

Mixed-effect Models for Truncated Longitudinal Outcomes with Nonignorable Missing Data

Sujuan Gao¹ and Rodolphe Thiébaud²

¹Indiana University and ²Universit de Bordeaux

Abstract: Mixed effects models are often used for estimating fixed effects and variance components in continuous longitudinal outcomes. An EM based estimation approach for mixed effects models when the outcomes are truncated was proposed by Hughes (1999). We consider the situation when the longitudinal outcomes are also subject to non-ignorable missing in addition to truncation. A shared random effect parameter model is presented where the missing data mechanism depends on the random effects used to model the longitudinal outcomes. Data from the Indianapolis-Ibadan dementia project is used to illustrate the proposed approach.

Key words: Longitudinal data, mixed-effects model, nonignorable missing, truncation.

1. Introduction

This paper is concerned with estimating the effects of putative risk factors on cognitive decline in the elderly which is the focus of many longitudinal studies, both epidemiological or clinical trials in nature. Many cognitive assessment instruments currently used in dementia studies have an upper ceiling due to the limited time available for testing and the fact that the instruments are also functioning as screening or diagnostic tools for dementia with much greater emphasis on sensitivity at the lower end of the instruments. A well known example of such an instrument is the Mini-Mental State Examination (MMSE) which has a ceiling of 30. However, more extensive and lengthy neuropsychological tests are shown to be normally distributed in a large cohort of subjects. But given a shorter version of a lengthy test, many who would have scored in the tails of the longer version are now scoring at either endpoint. Inference on cognitive function or decline is concerned with the “true” cognitive ability or change that would have been measured rather than the observed score on the shorter version with a ceiling effect. The ceiling or truncation effect in neuropsychological tests was also

observed in van Belle and Arnold (2000) although their focus was on measuring reliability.

Analysis of cognitive decline data is further complicated by the fact that selective groups of subjects maybe missing some measurements either by study design or by happenstance. For example, in large cohort studies of dementia, subjects previously diagnosed with cognitive impairment no dementia (CIND), an intermediate status between normal and dementia, were usually not re-screened with the cognitive test and may directly proceed to clinical evaluations. Other missing data due to death or nursing home placement may also raise the possibility of non-ignorable missing data in this situation.

When measurements with ceilings are used in regression analysis, it has been shown that the ordinary least square (OLS) estimator ignoring the ceiling is biased and inconsistent (Goldberger, 1981). There have been efforts to correct for the OLS bias in the regression model setting (Tsui, Jewell, and Wu, 1988). Hasselblad, Stead and Galke (1980) considered a univariate regression model with multiple truncation points using the EM algorithm. In the case of a single truncation point with normally distributed data, the model is sometimes referred to as the Tobit model in the econometric literature (Amemiya, 1984) after an earlier econometric application (Tobin, 1958). Little and Rubin (2002) called this type of truncation “non-ignorable missing data with known mechanism” since the truncation point is known in this situation. When such outcomes are measured repeatedly and factors associated with change in outcome over time is of interest, special methods are also needed to ensure unbiased inferences. Hughes (1999) proposed an EM based maximum likelihood approach for longitudinal outcomes with truncated outcomes. Publications adopting the Hughes model are mostly concerned with the modeling of viral load in HIV and related lab data with various special mixed-effect models (Wu, 2002)).

Lyles (2000) considered the mixed-effect model in Hughes (1999) with the additional problem when the outcomes are also subject to informative drop-out. The authors adopted the approach of Schluchter (1992) by joint modeling of outcome variable and a log-normal survival model for time remained in the study for each subject. Thiebaut, Jacqmin-Gadda, Babiker and Commenges (2005) and Pantazis, Gouloumi, Walker, and babiker (2005) also considered the joint modeling of bivariate longitudinal data with a log-normal survival model for time to dropout. In this paper, we propose to use a binary survival model proposed by Wu and Carroll (1988) and previously adopted by by Pulkstenis, Ten Have, and Landis (1998), and Ten Have, Kunselman, Pulkstenis, and Landis (1998) to model the incidence of dropping out in a shared random effect model approach. We present results from a simulation study and illustrate the proposed method using data from a community-based dementia study.

