SSC Annual Meeting, May 2006
Proceedings of the Survey Methods Section

SOME METHODS ON LONGITUDINAL DATA ANALYSIS

Grace Y. Yi [1]

ABSTRACT

Longitudinal data analysis has received extensive research interest. Missingness and covariate measurement error present major challenges to standard inference methods. There have been considerable advances on methodology for dealing with either missingness or measurement error in covariates. Relatively little attention has been directed to deal with both features simultaneously. In this article I will discuss inference frameworks for longitudinal data with missing observations and error-contaminated covariates. Some methods for further study are indicated.

KEY WORDS: Longitudinal data, Measurement error, Missing data.

RÉSUMÉ

L'analyse des données longitudinales a été l'objet de recherches extensives. Les données manquantes et l'erreur de mesure pour les covariables posent des problèmes majeurs en ce qui a trait aux méthodes d'inférence habituelles. Il y a eu des percées considérables en ce qui concerne la méthodologie permettant de de tenir compte des données manquantes et de l'erreur de mesure pour les covariables. Relativement peu d'intérêt a été porté au traitement simultané de ces deux problèmes. Dans cet article, je présenterai le cadre de travail pour l'inférence dans le contexte de données longitudinales avec des données manquantes et des covariables affectés par des erreurs de mesure. Certaines méthodes sont identifiées comme étant sujettes à plus d'études.

KEY WORDS: Données longitudinales; données manquantes; erreurs de mesure.

1. INTRODUCTION

In longitudinal studies, a response variable, often with a set of covariates, is observed on individuals repeatedly over time. It is of interest to understand the relationship between the response variable and the covariates. As data are often correlated, it is important to take account of possible association among repeated measurements. Common inference methods include random effects models, likelihood based approaches, and generalized estimating equations methods. For a comprehensive review, see Verbeke and Molenberghs (2000), for example.

Though longitudinal studies are designed to collect information for every subject on each occasion, missing observations often arise due to various reasons. This may be due, for example, to the nature of the study procedure or to causes that are not related to the responses. In many medical studies, the reason for missingness depends on the response values themselves. For example, the side effects of the treatment may make patients worse and thereby affect patient participation. In the literature there is substantial research on the impact missing data can have on inferences regarding treatment and other covariate effects.

On the other hand, covariate measurement error is a typical feature in data collection. Sometimes covariates of interest may be difficult to observe precisely due to physical location or cost. For example, the degree of narrowing of coronary arteries may reflect risk of heart failure, but physicians may measure the degree of narrowing in carotid arteries instead due to the less invasive nature of this method of assessment. Sometimes it may be difficult or impossible to measure covariates accurately due to the nature of the covariates. For example, the level of exposure to potential risk factors for cancer such as radiation can not be measured accurately (Pierce et al. 1992). In other situations, a covariate may represent an average of a certain quantity over time, and any practical way of measuring such a quantity necessarily features

[1] Grace Y. Yi, Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1, E-mail: yyi@uwaterloo.ca

measurement error. This is the case, for instance, in the Framingham Heart Study (Kannel et al. 1986) where one of the risk factors of interest for coronary artery disease is the average daily systolic blood pressure.

Missingness and covariate measurement error present major challenges to standard inference methods. To conduct valid inference, it is important to take these features into account; otherwise, erroneous or even misleading results may be produced. In this article I will discuss inference frameworks and some methods that feature both missingness and measurement error in covariates. Section 2 presents the notation and frameworks. Section 3 outlines some analysis methods. General discussion is included in Section 4.

2. NOTATION AND FRAMEWORK

Let Y_{ij} be the response variable for subject i at time point j, X_{ij} be the covariate vector subject to error, and Z_{ij} be the vector of error-free covariates, $i = 1, 2, ..., n$, and $j = 1, 2, ..., m$. Let W_{ij} be an observed version for X_{ij}. Denote Y_i = $(Y_{i1}, Y_{i2}, ..., Y_{im})'$, $X_i = (X_{i1}', X_{i2}', ..., X_{im}')'$, $Z_i = (Z_{i1}', Z_{i2}', ..., Z_{im}')'$, and $W_i = (W_{i1}', W_{i2}', ..., W_{im}')'$. Let R_{ij} be 1 if Y_{ij} is observed, and 0 otherwise. Denote by $R_i = (R_{i1}, R_{i2}, ..., R_{im})'$ the vector of missing data indicators, $i = 1, 2, ..., n$.

