.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*, Chapman and Hall, London.

Carroll, R. J., Spiegelman, C., Lan, K. K., Bailey, K. T. and Abbott, R. D. (1984). On errors-in-variables for binary regression models. *Biometrika*, **71**, 19-26.

Clayton, D. G. (1991). Models for the analysis of cohort and case–control studies with inaccurately measured exposure. In: *Statistical Models for Longitudinal Studies of Health.* Dwyer, J. H., Feinleib, M., Lipsert, P., et al. eds., pp. 301–333. Oxford University Press, New York.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, Ser. B **34**, 187-220.

Fuller, W. A. (1987). *Measurement Error Models*, John Wiley & Sons, New York.

Higgins, K. M., Davidian, M and Giltinan, D. M. (1997). A two-step approach to measurement error in time-dependent covariates in nonlinear mixed-effects models, with application to IGF-I Pharmacokinetics. *Journal of the American Statistical Association* **92**, 436-448.

Hu, P., Tsiatis, A. A. and Davidian, M. (1998). Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics* **54**, 1407-1419.

Liang, K. Y. and Liu, X. H. (1991). Estimating equations in generalized linear models with measurement error. In *Estimating Functions*, ed. V. P. Godambe, pp 47–63. Clarendon Press, Oxford.

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

Wang, C. Y. (1999). Robust sandwich covariance estimation for regression calibration estimator in Cox regression with measurement error. *Statistics and Probability Letters*, in press.

Wang, C. Y. and Pepe, M. S. (1999). Expected estimating equations to accommodate covariate measurement error. *Journal of the Royal Statistical Society*, Ser. B, conditionally accepted.

Wang, C. Y., Hsu, L., Feng, Z. D. and Prentice, R. L. (1997). Regression calibration in failure time regression with surrogate variables. *Biometrics*, **53**, 131–145.

Wang, N., Carroll, R. J. and Liang, K. Y. (1996). Quasilikelihood estimation in measurement error models with correlated replicates. *Biometrics*, **52**, 401–411.

Whitaker, R. C., Wright, J. A., Pepe, M. S., Seidel, K. D. and Dietz, W. H. (1997). Predicting adult obesity from childhood and parent obesity. *New England Journal of Medicine*, **13**, 869–873.

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

Xie, S. X., Wang, C. Y., Prentice, R. L. (1999). A risk set calibration method for failure time regression using a covariate reliability sample. Preprint.

Zucker, D. M., Zerbe, G. O. and Wu, M. C. (1995). Inference for the association between coefficients in a multivariate growth curve model. *Biometrics*, **413**, 413-424.

Appendix: Technical Proofs

Proof of Proposition 1:

First, the EEE is the same as the score obtained from the likelihood of Y_i given $\boldsymbol{W}_i$ since

$$\frac{\partial}{\partial \boldsymbol{\beta}} \ln\{P(Y_i|\boldsymbol{W}_i)\} = \frac{(\partial/\partial\boldsymbol{\beta}) \int P(Y_i|\boldsymbol{x})P(\boldsymbol{W}_i|\boldsymbol{x})P(\boldsymbol{x})d\boldsymbol{x}}{P(Y_i,\boldsymbol{W}_i)} = \mathrm{E}\{\boldsymbol{\psi}(Y_i,\boldsymbol{X}_i,\boldsymbol{\beta})|Y_i,\boldsymbol{W}_i\}.$$

If $Y_i|\boldsymbol{W}_i$ is normal, then the log-likelihood is $\ln\{P(Y_i|\boldsymbol{W}_i)\} = (-.5)\ln(2\pi) - .5\ln\{\mathrm{var}(Y_i|\boldsymbol{W}_i)\}$ $-.5\{Y_i - \mathrm{E}(\boldsymbol{\mathcal{X}}_i'|\boldsymbol{W}_i)\boldsymbol{\beta}\}^2/\mathrm{var}(Y_i|\boldsymbol{W}_i)$. Since β_0 is not a function of $\mathrm{var}(Y_i|\boldsymbol{W}_i)$, we have $(\partial/\partial\beta_0)\ln\{P(Y_i|\boldsymbol{W}_i)\} = \{Y_i - \mathrm{E}(\boldsymbol{\mathcal{X}}_i'|\boldsymbol{W}_i)\boldsymbol{\beta}\}/\mathrm{var}(Y_i|\boldsymbol{W}_i)$. Also, for $\ell = 1,\cdots,q$,

