

Pattern Mixture Models and Latent Class Models for the Analysis of Multivariate Longitudinal Data with Informative Dropouts

Etienne Dantan, Cécile Proust-Lima, Luc Letenneur, Helene Jacqmin-Gadda

► **To cite this version:**

Etienne Dantan, Cécile Proust-Lima, Luc Letenneur, Helene Jacqmin-Gadda. Pattern Mixture Models and Latent Class Models for the Analysis of Multivariate Longitudinal Data with Informative Dropouts. International Journal of Biostatistics, De Gruyter, 2008, 4 (1), 10.2202/1557-4679.1088 . hal-03237519

HAL Id: hal-03237519

<http://hal.univ-nantes.fr/hal-03237519>

Submitted on 21 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The International Journal of Biostatistics

Volume 4, Issue 1

2008

Article 14

Pattern Mixture Models and Latent Class Models for the Analysis of Multivariate Longitudinal Data with Informative Dropouts

Etienne Dantan, *INSERM, U897, Center of Epidemiology
and Biostatistics of Bordeaux*

Cécile Proust-Lima, *Université Victor Segalen Bordeaux II*

Luc Letenneur, *INSERM, U897, Center of Epidemiology
and Biostatistics of Bordeaux*

Helene Jacqmin-Gadda, *INSERM, U897, Center of
Epidemiology and Biostatistics of Bordeaux*

Recommended Citation:

Dantan, Etienne; Proust-Lima, Cécile; Letenneur, Luc; and Jacqmin-Gadda, Helene (2008)
"Pattern Mixture Models and Latent Class Models for the Analysis of Multivariate Longitudinal
Data with Informative Dropouts," *The International Journal of Biostatistics*: Vol. 4: Iss. 1,
Article 14.

DOI: 10.2202/1557-4679.1088

Available at: <http://www.bepress.com/ijb/vol4/iss1/14>

©2008 Berkeley Electronic Press. All rights reserved.

Pattern Mixture Models and Latent Class Models for the Analysis of Multivariate Longitudinal Data with Informative Dropouts

Etienne Dantan, Cécile Proust-Lima, Luc Letenneur, and Helene Jacqmin-Gadda

Abstract

Missing data and especially dropouts frequently arise in longitudinal data. Maximum likelihood estimates are consistent when data are missing at random (MAR) but, as this assumption is not checkable, pattern mixture models (PMM) have been developed to deal with informative dropout. More recently, latent class models (LCM) have been proposed as a way to relax PMM assumptions. The aim of this paper is to compare PMM and LCM in order to tackle informative dropout in a longitudinal study of cognitive ageing measured by several psychometric tests. Using a multivariate longitudinal model with a latent process, a sensitivity analysis was performed to compare estimates under the MAR assumption, from a PMM and from two LCM. In the PMM, dropout patterns are included as covariates in the multivariate longitudinal model. In the simple LCM, they are predictors of the class membership probabilities while, in the joint LCM, the dropout time is jointly modeled using a proportional hazard model depending on latent classes. We show that parameter interpretation is different in the two kinds of models and thus can lead to different estimated values. PMM parameters are adjusted on the dropout patterns while LCM parameters are adjusted on the latent classes. This difference is highlighted in our data set because the latent classes exhibit much more heterogeneity than dropout patterns. We suggest several complementary analyses to investigate the characteristics of latent classes in order to understand the meaning of the parameters when using LCM to deal with informative dropout.

KEYWORDS: dropout, joint model, latent class model, missing data, mixed model, pattern mixture model

1 Introduction

The main objective of longitudinal studies is to describe change of an outcome over time. Missing responses and especially dropouts frequently hamper the analyses of longitudinal data. It is now well known that maximum likelihood estimators computed from the available data are consistent when missing data are ignorable (Verbeke, 2000), that is when data are Missing At Random (MAR) (Little, 1987) and parameters in the dropout model and in the response model are functionally independent. However, the MAR assumption cannot be tested since it requires that the missingness probability does not depend on the unobserved values of the outcome. When the response process and the missingness process are not independent, data are called Missing Not At Random (MNAR) or informative, and the analysis requires joint modeling of the two processes. To date, three approaches have been proposed: selection models (Diggle, 1994), pattern mixture models (Little, 1993) and latent class models (Roy, 2003). Based on different factorizations of the joint distribution, they rely on strong and uncheckable hypotheses on the missingness process and on the distribution of the unobserved outcome. As a consequence, it is generally recommended to perform a primary MAR analysis followed by a sensitivity analysis under various MNAR assumptions (Verbeke, 2001; Thijs, 2002).

Selection models factor the joint distribution into the marginal distribution of the outcome and the distribution of the missingness probability given the outcome. They are called either outcome-dependent when the dropout probability at time t depends directly on the unobserved outcome, or random-effect dependent (also referred to as shared parameter models) when the dropout probability depends on the random-effects from the mixed model for the outcome (Little, 1995). In the beginning, selection models raised enthusiasm because they allowed direct estimation of the parameters from the marginal distribution of the outcome that are of primary interest. The enthusiasm waned when it was shown that these estimates were very sensitive both to misspecification of the complete distribution of the outcome or to the assumed shape of the dependency between the dropout process and the outcome process (Kenward, 1998). These results increased the attractiveness of the main alternative approach, the Pattern Mixture Models.

Pattern Mixture Models (PMM) factor the joint distribution into the marginal distribution of the missingness process and the conditional distribution of the outcome given the missingness pattern. Thus, evolution of the outcome is described conditionally on time of dropout and marginal parameters are not directly available. Simple pattern mixture analyses may be performed

by stratifying on the dropout pattern or by including the dropout pattern as covariate in the model. In most cases, some patterns have too few subjects or too few measurements per subject and PMM require specification of constraints to ensure parameter identifiability. These constraints are parametric relationships (most often, equality) between parameters associated with different dropout patterns (Little, 1993; Molenberghs, 1998) or grouping of several patterns. PMM are often preferred to selection models in sensitivity analysis because the constraints are explicit and more meaningful than the assumptions required in selection models, and software is readily available (PMM are estimated with the same software as used under the MAR assumption).

More recently, several papers have proposed to tackle MNAR data in longitudinal studies using Latent Class Models (LCM) (Roy, 2003; Lin, 2004; Beunckens, 2007). As in PMM, the idea underlying LCM is that the population is a mixture of sub-populations with different profiles of outcome evolution. However, in PMM, the sub-populations are known *a priori* since they are defined by the dropout patterns (possibly grouping some patterns) while class membership is unobserved in LCM. It is data-driven assuming only an association with the dropout. LCM have been recommended as an alternative to PMM for data sets with numerous or sparse patterns because it is typically expected that LCM will include fewer latent classes than dropout patterns. This avoids identifiability issues and noisy estimates that are difficult to interpret (Roy, 2003; Lin, 2004). However, the well-defined constraints required for PMM identifiability are replaced in LCM by an assumption of conditional independence between dropout and responses, given the latent classes which can be only partially checked. In Roy (2003), dropout time is included as covariate in the class membership probability (this will be referred to afterwards as simple LCM). In Lin et al. (2004) and Beunckens et al. (2007), the missingness process is modeled jointly with the outcome process through latent classes (referred to as joint LCM). Thus, these joint latent class models may also be considered as random-effect dependent selection models with qualitative random-effects linking the two processes. Nevertheless, they have the notable advantage that the correlation between repeated measures of the outcome and the correlation between the outcome and the missingness process are modeled separately. However, LCM share the drawbacks of mixture models regarding possible local maxima in the likelihood (Redner, 1984) and somewhat unclear interpretation of the latent classes (Bauer, 2003). Although LCM have been presented as a relaxing of PMM, the two approaches have never been compared on real data sets except in Beunckens et al. (2007), who briefly presented estimates with the two approaches but did not discuss the different parameter interpretations.