Figure 1 displays the four processes that are related to inference purposes. It is of essential concern to characterize the relationship between the true covariate process and the response process. In particular, we are interested in inference procedures about the associated parameter β. Since not all covariates (i.e., X here) are observable, only their surrogate values (i.e., W) may be used to estimate β, we need to adjust for the bias induced by using the observed version W. As some response components are missing, we also need to adjust for the bias induced by missingness. Therefore, a valid inference procedure involves two types of adjustment.

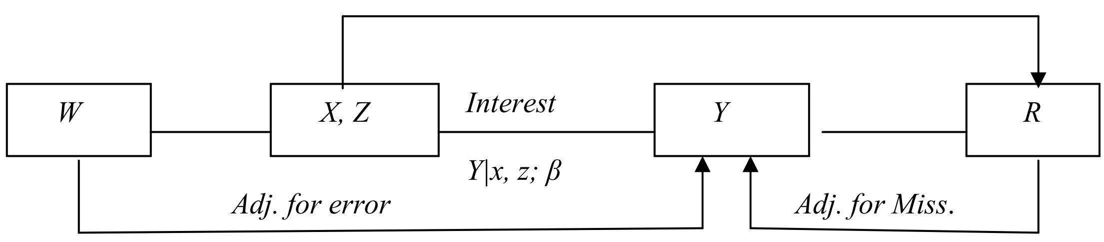

Figure 1: The relationship of the four processes: the response process (Y), the true covariate process (X, Z), the surrogate covariate process (W), and the missing data process (R).

2.1 Response Process

Random effects models are commonly employed to modulate the response process in connection with covariates to characterize subject-specific features (Laird and Ware 1982; Stiratelli, Laird and Ware 1984; Carroll et al. 2006, Ch. 11). Namely, a two-stage formulation may be taken. At stage 1, we assume that conditional on random effects u_i, each response component Y_{ij} follows a distribution of the density function

$$f(y_{ij} \mid u_i; \theta_{ij}, \varphi) = \exp\{(y_{ij}\theta_{ij} - b(\theta_{ij})) / a(\varphi) + c(y_{ij}, \varphi)\}, \tag{1}$$

where $a(.)$, $b(.)$, and $c(.,.)$ are known functions; φ is the dispersion parameter, and parameter θ_{ij} can be further modelled to facilitate within-subject variability. Let μ_{ij}^u be the conditional mean $E(Y_{ij} \mid u_i) = b'(\theta_{ij})$.

At stage 2, we assume that

$$g(\mu_{ij}^{u}) = X_{ij}'\beta_x + Z_{ij}'\beta_z + A_{ij}'u_i, \tag{2}$$

where $g(.)$ is the link function, and X_{ij}, Z_{ij} and A_{ij} are covariate vectors. The regression parameters β_x and β_z are called fixed effects because they are constant between subjects. Random effects are assumed following a certain distribution: $u_i \sim f(u_i \mid \delta)$ with δ being the associated parameter.

When primary interest lies in the population mean at each time point, we may alternatively consider the marginal regression model (e.g., Robins, Rotnitzky and Zhao 1995; Fitzmaurice, Molenberghs and Lipsitz 1995; Yi and Cook 2002). Let $\mu_{ij} = E(Y_{ij} \mid X_i, Z_i)$ and $v_{ij} = \text{var}(Y_{ij} \mid X_i, Z_i)$ be the conditional expectation and variance of Y_{ij}, respectively, given the covariates X_i and Z_i. Denote $\mu_i = (\mu_{i1}, \mu_{i2}, ..., \mu_{im})'$. Consider

