

On Estimating the Relationship between Longitudinal Measurements and Time-to-Event Data Using a Simple Two-Stage Procedure

Paul S. Albert and Joanna H. Shih
Biometric Research Branch
Division of Cancer Treatment and Diagnosis
National Cancer Institute
Bethesda, Maryland U.S.A. 20892

January 30, 2009

SUMMARY

Ye et al. (2008) proposed a joint model for longitudinal measurements and time-to-event data in which the longitudinal measurements are modeled with a semiparametric mixed model to allow for the complex patterns in longitudinal biomarker data. They proposed a two-stage regression calibration approach which is simpler to implement than a joint modeling approach. In the first stage of their approach, the mixed model is fit without regard to the time-to-event data. In the second stage, the posterior expectation of an individual's random effects from the mixed-model are included as covariates in a Cox model. Although Ye et al. (2008) acknowledged that their regression calibration approach may cause bias due to the problem of informative dropout and measurement error, they argued that the bias is small relative to alternative methods. In this article, we show that this bias may be substantial. We show how to alleviate much of this bias with an alternative regression calibration approach which can be applied for both discrete and continuous time-to-event data. Through simulations, the proposed approach is shown to have substantially less bias than the regression calibration approach proposed by Ye et al. (2008). In agreement with the methodology proposed by Ye et al., an advantage of our proposed approach over joint modeling is that it can be implemented with standard statistical software and does not require complex estimation techniques.

1 Introduction

Ye et al. (2008) proposed a two-stage regression calibration approach for estimating the relationship between longitudinal measurements and time-to-event data. Their approach was motivated by trying to establish such a relationship when the longitudinal measurements follow a complex semi-parametric mixed model with subject-specific random stochastic processes and the time-to-event data follow a proportional hazards model. Specifically, they proposed a semi-parametric model with additive errors for the longitudinal measurements X_{ij} of the form

$$\begin{aligned} X_{ij} &= \mathbf{Z}'_i \boldsymbol{\beta} + \varphi(t_{ij}) + \mathbf{U}_i(t_{ij}) \mathbf{b}_i + W_i(t_{ij}) + \epsilon_{ij} \\ &= X_i^*(t_{ij}) + \epsilon_{ij}, \end{aligned} \tag{1}$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients associated with fixed effect covariates \mathbf{Z}_i , $\varphi(t)$ is an unknown smooth function over time, \mathbf{b}_i is a vector of subject-specific random effects corresponding to covariates $\mathbf{U}_i(t)$ which is assumed normally distributed with mean $\mathbf{0}$ and variance $\Sigma_{\mathbf{b}}$. Further, $W_i(t_{ij})$ is a zero mean integrated Wiener stochastic process. We denote \mathbf{X}_i as all longitudinal measurements on the i th individual.

In Ye et al.'s approach, the relationship between the slope of the longitudinal process and a time-to-event outcome T_i is characterized by a Cox proportional hazards model with the slope at time t , denoted as $X_i^{*'}(t)$, being treated as a time-dependent covariate. The authors proposed a two-stage estimation procedure in which in the first stage, the mean of the posterior distribution of the slope at time t , $E[X_i^{*'}(t)|\mathbf{X}_i, \mathbf{Z}_i]$, is estimated using model (1) without regard to the time-to-event process T_i . In the second stage, $E[X_i^{*'}(t)|\mathbf{X}_i, \mathbf{Z}_i]$ replaces $X_i^{*'}(t)$ in the Cox model. Ye et al. (2008) proposed two approaches: (i) the ordinary regression calibration (ORC) approach in which $E[X_i^{*'}(t)|\mathbf{X}_i, \mathbf{Z}_i]$ is estimated using (1) with all available longitudinal measurements and (ii) the risk set regression calibration (RRC) approach in which these expectations are obtained by estimating model (1) after each event using only longitudinal measurements for subjects at risk at time t (i.e., subjects who have

an event before time t are removed from the estimation).

