Mixed-effect models for longitudinal responses with different types of dropout: an application to the Leiden 85-plus study

A. Spagnoli\textsuperscript{a}, J.J. Houwing-Duistermaat\textsuperscript{b} and M. Alf\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a}Dipartimento di Scienze Statistiche, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Roma, Italy; \textsuperscript{b}Department of Medical Statistics and Bioinformatics, LUMC, Leiden, The Netherlands

(Received 30 January 2014; accepted 30 January 2015)

Longitudinal studies on cognitive functioning in geriatric populations usually cover short follow-up times and may be influenced by different sources of selection: only a portion of the designed sample may agree to participate in the study, and only some of the participants may complete the study. Motivated by a real-life data example, we discuss a variance component model with two peculiar features. First, we account for differences in individual status when entering the study by defining a flexible association structure between baseline and subsequent responses, where individual characteristics influencing entrance and participation in the follow-up are jointly modelled. Second, since we may argue that death and non-participation could not be treated as equivalent reasons for dropout, we introduce a pattern mixture model that takes into account the information on the time spent in the study and the reasons for dropout. The model is applied to data on cognitive functioning from the Leiden 85+ study, and its performance is analysed through a large-scale simulation study.

Keywords: longitudinal studies; variance component models; baseline response; dropout; pattern mixture model

1. Introduction

Longitudinal data concern measurements from the same individuals, taken repeatedly over time. In the longevity study we discuss in this paper, the Mini-Mental Status Examination (MMSE) index \cite{12} is measured over time to assess the cognitive status in older adults. The MMSE index is calculated on a questionnaire whose items are designed to assess orientation, attention, immediate or short recall, language, and the ability to follow a simple command; each item receives a binary score (1 for a correct answer and 0 otherwise). MMSE is meant to measure the progression of cognitive functioning by comparing the individual value of the index with a reference cut-off defined according to a specific level of impairment, see, for example \cite{24}, or to population-based

\*Corresponding author. Email: marco.alf@uniroma1.it

\textcopyright 2015 Taylor & Francis
norms [7]; however, due to difficulties in establishing cut-off values, some studies suggest to consider the MMSE index as an ordinal variable see, for example [14]. We analyse data drawn from the Leiden 85-Plus Study, see [9] or [31], and treat the MMSE index as composed by (conditionally) independent binary responses. We adopt a variance component (VC) model to account for dependence between repeated measures. We should notice that at least two sources of selection may bias the study design: at the study start, only subjects in a good health state may decide to enter the study and, during the study, only subjects with permanent healthy status may continue participating. Therefore, heterogeneity at the start and during the study may not coincide and overlap only partially. We use a simple reparametrization to account for differences between the distributions of baseline and subsequent responses. As it is usual in longitudinal studies, a large number of subjects present incomplete sequences due to premature dropout; longevity studies, as the one we discuss here, have the peculiar feature that non-participation and death are potential causes for dropout. Since the inference may be biased due to potentially informative missing data, we have defined a pattern mixture (PM) model, where patterns are related to either the number of the available measurements or the type (non-participation/death) of dropout.

The structure of the paper follows: in Section 2, the Leiden 85+ study is described. In Section 3, we introduce a modified VC model to account for dependence between baseline and subsequent responses. In Section 4, we propose a PM model to handle potentially non-ignorable missing data in the perspective of a sensitivity analysis. Section 5 describes the application of these models to the Leiden 85+ data. The last section contains concluding remarks. In the supplementary material, the behaviour of the proposed model is investigated in a simulation study where different data-generation schemes are considered for the longitudinal response and the missingness process.

2. Motivating example

The analysed data come from the Leiden 85+ study carried out by the Leiden University Medical Centre. Between september 1997 and september 1999, 705 inhabitants of Leiden, The Netherlands, reached the age of 85 years; in the month after their 85th birthday, all of them were asked to participate in a study on cognitive functioning in the elderly. Of the 705 subjects who were eligible, 14 died before they could be enrolled, 92 refused to participate, 38 refused to provide a blood sample, yielding a total simple size of 561 subjects; out of these, only 541 subjects have complete covariate information. The following covariates were collected at entry time: gender, educational status and plasma Apolipoprotein E (APOE) genotype. For each subject, the educational level was determined by the number of years she/he went to school; primary education was defined by less than 7 years of schooling. The APOE genotype is supposed to be a genetic risk factor for dementia; in particular, \( \epsilon_4 \) allele should be linked to increased risk for dementia, whereas \( \epsilon_2 \) allele carriers should be relatively protected. The analysis of APOE genotypes was confined to the three largest groups: \( \epsilon_2, \epsilon_3, \epsilon_4 \). Subjects were visited yearly until the age of 90 at their own place of residence, and face-to-face interviews were conducted through a questionnaire whose items are designed to assess orientation, attention, language skills and the ability to perform simple actions. The score on each item is binary (1 for a correct answer); the MMSE index for a subject is obtained by summing the scores that she/he obtained on the 30 questionnaire items. Thus, the observed values are integers between 0 and 30 (the total number of items). Data on a given subject are missing in the case of a completely missing questionnaire; this could occur if the subject drops out (e.g. due to poor physical health), or because he/she dies. The purpose of the study is the identification of demographic and genetic factors influencing the dynamics of cognitive functioning in the elderly. We will discuss two issues; namely, we aim at modelling the influence of individual covariates on cognitive functioning, while controlling for potential sources of unobserved heterogeneity and initial selection (see Section 3). Furthermore,
we develop a PM model to study the potential influence of the observed patterns of dropout on the model parameter estimates (see Section 5).