2. A Longitudinal Dementia Study

The Indianapolis Study of Health and Aging is one of two longitudinal cohorts in the Indianapolis-Ibadan Dementia Project aimed at identifying risk factors for dementia, Alzheimer's disease and cognitive decline. The study population consists of 2212 African Americans age 65 and older living in Indianapolis, USA, at study baseline. Study participants were evaluated at study baseline and repeatedly evaluated at 2, 5, 8 and 11 years after baseline with a two-phase design at each evaluation wave. At the first phase (screening phase), study subjects were interviewed at their homes with a questionnaire designed to evaluate their cognitive function. In addition, demographic information, family history of illness, medical history of the subject, consumptions of alcohol and tobacco and blood samples were collected during the screening interview. At the second phase (clinical phase), selected subjects from a stratified sample based on screening results received full physical and neurological examinations to determine disease status. Subjects who received a full clinical evaluation were classified as demented, CIND, or normal. Demented subjects were then followed using a separate protocol. The CIND subjects were allowed to skip the screening phase and proceeded directly to clinical evaluation at the next follow-up evaluation.

At each screening phase study subjects were interviewed using the Community Screening Instrument for Dementia (CSID), a questionnaire designed for dementia screening in diverse cultural and educational backgrounds. The CSID consists of two parts: an interview with the study participant and an interview with an informant. The interview with the study participant assesses cognitive functioning, medical history, social involvement, and other putative risk factors. The interview with informants assesses the study participant's cognitive functioning, activities of daily living (ADL) and functioning at work and in social relationships.

The cognitive test of the CSID includes a number of test items measuring multiple cognitive domains including language, memory, orientation, judgment, comprehension and constructional praxis. Several neuropsychological tests including the animal fluency test and East Boston story were also included. In this paper, we consider a total cognitive score created by summing corrected answers from 40 questions in which lower scores indicate more cognitive impairment. These 40 questions were repeatedly administered at baseline and at each of the four follow-up waves. Therefore, there is an interest in investigating the patterns of cognitive decline and factors associated with cognitive decline. One particular factor of interest is the influence of education on cognitive decline. Many cross-sectional studies have reported low education is a risk factor for poorer cognitive function. However, longitudinal studies have been inconsistent with some finding

no effect suggesting the effect seen in cross-sectional studies was due to biases in cognitive assessment favoring highly educated individuals.

The investigation of education on cognitive decline is complicated in the Indianapolis data by two facts. The first is the truncation of the CSID scores at 40. In Table 1, we show significant difference in mean CSID cognitive scores at baseline in two groups defined by education level, namely, those who had 6 or less years of education (low education group) and those who had 7 or more years of education (high education). The cut-off of 6 years of education was chosen for this cohort in previous report (Hall, Gao, Unverzagt, and Hendrie, 2000). The percentages of subjects who scored at 40 also differed significantly in the two groups.

Table 1: Comparisons of baseline characteristics between the two groups defined by education levels

Characteristics	Low Education	High Education	<i>p</i> -value
Mean Cognitive Score (SD)	35.0(3.7)	37.1(3.0)	<0.0001
Score at Ceiling(%)	5.6	22.8	<0.0001

The second complication is the substantial proportions of missing data at each of the follow-up evaluations. Only 425 (21%) subjects out of a total of 2028 eligible subjects at baseline had complete information for all evaluations. The number of dropping out after baseline, 2, 5 and 8 years are 370, 469, 497 and 267, respectively. In Table 2, we show significant differences in percentages of subjects with missing data at each follow-up evaluation by education groups. Subjects with low education were more likely to drop out of the study than those in the high education group. Statistical inference ignoring these two facts may create a distorted conclusion regarding education’s effect on cognitive decline. In the follow section, we adopt a shared parameter model approach to model the longitudinal outcomes while accounting for both truncation and missing data. We note that our setting differ somewhat with the HIV laboratory test setting where a biomarker value below detectable range was considered censored. In our setting, we considered the test responses at 40 truncated due to the design of the questionnaire. We also assumed the test scores are measured without errors.