The advantage of these regression calibration approaches are that they do not require the complex joint modeling of the longitudinal and time-to-event processes. In the discussion of their paper, Ye et al. acknowledge that these approaches may result in biased estimation due to informative dropout and measurement error, and that improved performance will require incorporating informative dropout and the uncertainty of measurement error into the estimation. In this article we show that an alternative two-stage procedure can be formulated which reduces the bias considerably without requiring complex joint modeling of both processes. For simplicity, we develop the approach for a longitudinal model without the smooth function $\varphi(t)$ and the stochastic component $W_i(t)$ in (1), but the proposed approach applies more generally. In this approach, we approximate the conditional distribution of the longitudinal process given the event time, simulate complete follow-up data based on the approximate conditional model, and then fit the longitudinal model with complete follow-up on each patient (hence avoiding the problem of informative dropout in Ye et al.'s approach). Section 2 develops the approach for a discrete event time distribution followed by an approximation for the continuous event time distribution. The results of simulations which show the advantages of the proposed approach over ORC and RRC are provided in Section 3. A discussion follows in Section 4.

2 Modeling Framework

We begin by considering a discrete event time distribution. Define T_i to be a discrete event time which can take on discrete values t_j , $j = 1, 2, \dots, J$, and Y_{ij} to be a binary indicator of whether the i th patient is dead at time t_j . Then $J_i = \sum_{j=1}^J (1 - Y_{ij}) = J - Y_i$ where $Y_i = \sum_{j=1}^J Y_{ij}$ indicates the number of follow-up measurements before death or administrative censoring for the i th patient. Every patient will be followed until death or the end of follow-up at time t_J .

For illustrative purposes, we will consider a joint model for longitudinal and discrete time-to-event data in which the discrete event time distribution is modeled as a linear function of

the slope of an individual's longitudinal process on the probit scale. Specifically,

$$P(Y_{ij} = 1 | Y_{i(j-1)} = 0) = \Phi(\alpha_{0j} + \alpha_1 b_{i1}), \quad (2)$$

where $j = 1, 2, \dots, J$, Y_{i0} is taken as 0, α_{0j} governs the baseline discrete event time distribution and b_{i1} is the individual slope from the linear mixed model,

$$X_{ij} = X_i^*(t_j) + \epsilon_{ij}, \quad (3)$$

$$X_i^*(t_j) = \beta_0 + \beta_1 t_j + b_{i0} + b_{i1} t_j, \quad (4)$$

where $i = 1, 2, \dots, I$ and $j = 1, 2, \dots, J_i$. In (4), the parameters β_0 and β_1 are fixed-effect parameters characterizing the mean intercept and slope of the longitudinal process, respectively, $(b_{i0}, b_{i1})'$ is a vector of random effects which are assumed multivariate normal with mean $\mathbf{0}$ and variance $\Sigma_{\mathbf{b}} = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0, b_1} \\ \sigma_{b_0, b_1} & \sigma_{b_1}^2 \end{pmatrix}$, and ϵ_{ij} is a residual error term which is assumed normal with mean zero and variance σ_ϵ^2 . In (2)-(4), the event time and the longitudinal process are linked through b_{i1} , and the parameter α_1 governs the relationship between the slope of the longitudinal process and the event time distribution. Denote $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{iJ_i})'$, $\mathbf{b}_i = (b_{i0}, b_{i1})'$, and $\boldsymbol{\beta} = (\beta_0, \beta_1)'$. As in Ye et al., the normality assumption for \mathbf{b}_i is made for these joint models. Although, not the focus of this article, various articles have proposed methods with flexible semi-parametric random effects distributions and have demonstrated that inferences are robust to departures from normality (Song et al., 2002; Hsieh et al., 2006).