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{\beta}_\ell}\ln\{P(Y_i|\boldsymbol{W}_i)\} = {} & \frac{\mathrm{E}(X_{i\ell}|\boldsymbol{W}_i)\{Y_i - \mathrm{E}(\boldsymbol{\mathcal{X}}_i'|\boldsymbol{W}_i)\boldsymbol{\beta}\}}{\mathrm{var}(Y_i|\boldsymbol{W}_i)} \\ & + \frac{\{(\partial/\partial\boldsymbol{\beta}_\ell)\mathrm{var}(Y_i|\boldsymbol{W}_i)\}\left[\{Y_i - \mathrm{E}(\boldsymbol{\mathcal{X}}_i'|\boldsymbol{W}_i)\boldsymbol{\beta}\}^2 - \mathrm{var}(Y_i|\boldsymbol{W}_i)\right]}{2\{\mathrm{var}(Y_i|\boldsymbol{W}_i)\}^2}.\end{aligned}$$

Note that $\mathrm{var}(Y_i|\boldsymbol{W}_i) = \sigma_e^2 + \boldsymbol{\beta}(\boldsymbol{\Sigma}_x\boldsymbol{D}_i')(\boldsymbol{D}_i\boldsymbol{\Sigma}_x\boldsymbol{D}_i' + \boldsymbol{\Sigma}_{ui})^{-1}\boldsymbol{D}_i\boldsymbol{\Sigma}_x\boldsymbol{\beta}'$. Hence, if the estimate of σ_e^2 satisfies (4), then the second term in $\sum_{i=1}^n(\partial/\partial\boldsymbol{\beta}_\ell)\ln\{P(Y_i|\boldsymbol{W}_i)\}$, $\ell \geq 1$, vanishes, provided that $\text{var}(Y_i|\boldsymbol{W}_i)$ does not depend on i. This sufficient condition is fulfilled when $\boldsymbol{t}_i = \boldsymbol{t}_1$, *i.e.* $\boldsymbol{D}_i = \boldsymbol{D}_1$. In this case solving $\sum_{i=1}^n (\partial/\partial\boldsymbol{\beta})\ln\{P(Y_i|\boldsymbol{W}_i)\} = 0$, the expected estimating equations, is equivalent to solving $\sum_{i=1}^n \boldsymbol{\psi}\{Y_i, \mathrm{E}(\boldsymbol{\mathcal{X}}_i|\boldsymbol{W}_i), \boldsymbol{\beta}\} = 0$.

Table 1. Simulation results when Y given (X_1, X_2) is logistic

		$n=200$				$n=500$			
		Naive	RC	Refined RC	Pseudo EEE	Naive	RC	Refined RC	Pseudo EEE
$\boldsymbol{\beta}=(0,\ln(2),\ln(1.5))$, $\sigma_u=1$									
β_1	bias	-0.177	0.032	0.060	0.069	-0.198	-0.014	0.003	0.009
	s.d.	0.138	0.231	0.262	0.271	0.087	0.138	0.150	0.154
	mean(s.e.)	0.136	0.221	0.247	0.253	0.084	0.131	0.142	0.144
	95% cov.	0.714	0.948	0.960	0.960	0.328	0.934	0.944	0.944
β_2	bias	0.141	-0.037	-0.032	-0.031	0.123	-0.017	-0.010	-0.007
	s.d.	0.255	0.459	0.477	0.483	0.153	0.253	0.258	0.259
	mean(s.e.)	0.251	0.432	0.448	0.453	0.156	0.255	0.259	0.261
	95% cov.	0.928	0.952	0.960	0.966	0.878	0.966	0.966	0.966
$\boldsymbol{\beta}=(0,\ln(2),\ln(1.5))$, $\sigma_u=1.414$									
β_1	bias	-0.269	0.045	0.103	0.124	-0.286	-0.015	0.013	0.022
	s.d.	0.122	0.308	0.418	0.469	0.076	0.169	0.193	0.202
	mean(s.e.)	0.121	0.315	0.408	0.443	0.075	0.163	0.185	0.190
	95% cov.	0.396	0.960	0.968	0.968	0.054	0.932	0.954	0.960
β_2	bias	0.123	-0.072	-0.082	-0.089	0.100	-0.035	-0.026	-0.023
	s.d.	0.228	0.640	0.751	0.809	0.137	0.322	0.334	0.338
	mean(s.e.)	0.223	0.641	0.741	0.782	0.139	0.326	0.338	0.341
	95% cov.	0.918	0.966	0.976	0.980	0.890	0.976	0.980	0.982
$\boldsymbol{\beta}=(0,\ln(4),\ln(3))$, $\sigma_u=1$									
β_1	bias	-0.389	-0.029	0.113	0.143	-0.415	-0.091	0.010	0.030
	s.d.	0.173	0.285	0.413	0.437	0.108	0.169	0.216	0.222
	mean(s.e.)	0.172	0.273	0.380	0.393	0.107	0.163	0.209	0.214
	95% cov.	0.384	0.926	0.978	0.980	0.048	0.870	0.934	0.932
β_2	bias	-0.101	-0.118	-0.041	-0.028	0.080	-0.087	-0.014	0.001
	s.d.	0.298	0.517	0.564	0.576	0.187	0.298	0.313	0.315
	mean(s.e.)	0.295	0.489	0.532	0.540	0.183	0.288	0.304	0.306
	95% cov.	0.944	0.956	0.960	0.962	0.928	0.936	0.954	0.948