Table 2: Comparisons of percentages with missing data in the two groups defined by education levels. Data from the Indianapolis cohort of the Indianapolis-Ibadan Dementia Project

% Missing	Total Eligible (<i>n</i>)	Low Education	High Education	<i>p</i> -value
2 Year Follow-up	2028	22.5	17.6	0.0471
5 Year Follow-up	1658	34.4	27.4	0.0351
8 Year Follow-up	1240	61.8	39.3	< 0.0001
11 Year Follow-up	728	64.3	36.3	0.0002

3. Shared Random Effect Models

Let y be the true cognitive outcome that would be measured if the instrument used for assessing cognitive ability did not have an upper ceiling. y_{ij} denotes the j th measurement from the i th subject. Instead of y_{ij} , we observe a pair of random variables (Q_{ij}, T_{ij}) , where Q_{ij} is the truncated response and T_{ij} is the truncation indicator. Therefore, we have $Q_{ij} = y_{ij}$ if $T_{ij} = 0$; $Q_{ij} \leq y_{ij}$ if $T_{ij} = 1$. Note that Hughes (1999) considered truncation at both the floor and ceiling by having three levels of T_{ij} . In our example data setting, we have a uniform truncation point, *i.e.* $T_{ij} = T$.

3.1 The longitudinal model

We assume the following mixed-effect model for the true cognitive outcome, y_{ij} :

$$y_{ij} = X_{ij}\beta + Z_{ij}\gamma_i + e_{ij} \quad (3.1)$$

where X_{ij} is a $1 \times p$ vector of fixed effect covariates, β is a $p \times 1$ vector of fixed effect, Z_{ij} is a $1 \times q$ vector of random effect covariates, γ_i is a $q \times 1$ vector of random effect, $\gamma_i \sim N(0, D)$, and $e_{ij} \sim N(0, \sigma^2 I)$.

Note that y has a multivariate normal distribution with mean $X\beta$ and variance-covariance matrix of $V = ZDZ' + \sigma^2 I$, where X is the design matrix for the fixed effect where it has X_{ij} as its rows and Z is the design matrix for the random effect where Z_{ij} forms its rows. Our interest is in estimating the fixed effect parameter β , the variance-component matrix D and the random error parameter σ^2 using the observed data Q_{ij} and T_{ij} . An EM algorithm can be used to derive parameter estimates from (3.1) when there is no missing data. We refer the details of the EM algorithm to Hughes (1999) and concentrate instead to describing the methods for dealing with nonignorable missing data.

3.2 The drop out model

Let R_{ij} be an indicator variable for missing observation for the i th individual at the j th follow up wave: $R_{ij} = 0$, if an outcome is observed; $R_{ij} = 1$, if an outcome is missing. Using the framework of the shared random effect model, we assume the following model for the probability of a subject having missing data at evaluation wave j conditional on this individual being followed up at evaluation wave $j - 1$:

$$\eta[\text{Prob}(R_{ij} = 1 | R_{ij-1} = 0)] = W_{ij}\alpha + \delta U_{ij}\gamma_i \quad (3.2)$$

where η is a link function for a generalized linear model, W_{ij} is a $1 \times r$ vector of covariates, α is a $r \times 1$ vector of fixed effect, δ is a parameter for the random

effect, U_{ij} is a $1 \times q$ vector of covariates for the random effect, and γ_i is the same subject-specific random effect defined in the longitudinal outcome model (3.1).