In many longitudinal studies, repeated measures of several correlated outcomes are often collected. In some cases, these outcomes are so correlated that they may be viewed as several noisy markers of the same latent variable that is the actual variable of interest. When studying risk factors of cognitive ageing, for instance, association with the latent cognitive process is of greater interest than association with specific psychometric tests used to measure cognition. Recently, a latent process model has been proposed to analyze such multivariate longitudinal data (Proust, 2006) and has been applied to the Paquid cohort study for investigating nutritional risk factors in cognitive ageing under the MAR assumption (Letenneur, 2007). However, prospective cohorts of elderly subjects contain many dropouts that have been found associated with cognitive decline (Jacqmin-Gadda, 1997).

In the present paper, we conducted a sensitivity analysis of these results comparing pattern mixture and two latent class analyses (simple and joint LCM) to deal with dropouts. The latter was performed using latent class extension of the Proust et al. model for jointly modeling multivariate longitudinal data and time-to-event (Proust-Lima, submitted). The main goal of this work is to highlight and discuss the different interpretations of the parameters in these models that may lead to apparently inconsistent results.

In the next section, we present the data set. Section 3 is devoted to a brief description of the latent process model for multivariate longitudinal data (Proust, 2006). The three methods for handling informative dropout are reviewed in section 4. Data analysis results are presented in section 5 and the relative merits of the two approaches are discussed in section 6.

2 Data

Data come from the French prospective cohort study PAQUID on functional and cerebral aging (Letenneur, 1994). This cohort included at baseline 3,777 community dwellers aged 65 and over randomly recruited from electoral rolls in two administrative areas of southwest France. Participants were visited at home by a psychologist for the baseline interview in 1988-1989 and one (V1), three (V3), five (V5), eight (V8), ten (V10) and thirteen years (V13) after the baseline assessment. At each visit, a set of psychometric tests was administered, including evaluation of global mental status by the Mini-Mental State Examination (MMSE) (Folstein, 1975), visual memory by the Benton's Visual Retention Test (BVRT) (Benton, 1965) and verbal fluency by the Isaacs Set Test shortened to 15 seconds (IST) (Isaacs, 1973). Diagnosis of dementia was assessed using DSM-III R criteria (American Psychiatric Association, 1987).

The present analysis is based on a sub-sample of 1,640 non-demented subjects who participated in a nutritional study at the three-year follow-up (Letenneur, 2007) and completed at least one of the three psychometric tests at one of the next visits. Evolution of cognitive performance measured by the MMSE, the BVRT and the IST was analyzed over a 10-year period between the three-year follow-up (considered as the baseline visit, i.e. $t=0$) and the 13-year follow-up ($t=10$). For this sensitivity analysis, we focused only on educational level (without the first French diploma from primary school versus this diploma or higher level) and sex, two factors whose impact on cognitive decline in the elderly is still debated (Wiederholt, 1993; Elias, 1997; Le Carret, 2003). The analyses were adjusted for age in 4 age-groups [65-70[, [70-75[, [75-80[and $[\geq 80$].

In longitudinal studies, missing data can be monotone when a missing data is not followed by any measurement, or intermittent when a missing data is followed by at least one measurement. Monotone missing data are also called dropouts. We considered that one subject dropped out at time t if none of the three psychometric tests was completed at time t and until the end of the study. Five dropout profiles were defined: dropout at V5, V8, V10 and V13 and no dropout. In this analysis, we studied only the sensitivity to informative dropouts; intermittent missing data were taken to be ignorable. Characteristics of subjects according to dropout time are described in Table 1 as percentages.

3 The latent process model for multivariate longitudinal data

The analysis under the MAR assumption was carried out using the nonlinear latent process model for multivariate longitudinal outcomes proposed by Proust et al. (2006) which is outlined in figure 1 and briefly described below. Let define Y_{ijk} the outcome k measured at time t_{ijk} for subject i with $k = 1, \dots, K$, $i = 1, \dots, N$ and $j = 1, \dots, n_{ik}$. In our application, the outcomes are the 3 psychometric tests ($K=3$). We assume that a transformation of Y_{ijk} is a measure with error of a continuous latent process $\Lambda_i(t)$ representing the latent cognitive level in our application.

$$h_k(Y_{ijk}; \eta_k) = \Lambda_i(t_{ijk}) + \alpha_{ik} + \epsilon_{ijk} \quad (1)$$

where h_k is a flexible monotonous increasing transformation that depends on test-specific parameters η_k to be estimated and serves as a link function be-

Table 1: Distribution of characteristics of subjects in Paquid sample according to dropout patterns (%) and p-value for the Chi-square test of independence.

| | Visit of dropout | | | | | p-value |
|------------------------------------|------------------|---------------|----------------|----------------|-----------------------|---------|
| | V5 (N=236) | V8 (N=256) | V10 (N=173) | V13 (N=261) | No dropout (N=714) | |
| Intermittent missing data | 0.0 | 1.2 | 4.6 | 21.5 | 12.7 | |
| Diagnosed as demented ² | 0.4 | 8.2 | 9.3 | 7.7 | 15.4 | < 0.001 |
| Men | 49.6 | 43.4 | 54.3 | 39.8 | 37.2 | < 0.001 |
| No diploma ³ | 33.9 | 27.3 | 26.6 | 25.7 | 21.3 | 0.003 |
| Age at V3 | | | | | | < 0.001 |
| 65-70 | 4.7 | 9.0 | 7.5 | 8.1 | 16.9 | |
| 70-75 | 25.0 | 21.5 | 28.3 | 31.0 | 42.9 | |
| 75-80 | 19.5 | 22.3 | 22.6 | 24.9 | 26.9 | |
| ≥80 | 50.8 | 47.3 | 41.6 | 36.0 | 13.3 | |

¹ % of subjects with intermittent missing data among subjects dropped out at each visit (e.g. 4.6% of the 173 subjects dropped out at V10 had missed at least one visit before (at V3 or V5))

² before dropout

³ no first diploma from primary school

tween the measured score and the latent process. We chose the Beta cumulative density function (CDF) for h_k because it depends only on two parameters and offers large flexibility in the shape (Proust, 2006). As a Beta CDF is defined in $[0 - 1]$, each marker was rescaled to the unit interval. The subject-and-marker-specific random intercept α_{ik} was introduced to allow variability in the individual performance to each psychometric test conditionally on the latent process value. This random intercept and the independent error ϵ_{ijk} followed respectively the independent Gaussian distributions $\mathcal{N}(0, \sigma_{\alpha_k}^2)$ and $\mathcal{N}(0, \sigma_{\epsilon_k}^2)$.

Change over time of the latent process is described by a linear mixed model:

$$\Lambda_i(t) = X_{1i}^T(t)\beta + Z_i^T(t)u_i \quad (2)$$

where $X_{1i}(t)$ is a vector of possibly time-dependent covariates associated with the vector of fixed-effects β , $Z_i(t)$ is a subvector of $X_{1i}(t)$ and u_i is a vector of subject-specific random-effects with $\mathcal{N}(0, B)$ distribution, where B is an unstructured positive definite matrix. In our application, the latent process evolution is assumed to be a linear function of time.