3. Modelling approach

In the Leiden 85+ study, we observe a set of longitudinal MMSE measurements, $y_{it}$, recorded on $i = 1, \ldots, n$ subjects at time points $t = (1, \ldots, T)$, together with a set of $p$ covariates $x_{it} = (x_{it1}, \ldots, x_{ipt})'$. The MMSE score is calculated as $y_{it} = \sum_{j=1}^{30} y_{ijt}$, where $y_{ijt}$ represents the score on the $j$th item for the $i$th subject at occasion $t$. When proposing a model for such data, we must bear in mind that the MMSE values are collected on the same subjects across time and, therefore, they are likely dependent. Since not all individuals who were eligible agreed to participate in the study, and not all of those who decided to participate entered the study in the same health status, it is important to account for initial differences in the selected sample and describe post-baseline dynamics in the observed MMSE values. To model association between repeated MMSE measures, we assume that, conditional on a subject-specific vector of random parameters $b_i$ and on the baseline value (the MMSE score at age 85, $Y_{i1}$), the responses $Y_{it}$ are independent Binomial random variables with index $m = 30$. The canonical parameter $\theta_{it}$ is defined by the following regression model:

$$\theta_{it} = x_{it}'\beta^* + z_{it}'b_i + \alpha y_{i1} \text{ for } t = 2, \ldots, T. \quad (1)$$

The terms $b_i, i = 1, \ldots, n$, are used to model unobserved individual-specific (time-invariant) heterogeneity shared by each lower level unit (time points) within the same individual, while $\beta^*$ is a $p$-dimensional vector of fixed regression parameters. Those effects that vary across individuals are collected in the design vector $z_{it} = (z_{it1}, \ldots, z_{itq})$.

These modelling assumptions deserve a few space of discussion. First, we have used the baseline measure to account for initial differences in the selected sample. According to Fitzmaurice et al. [11], we may use different strategies to handle the baseline response; however, no general approach is available, and the strategy should be chosen as a function of the study purposes. The Leiden 85+ study covers a short temporal range and all MMSE measurements have been taken after treatment or intervention (inclusion in the sample). In our perspective, the baseline value should be considered both as a dependent variable and as a covariate for subsequent responses; the baseline value accounts for potential sources of heterogeneity at the study entrance that are only partially persistent during the follow-up (for these reasons they cannot be summarized by common latent effects). For these purposes we use a modified model, see [1], for the first occasion:

$$\theta_{i1} = x_{i1}'\beta + z_{i1}'b_{i1}. \quad (2)$$

That is, we allow the fixed parameter vector to vary according to the different model structures at $t = 1$ and $t \geq 2$. Second, the (conditional) binomial assumption implies that item-specific complexity is assumed to be (at least approximately) constant across items and time; obviously, this is a working approximation based on the following considerations. The MMSE score, not the single items, is of interest here; by using appropriate thresholds, for example, the score could be discretized and models for the analysis of ordered response variables could be used. However, as pointed out by Woodford and George [33], there is no general agreement on the choice of such thresholds, both to define impairment (that is a binary response) or levels of impairment (that is an ordinal response). By looking at empirical experience in the field, see, for example [20], unobserved individual heterogeneity is by far the most important source of variability; thus, the binomial assumption is not that far from empirical experience. Obviously, we could have employed a nonlinear (e.g. with floor and ceiling effects) model for an approximately Gaussian response, but this approach would have needed the choice of an appropriate transformation of
the response variable which is not trivial and could result in parameter estimates that are quite complex to be interpreted. For all these reasons, we have decided to adopt a mixed-effect logit model for (approximately) binomial longitudinal responses, even at the cost of a more complex estimation algorithm. As in many longitudinal studies and, in particular, in studies on the elderly, also in this case partial missingness, i.e. missing data on one or more items, could be of concern. However, we do not have access to individual data on each item and, therefore, we will proceed by guessing that, if present, partial missingness does not substantially influence the general score.

Given the assumption of conditional independence, the joint (marginal) distribution of the individual sequence is

$$ f(y_i) = \int \prod_{t=2}^{T} f(y_{it}|y_{i1}; b_i)f(y_{i1}|b_{i1})g(b_i, b_{i1}) \, db_i \, db_{i1}. \quad (3) $$

The specification of the baseline distribution $f(y_{i1}|b_{i1})$ helps overcome the bias arising in the so-called analysis of covariance models, see, for example [30]. For modelling purposes, we should make the link between the random effects in submodels (1) and (2) explicit. We may use, see, for example [22], shared random parameters $b_{i1} = \Lambda b_i$ where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ leading to $\text{cov}(b_{i1}) = \Lambda \Sigma \Lambda'$, where $\text{cov}(b_i) = \Sigma$. In this case, the same heterogeneity sources affect the first and the subsequent occasions but for scale change, determined by the matrix $\Lambda$; this assumption does imply unit correlation between the observed outcomes (on the link function scale). Correlated random effects represent a further alternative, see, for example [1]: they control for potential overdispersion in the univariate profiles and for association between the longitudinal and the baseline responses. This structure is the more general one, but leaving the association between $b_i$ and $b_{i1}$ completely free could produce weak identifiability of the random effect distribution, since $b_{i1}$ is estimated using the observed baseline values only. We propose to adopt a simpler and more effective parametrization:

$$ b_{i1} = b_i + \phi_i, \quad (4) $$

where $\phi_i \perp \perp b_i$ represents additional, unobserved, heterogeneity at study entrance; here, the correlation estimate could be far from unit. This can be viewed as a particular specification of the framework discussed by Creemers et al. [6]. Unobserved heterogeneity sources for the baseline values is split into two components: the first ($b_i$) is common to subsequent occasions as well, while the second ($\phi_i$) accounts for unobserved heterogeneity at baseline which is only partially persistent in subsequent occasions. As suggested by a referee, the baseline-specific unobserved heterogeneity may also be associated with measurement error at baseline. By considering the additional term $\phi_i$, we may give a formal justification for Model (1). Let us suppose that the canonical parameter at time $t = 1$, is defined by

$$ \theta_{i1} = x'_{i1} \beta + z'_{i1}(b_i + \phi_i). $$

Since the selection at entrance may have a reduced effect on subsequent responses, for $t = 2, \ldots, T$ the following model may hold:

$$ \theta_{it} = x'_{it} \beta + z'_{it}(b_i + \alpha \phi_i), $$

where $\alpha$ measures the effect of $\phi_i$ on subsequent responses. Obviously, we may generalize by assuming a time-varying effect $\alpha_i$; however, we aim at making things simple and assume that, at least approximately, $\alpha_i = \alpha, \forall t = 1, \ldots, T$. It is worth to notice that considering the terms $(b_i + \phi_i)$ and $(b_i + \alpha \phi_i)$ would lead to a quite complex specification; in fact, while we are used
to estimate shared, $\alpha = 1$, proportional, $b_i$ and $\alpha b_i$, or generally correlated random effect models, the use of the $\alpha$ multiplier only for a part of the random effect (that is $\phi$) would make things quite complex to be handled from a computational perspective. Therefore, to improve computational handiness, we propose to approximate $z_i'\phi_i$ by the conditional mean $E(z_i'\phi_i \mid y_{i1})$. For this purpose, let us consider a longitudinal Gaussian response model:

$$y_{i1} = x_{i1}'\beta + z_{i1}'(b_i + \phi_i) + \epsilon_i.$$ 

In this case, we have that $E(z_i'\phi_i \mid y_{i1}) = y_{i1} - x_{i1}'\beta$. Let us suppose the following regression model holds at occasion $t > 1$

$$\theta_{it} = x_{it}'\beta + z_{it}'(b_i + \alpha \phi_i) = x_{it}'\beta + z_{it}'b_i + \alpha z_{it}'\phi_i.$$ 

Based on previous results, the canonical parameter can therefore be approximated by

$$\theta_{it} \simeq x_{it}'\beta + z_{it}'b_i + \alpha(y_{i1} - x_{i1}'\beta).$$

(5)

In the present study, all the observed covariates are time constant or null (deviation from the age at entry) at $t = 1$; therefore, $x_{it} = x_{i1} = x_i$, and $z_{it} = z_{i1} = z_i$. Thus, expression (5) changes to

$$\theta_{it} \simeq x_{it}'\beta(1 - \alpha) + z_{it}'b_i + \alpha y_{i1} = x_{it}'\beta^* + z_{it}'b_i + \alpha y_{i1},$$

where $\beta^* = \beta(1 - \alpha)$, motivating the use of different parameter vectors in expressions (1) and (2). This model structure can be summarized by

$$\theta_{it} = d_{it}(x_{it}'\beta^* + z_{it}'b_i + \alpha y_{i1}) + (1 - d_{it})(x_{i1}'\beta + z_{i1}'b_i)$$

$$= d_{it}(x_{it}'\beta^* + \alpha y_{i1}) + (1 - d_{it})(x_{i1}'\beta + z_{i1}'\phi_i) + z_{it}'b_i,$$

(6)

where the dummy variable $d_{it} = 0$ if $t = 1$, $d_{it} = 1$ otherwise, $i = 1, \ldots, n$. The estimated effects for $\beta^*$ may be lower than the estimates for $\beta$, for all covariates with non-null values at $t = 1$ due to this conditional structure. A similar approximation, within a different context, has been used in the approximate conditional model of Wu and Carroll [34], see also Aitkin and Alfó [1] and Alfó and Maruotti [2].

Maximum Likelihood (ML) estimation of model parameters can be achieved using a standard algorithm for random coefficient models, after choosing a specification for the distribution of the random parameters.

4. PM models for multiple causes of dropout

Longitudinal studies may have partial respondents, since some units do not answer on some occasions; for this reason, we may refer to the taxonomy introduced by Rubin [28]. The sequence for the $i$th individual $y_i = (y_{i1}, \ldots, y_{iT}), i = 1, \ldots, n$, can be written as $y_i = (y_{i}^{obs}, y_{i}^{miss})$, where $y_{i}^{obs}$ and $y_{i}^{miss}$ represent the observed and the unobserved components of the sequence, respectively. Let $R_{it}$ denote the missing data indicator, where $R_{it} = 1$ if the $i$th unit drops out at any point in $(t - 1, t)$, and $R_{it} = 0$ otherwise, $t = 1, \ldots, T$. Data are missing not a random (MNAR) if, conditional on observed covariates and observed responses $y_{i}^{obs}$, the probability of dropout still depends on $y_{i}^{miss}$.