NOTE: Nuisance parameters are $\boldsymbol{\mu}_x = (0,0)'$, $\text{var}(X_1) = 1$, $\text{var}(X_2) = 0.25$, $\text{corr}(X_1, X_2) = -0.1$. For $j = 1, \ldots, 4$, t_{ij} were normally distributed with mean $j-1$, standard deviation 0.1. The pseudo EEE were obtained by using a usual Riemann integral from 50 partitions for X_1 and X_2 axes, respectively. The results were from 500 runs.

Table 2. Simulation results in Cox regression

		$\sigma_u = 1$				$\sigma_u = 1.414$			
		$n = 200$		$n = 400$		$n = 200$		$n = 400$	
		Naive	RC	Naive	RC	Naive	RC	Naive	RC
$\boldsymbol{\beta} = \{\ln(2), \ln(2)\}$, censoring rate = 0.5									
β_1	bias	-0.163	0.009	0.172	-0.013	-0.253	0.030	0.260	-0.013
	s.d.	0.096	0.163	0.064	0.101	0.086	0.298	0.058	0.140
	mean(s.e.)	0.091	0.148	0.064	0.101	0.081	0.263	0.057	0.137
	95% cov.	0.560	0.948	0.220	0.942	0.150	0.952	0.010	0.946
β_2	bias	0.016	-0.042	0.009	-0.032	-0.053	-0.113	-0.058	-0.062
	s.d.	0.177	0.317	0.121	0.196	0.159	0.534	0.108	0.266
	mean(s.e.)	0.172	0.295	0.120	0.198	0.153	0.521	0.107	0.270
	95% cov.	0.942	0.950	0.952	0.946	0.920	0.972	0.914	0.954
$\boldsymbol{\beta} = \{\ln(2), \ln(2)\}$, censoring rate = 0.1									
β_1	bias	-0.176	-0.010	-0.187	-0.031	-0.270	-0.010	-0.279	-0.041
	s.d.	0.074	0.124	0.050	0.080	0.065	0.152	0.045	0.111
	mean(s.e.)	0.071	0.117	0.050	0.079	0.063	0.231	0.044	0.111
	95% cov.	0.314	0.938	0.046	0.912	0.024	0.928	0.000	0.874
β_2	bias	-0.003	-0.058	-0.013	-0.055	-0.077	-0.122	-0.088	-0.098
	s.d.	0.137	0.242	0.093	0.151	0.122	0.469	0.085	0.209
	mean(s.e.)	0.131	0.223	0.092	0.150	0.117	0.456	0.082	0.211
	95% cov.	0.938	0.944	0.938	0.940	0.880	0.972	0.798	0.954
$\boldsymbol{\beta} = \{\ln(4), \ln(4)\}$, censoring rate = 0.5									
β_1	bias	-0.426	-0.121	-0.448	-0.161	-0.623	-0.151	-0.640	-0.214
	s.d.	0.120	0.199	0.075	0.120	0.103	0.410	0.065	0.175
	mean(s.e.)	0.106	0.175	0.073	0.120	0.091	0.366	0.063	0.169
	95% cov.	0.054	0.832	0.000	0.660	0.000	0.764	0.000	0.606
β_2	bias	-0.093	-0.183	-0.120	-0.193	-0.270	-0.344	-0.292	-0.300
	s.d.	0.203	0.345	0.138	0.222	0.181	0.758	0.122	0.317
	mean(s.e.)	0.187	0.317	0.130	0.214	0.165	0.708	0.115	0.307
	95% cov.	0.882	0.922	0.820	0.856	0.600	0.960	0.266	0.894

NOTE: Parameters are $\boldsymbol{\mu}_x = (0, 0)'$, $\text{var}(X_1) = 1$, $\text{var}(X_2) = 0.25$, $\text{corr}(X_1, X_2) = -0.1$. For $j = 1, \ldots, 4$, t_{ij} were normally distributed with mean $j - 1$, standard deviation 0.1. The results were from 500 runs.