The log-likelihood of the model may be decomposed as the log-likelihood for the transformed outcomes $\tilde{Y}_{ijk} = h_k(Y_{ijk}; \eta_k)$ plus the jacobian $J(y_i, \theta)$ of the transformations:

$$l(\theta) = \sum_{i=1}^N f(\tilde{y}_i; \theta) + \sum_{i=1}^N \ln(J(y_i, \theta)) \quad (3)$$

where $\tilde{y}_i = (\tilde{y}_{i11}, \dots, \tilde{y}_{in_{i1}1}, \dots, \tilde{y}_{in_{iK}K})^T$.

Maximum likelihood estimators of the whole set of parameters θ are computed using the Marquardt algorithm (Marquardt, 1963), a Newton-Raphson-like algorithm. The estimation procedure was implemented in a Fortran90 program (Proust, 2006). The code for the estimation the latent process model for multivariate longitudinal data and its user's guides files are available at the following url: <http://biostat.isped.u-bordeaux2.fr> (program NLMULTIMIX).

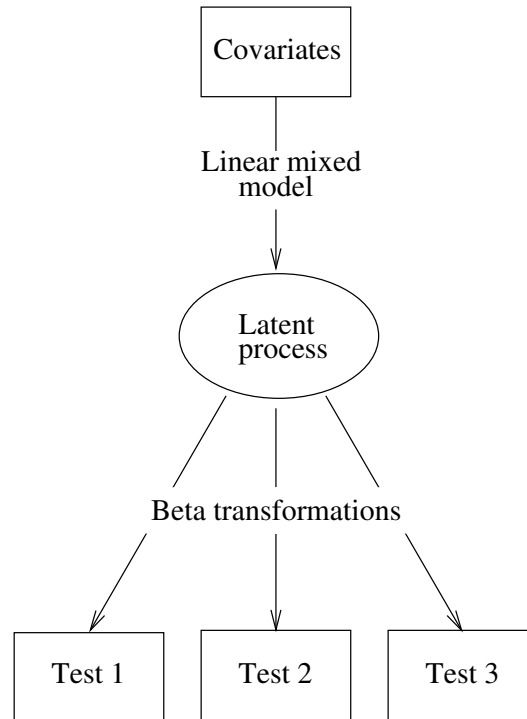


Figure 1: Diagram of the latent process model for multivariate longitudinal data estimated under the MAR assumption.

4 Methods for handling informative dropout

4.1 Pattern mixture approach

PMM factorizes the joint distribution of the response variable Y and the dropout process D given covariates X as $[Y, D|X] = [Y|D, X][D|X]$. The dropout distribution $[D|X]$ is often estimated as the proportion of each dropout pattern for each combination of covariate values, while the estimation of the conditional distribution $[Y|D, X]$ requires a model for the evolution of Y given dropout patterns D and covariate X , which can be estimated separately by maximum likelihood. With the primary model defined by (1) and (2) and denoting $D_{il} = 1$ if subject i dropped out at visit V_l , a flexible model for $[Y|D, X]$ is:

$$\Lambda_i(t) | D_{il} = 1 = X_{1i}^T(t)\beta_l + Z_i^T(t)u_{il} \quad \text{with } u_{il} \sim \mathcal{N}(0, B_l) \quad (4)$$

All the parameters are pattern-specific and estimates are obtained through stratified analyses on the dropout patterns. Some dropout patterns can be gathered if necessary to ensure parameter identifiability. However, this leads to a very large number of parameters that are estimated on subsamples of unequal sizes and unequal numbers of measurements per subject, thereby inducing a large variability. More parsimonious models are obtained by introducing the dropout patterns as covariates in the primary model and assuming that some parameters are common over the patterns. Some authors have given examples of meaningful assumptions to help to specify the constraints (Little, 1993; Kenward, 1998). As these models are nested in the stratified model, model selection may be guided by the likelihood ratio test. In the present work, estimates from the PMM were obtained using the estimation program developed for the latent process model.

4.2 Simple latent class model

The latent class model assumes that the population is divided into G unobserved sub-populations with distinct profiles of evolution for the latent process. Denoting C the latent class variable, the joint distribution $[Y, D|X]$ is decomposed as $[Y, D|X] = \sum_C [Y|C, X][C|D, X][D|X]$. As for PMM, $[D|X]$ may be estimated empirically while parameters from $[Y|C, X][C|D, X]$ are estimated by maximum likelihood conditioning on dropout patterns. Note that interpretation of the regression parameters in the conditional distribution $[Y|C, X]$ becomes unclear if the same covariates are included in the class membership

probability $[C|D, X]$. Thus, in the following, we assumed $[C|D, X] = [C|D]$. We return to this assumption in section 4.4.

More specifically, by denoting $c_{ig}(g = 1, \dots, G)$ the latent class membership variables, which equal 1 if subject i belongs to latent class g and 0 otherwise, the latent process evolution is defined given the latent class by:

$$\Lambda_i(t) |_{c_{ig}=1} = X_{1i}^T(t)\beta_g + Z_i^T(t)u_{ig}, \quad t \geq 0 \quad (5)$$

$$u_{ig} \sim \mathcal{N}(0, \omega_g^2 B)$$

Similarly to PMM, more parsimonious models are obtained by assuming that some parameters are common over classes. As in Roy (2003), the class membership probability is defined using a multinomial logistic regression with the dropout pattern included as a categorical variable:

$$\pi_{ig} = P(c_{ig} = 1|D_i) = \frac{e^{\xi_{0g} + D_i^T \xi_{1g}}}{\sum_{l=1}^G e^{\xi_{0l} + D_i^T \xi_{1l}}} \quad (6)$$

where ξ_{0g} is the intercept for class g and ξ_{1g} is the vector of class-specific parameters associated with the vector of indicator variables for dropout patterns: $D_i^T = (D_{i5}, D_{i8}, D_{i10}, D_{i13})$ (reference: no dropout). For identifiability, $\xi_{01} = 0$ and $\xi_{11} = 0$.

The log-likelihood of the simple LCM is:

$$l(\theta) = \sum_{i=1}^N \ln \left(\sum_{g=1}^G \pi_{ig} f(\tilde{y}_i | c_{ig} = 1; \theta) \right) + \sum_{i=1}^N \ln(J(y_i; \theta)) \quad (7)$$

4.3 Joint latent class model

In the joint latent class model, the joint distribution $[Y, D|X]$ is decomposed as $[Y, D|X] = \sum_C [Y|C, X][D|C, X][C|X]$. The dropout process is thus jointly modeled and depends on Y through the latent class variable. Parameters for the 3 distributions are estimated simultaneously by maximizing the joint likelihood. As for the simple LCM, the model for $[Y|C, X]$ is defined by (1) and (5). To facilitate parameter interpretation in $[Y|C, X]$, we do not include any covariate in the class membership probability ($[C|X] = [C]$).