The shared random effect model has an attractive appeal to biomedical researchers who generally believe that there may be some latent yet to be measured quantity underlying a person's susceptibility to both cognitive decline and missing data due to adverse outcome (nursing home entrance or death). In addition, the shared random effect model does not explicitly assume that missingness depends on *the* unobserved outcome; rather, it depends on a latent variable that is inherent in all outcomes from the same subject.

3.3 The Joint Likelihood

Without loss of generalizability, we assume that for each completely followed up subject, the first m_i of the total n_i observed outcomes are not truncated. We also assume that for subjects who were lost to follow-up, the first m_i observations were not truncated ($m_i \leq n_i$), the next $n_i - m_i$ observations were truncated and the last observation was unobserved. This can be easily achieved by re-arranging the observations for each subject along with all relevant covariates.

Let $p_{ij} = \text{Prob}(R_{ij} = 1)$. Assuming that y_{ij} and R_{ij} are independent given γ_i , the joint likelihood function of observed y_{ij} and R_{ij} can be written as

$$L = \prod_{i=1}^n \int \prod_{j=1}^{m_i} (1 - p_{ij}) f(y_{ij} | X_{ij}, Z_{ij}, \gamma_i) \times \prod_{j=m_i+1}^{n_i} (1 - p_{ij}) \left[\int_T^\infty f(y_{ij} | X_{ij}, Z_{ij}, \gamma_i) dy_{ij} \right] \prod_{R_{ij}=1} p_{in_i+1} d\gamma_i \quad (3.3)$$

Note that the joint likelihood can be divided into three parts. The first product in (3.3) is the likelihood function for the regular mixed-effect model for measures with observed outcome and no truncation; the second product involves those subjects with observed, but truncated outcomes; the third product is the probability of missing data conditional on random effects. Various link functions may be considered in the missing data model. We consider, for example, when η is a logit link function.

Since the missing data model is built on the incidence function of the missing data variable, we use the following recursive relationship to define the marginal probability of R_{ij} :

$$p_{ij} = \eta_{ij}(1 - p_{ij-1}) + p_{ij-1}, j = 2, \dots, n_i, \text{ and } p_{i1} = 0$$

Assuming the shared random effect parameter follows a normal distribution, the maximum likelihood estimates can be derived using numerical integration

techniques offered in the NL MIXED procedure in the SAS software package. Programming codes for the implementation are available upon request. In the following section we investigate the empirical properties of the shared random effect models and compare the proposed approach to other methods in simulation studies.

4. Simulation Results

We conducted a simulation to compare three different approaches in data with truncation and missing data. The first method is a so-called naive approach where a regular mixed-effect model is fit for the longitudinal data ignoring both the truncation and missing data. The second method considers the truncation process but ignores the missing data and the third method utilizes the shared random-effect models to obtain maximum likelihood estimates for the joint likelihood function. We designed the simulation studies using the data structure of the Indianapolis Study of Health and Aging. Number of baseline eligible cohorts and covariates were fixed to be those in the Indianapolis data set. We considered three covariates in the simulation. One is an indicator variable for age: 0 for those 75 and younger and 1 for those over 75 years at each evaluation. This variable is a time-dependent variable. The Second covariate included in the simulation is the dichotomized education variable discussed in Section 2. The third covariate is follow-up time in years since baseline.

We simulated longitudinal true outcomes according to the following model:

$$\begin{aligned} \text{score} = & \beta_0 + \beta_1 \text{age group} + \beta_2 \text{low education} \\ & + \beta_3 \text{time} + \beta_4 \text{low education} * \text{time} + \gamma_i + e_{ij} \end{aligned} \quad (4.1)$$

where $\gamma_i \sim N(0, \sigma_g^2)$, and $e_{ij} \sim N(0, \sigma_e^2)$. All generated scores above 40 were then truncated to 40.