$$g(\mu_{ij}) = X_{ij}'\beta_x + Z_{ij}'\beta_z, \tag{3}$$

where $\beta = (\beta_x', \beta_z')'$ is the vector of regression parameters, and $g(.)$ is a known function. Model (3) implies that the dependence of mean μ_{ij} and variance v_{ij} on the subject level covariates X_i and Z_i is completely reflected by the time specific covariates X_{ij} and Z_{ij}. Pepe and Anderson (1994) justified this form from the viewpoint of formulating unbiased estimating functions. This assumption has been widely made in modeling longitudinal data. See Diggle and Kenward (1994), Fitzmaurice et al. (1995), and Yi and Thompson (2005), for example.

2.2 Missing Data Process

Different missing data mechanisms have been classified according to the dependence of the missing data process on the response process. Little (1995) discussed a unified framework for modelling the response and missing data processes simultaneously where random coefficients may be shared by both processes. Vandenhende and Lambert (2002) described a method which models the response and missing data processes through a copula model. Alternatively, Little (1993) and Little and Rubin (2002) distinguished two classes of models with missing data: selection models and pattern-mixture models, based on the factorization form of the joint probability function $f(y_i, r_i \mid x_i, z_i)$ of response vector Y_i and the missingness indicator vector R_i. Pattern-mixture models come with the factorization

$$f(y_i, r_i \mid x_i, z_i) = f(y_i \mid r_i, x_i, z_i)\ f(r_i \mid x_i, z_i),$$

whereas selection models assume the form

$$f(y_i, r_i \mid x_i, z_i) = f(r_i \mid y_i, x_i, z_i)\ f(y_i \mid x_i, z_i).$$

Under the setup of selection models, we may classify missing data mechanisms into three categories. Given the covariates X_i and Z_i, if the conditional distribution $f(r_i \mid y_i, x_i, z_i)$ of the missing data indicator vector does not depend on the responses, then the missing data mechanism is called missing completely at random (MCAR); if the conditional distribution $f(r_i \mid y_i, x_i, z_i)$ depends on the observed response components Y_i^{obs}, i.e., $f(r_i \mid y_i, x_i, z_i) = f(r_i \mid y_i^{obs}, x_i, z_i)$, then the resulting mechanism is called missing at random (MAR); otherwise, not missing at random (NMAR) mechanism arises where $f(r_i \mid y_i, x_i, z_i)$ typically depends on unobserved response components (Little and Rubin 2002).

If missing data follow MCAR, the complete data analysis based on the observed data may be performed since it leads to consistent estimators. For many applications dependence of missing data on the observed responses is perhaps of most interest. For example, when patients feel worse after taking a certain drug, they may decide to quit the drug and withdraw from the study. Hence, it may be plausible to assume an MAR mechanism. The rationale of using MAR mechanisms may be found, for instance, in Robins et al. (1995). NMAR mechanisms are often addressed for sensitivity analyses due to the lack of information on identifying or estimating parameters associated with the response and missing data indicator models (e.g., Rotnitzky, Robins and Scharfstein 1998).

2.3 Measurement Error Process

A classical additive measurement error model is widely adopted to link the true covariates and their observed values on the subject of measurement error. That is, conditional on X_{ij} and Z_{ij},

$$W_{ij} = X_{ij} + e_{ij}, \qquad j = 1, 2, ..., m, \tag{4}$$

where the error terms e_{ij}'s are assumed to follow $N(0, \Sigma_e)$ with Σ_e = diag (σ_k^2, $k = 1, 2, ..., p$) (e.g., Li and Lin 2000). Here p is the dimension of X_{ij}.

In some situations (e.g., pesticide studies), it may be plausible to adopt the Berkson model to modulate the true covariates and their observed version:

$$X_{ij} = W_{ij} + e_{ij}, \qquad j = 1, 2, ..., m, \tag{5}$$

where the error terms e_{ij}'s are assumed to have mean 0. Those two models differ intrinsically in the relationship of variance. Model (4) leads to var(W_{ijk}) > var(X_{ijk}) while model (5) gives var(W_{ijk}) < var(X_{ijk}) for each component with $k = 1, 2, ..., p$. More detailed discussion on these two models may be found in Carroll et al. (2006).