In the present work, we model the time to dropout using a proportional hazard model. Let define δ_i the dropout indicator that equals 1 if subject i dropped out and 0 if subject i was seen at the last visit V13. The dropout time is not observed continuously but only at discrete visit times. We denote T_{oi} the time at the last observation and T_{mi} the time at the first missing value

(next planned visit). Thus, if $\delta_i = 0$, $T_{oi} = 10$. The risk function may be defined as follows:

$$\lambda(t \mid c_{ig} = 1, X_{2i}; \gamma_{0g}, \gamma_1) = \lambda_0(t)e^{\gamma_{0g} + X_{2i}\gamma_1} \quad \text{with } \gamma_{01} = 0 \quad (8)$$

where γ_{0g} is the logarithm of the relative risk of dropout in latent class g compared to latent class 1 adjusted for covariates X_{2i} . The vector of covariates X_{2i} is associated with the vector of parameters γ_1 . In this work, we assume that the impact of X_{2i} on the risk of dropout does not depend on latent classes. A 5-step function was used for the baseline risk function $\lambda_0(t)$.

To make sure that probabilities remain in $[0, 1]$, the class membership probability is defined by:

$$\pi_{ig} = P(c_{ig} = 1) = \frac{e^{\xi_{0g}}}{\sum_{l=1}^G e^{\xi_{0l}}} \quad \text{with } \xi_{01} = 0 \quad (9)$$

The individual contribution to the likelihood is decomposed according to the latent classes using the conditional independence of dropout time and outcomes. Then the density $f(y_i \mid c_{ig} = 1)$ is computed using the Jacobian of the Beta transformations as in (3). Thus, the log-likelihood is obtained by:

$$\begin{aligned} l(\theta) = & \sum_{i=1}^N \ln \left(\sum_{g=1}^G \pi_{ig} f(\tilde{y}_i \mid c_{ig} = 1; \theta) [S(T_{oi} \mid c_{ig} = 1) - S(T_{mi} \mid c_{ig} = 1)]^{\delta_i} \right. \\ & \left. \times S(T_{oi} \mid c_{ig} = 1)^{1-\delta_i} \right) + \sum_{i=1}^N \ln(J(y_i; \theta)) \end{aligned} \quad (10)$$

where S denotes the survival function. For both simple and joint latent class models, parameters were estimated by maximizing (7) or (10) for several numbers of latent classes G by using the Marquardt algorithm (Proust-Lima, submitted; Marquardt, 1963) The optimal number of classes was selected according to the Bayesian Information Criterion (BIC) (Schwartz, 1978). As local maxima are possible with mixture models, each model was estimated by using different sets of initial parameters to ensure convergence to the global maximum. When parameters are estimated, subjects may be classified in the latent classes using the posterior probabilities to belong to each latent class given the data. These probabilities are computed using the Bayes theorem. The estimation procedure was implemented in Fortran90 and codes for methods dealing with informative dropouts can be provided on request.

4.4 Parameter interpretation

In PMM and LCM, parameters are estimated from the conditional distribution $[Y|D, X]$ or $[Y|C, X]$. These parameters provide interesting information about the heterogeneity of the outcome distribution, but in most cases their interpretation is different from those in the marginal distribution $[Y|X]$. The latter may be computed by using

$$\left. \begin{array}{l} \text{PMM: } [Y|X] = \sum_D [Y|D, X][D|X] \\ \text{Simple LCM: } [Y|X] = \sum_C [Y|C, X][C|X] = \sum_C [Y|C, X] \sum_D [C|D][D|X] \\ \text{Joint LCM: } [Y|X] = \sum_C [Y|C, X][C] \end{array} \right\} \quad (11)$$

To emphasize the different interpretations, let assume a conditional model for Y including one covariate, X_{11i} , with a common effect over the groups (defined by dropout patterns in PMM or latent classes in LCM) and one covariate X_{12i} with a group-specific effect. Whatever the kind of model (PMM or LCM), the group-specific expectation of the latent process in group g is:

$$E(\Lambda_i | X_{11i}, X_{12i}, g) = X_{11i}\beta_1 + X_{12i}\beta_{2g}. \quad (12)$$

Formula (12) shows that β_{2g} is a measure of the impact of X_{12i} in each group and β_1 is a measure of the impact of X_{11i} adjusted for the differential effect of X_{12i} across the groups. For instance, if X_{12i} is a vector of 1, β_1 is the effect of X_{11i} adjusted for the group difference on the intercept. The meaning of β_1 depends on the meaning of the groups, so it is different in a PMM where the groups are the dropout patterns and in a LCM where the groups are the latent classes.

To obtain the marginal effect of X_{11i} adjusted for the other covariates, the marginal expectation needs to be computed:

$$E(\Lambda_i | X_{11i}, X_{12i}) = X_{11i}\beta_1 + X_{12i} \sum_g \beta_{2g} P(g | X_{11i}, X_{12i}) \quad (13)$$

and the marginal effect of X_{11i} is:

$$\begin{aligned} E(\Lambda_i | X_{11i} = 1, X_{12i}) - E(\Lambda_i | X_{11i} = 0, X_{12i}) &= \beta_1 + \\ X_{12i} \times \sum_g \beta_{2g} (P(g | X_{11i} = 1, X_{12i}) - P(g | X_{11i} = 0, X_{12i})) &\quad (14) \end{aligned}$$

Thus, β_1 can be interpreted as the marginal effect of X_{11i} only if the group membership probability does not depend on X_{11i} . This probability is $P(D_{ig} = 1|X_i)$ in the PMM, $P(c_{ig} = 1|X_i) = \sum_l Pr(c_{ig} = 1|D_{il} = 1)P(D_{il} = 1|X_i)$ in the simple LCM and $P(c_{ig} = 1|X_i)$ in the joint LCM. As a consequence, even if the estimation of conditional parameters in PMM and LCM does not require the specification of $[D|X]$, additional assumptions about the dropout pattern distribution are required for marginal inference. More specifically, parameters common over classes retain the marginal interpretation only under the strong and generally violated assumption that the dropout probabilities do not depend on X . Furthermore, common parameters in a joint LCM retain the marginal interpretation only when the class membership probability does not depend on X . In the other cases, the difference in the marginal means (13) depends on the value of all the covariates with group-specific effects.

5 Application

5.1 Analysis under the MAR assumption

Parameters from the model for multivariate longitudinal data defined by (1) and (2) including time from the baseline visit (V3), sex, educational level, age and the 3 interactions with time as explanatory variables were estimated under the MAR assumption. Estimates are given in Table 2.

The difference between men and women with regard to initial cognitive level was not significant ($p = 0.07$), but men had a slower cognitive decline than women ($p = 0.01$). Subjects without a diploma had a lower initial level of cognition ($p < 0.01$), but the impact on cognitive evolution was not significant ($p = 0.56$). Older subjects had a lower initial cognitive level ($p < 0.01$) and a sharper cognitive decline ($p < 0.01$).