We also simulated missing data by using the following missing data model:

$$\begin{aligned} & \text{logit}[\text{Prob}(R_{ij} = 1 | R_{ij'} = 0, j' < j)] \\ = & \alpha_0 + \alpha_1 \text{age group} + \alpha_2 \text{low education} + \alpha_3 \text{time} + \delta \gamma_i \end{aligned} \quad (4.2)$$

where $\gamma_i \sim N(0, \sigma_g^2)$. δ is the parameter controlling the degree and the direction of the "closeness" between the longitudinal model and the missing data process. We consider three scenarios in this paper. $\delta=0$ indicates that missing data is not linked to the outcomes, hence the method considering truncation only is adequate. We also consider the scenario when $\delta = -1$ where those subjects with higher cognitive scores are less likely to drop out, consistent with our dementia data. The last scenario we consider is when $\delta = 1$ where those with higher cognitive scores are more likely to drop out. Other parameter values used for

Table 3: Parameter estimates and percent bias (second entry) of model parameters assuming various missing data models in simulations

Parameters	$\delta=0$			$\delta=1$		
	Naive	Truncation	Missing	Naive	Truncation	Missing
Model (1)						
β_0	0.5225 47.75	1.0022 0.22	1.0024 0.24	0.5228 47.72	1.0051 0.51	1.0039 0.39
β_1	-1.6527 17.36	-1.9993 0.04	-2.0007 0.04	-1.6254 18.73	-1.9636 1.82	-2.0019 0.09
β_2	-0.7707 22.93	-0.9970 0.30	-0.9943 0.57	-0.7511 24.89	-0.9887 1.13	-0.9940 0.60
β_3	-0.2438 18.75	-0.3002 0.07	-0.3004 0.12	-0.1965 34.51	-0.2634 12.19	-0.2901 3.31
β_4	-0.2257 12.85	-0.2012 0.59	-0.2016 0.81	-0.2036 1.82	-0.1776 11.21	-0.1980 1.02
σ_g^2	–	0.9915	0.9918	–	0.8999	1.1077
σ_e^2	–	0.85	0.82	–	10.02	0.87
	–	0.9991	0.9990	–	1.0061	0.9804
	–	0.09	0.10	–	0.61	1.96
Model (2)						
α_0	–	–	-1.6585	–	–	-1.5573
	–	–	17.07	–	–	22.13
α_1	–	–	0.7041	–	–	0.7037
	–	–	29.59	–	–	29.63
α_2	–	–	0.7242	–	–	0.7084
	–	–	27.58	–	–	29.16
α_3	–	–	-0.0930	–	–	-0.1140
	–	–	161.97	–	–	175.97
δ	–	–	0.0032	–	–	0.5104
	–	–	–	–	–	34.58

generating the simulations are fixed to be: $\beta_0 = 1.0$, $\beta_1 = -2.0$, $\beta_2 = -1.0$, $\beta_3 = -0.3$, $\beta_4 = -0.2$, $\sigma_g^2 = 1.0$ and $\sigma_e^2 = 1.0$ for the longitudinal model (1), and $\alpha_0 = -2.0$, $\alpha_1 = 1.0$, $\alpha_2 = 1.0$, $\alpha_3 = 0.15$ for the missing data model (2). These parameters were chosen to make the truncation and missing data patterns similar to our dementia example data. Approximately 35% of subjects have truncated scores at baseline and the percentages decline with time. Drop-out ranged from about 25% at the first follow-up to about 50% at the last follow-up.

We present simulation results in Tables 3, 4 and 5. For space's sake, we omitted results for the $\delta=-1$ scenario which are similar to $\delta=1$. In Table 3, we present parameter estimates, and percent bias defined as $|\text{estimates}-\text{true parameter value}|/(\text{true parameter value})$. The naive estimates were shown to be consistently biased from the true parameters. The truncation only approach was adequate when $\delta = 0$ (missing at random) with minimum bias. When $\delta \neq 0$, the approach considering truncation and ignoring the missing data also showed some bias. If we focus on the parameter for interaction, β_4 , the truncation only method showed about 11% relative bias in estimating β_4 where the shared random effect approach had bias smaller than 1%. It should also be noted that the shared random effect model approach showed considerable bias in estimating parameters of

the missing data model (2), although the β estimates for model (1) do not seem to be affected as a consequence. It is worth noting that the shared random-effect model approach underestimated the magnitude of the δ parameter.