For estimating the relationship between the slope of the longitudinal process and the time-to-event process, the calibration approach of Ye et al. (2008) reduces to first, estimating $E[b_{i1} | \mathbf{X}_i, \boldsymbol{\beta}]$ using (3) and (4), and second, replacing b_{i1} by $E[b_{i1} | \mathbf{X}_i, \widehat{\boldsymbol{\beta}}]$ in estimating (2). As recognized by Ye et al., this methodology introduces bias in two ways. First, there is the problem of informative dropout, whereby b_{i0} and b_{i1} can depend on the event time T_i (which will occur if $\alpha_1 \neq 0$ in (2)). Ignoring this informative dropout may result in substantial bias. Second, not accounting for the measurement error in $E[b_{i1} | \mathbf{X}_i, \widehat{\boldsymbol{\beta}}]$ relative to true values of b_{i1} will result in attenuated estimation of α_1 .

We propose a simple approach which reduces these two sources of bias. We first focus on the problem of informative dropout. The bias from informative dropout is a result of differential follow-up whereby the response process is related to the length of follow-up (i.e., in (2)-(4), when α_1 is positive, patients who die early are more likely to have large positive slopes). There would be no bias if all J follow-up measurements were observed on all patients. Thus, we recapture these missing measurements by generating data from the conditional distribution of \mathbf{X}_i given T_i , denoted as $\mathbf{X}_i|T_i$. Since $\mathbf{X}_i|T_i$ under (2)-(4) does not have a tractable form, we propose a simple approximation for this conditional distribution.

Under model (2)-(4), the distribution of $\mathbf{X}_i|T_i$ can be expressed as

$$P(\mathbf{X}_i|T_i) = \int h(\mathbf{X}_i|\mathbf{b}_i, T_i)g(\mathbf{b}_i|T_i)d\mathbf{b}_i. \quad (5)$$

Since T_i and the values of \mathbf{X}_i are conditionally independent given \mathbf{b}_i , $h(\mathbf{X}_i|\mathbf{b}_i, T_i) = h(\mathbf{X}_i|\mathbf{b}_i)$, where $h(\mathbf{X}_i|\mathbf{b}_i)$ is the product of J_i univariate normal density functions each with mean $X_i^*(t_j)$ ($j = 1, 2, \dots, J_i$) and variance σ_ϵ^2 . The distribution of $\mathbf{X}_i|T_i$ can easily be obtained with standard statistical software if we approximate $g(\mathbf{b}_i|T_i)$ by a normal distribution. Under the assumption that $g(\mathbf{b}_i|T_i)$ is normally distributed with mean $\boldsymbol{\mu}_{T_i} = (\mu_{0T_i}, \mu_{1T_i})'$ and variance $\Sigma_{bT_i}^*$, and by rearranging mean structure parameters in the integrand of (5) so that the random effects have mean zero, $\mathbf{X}_i|T_i$ corresponds to the following mixed model

$$X_{ij}|(T_i, b_{i0T_i}^*, b_{i1T_i}^*) = \beta_{0T_i}^* + \beta_{1T_i}^* t_j + b_{i0T_i}^* + b_{i1T_i}^* t_j + \epsilon_{ij}^*, \quad (6)$$

where $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J_i$, and the residuals ϵ_{ij}^* are assumed to have independent normal distributions with mean zero and variance σ_ϵ^{*2} . Further, the fixed-effects parameters $\beta_{0T_i}^*$ and $\beta_{1T_i}^*$ are intercept and slope parameters for patients who have an event at time T_i or who are censored at time $T_i = t_j$. In addition, the associated random effects $\mathbf{b}_{iT_i}^* = (b_{i0T_i}^*, b_{i1T_i}^*)'$ are multivariate normal with mean $\mathbf{0}$ and variance $\Sigma_{bT_i}^*$ for each T_i . Thus, this flexible conditional model involves estimating separate fixed effect intercept and slope parameters for each potential event-time and for subjects who are censored at time t_j .

Likewise, separate random effects distributions are estimated for each of these discrete time points. For example, the intercept and slope fixed-effect parameters for those patients who have an event at time $T_i = 3$ are β_{03}^* and β_{13}^* , respectively. In addition, the intercept and slope random effects for those patients who have an event at $T_i = 3$, $\mathbf{b}_{iT_i}^* = (b_{i03}^*, b_{i13}^*)'$, are multivariate normal with mean $\mathbf{0}$ and variance Σ_{b3}^* . Model (6) can be fit with standard R code which is available from the first author.