5.2 Pattern mixture models

Pattern mixture analyses were carried out by putting together subjects dropped out at V5 (with only one measure) and V8 to ensure parameter identifiability. This gave 4 dropout patterns (V5/V8, V10, V13, no dropout). The primary model stratified on the 4 dropout patterns, that is with all parameters specific to the pattern (4×27 parameters), was estimated and the log-likelihood was computed by summing the log-likelihoods of the 4 pattern-specific models. Then, we sought a more parsimonious model including dropout pattern and interactions, with dropout pattern as covariates in the mixed model for the la-

Table 2: Estimates from the model for multivariate longitudinal data under the MAR assumption and in the PMM, simple LCM and joint LCM.

| | MAR | | PMM | | Simple LCM | | Joint LCM | |
|----------------------|---------|-------|---------|-------|------------|-------|-----------|-------|
| | β | se | β | se | β | se | β | se |
| intercept | 0.643 | 0.009 | 0.659 | 0.009 | | | | |
| t ¹ | -0.077 | 0.011 | -0.071 | 0.010 | | | | |
| class 1 | | | | | 0.683 | 0.009 | 0.696 | 0.010 |
| class 1 * t | | | | | -0.045 | 0.009 | -0.041 | 0.010 |
| class 2 | | | | | 0.628 | 0.010 | 0.624 | 0.016 |
| class 2 * t | | | | | -0.166 | 0.034 | -0.099 | 0.021 |
| class 3 | | | | | 0.638 | 0.019 | 0.641 | 0.019 |
| class 3 * t | | | | | -0.551 | 0.004 | -0.494 | 0.044 |
| Men | 0.008 | 0.005 | 0.012 | 0.004 | 0.008* | 0.004 | 0.004* | 0.004 |
| Men * t | 0.021 | 0.008 | 0.024 | 0.008 | 0.012* | 0.006 | 0.008* | 0.006 |
| No Diploma | -0.104 | 0.005 | -0.103 | 0.005 | -0.096 | 0.005 | -0.096 | 0.005 |
| No Diploma * t | -0.006* | 0.010 | -0.006* | 0.009 | -0.016 | 0.008 | -0.018 | 0.008 |
| Age ² | | | | | | | | |
| 70-75 | -0.014* | 0.008 | -0.011* | 0.007 | -0.009* | 0.006 | -0.011* | 0.007 |
| 75-80 | -0.048 | 0.008 | -0.041 | 0.008 | -0.036 | 0.007 | -0.039 | 0.007 |
| ≥80 | -0.098 | 0.008 | -0.078 | 0.008 | -0.073 | 0.007 | -0.084 | 0.007 |
| Age * t ² | | | | | | | | |
| 70-75 * t | -0.023* | 0.012 | -0.019* | 0.011 | -0.014* | 0.009 | -0.017* | 0.009 |
| 75-80 * t | -0.067 | 0.013 | -0.058 | 0.012 | -0.044 | 0.010 | -0.049 | 0.010 |
| ≥80 * t | -0.117 | 0.014 | -0.094 | 0.014 | -0.067 | 0.011 | -0.085 | 0.011 |
| Dropout ³ | | | | | | | | |
| V5/V8 | | | -0.047 | 0.006 | | | | |
| V10 | | | -0.022 | 0.007 | | | | |
| V13 | | | -0.023 | 0.007 | | | | |
| Dropout * t | | | | | | | | |
| V5/V8 * t | | | -0.081 | 0.026 | | | | |
| V10 * t | | | -0.048 | 0.016 | | | | |
| V13 * t | | | -0.056 | 0.012 | | | | |
| Beta parameters | 1.173 | 0.028 | 1.146 | 0.029 | 0.983 | 0.034 | 0.985 | 0.034 |
| | -2.690 | 0.037 | -2.665 | 0.037 | -2.496 | 0.038 | -2.497 | 0.038 |
| | 0.325 | 0.021 | 0.306 | 0.021 | 0.199 | 0.026 | 0.202 | 0.025 |
| | -2.389 | 0.025 | -2.383 | 0.025 | -2.315 | 0.024 | -2.318 | 0.024 |
| | 0.470 | 0.024 | 0.445 | 0.024 | 0.320 | 0.028 | 0.323 | 0.028 |
| | -2.377 | 0.029 | -2.361 | 0.028 | -2.264 | 0.028 | -2.266 | 0.028 |

¹ t: unit = 10 years

² Likelihood Ratio Test with 3 df significant at $\alpha = 0.01$, for each approach

³ Reference: No dropout

* Non significant at $\alpha = 0.05$

tent process (4) and allowing pattern-specific covariance matrix for the random effects. According to the likelihood ratio statistic, none of the interactions between the dropout patterns and the covariates (age, sex and educational level) was significant, but the random-effect variance was significantly different over dropout patterns. Thus, the best model included dropout pattern as a simple effect and with an interaction with time. Estimates of fixed effect are displayed in Table 2.

The likelihood ratio test revealed that dropout patterns were significantly associated with cognition ($\chi^2 = 187.4$ for 6 df, $p < 0.01$): early dropout was associated with a lower initial cognitive level and a sharper cognitive decline. Subjects dropped out at V10 and V13 presented similar profiles of cognitive evolution. Regarding the covariates of interest (sex, educational level and age), PMM results were close to those obtained under the MAR assumption, except for the sex effect on initial level which appeared to be more significant: men exhibited a significantly higher mean initial level than women ($p = 0.01$).

5.3 Latent class models

The two LCM were estimated by including the same covariates as under the MAR assumption (age, sex, educational level and their interaction with time) in the mixed model for the latent process (5) and assuming class-specific intercepts and class-specific slopes with time. Thus, the mixed model in the LCM was identical to the mixed model in the PMM estimated above, except that dropout patterns were replaced by latent classes. We successively estimated models with 1, 2, 3 and 4 latent classes. The 3-latent class model was selected according to the BIC.

Table 3 displays parameter estimates of the multinomial logistic model for class membership probabilities defined by (6) for the simple LCM. Dropout patterns were globally significantly associated with the latent classes ($\chi^2 = 118.7$ for 8 df, $p < 0.0001$). Subjects who did not drop out from the study had a very low probability to be in class 2 or 3: the odds (or ratio of the probabilities) of being in class 2 (respectively 3) compared to class 1 was $\exp(-2.08) = 0.12$ (respectively $\exp(-3.61) = 0.027$). Moreover, the odds ratio of class 2 versus 1 clearly increased as the time of follow-up decreased compared to subjects without dropout. The odds for being in class 3 versus 1 was maximum for subjects dropped out at V8 and V13.

In the joint LCM, dropout profiles were no longer included as covariates, but the risk of dropout was modeled jointly using a proportional hazard model (8) as a function of the latent classes and covariates. Estimates from this proportional hazard model are displayed in Table 4. The risk of dropout was globally associated with the latent classes ($\chi^2 = 72.2$ for 2 df, $p < 0.0001$). More specifically, adjusted for age, sex and educational level, the risk of dropout was higher in class 2 (hazard ratio=2.98, $p < 0.01$) and, to a lesser extent, in class 3 compared to class 1 (hazard ratio=2.34, $p < 0.01$). Men ($p < 0.01$), subjects with low educational level ($p = 0.04$), and older subjects ($p < 0.01$) had a higher risk of dropout.