In Table 4, we present standard error estimates using the information matrices of each likelihood function for the three methods, compared to the empirical standard error estimates based on the estimates in the simulations. Our results demonstrate that these standard error estimates are fairly consistent with the empirical standard error estimates for all three methods.

Table 4: Standard error estimates and empirical standard errors from simulations assuming various missing data models

Parameters	$\delta=0$			$\delta=1$		
	Naive	Truncation	Missing	Naive	Truncation	Missing
Model (1)						
β_0	0.0382	0.0510	0.0510	0.0362	0.0496	0.0506
	0.0303	0.0498	0.0498	0.0297	0.0489	0.0494
β_1	0.0654	0.0832	0.0833	0.0617	0.0804	0.0824
	0.0720	0.0811	0.0812	0.0720	0.0834	0.0847
β_2	0.0975	0.1226	0.1226	0.0922	0.1184	0.1209
	0.1071	0.1229	0.1230	0.1093	0.1252	0.1259
β_3	0.0064	0.0080	0.0082	0.0057	0.0076	0.0077
	0.0069	0.0077	0.0080	0.0065	0.0076	0.0077
β_4	0.0311	0.0372	0.0373	0.0240	0.0295	0.0289
	0.0376	0.0376	0.0374	0.0290	0.0297	0.0292
σ_g	–	0.0684	0.0685	–	0.0670	0.0694
	–	0.0655	0.0654	–	0.0660	0.0671
σ_e	–	0.0387	0.0387	–	0.0396	0.0375
	–	0.0387	0.0387	–	0.0397	0.0377
Model (2)						
α_0	–	–	0.0848	–	–	0.0944
	–	–	0.0822	–	–	0.0913
α_1	–	–	0.0766	–	–	0.0937
	–	–	0.0714	–	–	0.0847
α_2	–	–	0.1070	–	–	0.1294
	–	–	0.1033	–	–	0.1271
α_3	–	–	0.0110	–	–	0.0125
	–	–	0.0097	–	–	0.0107
δ	–	–	0.0411	–	–	0.0562
	–	–	0.0356	–	–	0.0518

We then constructed 95% confidence intervals using the derived parameter estimates and standard error estimates based on asymptotic normality assumption for each of the parameters in the two models. In Table 5, we present estimated coverage probabilities for parameters in the two models using the three approaches. The naive method provided poor coverage for almost all parameters under all three scenarios, while the approach considering truncation but ignoring missing data provided adequate coverage when $\delta = 0$, as expected. However, when $\delta \neq 0$, the truncation only approach had poor coverage on some parameters, especially on time-dependent covariates, such as time and time by education

interaction. In addition, the truncation only approach also had poor coverage of the within-person variance parameter, σ_g^2 .

Table 5: Estimated coverage probabilities of 95% confidence intervals for model parameters using asymptotic normality assumptions from simulations with various missing data models

Parameters	$\delta=0$			$\delta=1$		
	Naive	Truncation	Missing	Naive	Truncation	Missing
Model (1)						
β_0	0	95.6	95.6	0	94.6	94.6
β_1	0.2	95.8	96.2	0	91.4	94.6
β_2	33.6	94.8	94.6	26.8	92.6	93.2
β_3	0	95.6	95.6	0	0	75.8
β_4	81.6	94.6	94.6	89.0	88.2	94.8
σ_g	–	94.6	94.8	–	67.8	95.2
σ_e	–	95.6	95.6	–	94.8	94.8
Model (2)						
α_0	–	–	2.0	–	–	0.6
α_1	–	–	1.6	–	–	8.4
α_2	–	–	26.0	–	–	38.0
α_3	–	–	0	–	–	1.3
δ	–	–	98.0	–	–	0.2

The shared random-effect model approach provided adequate coverage for all parameters in the longitudinal model with the exception of the time covariate, where the coverage was below the 95% nominal level. Coverage for the two variance component parameters are also excellent. However, the shared random-effect approach gave very poor coverage to the missing data model parameters indicating that inferences on the missing data model parameters are not reliable and better estimation approaches may be needed.