A similar approximation of the conditional distribution of the longitudinal process given dropout time has been proposed for estimating mean change over time in longitudinal measurements subject to informative dropout (Wu and Bailey, 1989; Wu and Follmann, 1999). In this article, we use the approximation to construct complete longitudinal datasets which in turn are used to estimate the mean of the posterior distribution of an individual's random effects given the data. Specifically, multiple complete longitudinal datasets can then be constructed by simulating X_{ij} values from (6) where the parameters are replaced by their estimated values. Since the simulated datasets have complete follow-up on each individual, the bias in estimating $E[b_{i1}|\mathbf{X}_i, \beta]$ caused by informative dropout is much reduced.

We provide a correction to account for the measurement error in using $E[b_{i1}|\mathbf{X}_i, \hat{\beta}]$, denoted as \hat{b}_{i1} , instead of using the actual random slope b_{i1} . As in Carroll et al. (1984) who adjust for measurement error in a covariate, we note that

$$\begin{aligned} P(Y_{ij} = 1 | Y_{i(j-1)} = 0, \mathbf{X}_i) &= \int \Phi(\alpha_{0j} + \alpha_1 b_{1i}) g(b_{1i} | \mathbf{X}_i) db_{1i} \\ &= \Phi\left(\frac{\alpha_{0j} + \alpha_1 \hat{b}_{i1}}{\sqrt{1 + \alpha_1^2 \text{Var}(\hat{b}_{i1} - b_{i1})}}\right), \end{aligned} \quad (7)$$

where $\text{Var}(\hat{b}_{i1} - b_{i1})$ measures the error of estimation in \hat{b}_{i1} relative to b_{i1} which is the the 1-1 element in the matrix $\text{Var}(\hat{\mathbf{b}}_i - \mathbf{b}_i)$, given by

$$\text{Var}(\hat{\mathbf{b}}_i - \mathbf{b}_i) = \Sigma_{\mathbf{b}} - \Sigma_{\mathbf{b}} \mathbf{R}_i' \left\{ \mathbf{W}_i - \mathbf{W}_i \mathbf{F}_i \mathbf{Q} \mathbf{F}_i' \mathbf{W}_i \right\} \mathbf{R}_i \Sigma_{\mathbf{b}},$$

where $\mathbf{Q} = \sum_{i=1}^I (\mathbf{F}_i' \mathbf{W}_i \mathbf{F}_i)^{-1}$, and where $\mathbf{W}_i = \mathbf{V}_i^{-1}$, where \mathbf{V}_i is the variance of \mathbf{X}_i (Laird and Ware, 1982; Verbeke and Molenberghs, 2000). Further, \mathbf{F}_i and \mathbf{R}_i are vectors of fixed and random effects for the i th subject. This variance formula incorporates the error in

estimating the fixed effects in the longitudinal model. Expression (7) follows from that fact that $E\left[\Phi(a + V)\right] = \Phi\left[(a + \mu)/\sqrt{1 + \tau^2}\right]$, where $V \sim N(\mu, \tau^2)$.

Only individuals who have at least two longitudinal measurements provide useful information in assessing the relationship between an individual's slope and their time-to-event data, so we assume that all individuals in the analysis have at least two follow-up times. Thus, $\alpha_{01} = \alpha_{02} = -\infty$ and the regression parameters in the discrete-time model α_{0j} ($j = 3, 4, \dots, J$) and α_1 can be estimated by maximizing the likelihood

$$L = \prod_{i=1}^I \left[\prod_{j=1}^{J_i} \{1 - P(Y_{ij} = 1 | Y_{i(j-1)} = 0, \mathbf{X}_i)\}^{(1-Y_{ij})} \right] P(Y_{i(J_i+1)} = 1 | Y_{iJ_i} = 0, \mathbf{X}_i)^{J_i < J}. \quad (8)$$