For some practical problems more complex models may be required. For example, Li, Shao and Palta (2005) consider a latent model to analyze data arising from a sleep cohort study. The true covariate (X) is the severity of sleep-disordered breathing (SDB), and the observed surrogate (W) is the apnea-hypopnea index (AHI) which records the number of breathing pauses per unit time of sleep. If SDB is positive, the observed AHI can be larger or smaller but it cannot be negative; if SDB is zero, the AHI can only be larger than or equal to the true value. To feature this complicated structure of the measurement error, one may assume a latent variable model:

$$W = \max(0, V + e), \text{ and } X = \max(0, V),$$

where V is a continuous latent variable which links X and W, and e is the measurement error on the latent scale having the distribution $N(0, \sigma_e^2)$.

3. INFERENCE PROCEDURES

In the literature, there are numerous methods developed to handle missingness for longitudinal data. Common strategies include likelihood-based analyses (e.g., Dempster, Laird and Rubin 1977; Diggle and Kenward 1994; Molenberghs, Kenward and Lesaffre 1997; Yi and Thompson 2005), imputation methods (e.g., Little and Rubin 2002; Cook, Zeng, and Yi 2004), and marginal methods (e.g., Robins et al. 1995; Fitzmaurice et al. 1995; Yi and Cook 2002). Under the framework of random effects models, authors including Wang and Davidian (1996), Higgins, Davidian and Giltinan (1997), Tosteson, Buonaccorsi and Demidenko (1998), Ko and Davidian (2000), and Li, Zhang and Davidian (2004) studied measurement error problems in longitudinal studies. However, relatively little attention has been directed to deal with missingness and measurement error in covariates simultaneously. In the following I will discuss inferential procedures that take account of both features.

3.1 Structural Modeling Strategy

With a random effects model (1) together with (2) we may employ a likelihood-based method to perform inference. If a classical additive error model (4) is assumed, then one may consider the likelihood of the observed data

$$L_i(y_i, w_i, r_i \mid z_i) = \int f(r_i \mid y_i, x_i, z_i)\, f(y_i \mid x_i, z_i, u_i)\, f(w_i \mid x_i, z_i)\, f(x_i \mid z_i)\, f(u_i)\, du_i\, dx_i \tag{6}$$

where the assumptions that $f(r_i \mid y_i, w_i, x_i, z_i, u_i) = f(r_i \mid y_i, x_i, z_i)$ and $f(y_i \mid w_i, x_i, z_i, u_i) = f(y_i \mid x_i, z_i, u_i)$ are made. The first assumption on missing data indicators R_i implies that the missing data process is governed by the response and the true covariate processes. Observed measurements W_i do not control missing data mechanisms. The second assumption pertains to non-differential measurement mechanism (Carroll et al. 2006, p.36). It says that, conditional on the true covariates X_i and Z_i as well as random effects u_i, the observed measurements W_i do not contribute additional information on the inference on the responses Y_i. This mechanism is plausible for many practical problems, especially for observational studies. The above formulation also requires $f(w_i \mid x_i, z_i, u_i) = f(w_i \mid x_i, z_i)$ and $f(x_i \mid z_i, u_i) = f(x_i \mid z_i)$.

This approach provides a modeling framework to accommodate different missing data mechanisms. The method is appealing because it is flexible, efficient, and reliable in dealing with problems concerning covariate measurement error (Stefanski and Carroll 1990; Schafer and Purdy 1996). However, robustness to model assumptions is a major concern in this context. Typically, specification of the conditional distribution of X_i, given Z_i, is generally difficult since X_i is often not observable. Furthermore, likelihood methods are often computationally demanding because of the integrals involved.