5. The Dementia Data

We also used the three estimating approaches on the dementia data described in Section 2. We present parameter estimates, standard error estimates and p-values by the three methods in Table 6. Note that the shared random-effect model estimated $\hat{\delta}=-0.3416$. Based on the simulation results, we believe that this may be an underestimate of the true parameter value and the estimate seems to confirm our observation that the high cognitive functioning subjects are less likely to have missing outcomes. We illustrate the estimated trend of cognitive decline for the two groups defined by education levels using the three methods in Figure 1. The truncation only method provided very similar result with the shared random-effect model approach, differing slightly toward the end of the follow-up period by estimating lesser declines for both groups. This difference is expected because the truncation approach assumes missing at random for those subjects dropped

out during follow-up while the shared random-effect model approach assumes that the drop-outs are more likely to be worse in cognitive functioning. The naive approach, however, differed from the other two approaches by estimating a much lesser decline especially in the high education group since it neglected the truncation in the scores.

Table 6: Results from the dementia data using the three different approaches

Parameters	Naive		Truncation		Missing	
	Est(StdErr)	<i>p</i> -value	Est(StdErr)	<i>p</i> -value	Est(StdErr)	<i>p</i> -value
Model (1)						
β_0	0.4053(0.0232)	<0.0001	0.6149(0.0278)	< 0.0001	0.6140(0.0279)	< 0.0001
β_1	-0.4683(0.0360)	< 0.0001	-0.5607(0.0425)	< 0.0001	-0.5661(0.0426)	< 0.0001
β_2	-0.4689(0.0554)	< 0.0001	-0.5924(0.0648)	< 0.0001	-0.5911(0.0650)	< 0.0001
β_3	-0.0722(0.0032)	< 0.0001	-0.0895(0.0038)	< 0.0001	-0.0946(0.0039)	< 0.0001
β_4	-0.0384(0.0110)	0.0005	-0.0274(0.0127)	0.0303	-0.0280(0.0126)	0.0264
σ_g^2	–	–	0.4704(0.0264)	< 0.0001	0.4780(0.0264)	< 0.0001
σ_e^2	–	–	0.7814(0.0192)	< 0.0001	0.7660(0.0189)	< 0.0001
Model (2)						
α_0					-0.9781(0.0608)	< 0.0001
α_1					0.2773(0.0581)	< 0.0001
α_2					0.2890(0.0770)	0.0002
α_3					-0.2603(0.0091)	< 0.0001
δ					-0.3416(0.0444)	< 0.0001

For our dementia data, all three approaches found the interaction between education level and time since baseline significant, providing evidence supporting an education effect on cognitive decline in this cohort. Therefore, application of the shared random-effect model did not alter conclusions reached by the naive method or the truncation only method. However, it is conceivable in other data set with a less strong interaction that the application of the shared random-effect model approach may lead to different conclusion reached using the other two methods.

6. Conclusion

We propose to use the shared random-effect model for longitudinal data with both truncation and missing data. We showed that maximum likelihood approach under the correctly specified models can provided unbiased estimates for the parameters in the longitudinal outcome model. More research is needed to provide adequate estimation and inference procedures for the missing data model parameters.

Acknowledgments

This research was supported by NIH grants: R01 AG15813, R01 AG09956

and P30 AG10133. We thank Dr. H el ene Jacqmin-Gadda for her review and comments on an earlier version of the manuscript.