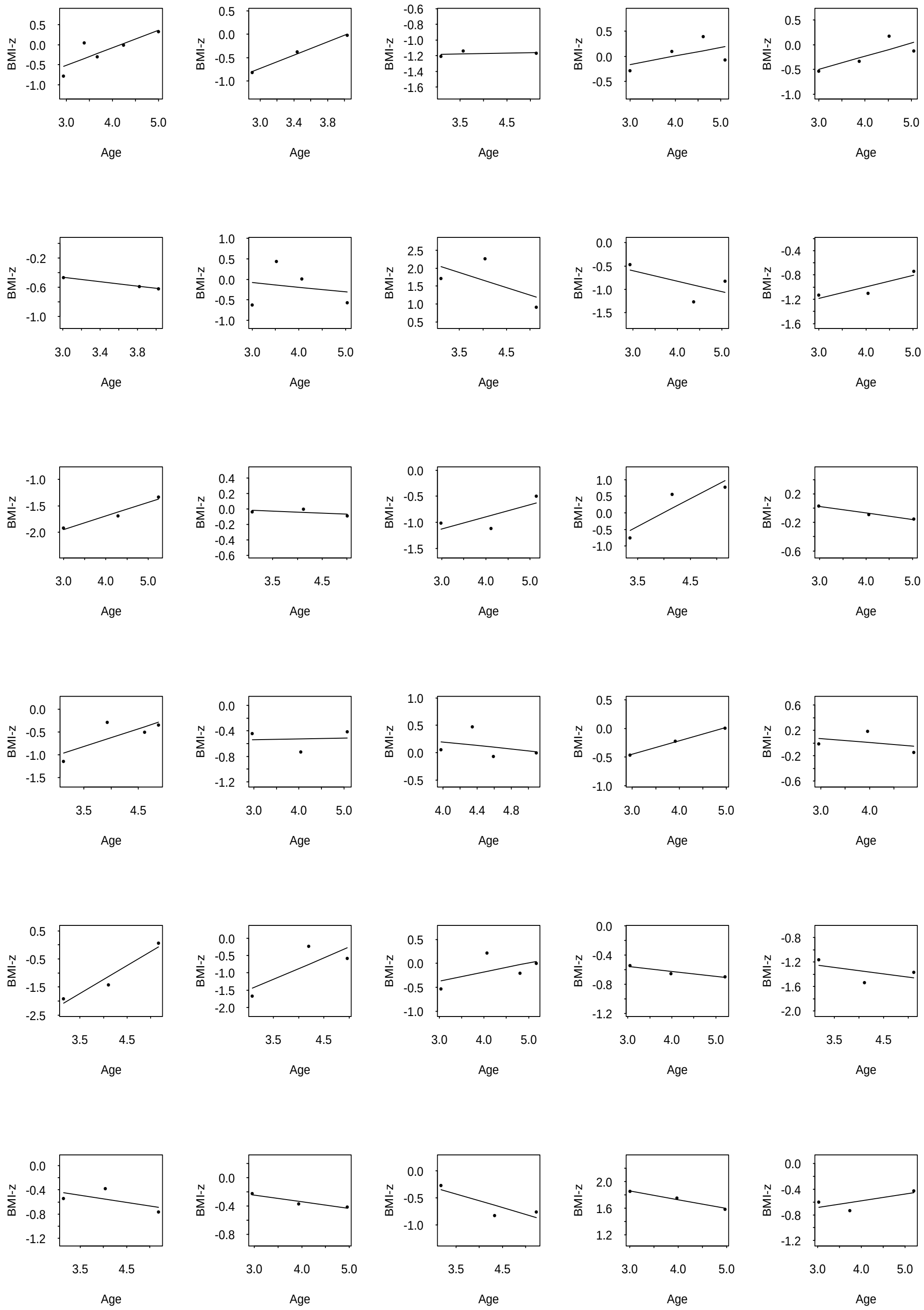

Figure 1: *Longitudinal BMI-z values for a simple random sample of 30 subjects*