In the analysed study, the observed missing data pattern is monotone: once a subject leaves, she/he does not reenter. Since we are considering a longitudinal study on cognitive functioning in the elderly, death is a potential cause for dropout which cannot be confused with non-participation. According to Harel and Demirtas [15], death is, in gerontological studies, a natural
end point and an individual who has been followed over time and dies during the study is often associated with a complete case. We will, however, avoid the philosophical question whether death must be considered as leading to a missing value. Rather, let us observe, according to Baltes [3], that while non-participation affects the characteristics of the sample, death modifies both the sample and the population under investigation. Death is an absorbing state, and the probability that an individual dies may depend on previous MMSE values, but not on current or future values since they simply do not exist. When time is continuous, there could be a gap between the last available longitudinal measurement and the time of death: therefore, the dropout process may still depend on unobserved longitudinal outcomes after the last available one. When non-participation (non-death dropout) is considered, the probability that the \( i \)th individual is observed at occasion \( t \) may depend on (potentially unobserved) current and future MMSE values, that is, the missing data mechanism could be MNAR. Several studies, see [3, 16, 29], have shown a strong relationship between non-participation and death in longitudinal studies on ageing, since the two reasons for dropout are highly associated with ageing-related changes.

We have limited knowledge on the link between the longitudinal and the missing data process, and since inferences can be biased if we do not account for potentially informative missing data, modified modelling approaches should be used. Obviously, we cannot test for informativeness, since, as detailed by Demirtas and Schafer [8], informative and non-informative missing data-generating process may produce incomplete trajectories that are virtually indistinguishable. Several authors have discussed modelling approaches for longitudinal data in the presence of death and other reasons of missingness. Kurland et al. [19] give a detailed overview of the topic, and discuss the major approaches that can be used to handle longitudinal data in the presence of death and dropout. They illustrate both marginal and (fully/partially) conditional models, stressing that the choice of the modelling tools has a great influence on the interpretation of longitudinal data truncated by death. Therefore, one should consider which methods to accommodate death is more consistent with her/his research aims. Kurland and Heagerty [18] discuss conditional (on time to dropout) and partly conditional (on survival status) regression models. As far as the marginal models, we may cite Dufouil et al. [10] who provide a re-weighting approach where dropout is treated as a true event and death is considered as a censoring event; Rajan and Leurgans [27] propose to use weighted estimating equations to estimate the marginal means, where weights are based on the probability to be alive and willing to participate. We should take into account that marginal representations suffer from a well-known drawback, often referred to as the immortal population, see [10]. However, in gerontological populations, participants at age 85 are very different from those at 90; for this reason, a conditional representation seems more appropriate, since the marginal model does not represent the behaviour of any individuals in the analysed population. In this context, Pauler et al. [25] describe a pattern mixture model where patterns are defined by the time of dropout and the survival time. Frangakis et al. [13] consider principal stratification based on potential outcomes (e.g. mortality under both levels of treatment). In a partially conditional perspective, le Cessie et al. [21] use a multi-state model where death is treated differently from other disease states; within each state, a weighted model is used to provide consistent estimates of (state-specific) marginal means, with independence working correlation to avoid the implicit imputation of data for deceased subjects, which is indeed a feature of random effect models. For all these reasons, we have defined a PM model, see [23, 32], where separate patterns are defined as a function of the time to and the type of dropout. The PM model does not need a formal specification of the dropout mechanism, but rather stratifies the observed sample according to the observed dropout patterns. This may help explore potential interactions between patterns of dropout (summarized by survival time/status) and observed covariates, and be used in a sensitivity analysis perspective. In other applications, joint models, see Little and Rubin [23], Hogan and Laird [17] can be adopted as well, for example, in longi-
tudinal studies on palliative care where survival and quality of life response are both of interest, since the research focus is on predicting survival with acceptable quality of life.

Let us denote by $T$ the expected number of observations per individual. Since we observe monotone missing data patterns only, we define:

$$S_i = T - \sum_{t=1}^{T} R_{it} = \sum_{t=1}^{T} (1 - R_{it})$$

to indicate the number of observed responses for the $i$th individual, $S_i \leq T$. According to standard survival notation, let $\Delta_i$ represent the indicator variable for the reason the $i$th subject drops out; in the Leiden 85+ study, $\Delta_i \in \{0, 1, 2\}$, corresponding to a complete sequence, death or non-participation, respectively. The value of $R_{it}$ depends on both the willingness to stay in the study and on the survival status (ie being alive). Let $P_{it} = 1$ and $A_{it} = 1$ indicate that the generic $i$th unit is willing to participate at the $t$th visit and that she/he is alive at that time. We may notice that $R_{it} = 0$ iff $P_{it} \times A_{it} = 1$; conversely, $R_{it} = 0$ iff $P_{it} \times A_{it} = 0$, that is either $P_{it} = 0$ or $A_{it} = 0$.

Let use denote by $s_i$ and $\delta_i$ the observed values for $S_i$ and $\Delta_i$, respectively. In a PM framework, we obtain the following expression for the conditional log-likelihood:

$$\ell(s_i, \delta_i) = \sum_{i=1}^{n} \log \left\{ \int f(y_i | s_i, \delta_i, b_i, b_{1i}) g(b_i, b_{1i} | s_i, \delta_i) db_i db_{1i} \right\}, \quad (7)$$

where the pattern $(s_i, \delta_i)$ influences the longitudinal outcome through the conditional distributions $g(b_i, b_{1i} | s_i, \delta_i)$ and $f(y_i | s_i, \delta_i, b_i, b_{1i})$. That is, unlike dropout models based on shared/correlated random parameters, we do not assume the longitudinal responses and the missingness indicator are conditionally independent. Rather, the conditioning in the primary model may affect all the fixed parameter estimates in a full interaction model.

5. Application: Leiden 85+ study data

As it has been previously outlined, in the Leiden 85+ study cognitive functioning has been measured using the MMSE index, recorded once a year from the age of 85 up to the age of 90. The aim is at exploring the MMSE dynamics as a function of a set of covariates including gender (female = 0), educational level (primary = 0) and APOE genotype (APOE33 = 0). Let us remind that low MMSE values indicate a poorer condition; Figure 1 shows a boxplot for the MMSE distribution by age of participants.