Alternatively, if the Berkson error model is used, we may proceed

$$L_i(y_i, w_i, r_i \mid z_i) = \int f(r_i \mid y_i, x_i, z_i)\, f(y_i \mid x_i, z_i, u_i)\, f(x_i \mid w_i, z_i)\, f(w_i \mid z_i)\, f(u_i)\, du_i\, dx_i$$

$$\propto \int f(r_i \mid y_i, x_i, z_i)\, f(y_i \mid x_i, z_i, u_i)\, f(x_i \mid w_i, z_i)\, f(u_i)\, du_i\, dx_i. \qquad (7)$$

Similar assumptions to those in (6) are required in this formulation. However, the conditional distribution of the true covariates X_i, given Z_i, is not needed in this development.

3.2 Functional Modeling Strategy

When primary interest lies in the marginal mean parameters β in model (3), the inverse probability weighted generalized estimating equations (IPWGEE) approach (Robins et al. 1995) is commonly invoked to conduct inference about the parameters under MAR mechanisms. This marginal approach is widely viewed as attractive because it does not require complete specification of the joint distribution of the longitudinal responses but rather is based only on specification of the first two moments.

For $i = 1, 2, ..., n$, let $D_i = \partial \mu_i' / \partial \beta$ be the matrix of the derivatives of the mean vector μ_i with respect to β, and $\Delta_i(\alpha) =$ diag($I(R_{ij} = 1) / \pi_{ij}$, $j = 1, 2, ..., m$) be the weight matrix accommodating missingness, where $I(.)$ is the indicator function, $\pi_{ij} = P(R_{ij} = 1 \mid Y_i, X_i, Z_i)$, and α is the vecter of parameters associated with the missing data process. Let $V_i = G_i^{1/2} C_i G_i^{1/2}$ be the covariance matrix of Y_i, where $G_i =$ diag(v_{ij}, $j = 1, 2, ..., m$), and $C_i = [\rho_{i;jk}]$ is the correlation matrix with $\rho_{i;jk}$ being the correlation coefficient of response components Y_{ij} and Y_{ik} for $j \neq k$ and $\rho_{i;jj} = 1$. For $i = 1, 2, ..., n$, define $U_i = D_i V_i^{-1} \Delta_i(\alpha)(Y_i - \mu_i)$ and

$$U(\beta, \alpha) = \sum_{i=1}^{n} U_i. \qquad (8)$$

In the absence of measurement error, i.e., covariates X_i are precisely measured, $U(\beta, \alpha)$ are unbiased estimation functions for β. Estimator $\hat{\beta}$ of β can be obtained by solving $U(\beta, \alpha) = 0$ with α replaced by its estimate $\hat{\alpha}$ that is obtained from the model for the missing data process. When solving $U(\beta, \alpha) = 0$, one may use moment estimates to get an estimate for the correlation matrix C_i. Or alternatively, we may use working independence matrix G_i to replace V_i to obtain a consistent estimate for β with possible loss in efficiency (e.g., Sutradhar and Das 1999).

If covariates X_i are subject to measurement error, modeled by (4), for instance, then solving $U(\beta, \alpha) = 0$ no longer yields a consistent estimator of β. Instead, a modified version $U^*(\beta, \alpha)$, expressed by the observed data Y_{ij}, W_{ij}, and Z_{ij}, is needed (Nakamura 1990) such that $E[U^*(\beta, \alpha)] = E_{(R,Y,X,Z)}\,[E_{W|(R,Y,X,Z)}(U^*(\beta, \alpha))] = 0$. It suffices to find functions $U^*(\beta, \alpha)$ such that $E_{W|(R,Y,X,Z)}\,[U^*(\beta, \alpha)] = U(\beta, \alpha)$, because $U(\beta, \alpha)$ are unbiased estimating functions expressed by Y_{ij}, X_{ij}, and Z_{ij}. For various regression models Yi (2005) discussed this method to analyze longitudinal data when both missing response components and measurement error in covariates are present.