Thus, we propose the following algorithm for estimating α_{0j} ($j = 3, 4, \dots, J$) and α_1 with a two-stage procedure:

1. Estimate model (6) with all available longitudinal measurements using linear mixed-modeling software such as lme in R.
2. Simulate complete longitudinal pseudo measurements (i.e., X_{ij} for $i = 1, 2, \dots, I$ and $j = 1, 2, \dots, J$) using model (6) with model parameters estimated from step 1. Specifically, these measurements are simulated by first simulating values of $b_{iT_i}^*$ from a normal distribution with mean $\mathbf{0}$ and variance $\Sigma_{bT_i}^*$ and ϵ_{ij}^* from a normal distribution with mean 0 and variance σ_ϵ^{*2} , where the variance parameters are estimated in step 1.
3. Estimate model (3) and (4) (without regard to the event time distribution (2)) with complete longitudinal measurements simulated in step 2 using linear mixed modeling software.
4. Estimate α_{0j} ($j = 1, 2, \dots, J$) and α_1 (denoted as $\widehat{\alpha}_{0j}$ and $\widehat{\alpha}_1$, respectively) using (7) and (8) with \widehat{b}_{i1} obtained from step 3.
5. Repeat steps 2 to 4 M times and average $\widehat{\alpha}_{0j}$ and $\widehat{\alpha}_1$ to get final estimates.

The approach can be generalized for continuous event-time distributions where, T_i is the continuous event time for the i individual, all individuals are followed up to time T_E , and where patients are administratively censored at the end of the study when $T_i > T_E$. In addition, the Cox model, $\lambda(t, b_{i1}) = \lambda_0(t) \exp(\alpha b_{i1})$ is used to relate the longitudinal measurements to time-to-event data. We can approximate this conditional distribution by first discretizing the follow-up interval into K equally spaced intervals. We define d_i as a discretized version of the continuous event time distribution, whereby, $d_i = k$ when $T_i \in \left((k-1)T_E/K, kT_E/K \right]$, $k = 1, 2, \dots, K$, and where $d_i = K+1$ when patient i 's event time is administratively censored at time T_E . The conditional distribution of the longitudinal measurements given the continuous event time, $\mathbf{X}_i|T_i$, can be approximated by the distribution of the longitudinal measurements given the discretized version d_i , $\mathbf{X}_i|d_i$, where, as for the discrete event time model, this conditional distribution can be approximated by a linear mixed model

$$X_{ij}|(d_i, b_{i0d_i}^*, b_{i1d_i}^*) = \beta_{0d_i}^* + \beta_{1d_i}^* t_j + b_{i0d_i}^* + b_{i1d_i}^* t_j + \epsilon_{ij}^*, \quad (9)$$

where $i = 1, 2, \dots, I$ and $j = 1, 2, \dots, J_i$, and where J_i is the number of follow-up measurements before death or administrative censoring for the i th patient. Similar to (6), $\beta_{0d_i}^*$ and $\beta_{1d_i}^*$ are intercept and slope parameters for patients with a discretized event time of d_i . Also, $\mathbf{b}_{id_i}^* = (b_{i0d_i}^*, b_{i1d_i}^*)'$ are assumed to be normally distributed with mean $\mathbf{0}$ and variance $\Sigma_{bd_i}^*$. For continuous event times, we apply the previous algorithm for discrete-time data except that in Step 1 we fit model (9) for a reasonably large K , and in Step 3, we fit a Cox model without a measurement error correction instead of the discrete-time model.

Asymptotic standard errors from the discrete or continuous event time models cannot be used for inference since they fail to account for the missing data uncertainty in our procedure. The bootstrap (Efron and Tibshirani, 1993) can be used for valid standard error estimation.