As it can be easily noticed, the MMSE median does not show relevant changes with increasing age, suggesting that subjects in good health tend to remain in the study, masking the decrease in cognitive functioning we do expect with increasing age of the participants. We have considered several models based on different association structures between the random effects in the post-baseline and in the baseline outcomes. Based on the analysis of the model fit, we will discuss random intercept models only, that is $b_i = b_{1i}$ and $b_{1i} = b_{1i}$.

The linear predictor is defined as follows:

$$\theta_{it} = x_i^\prime \beta^* + b_i + \alpha y_{1i}, \quad t = 2, \ldots, T_i,$$
$$\theta_{1i} = x_{1i}^\prime \beta + b_{1i},$$

where the design vectors are defined as $x_i = (1, \text{Age}_i - 85, \text{Gender}_i, \text{Educ}_i, \text{APOE}_i)$ and $x_{1i} = (1, 0, \text{Gender}_i, \text{Educ}_i, \text{APOE}_i)$, where Age$_i$ = 85, Gender, Educ and APOE denote time-constant factors, and $y_{1i}$ represents the MMSE values observed at $t = 1$. For the sake of brevity, we will focus on random intercept models based either on shared (model A: $b_{1i} = \lambda b_i$) or combined
Figure 1. Leiden 85 Plus Study: boxplot of MMSE distribution by age.

(shared plus independent, model B: $b_{i1} = b_i + \phi_i$) random effects. The proposed approach is fully parametric, as we consider Gaussian random intercepts, $b_i \sim \text{N}(0, \sigma^2_b)$ and $\phi_i \sim \text{N}(0, \sigma^2_\phi)$; ML estimates have been obtained through an EM algorithm using 15-point adaptive Gaussian Quadrature, see, for example [26]. For this purpose, and throughout the paper, we have used the glmer function from the library lme4, [4,5] available in the R environment for statistical computing. Table 1 summarizes the parameter estimates; we report in bold significant estimates ($\alpha = 0.05$).

By looking at estimates for model (A), we may notice that age negatively influences cognitive functioning, while subjects with higher school level are less cognitively impaired. Subjects with higher levels of the MMSE index at the beginning of the study have higher values of the index throughout the follow-up, while, as far as genetic factors are considered, the effect of APOE(34–44) is negative and significant. That is, subjects with allele $\epsilon_4$ have lower MMSE values at the beginning, see the significant effect of APOE$_{(34–44)\text{bas}}$, and throughout the study. These results are consistent with the literature, since allele $\epsilon_4$ is supposed to be a potential risk factor for dementia. Looking at APOE$_{(34–44)\text{bas}}$ and APOE$_{(34–44)}$ effect estimates and considering the corresponding magnitude, we may guess that the effect of allele $\epsilon_4$ is, at least approximately, time invariant (the same is true for the educational effect), even if the analysed time window is not that wide to make any inference on this point.

In general, the results obtained through models (A) and (B) are coherent in terms of parameter interpretation; if we take into account the combined (shared plus independent) random effect structure, we observe some changes in the random effect variance and in the size of parameter estimates for MMSE$_{85}$, APOE$_{(34–44)}$, Educ, which can be related to the $(1 - \alpha)$ multiplier described in Equation (5). Looking at model (B), we observe that $\hat{\sigma}^2_{b1} = \hat{\sigma}^2_\phi + \hat{\sigma}^2_\phi$ is quite higher...
Table 1. Maximum-likelihood parameter estimates.

| Variable                  | Model (A) |          |          | Model (B) |          |          |
|---------------------------|-----------|----------|----------|-----------|----------|----------|
|                           | Coeff.    | Std. Err.| Coeff.   | Std. Err. |
| Cons.                     | 1.24      | 0.14     | -0.43    | 0.17      |
| Age                       | -0.18     | 0.01     | -0.19    | 0.01      |
| Gender                    | 0.18      | 0.12     | 0.12     | 0.10      |
| Educ.                     | 0.94      | 0.12     | 0.69     | 0.10      |
| MMSE$_{85}$               | 0.01      | 0.00     | 0.09     | 0.01      |
| APOE$_{(22-23)}$          | 0.03      | 0.15     | 0.04     | 0.12      |
| APOE$_{(24)}$             | -0.05     | 0.38     | 0.04     | 0.31      |
| APOE$_{(34-44)}$          | -0.70     | 0.14     | -0.46    | 0.12      |
| Cons$_{bas}$              | 0.30      | 0.12     | 1.98     | 0.16      |
| Gender$_{bas}$            | 0.17      | 0.13     | 0.19     | 0.12      |
| Educ$_{bas}$              | 1.11      | 0.13     | 1.06     | 0.12      |
| APOE$_{(22-23)}_{bas}$    | 0.01      | 0.16     | 0.02     | 0.16      |
| APOE$_{(24)}_{bas}$       | -0.29     | 0.40     | -0.29    | 0.39      |
| APOE$_{(34-44)}_{bas}$    | -0.76     | 0.15     | -0.77    | 0.14      |
| $\sigma^2_b$             | 1.53      |          | 0.95     |          |
| $\sigma^2_\phi$          |           |          | 0.51     |          |
| $\ell$                   | -2541     |          | -2282    |          |
| AIC                       | 5112      |          | 4597     |          |

Note: Models (A)–(B) refer to common and, respectively, partially common random effects for the baseline and later occasion outcomes.