Alternatively, we may consider a functional estimation method by adapting the simulation-extrapolation (SIMEX) approach described in Cook and Stefanski (1994). The SIMEX method consists of two steps — a simulation step and a subsequent extrapolation step. The simulation step establishes the naive estimates for the cases when the variance of the error term for each measurement W_{ijk}, $k = 1, 2, ..., p$, is $(1+\lambda)\sigma_k^2$. Given $\lambda > 0$, which takes values in $\Lambda = \{\lambda_1, \lambda_2\ \ldots, \lambda_M\}$, say, for each $k = 1, 2, \ldots, p$, generate a large number, B, of simulated data, $W_{ijk}(b,\lambda)$, $b = 1, 2, \ldots, B$, by adding to each observed measurement W_{ijk} a random variable with mean 0 and variance $\lambda\sigma_k^2$ according to

$$W_{ijk}(b,\lambda) = W_{ijk} + \sqrt{\lambda}\ \sigma_k\, U_{ijkb},$$

where U_{ijkb}'s are independent $N(0,1)$ observations. Estimating functions (8) are then applied with X_{ijk} being replaced by $W_{ijk}(b,\lambda)$. Denote the resulting estimator as $\tilde{\theta}_b(\lambda)$. This procedure is repeated B times for each choice of λ. Let

$$\hat{\theta}(\lambda) = B^{-1}\sum_{b=1}^{B}\tilde{\theta}_b(\lambda) \qquad \text{for each } \lambda \in \Lambda.$$

In the extrapolation step, for each component $\tilde{\theta}_k(\lambda)$ of $\hat{\theta}(\lambda)$, regress the estimate $\tilde{\theta}_k(\lambda)$ on λ and extrapolate the resulting predicted value to $\lambda = -1$ to obtain the SIMEX estimator $\tilde{\theta}_k$. Analogously, we may obtain the variance estimates for the SIMEX estimators.

The SIMEX approach is attractive because it does not require modeling the covariate process, and hence the resultant estimators are robust to possible misspecification of the distribution of covariates. Implementation of the SIMEX method is computationally straightforward. However, we note that this method only leads to approximately consistent estimators in general, because an approximate (rather than an exact) extrapolation function is used in the extrapolation step.

4. DISCUSSION

Missing values and measurement error in covariates represent major challenges in longitudinal data analysis. Though there has been substantial research on inference methods to deal with either missingness or covariate measurement error, there has been relatively little work on accounting for both features simultaneously. In this article I give an overview on the modeling frameworks and discuss some inference methods. The discussion here focuses on handling longitudinal data with missing responses together with mismeasured covariates. In practice, covariates may be missing as well. We may extend the discussion to accommodating missing covariates and missing responses as well as covariate measurement error. An overview on missing covariates may be found in Ibrahim et al. (2005).

ACKNOWLEDGEMENT

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Carroll, R.J., Ruppert, D., Stefanski, L.A., and Crainiceany, C.M. (2006). *Measurement Error in Nonlinear Models,* 2nd ed., Chapman & Hall.

Cook, J. and Stefanski, L.A. (1994). A simulation extrapolation method for parametric measurement error models. *Journal of the American Statistical Association,* **89,** 464-467.

Cook, R.J., Zeng, L., and Yi, G.Y. (2004). Marginal analysis of incomplete longitudinal binary data: a cautionary note on LOCF imputation. *Biometrics*, **60,** 820-828.

Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Ser. B*, **39**, 1-38.

Diggle, P. and Kenward, M.G. (1994). Informative drop-out in longitudinal data analysis (with discussion). *Applied Statistics*, **43**, 49-93.

Fitzmaurice, G.M., Molenberghs, G., and Lipsitz, S.R. (1995). Regression models for longitudinal binary responses with informative drop-outs. *Journal of the Royal Statistical Society, Ser. B*, **57**, 691-704.

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-1 pharmacokinetics. *Journal of the American Statistical Association*, **92**, 436-448.

Ibrahim, J.G., Chen, M.H., Lipsitz, S., and Herring, A. (2005). Missing-data methods for generalized linear models: a comparative review. *Journal of the American Statistical Association*, **100**, 332–346.

Kannel, W.B., Neaton, J.D., Wentworth, D., Thomas, H.E., Stamler, J., Hulley, S.B., and Kjelsberg, M.O. (1986). Overall and coronary heart disease mortality rates in relation to major risk factors in 325,348 men screened for MRFIT. *American Heart Journal*, **112**, 825-836.