References

- Amemiya, T. (1984). Tobit models: a survey. *Journal of Econometrics* **24**, 3-61.
- Goldberger, A. S. (1981). Linear regression after selection. *Journal of Econometrics* **15**, 357-366.
- Hall, K. S., Ogunniyi, A. O., Hendrie, H. C., Osuntokun, B. O., Hui, S. L., Musick, B. S., Rodenberg, C. A., Unverzagt, F. W., Guerje, O. and Baiyewu, O. A. (1996) A cross-cultural community based study of dementias: methods and performance of the survey instrument Indianapolis, U.S.A and Ibadan, Nigeria. *International Journal of Methods in Psychiatric Research* **13**, 129-142.
- Hall, K. S., Gao, S., Unverzagt, F. W. and Hendrie, H. C. (2000) Low education and childhood rural residence: risk for Alzheimer's disease in African Americans. *Neurology* **54**(1), 95-99.
- Hasselblad, V., Stead, A. G. and Galke, W. (1980). Analysis of coarsely grouped data from the lognormal distribution. *Journal of the American Statistical Association* **75**, 771-778.
- Hughes, J. P. (1999). Mixed effects models with censored data with application to HIV RNA levels. *Biometrics* **55**, 625-629.
- Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd Edition, John Wiley.
- Lyles, R. H., Lyles, C. M. and Taylor, D. J. (2000). Random regression models for human immunodeficiency virus ribonucleic acid data subject to left censoring and informative drop-outs. *Journal of Royal Statistical Society C*, **49**, 485-497.
- Pantazis, N., Gouloumi, G., Walker, A. S., Babiker, A. G. (2005) Bivariate modeling of longitudinal measurements of two human immunodeficiency type 1 disease progression markers in the presence of informative drop-outs. *Applied Statistics* **54**, 405-423.
- Pulkstenis E. P, Have T. R. T. and Landis J. R. (1998) Model for the analysis of binary longitudinal pain data subject to informative dropout through remedication. *Journal of American Statistical Association*, **93**, 438-450.
- Schluchter, M. D. (1992). Methods for the analysis of informatively censored longitudinal data. *Statistics in Medicine* **11**, 1861-1870.
- Ten Have T. R., Kunselman, A. R., Pulkstenis, E. P., Landis, J. R. (1998) Mixed effect logistic regression models for longitudinal binary response data with informative drop-out. *Biometrics*, **54**, 367-383.

- Thiebaut, R., Jacqmin-Gadda, H., Babiker, A., Commenges, D. and the CASCADE Collaboration. (2005) Joint modeling of bivariate longitudinal data with informative drop out and left censoring, with application to the evolution of CD4+ cell count and HIV RNA viral load in response to treatment of HIV infection. *Statistics in Medicine*, **24**, 65-82.
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica* **26**, 24-36.
- Tsui, K. L., Jewell, N. P. and Wu, C. F. J. (1988). A nonparametric approach to the truncated regression problem. *Journal of the American Statistical Association* **83**, 785-792.
- van Belle, G. and Arnold, A. (2000). Reliability of cognitive tests used in Alzheimer's disease. *Statistics in Medicine* **19**, 1411-1420.
- Wu, M. C., Carroll, R. J. (1988) Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* **44**, 175-188.
- Wu, L. (2002). A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *Journal of American Statistical Association* **97**, 955-964.

Received February 12, 2007; accepted October 4, 2007.

Sujuan Gao
Division of Biostatistics
Department of Medicine
Indiana University School of Medicine
Indianapolis, U.S.A
sgao@iupui.edu

Rodolphe Thiébaud
INSERM E0338 & U593- ISPED
Case 11 - Bureau 20
Universit de Bordeaux 2 - Victor Segalen
146 Rue Leo Saignat 33076 Bordeaux Cedex. France
rodolphe.thiebaut@isped.u-bordeaux2.fr