Table 2. Leiden 85 + Study: number of subjects participating in each stage.

| Follow-up age | Total | Present at the next visit (%) | Does not participate to the next visit (%) | Died before the next visit |
|---------------|-------|--------------------------------|--------------------------------------------|----------------------------|
| 85            | 541   | 484 (89.46)                    | 9 (1.66)                                   | 48 (8.87)                  |
| 86            | 484   | 422 (87.19)                    | 3 (0.62)                                   | 59 (12.19)                 |
| 87            | 422   | 373 (88.39)                    | 2 (0.47)                                   | 47 (11.14)                 |
| 88            | 373   | 318 (85.25)                    | 6 (1.61)                                   | 49 (13.14)                 |
| 89            | 318   | 266 (83.65)                    | 15 (4.72)                                  | 37 (11.63)                 |

As it can be easily noticed by looking at Table 2, the number of subjects who drop out between two consecutive time occasions is approximatively constant; thus, the corresponding proportion increases with age. To summarize, 240 (44.4%) participants died before the end of the follow-up, while 35 (6.5%) stop participating in the follow-up before its end.

Due to the high incidence of death/dropout, we may wonder whether the regression results reflect a selection effect where subjects in poorer health conditions (with lower MMSE values) tend to drop out. To graphically visualize potential selection bias due to non-participation and/or death, we report in Figure 2 the time trajectories of the mean MMSE score, by pattern of participation.

As it is clear, subjects who die have, throughout the study, a lower mean MMSE score than those that complete the study or drop out for other reasons. Curves referring to ‘completers’ and to ‘dropout’ show very similar values, with a slight acceleration in the decline of
Figure 2. Leiden 85+ Study: mean MMSE score by age of participants and type of dropout.

Table 3. Leiden 85+ Study: demographic and genetic characteristics of participants subjects by type of dropout.

| Variable     | Total | Completed (%) | Died (%) | Did not participate (%) |
|--------------|-------|---------------|----------|-------------------------|
| Gender       |       |               |          |                         |
| Male         | 180   | 74 (41.11)    | 97 (53.89)| 9 (5)                   |
| Female       | 361   | 192 (53.18)   | 143 (39.61)| 26 (7.20)              |
| Educational  |       |               |          |                         |
| Primary      | 351   | 166 (47.29)   | 162 (46.15)| 23 (6.55)              |
| Secondary    | 190   | 100 (52.63)   | 78 (41.05)| 12 (6.32)              |
| APO-E        |       |               |          |                         |
| 22–23        | 96    | 54 (56.25)    | 35 (36.46)| 7 (7.29)                |
| 24           | 12    | 6 (50)        | 6 (50)   | 0 (0)                   |
| 33           | 319   | 162 (50.78)   | 136 (42.63)| 21 (6.58)             |
| 33–34        | 114   | 44 (38.60)    | 63 (55.26)| 7 (6.14)               |
| $S$          | 541   | 6             | 2.9 (1.4) | 3.4 (1.7)              |

MMSE values in the period just before the dropout (this is common to those exiting by death). The trajectories of mean MMSE scores by age of participants suggest that the baseline values of MMSE may well predict the follow-up status of the subjects: it is likely that subjects who have a low (high) MMSE score at the beginning of the study have low (high) MMSE values throughout the follow-up, see the MMSE$_{85}$ estimate in Table 1. We report in Table 3
the characteristics of the subjects by pattern of participation; the survival status seems to be influenced by gender and APOE genotype, with males and allele $e_4$ carriers having a higher mortality rate.

Since completers have 6 visits while those who drop out have, on average, 2.9 and 3.4 visits (death and non-participation), model parameter estimates may be influenced by those who stay longer in the study. To account for this differential participation in the study, we define a class of pattern mixture models where all the available information on the observed dropout patterns is considered. In the following, we will consider only models with combined (shared plus independent) random effects, that is model (B) in Table 1. The first model (in the following model B1) only considers $S_i$ to approximate the individual propensity to remain in the study. In this case, the linear predictor is defined as follows:

$$\theta_{it} = x_{it}'\beta^* + (s_i \times x_{it}')\gamma^* + b_i + \alpha y_{i1}, \quad t = 2, \ldots, T_i,$$

$$\theta_{i1} = x_{i1}'\beta + b_i + \phi_i,$$

where, as we have remarked before, the design vectors are defined as $x_{it} = (1, \text{Age}_{it} - 85, \text{Gender}_{it}, \text{Educ}_{it}, \text{APOE}_{it})$ and $x_{i1} = (1, 0, \text{Gender}_{it}, \text{Educ}_{it}, \text{APOE}_{it})$, where $\text{Age}_{i1} = 85$, Gender, Educ and APOE denote time-constant factors. In the previous equation, $y_{i1}$ represents the MMSE values observed at $t = 1$ and $s_i$ is the number of visits available for each subject. Corresponding parameter estimates are given in Table 4, model (B1); significant parameter estimates ($\alpha = 0.05$) are reported in bold.

The estimates for $\text{Educ}_{i\text{bas}}, \text{MMSE}_{85}, \text{APOE}_{i\text{bas}}$ are similar in terms of magnitude to parameter estimates from model (B). The age effect is increased (in absolute value) and this is probably due to the effect of the interaction term ($\text{Age} \times s$) which is positive and significant. As far as the main effects are considered, estimates for APOE$_{34-44}$ and Educ are no longer significant in B1, but still significant in the baseline model (see APOE$_{34-44\text{bas}}$ and $\text{Educ}_{i\text{bas}}$).