Ko, H. and Davidian, M. (2000). Correcting for measurement error in individual-level covariates in nonlinear mixed effects models. *Biometrics*, **56**, 368-375.

Laird, N.M. and Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**, 963-974.

Li, Y. and Lin, X. (2000). Covariate measurement errors in frailty models for clustered survival data. *Biometrika,* **87**, 849-866.

Li, E., Zhang, D., and Davidian, M. (2004). Conditional estimation for generalized linear models when covariates are subject-specific parameters in a mixed model for longitudinal parameters. *Biometrics*, **60**, 1-7.

Li, L., Shao, J., and Palta, M. (2005). A longitudinal measurement error model with a semicontinuous covariate. *Biometrics*, **61**, 824-830.

Little, R.J.A. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association,* **88**, 125-134.

Little, R.J.A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, **90**, 1112-1121.

Little, R.J.A. and Rubin, D.B. (2002). *Statistical Analysis with Missing Data*. 2nd Ed., John Wiley & Sons, Inc.

Molenberghs, G., Kenward, M.G., and Lesaffre, E. (1997). The analysis of longitudinal ordinal data with nonrandom drop-out. *Biometrika*, **84**, 33-44.

Nakamura, T. (1990). Corrected score functions for errors-in-variables models: methodology and application to generalized linear models. *Biometrika*, **77**, 127-137.

Pepe, M. S. and Anderson, G. L. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in Statistics, Simulation and Computation*, **23**, 939-951.

Pierce, D.A, Stram, D.O., Vaeth, M., and Schafer, D. (1992). Some insights into the errors in variables problem provided by consideration of radiation dose-response analyses for the A-bomb survivors. *Journal of the American Statistical Association*, **87**, 351-359.

Robins, J.M., Rotnitzky, A., and Zhao, L.P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, **90**, 106-121.

Rotnitzky, A., Robins, J.M., and Scharfstein, D.O. (1998). Semiparametric regression for repeated outcomes with nonignorable nonresponse. *Journal of the American Statistical Association*, **93**, 1321-1339.

Schafer, D.W. and Purdy, K.G. (1996). Likelihood analysis for errors-in-variables regression with replicate measurements. *Biometrika,* **83**, 813-824.

Stefanski, L.A. and Carroll, R.J. (1990). Structural logistic regression measurement models, in *Proceedings of the Conference on measurement Error Models,* P.J. Brown & W.A. Fuller, eds. Wiley, New York.

Stiratelli, R., Laird, N., and Ware, J.H. (1984). Random-effects models for serial observations with binary response. *Biometrics,* **40**, 961-971.

Sutradhar, B.C. and Das, K. (1999). On the efficiency of regression estimators in generalised linear models for longitudinal data. *Biometrika,* **86**, 459-465.

Tosteson, T., Buonaccorsi, J., and Demidenko, E. (1998). Covariate measurement error and the estimation of random effect parameters in a mixed model for longitudinal data. *Statistics in Medicine,* **17**, 1959-1971.

Vandenhende, F. and Lambert, P. (2002). On the joint analysis of longitudinal responses and early discontinuation in randomized trials. *Journal of Biopharmaceutical Statistics,* **12**, 425-440.

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

Wang, N. and Davidian, M. (1996). A note on covariate measurement error in nonlinear mixed models. *Biometrics*, **83**, 801-812.

Yi, G.Y. and Cook, R.J. (2002). Marginal methods for incomplete longitudinal data arising in clusters. *Journal of the American Statistical Association*, **97**, 1071-1080.

Yi, G.Y. (2005). Robust methods for incomplete longitudinal data with mismeasured covariates. *Far East Journal of Theoretical Statistics,* **16**, 205-234.

Yi, G.Y. and Thompson, M.E. (2005). Marginal and association regression models for longitudinal binary data with drop-outs: a likelihood-based approach. *The Canadian Journal of Statistics,* **33**, 3-20.