3 Simulations

We evaluated the procedure for both discrete and continuous time to event data with a simulation study. For discrete event-time data, we assume that there are potentially five follow-up times $J = 5$ at discrete times $t_j = j$ ($j = 1, \dots, 5$) and $I = 300$ subjects having at least 2 or more longitudinal measurements (i.e., $\alpha_{0_1} = \alpha_{0_2} = -\infty$) with $\alpha_{0_j} = 0.50$ for $j = 3, 4$, and 5. Table 1 shows the mean and standard deviation for various estimators of α_1 . These values are provided for estimators in which b_{i1} is assumed known, estimators which use complete simulated data, ORC and RCC, and our proposed approach with and without measurement error correction for different numbers of simulated datasets (M). The results show that ORC and RCC have a approximately 10% bias, while the proposed approach is unbiased. We also found that choosing $M = 10$ provided a good balance between efficiency and computational efficiency. Further, not incorporating the measurement error correction in the proposed approach had little effect on the results. We found this to be the case even when we increased the measurement error above 1, suggesting that this adjustment is not particularly important for the simple model in which longitudinal measurements and survival are linked through a random slope parameter. The measurement error correction may be more important for a more complex model such as that presented by Ye et al. (2008).

For continuous time-to-event data, the simulation was conducted with an exponential survival distribution with a mean of 5 years when $b_i = 0$, administrative censoring after 5 years, and $\alpha = 0.5$. We also assume that longitudinal measurements are taken at $t_1 = 0$, $t_2 = 0.125$, $t_3 = 0.25$, $t_4 = 0.75$, $t_5 = 1$, $t_6 = 2$, $t_7 = 3$, $t_8 = 4$, and $t_9 = 5$ ($J = 9$), with survival times being categorized into one-year intervals with $K = 5$. Table 2 shows the results of these simulations with $I = 300$, $\alpha = 0.5$, $M = 10$, and different values of the measurement error σ_ϵ . We present the mean (standard deviation) of parameter estimates with complete longitudinal data, ORC, RRC, and the the proposed approach. Although the proposed approach has increasing bias as σ_ϵ becomes large, this approach has less bias than both ORC and RCC for all values of σ_ϵ . Further, we conducted an additional simulation in which measurements t_3 to t_9 were missing with probability 0.5, creating datasets with

fewer observations on each subject. Results were essentially the same as reported in Table 2, suggesting that our approach does well even with shorter sequences of longitudinal data (data not shown).

4 Discussion

This article proposes a simple regression calibration approach for estimating the relationship between longitudinal measurements and time-to-event data which accounts for informative dropout in the longitudinal process. The approach is not completely unbiased since since the conditional distribution of the longitudinal process given the event time is approximated by a multivariate normal distribution. Particularly when the longitudinal and time-to-event processes are strongly linked, there may be small amounts of departure from normality. The effect of this lack of normality on bias appears to increase as the measurement error increases. However, in most situations, the bias is substantially smaller than the ORC and RCC approaches proposed in Ye et al. The simulation results demonstrate that, in general, the proposed approach results in estimates with increased variance relative to ORC and RRC. More precise estimation is possible under a more parsimonious parameterization. For example, $\beta_{1T_i}^*$ in (6) may be modeled as linear in T_i .

The proposed approach could be applied to a setting in which the two processes are linked through the true value of the longitudinal processes and time-to-event distribution. Further, the approach could be extended to allow for a more complex stochastic processes mean structure for the longitudinal process and for a semi-parametric fixed-effect structure as proposed by Ye et al. (2008). This would involve fitting model (6) or (9) with a different smooth curve $\varphi(t)$ and stochastic process $W_i(t)$ for each discretized dropout time. Such a model could be fit within the framework proposed by Zhang et al. (1998).

Our setup assumes that event times are only administratively censored after a fixed follow-up at the end of the study. For the case in which patients are censored prematurely dropout times can be imputed based on a model fit using only patients who had the potential to be followed over the entire study duration.