Since $S_i$ does not provide information regarding the reason for dropout, we have implemented a further PM model (in the following model B2) where the variable ‘pattern of dropout’ ($\Delta_i$), with the levels completer ($\Delta_i = 0$), death ($\Delta_i = 1$), non-participation ($\Delta_i = 2$), has been considered. Throughout the text the level $\Delta_i = 0$ has been used as the reference category. In this case, the linear predictor is defined as follows:

$$\theta_{it} = x_{it}'\beta^* + (\Delta_i \times x_{it}')\gamma^* + b_i + \alpha y_{i1}, \quad t = 2, \ldots, T_i,$$

$$\theta_{i1} = x_{i1}'\beta + b_i,$$

where, as we have remarked before, the design vectors are defined as $x_{it} = (1, \text{Age}_{it} - 85, \text{Gender}_{it}, \text{Educ}_{it}, \text{APOE}_{it})$ and $x_{i1} = (1, 0, \text{Gender}_{it}, \text{Educ}_{it}, \text{APOE}_{it})$, where $\text{Age}_{i1} = 85$, Gender, Educ and APOE denote time-constant factors. In the previous equation, $y_{i1}$ represents the MMSE values observed at $t = 1$, and $\Delta_i$ is the indicator for the type of dropout. The parameter estimates are given in Table 4, model (B2). Here, the interaction of $\Delta_i$ with covariates describes the departure of estimated effects for subjects who die or do not participate in the study when compared with the subsample of completers. This model seems to suggest that the effect of age varies by the pattern of dropout: this is evident when the dynamics of MMSE values for those who leave the study by death is compared with the remaining sample of subjects. This effect would have not been observed should we have considered $S_i$ only. Parameter estimates obtained through model (B2) are quite similar to those obtained through model (B1), but for the effects of APOE$_{34-44}$, and for Educ, which are now significant. Looking at the penalized likelihood criteria, model (B1) seems to be preferred; however, this comparison should be handled with care since it is based on the data we have observed only. When using a PM model, we implicitly assume that the
individual-specific pattern (either summarized by the number of available measurements before dropout, \( s_i \), or by the type of dropout, \( \Delta_i \)) represents the individual propensity to drop out. This may not be always true and we have absolutely no chance to verify this assumption. Therefore, by comparing the fit of model (B1) with that of model (B2) we may say that, probably, the information supplied by \( s_i \) is more relevant, more subject-specific, than the one provided by the type of dropout, given by \( \Delta_i \), but this does not mean that the former must be preferred with respect to the latter, for at least two reasons. First, we do not control for all the factors that may play a role in this context, that is we have very limited information on the missing data-generating process. Second, we implicitly assume that an individual with, say, \( s_i = 3 \) would have given the same information of a completer had he/she the chance to stay longer in the study. Obviously, the same can be said about the pattern defined by the type of dropout. We rather suggest that both model specifications could give some more insight on the effects of a potentially informative missingness and that, for this reason, they should be used together to have a clearer picture of the study at hand.
6. Simulation study

A large-scale simulation study has been carried out to investigate the performance of the proposed approaches. Due to reasons of space, the simulation design and results have been made available as supplementary material. The results obtained (in all the scenarios we have considered) suggest that ignoring the dropout mechanism could lead to wrong conclusions. The proposed PM model specifications do not seem to be completely satisfactory, as in a non-negligible portion (greater than the nominal 0.05 level) of the data sets simulated under an MAR scenario we observe statistically significant changes in model parameter estimates when stratifying on the time to or the type of dropout. However, results suggest that the proposed modelling approach, especially when we consider the type of dropout and not just the number of available measurements for each unit, may represent a reliable screening tool to build a more complex model to handle non-ignorable missing data.

7. Concluding remarks

The paper describes the application of mixed-effect models to the Leiden 85+ Study. We take into account the dependence between repeated measures recorded on the same subject by adopting a particular association structure between the random effects in the first and the later occasions, assuming that baseline values may influence subsequent response values. This association structure allows to distinguish the part of unobserved heterogeneity that is common to the first and the later occasions and the one specific to the first occasion only, which is used to represent initial selection. The results show that adopting a model with a combined association structure provides reliable estimates for the covariates effects, and better fit when compared with ‘standard’ shared random effect models. Since non-participation and death can both influence parameter estimates, we discuss a class of PM models where the distinction between reasons of dropout (death and non-participation), and the propensity of the subjects to stay in the study are considered. We can observe a significant genotype effect for APOE(34–44), at least in the baseline submodel, which does not seem to depend on the pattern of dropout; instead, age effect estimates seem to be influenced by the missing data process, with a steeper decline for non-completers. The results obtained in the empirical application and the simulation study show that the choice of a synthetic measure to summarize the observed patterns of dropout may not be appropriate when different dropout events are observed. The proposed approach might be useful to study parameter sensitivity by looking at all available information regarding the dropout pattern of study participants.

Acknowledgements

We gratefully acknowledge Dr Ton de Craen and Dr Rudi Westendorp of the Leiden University Medical Centre, for kindly providing the analysed data. We are grateful to the Editor and two Referees whose comments and recommendations led to a substantial improvement of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Supplemental data and research materials

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/02664763.2015.1014887.
References

[1] M. Aitkin and M. Alfó, *Variance component models for longitudinal count data with baseline information: Epilepsy data revisited*, Stat. Comput. 16 (2006), pp. 31–38.

[2] M. Alfó and A. Maruotti, *A selection model for longitudinal binary responses subject to non-ignorable attrition*, Stat. Med. 28 (2009), pp. 2435–2450.

[3] P.B. Baltes, *Longitudinal and cross-sectional sequences in the study of age and generation effects*, Hum. Dev. 11 (1968), pp. 145–171.