Estimates from the model for multivariate longitudinal data in the two

Table 3: Estimates from the multinomial logistic model for the class membership probabilities in the simple LCM.

| Covariates | Estimate | Standard error | p-value |
|--------------------------------------|----------|----------------|---------|
| Probability to be in Class 2: | | | |
| Intercept | -2.08 | 0.41 | < 0.001 |
| Dropout at V5 ¹ | ∞ | — | — |
| Dropout at V8 | 2.21 | 0.45 | < 0.001 |
| Dropout at V10 | 1.74 | 0.40 | < 0.001 |
| Dropout at V13 | 1.74 | 0.31 | < 0.001 |
| Probability to be in Class 3: | | | |
| Intercept | -3.62 | 0.33 | < 0.001 |
| Dropout at V5 | -0.75 | 14.82 | 0.960 |
| Dropout at V8 | 2.50 | 0.46 | < 0.001 |
| Dropout at V10 | 0.93 | 0.68 | 0.177 |
| Dropout at V13 | 1.66 | 0.40 | < 0.001 |

¹ Reference: No dropout

Table 4: Estimates from the proportional hazard model for dropout in the joint LCM.

| | Estimate | Standard error | p-value |
|-------------------------|----------|----------------|---------|
| Class 2 ¹ | 1.094 | 0.157 | < 0.001 |
| Class 3 ¹ | 0.852 | 0.239 | < 0.001 |
| Men ² | 0.432 | 0.075 | < 0.001 |
| No Diploma ³ | 0.174 | 0.083 | 0.036 |
| Age ^{4,5} | | | |
| 70-75 | 0.258 | 0.147 | 0.079 |
| 75-80 | 0.483 | 0.151 | < 0.001 |
| ≥ 80 | 1.349 | 0.148 | < 0.001 |

¹ Reference: Class 1

² Reference: Women

³ Reference: Diploma

⁴ Reference: 65-70 years

⁵ Likelihood Ratio Test with 3 df significant at $\alpha = 0.001$

LCM are displayed in Table 2. Regarding covariates of interest, conclusions of the PMM and the two LCM were qualitatively similar for the effect of age and educational level on initial level. However, there were large differences regarding the impact of educational level on cognitive decline: educational level was not associated with cognitive decline in the PMM ($\beta = -0.006, p = 0.49$) as under the MAR assumption ($\beta = -0.006, p = 0.56$), while both LCM revealed that subjects without diploma had a significantly sharper decline ($\beta = -0.016, p = 0.03$ in the simple LCM and $\beta = -0.018, p = 0.02$ in the joint LCM). Conclusions about sex effect also differed according to the analysis. The association between sex and initial cognitive level was not significant (except in the PMM where $\beta = 0.012, p = 0.01$ and the effect was close to significance for simple LCM ($\beta = 0.008, p = 0.006$)). Men exhibited a significantly slower decline under the MAR assumption ($\beta = 0.021, p = 0.01$) and in the PMM analyses ($\beta = 0.024, p < 0.01$) which was marginally significant in the simple LCM ($\beta = 0.012, p = 0.05$) but not significant in the joint LCM ($\beta = 0.008, p = 0.19$). In general, the two LCM led to similar estimates for the association with covariates except that the sex effect tended to be more pronounced in the simple LCM.

5.4 Description of the latent classes

As explained in section 4.4, regression parameters in LCM were adjusted for the latent classes. To investigate the differences between the latent classes and the dropout patterns, Figure 2 shows the mean evolution of the latent cognitive process for each dropout pattern predicted by the PMM and the mean evolution in each latent class estimated from the two LCM. In the PMM, there were only slight differences of evolution between dropout patterns, whereas in the LCM the three classes displayed dramatically different profiles. From both LCM, subjects in class 1 had a high initial cognitive level and a very slow decline over time. This class represented about 62.9% of the sample in the simple LCM and 69.1% in the joint LCM. Subjects in class 2 had a low initial cognitive level but their decline was hardly more pronounced. In the simple and joint LCM respectively, 33.4% and 27.2% of subjects were *a posteriori* classified in class 2. The initial cognitive level in class 3 was close to that in class 2, but these subjects exhibited a dramatic decline over time. With both LCM, 3.7% of subjects belonged to class 3. Figure 2 clearly shows that the latent classes exhibit much more heterogeneity than the dropout patterns.

Comparison of subject characteristics from the 5 dropout patterns (Table 1) with those of subjects *a posteriori* classified in the 3 classes by the two LCM (Table 5) clarifies the different evolution profiles. Chi-square tests

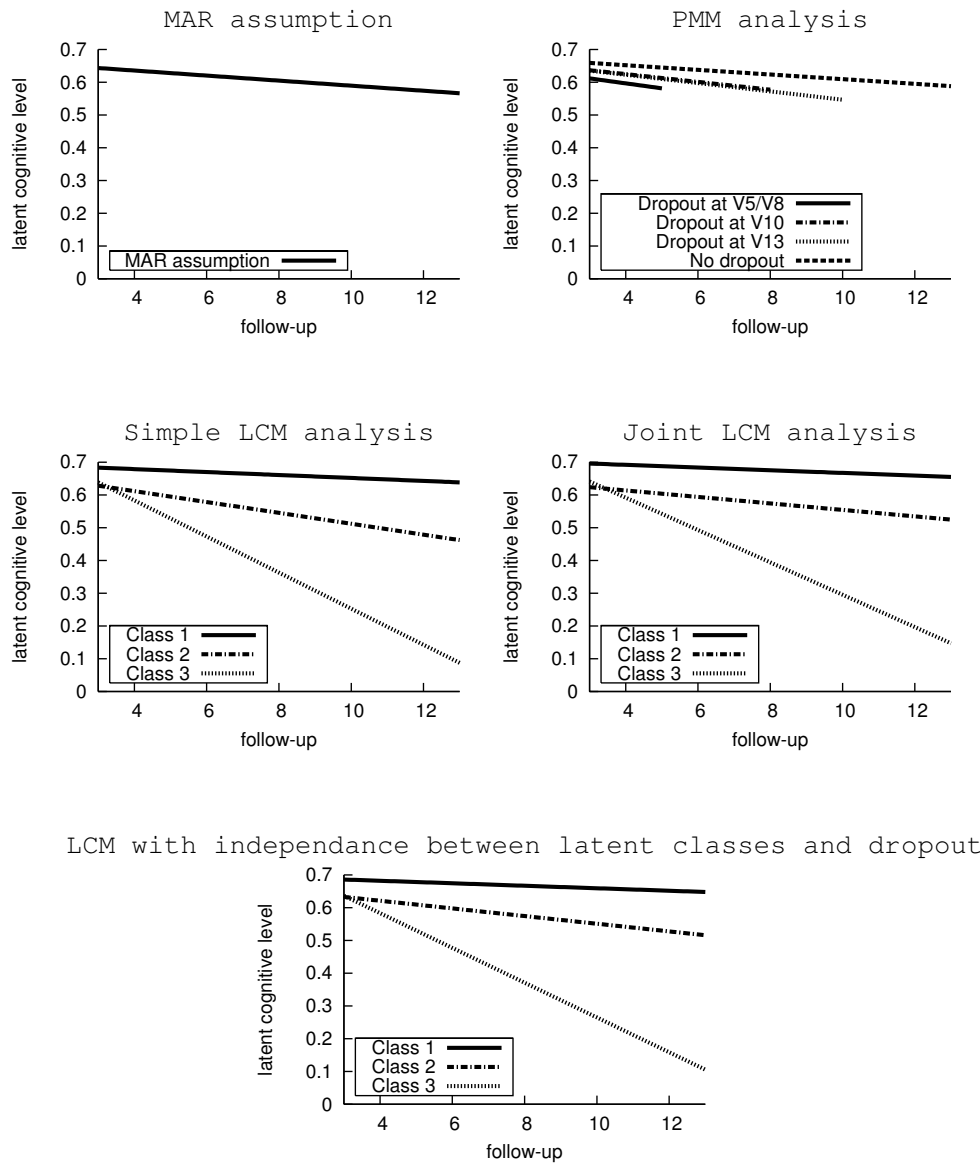


Figure 2: Estimated mean evolution of the latent process under the MAR assumption, given the dropout patterns by the PMM analysis, given the latent classes by the simple LCM, the joint LCM and the LCM with independence between the latent classes and the dropout for women with diploma and aged between 65 and 70 years-old.