REFERENCES

- Carroll, R.J., Spiegelman, C.H., Lan, K.K.G., Bailey, K.T., and Abbott, R.D. (1984). On errors-in-variables for binary regression models. *Biometrika* **71**, 19-25.
- Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap* Chapman and Hall: New York.
- Hsieh, F., Tseng, Y.K., and Wang, J.L. (2006). Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics* **62**, 1037-1043.
- Laird, N.M. and Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics* **38**, 963-974.
- Song, X., Davidian, M., and Tsiatis A. A. (2002) A semiparametric likelihood approach for joint modeling of longitudinal data and time-to-event data. *Biometrics* **58**, 742-753.
- Tsiatis, A.A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14**, 809-834.
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer Verlag.
- Wu, M.C. and Bailey, K.R. (1989) Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics* **45**, 939-955.
- Wu, M.C. and Follmann, D.A. (1999). Use of summary measurements to adjust for informative missingness repeated measures data with random effects. *Biometrics* **55**, 75-84.
- Zhang, D., Lin, X., Raz, J., and Sowers, M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association* **93**, 710-719.
- Ye, W., Lin, X., and Taylor, J.M.G. (2008). Semiparametric modeling of longitudinal measurements and time-to-event data-A two stage regression calibration approach. *Biometrics* **64**, 1238-1246.

Table 1: Estimates of α_1 from model (2)-(4) when $\beta_0 = 1$, $\beta_1 = 3$, $\sigma_{b_0} = 1$, $\sigma_{b_1} = 1$, and $\sigma_{b_0, b_1} = 0$. We assume that $\sigma_\epsilon = 0.75$, $\alpha_0 = -1.5$, $\alpha_1 = 0.50$, $J = 5$, and $I = 300$. Further, we assume that $t_j = j$ and all individual's who are alive at $t_5 = 5$ are administratively censored at that time point. The means (standard deviations) from 1000 simulations are presented.

Estimator	$\hat{\alpha}_1$	SD
Known ¹	0.50	0.076
Complete ²	0.50	0.077
ORC	0.45	0.078
RRC	0.45	0.075
Prop M=3 w/o MC	0.49	0.097
Prop M=3 w/ MC	0.49	0.098
Prop M=10 w/o MC	0.49	0.090
Prop M=10 w/ MC	0.49	0.092
Prop M=20 w/o MC	0.49	0.089
Prop M=20 w/ MC	0.49	0.090
Prop M=50 w/o MC	0.49	0.088
Prop M=50 w/ MC	0.49	0.089
Prop M=100 w/o MC	0.49	0.087
Prop M=100 w/ MC	0.49	0.089

¹ Model (2) fit with b_{i1} assumed known.

² Model (2) fit with \hat{b}_{i1} replacing b_{i1} . The empirical Bayes estimates \hat{b}_{i1} are obtained by fitting (3) and (4) with complete longitudinal measurements.

Table 2: Estimates of α from model (3)-(4) and $\lambda(t, b_{i1}) = \lambda_0(t) \exp(\alpha b_{i1})$ where $I = 300$ and $M = 10$. We also assume that $\beta_0 = 1$, $\beta_1 = 3$, $\sigma_{b_0} = 1$, $\sigma_{b_1} = 1$, and $\sigma_{b_{i0}, b_{i1}} = 0$. The means (standard deviations) from 1000 simulations are presented.

Parameters		Estimators of α			
σ_ϵ	α	Complete	ORC	RRC	Proposed
0.2	0.25	0.25(0.072)	0.22(0.069)	0.21 (0.071)	0.24 (0.084)
	0.50	0.50(0.077)	0.43(0.072)	0.42(0.072)	0.48 (0.088)
0.5	0.25	0.25(0.072)	0.19(0.066)	0.19 (0.067)	0.24(0.104)
	0.50	0.50(0.077)	0.37(0.069)	0.36(0.068)	0.48(0.109)
1.0	0.25	0.25(0.072)	0.14(0.067)	0.16 (0.063)	0.23(0.149)
	0.50	0.50(0.078)	0.32(0.067)	0.30(0.064)	0.46(0.155)