[4] D. Bates, M. Maechler, B. Bolker, and S. Walker, *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, 2014a.

[5] D. Bates, M. Maechler, B.M. Bolker, and S. Walker, *lme4: Linear mixed-effects models using eigen and s4*, (2014b). Available at ArXiv e-print; submitted to *Journal of Statistical Software*.

[6] A. Creemers, N. Hens, M. Aerts, G. Molenberghs, G. Verbeke, and M.G. Kenward, *Generalized shared-parameter models and missingness at random*, Stat. Model. 11(4) (2011), pp. 279–310.

[7] R. Crum, J. Anthony, S. Bassett, and M. Folstein, *Population-based norms for the mini-mental state examination by age and educational level*, J. Am. Med. Assoc. 63 (1993), pp. 81–92.

[8] H. Demirtas and J.L. Schafer, *On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out*, Stat. Med. 22 (2003), pp. 2553–2575.

[9] A.B.-v. der Wiel, E. van Exel, A.J.M. de Craen, J. Gussekloo, A.M. Lagaay, D.L. Knook, and R.G.J. Westendorp, *A high response is not essential to prevent selection bias: Results from the Leiden 85-plus study*, J. Clin. Epidemiol. 55 (2002), pp. 1119–1125.

[10] C. Dufouil, C. Brayne, and D. Clayton, *Analysis of longitudinal studies with death and drop-out: A case study*, Stat. Med. 23 (2004), pp. 2215–2226.

[11] G. Fitzmaurice, N. Laird, and J. Ware, *Applied Longitudinal Analysis*, John Wiley and Sons, Hoboken, NJ, 2004.

[12] M.F. Folstein, S.E. Folstein, and P.R. McHugh, *Mini-mental state: A practical method for grading the cognitive state of patients for the clinician*, J. Psychiatry Res. 12 (1975), pp. 189–198.

[13] C.E. Frangakis, D.B. Rubin, M.-W. An, and E. MacKenzie, *Principal stratification designs to estimate input data missing due to death*, Biometrics 63 (2007), pp. 641–649.

[14] H.J. Gadda, C.P. Lima, and H. Amièva, *Semi-parametric latent process model for longitudinal ordinal data: Application to cognitive decline*, Stat. Med. 29 (2010), pp. 2723–2731.

[15] O. Harel and H. Demirtas, *Letter to the Editor: Joint modelling of missing data due to non-participation and death in longitudinal aging study*, Stat. Med. 30 (2011), pp. 2263–2265.

[16] S. Hofer and M. Sliwinski, *Design and analysis of longitudinal studies of aging*, in *Handbook of the Psychology of Aging*, J.E. Birren, K.W. Schaie, R.P. Abeles, M. Gatz, and T.A. Salthouse, eds., Elsevier Academic Press, Burlington, MA, 2006, pp. 15–37.

[17] J.W. Hogan and N.M. Laird, *Mixture models for the joint distribution of repeated measures and event times*, Stat. Med. 16 (1997), pp. 239–257.

[18] B.F. Kurland and P.J. Heagerty, *Directly parameterized regression conditioning on being alive: Analysis of longitudinal data truncated by deaths*, Biostatistics 6 (2005), pp. 241–258.

[19] B. Kurland, L.L. Johnson, B.L. Egleston, and P.H. Diehr, *Longitudinal data with follow-up truncated by death: Match the analysis method to research aims*, Stat. Sci. 24 (2009), pp. 211–222.

[20] F. Lagona and Z. Zhang, *A missing composite covariate in survival analysis: A case study of the Chinese longitudinal health and longevity survey*, Stat. Med. 29 (2010), pp. 248–261.

[21] S. le Cessie, E.G.E. de Vries, C. Buijs, and W.J. Post, *Analyzing longitudinal data with patients in different disease states during follow-up and death as final state*, Stat. Med. 28 (2009), pp. 3829–3843.

[22] Y. Lee and J.A. Nelder, *Statistical Analysis with Missing Data*, 2nd ed., Wiley, New York, 2002.

[23] R. Little and D. Rubin, *Statistical Analysis with Missing Data*, 2nd ed., Wiley, New York, 2002.

[24] K.B. Rajan and S.E. Leurgans, *Joint modeling of missing data due to non-participation and death in longitudinal aging studies*, Stat. Med. 29 (2010), pp. 2260–2268.

[25] K.W. Schaie, G.V. Labouvie, and T.J. Barrett, *Selective attrition effects in a fourteen-year study of adult intelligence*, J. Gerontol. 28 (1973), pp. 328–334.
[30] G. Verbeke, S. Fieuws, E. Lesaffre, B. Kato, M. Foreman, P. Broos, and K. Milisen, *A comparison of procedures to correct for base-line differences in the analysis of continuous longitudinal data: A case-study*, J. Roy. Statist. Soc. Ser. C 44 (2006), pp. 93–101.

[31] P. van Vliet, R.G. Westendorp, D. van Heemst, A.J.M. de Craen, and A.M. Oleksik, *Cognitive decline precedes late-life longitudinal changes in vascular risk factors*, J. Neurol. Neurosurg. Psychiatry 81 (2010), pp. 1028–1032.

[32] K.J. Wilkins and G.M. Fitzmaurice, *A marginalized pattern-mixture model for longitudinal binary data when nonresponse depends on unobserved responses*, Biostatistics 8 (2007), pp. 297–305.

[33] H.J. Woodford and J. George, *Cognitive assessment in the elderly: A review of clinical methods*, QJM 100 (2007), pp. 469–484.

[34] M.C. Wu and R.J. Carroll, *Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process*, Biometrics 44 (1988), pp. 175–188.