Table 5: Characteristics of the subjects according to the posterior latent classes for the simple LCM and the joint LCM (%).

| | Class 1 | Class 2 | Class 3 | p-value |
|-------------------|----------|---------|---------|---------|
| Simple LCM | (N=1032) | (N=548) | (N=60) | |
| Dropout at | | | | < 0.001 |
| V5 | 0.0 | 43.0 | 0.0 | |
| V8 | 11.7 | 21.2 | 31.7 | |
| V10 | 10.7 | 10.6 | 8.3 | |
| V13 | 14.5 | 16.6 | 33.3 | |
| No dropout | 63.1 | 8.6 | 26.7 | |
| Demented | 7.4 | 10.6 | 56.7 | < 0.001 |
| Women | 58.4 | 54.7 | 75.0 | 0.009 |
| Diploma | 76.6 | 70.3 | 81.7 | < 0.001 |
| Joint LCM | (N=1023) | (N=556) | (N=61) | |
| Dropout at | | | | < 0.001 |
| V5 | 5.8 | 31.8 | 0.0 | |
| V8 | 10.5 | 24.7 | 18.0 | |
| V10 | 9.1 | 13.1 | 11.5 | |
| V13 | 14.4 | 16.7 | 34.4 | |
| No dropout | 60.2 | 13.7 | 36.1 | |
| Demented | 7.4 | 10.3 | 57.4 | < 0.001 |
| Women | 56.0 | 58.8 | 78.7 | 0.002 |
| Diploma | 76.7 | 70.3 | 80.3 | 0.012 |

showed that dementia diagnosis, sex, educational level and age were significantly associated with dropout patterns ($p < 0.01$ for each covariate). There was more dementia diagnosis among subjects with longer follow-up. Men, subjects without diploma and older subjects appeared to drop out earlier.

The posterior classification obtained with the simple LCM (upper part of Table 5) confirmed that class 1 corresponded to subjects with a long follow-up (63% of subjects did not drop out), while class 2 was associated with the shortest follow-up profile (64.2% dropped out at V5 or V8) and class 3 had a medium follow-up (nobody dropped at V5 but only 26.7% did not drop out). The percentage of positive diagnosis of dementia was relatively low in classes 1 and 2 but was superior to 50% in class 3. The sex distribution was similar for classes 1 and 2, but there was an imbalance in class 3 which included 75% of women. Although the difference was less glaring for diploma distribution, the percentage of qualified subjects was higher in class 3 (81%). Table 5 shows that the latent classes obtained by the two LCM had similar characteristics.

In addition, more than 87% of subjects were classified in the same class by the two LCM. Thus, the two latent class analyses highlight heterogeneous profiles of cognitive evolution in our sample that are only partially associated with the dropout patterns but highly associated with dementia diagnosis.

5.5 Conditional independence assumption

The joint latent class model is based on the assumption that cognitive evolution and dropout time are conditionally independent in view of the latent classes (Lin, 2004; Roy, 2007). This assumption was evaluated by estimating the association between cognitive scores and dropout time after adjustment for the posterior latent classes in the longitudinal model. The strength of association when adjusting for posterior classes was markedly reduced ($\chi^2 = 22.6$ (6 df) with adjustment versus $\chi^2 = 94.5$ (6 df) without), though it remained highly significant ($p < 0.001$). This suggested that the conditional independence assumption was not valid.

5.6 LCM assuming independence with dropout

We compared the latent classes obtained from the joint LCM with those obtained with the same model but assuming independence between latent classes and dropout ($\gamma_{0g} = 0, \forall g$ in (8)). The estimated cognitive evolutions in the 3 latent classes identified by the two models were close (Figure 2). Moreover, 70% of the subjects were classified in the same class by both models. Thus, the assumption regarding the link between cognitive decline and dropout had little impact on the latent classes in this data set. This result confirms that latent classes reflect the heterogeneity of cognitive evolution rather than the association between dropout and cognition.

5.7 Estimated Beta transformations

As suggested by a referee, the Beta CDF parameters could influence parameters of the longitudinal model. By comparing Beta CDF parameters in table 2, we observed slight differences between each approach. However, these differences had little impact on the shape of the estimated transformations as displayed in figure 3, so it was unlikely that they influenced estimates of regression parameters.

Nevertheless, to be sure that these slight differences did not explain the discrepancies between PMM and LCM due to covariate effects, we re-estimated

the 3 models by handling informative dropouts, with the Beta CDF parameters set at the values estimated in the MAR analysis. Estimates were close to those obtained in table 2 while the differences between the 3 methods in the estimated sex and educational level effects remained (results not shown). Finally, we computed the estimated marginal evolution of the latent process and each cognitive score in its natural scale for men and women aged between 65 and 70 with high educational level. As the Beta transformation was not linear, we computed $E\{h_k^{-1}(\tilde{Y})|g\}$ for each dropout pattern or each latent class g by simulation, as explained in Proust-Lima et al (2006). Then, the marginal expectations were obtained as the mean of the group-specific expectation weighted by the proportion of each pattern or each class given the covariates (11). Figure 4 shows that the trend is similar for the latent process and for the cognitive scores on their natural scales.

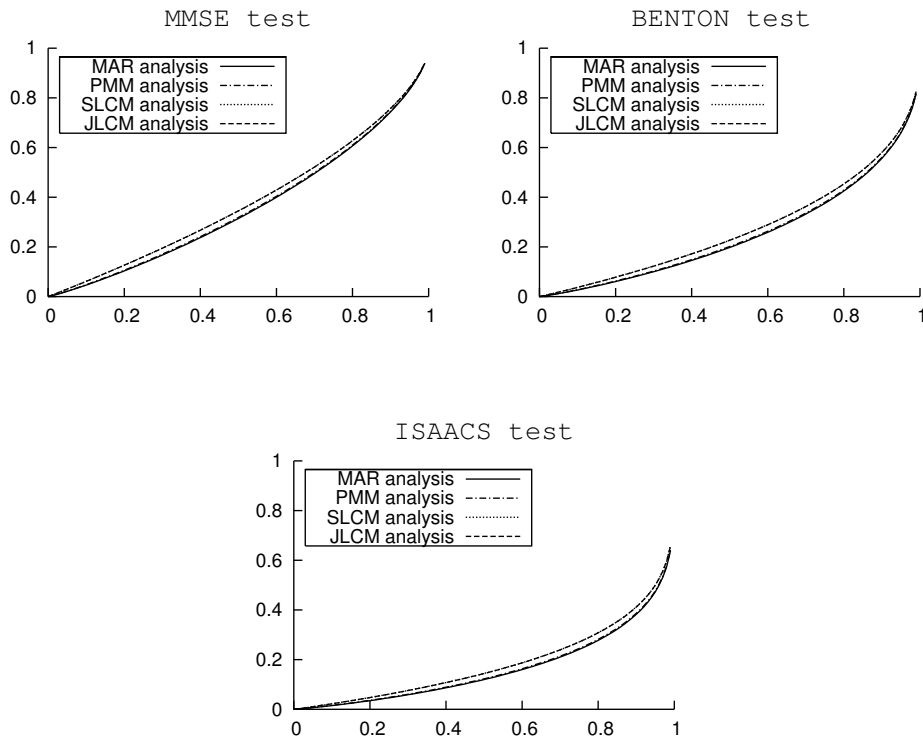


Figure 3: Beta distribution of each psychometric test under the MAR assumption and in the PMM, simple LCM and joint LCM.

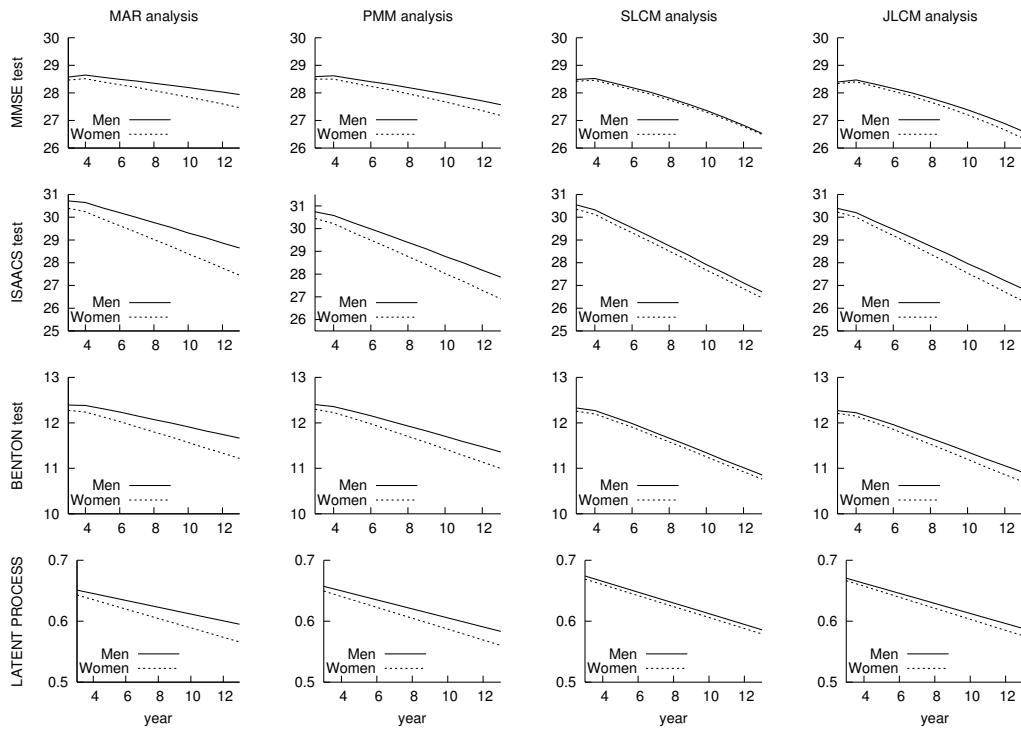


Figure 4: Comparison of men and women evolutions of the 3 psychometric tests and the latent evolution for each approach, for subjects with high educational level and aged between 65 and 70.

6 Discussion

By comparing three methods (PMM and two kinds of LCM) to account for informative dropout in a longitudinal study of cognitive aging, we show that LCM and PMM may lead to different parameter estimates due to different interpretations of the parameters. In the PMM, however the parameters are interpreted with adjustment for the dropout pattern. In the two LCM, the parameters are adjusted for the latent classes, which may reflect more than dropout patterns.

For instance, under the MAR assumption, we found that men tended to have a better initial cognitive level and a slower decline than women, and that these differences were reinforced when adjusting for dropout pattern in the PMM. Indeed, since the dropout rate was higher for men and for subjects with low cognitive level (see Table 1), the difference between men and women who dropped out at the same time tended to be higher than the unadjusted difference.

On the other hand, as the latent class approaches provide parameters adjusted for the latent classes, it is essential to perform complementary analyses to investigate the characteristics of the latent classes in order to better understand the meaning of the estimated parameters. We suggest several analyses to describe the classes and evaluate the link with dropout patterns. In the Paquid data set, the three classes discriminated subjects who kept a high cognitive level over the 10 years of follow-up and had a low rate of dropout, subjects with a poor initial cognitive level who tended to drop out earlier from the study and, those with pathological decline. The latter group included more than 50% of subjects diagnosed as demented before the end of their follow-up and 75% of women. This result is in agreement with previous findings showing a higher risk of dementia among women in the oldest ages (Commenges, 1998). Thus, it is clear that these 3 classes do not only represent dropout patterns but are the consequence of other sources of heterogeneity in the data. In the joint LCM, when adjusting for these 3 classes and especially on the pathological class, which is highly associated with sex, we found no residual association between sex and cognitive level, suggesting no difference between men and women in normal cognitive aging. However, in the simple LCM, the residual sex effect was borderline significant. This may be due to the fact that the latent classes from the simple LCM are more strongly associated with the dropout patterns and slightly less with sex (see Table 1). Furthermore, when adjusting for the latent classes, subjects without diploma exhibited a higher rate of decline which was not observed in the PMM and MAR analyses.

The slight differences observed between the estimates from the simple LCM and the joint LCM were due to the adjustment for sex, educational level and age in the proportional hazard model of the latter. Without these adjustments, estimates were nearly identical to those from the simple LCM (results not shown). Indeed $\sum_C [Y|C, X][C|D][D] = \sum_C [Y|C, X][D|C][C]$. The minor differences arose from the use of a Cox model instead of a logistic one. Moreover, if covariates X are included both in the class membership probability and in the dropout probability, $\sum_C [Y|C, X][C|D, X][D|X] = \sum_C [Y|C, X][D|C, X][C|X]$. Parameters from the conditional distribution $[Y|C, X]$ have the same meaning (and similar values) in the two LCM, but their interpretation is unclear because X impacts both the class membership and the distribution of Y given the class. However, an interesting advantage of the joint model is that it allows adjustment for covariates when modeling dropout given the class. Indeed, conditionally on the covariates included in the mixed model, if data are missing at random, the risk of dropout does not depend on the outcome, data are missing at random.

It should be underlined that this work is based on data from a large ob-

servational cohort study including a representative sample of the elderly population. For instance, the sample includes both demented and non-demented subjects with a wide range of educational levels. This heterogeneity explains the large discrepancy observed between the estimates from the two approaches and highlights the different parameter interpretations. The latter may be partly hidden with more homogeneous data sets, such as those in clinical trials. Implementing investigative methods to deal with informative missing data in observational cohort studies is especially useful since missing data are more frequent in these studies.

In the Paquid cohort, some dropouts are due to death. We did not distinguish death from other sources of dropout. Indeed, separating death from dropout would have greatly increased the number of dropout patterns and created sparse patterns. Moreover, it is unclear how to classify subjects who missed a visit and died thereafter. A proper way to take death into account would be to jointly model time-to-dropout and time-to-death and then compute the estimated evolution given that the subject is alive. However this is beyond the scope of this paper.

One could argue that interpretation of conditional parameters is not essential given that the main interest lies in marginal inference. However, in the literature, parameters from the conditional distribution $[Y|D, X]$ or $[Y|C, X]$ are often considered as retaining a marginal interpretation if they are not group-specific. Section 4.4 shows that this is true only under restrictive assumptions and that, in most cases, computation of marginal parameters is difficult since it depends on the values of the other covariates. Moreover, computation of the variance of these marginal estimates may be untractable. For example, we computed the marginal gender effect on the initial level using formula (13). The marginal estimates were 0.009 for the PMM (close to the MAR estimate), 0.005 for the simple LCM and 0.004 for the joint LCM. The latter was identical to the conditional estimate as the model assumes that the class-membership probabilities do not depend on covariates, but this makes the marginal estimate difficult to compare to the two others. To circumvent these problems, interesting hybrid approaches between selection models and PMM (Wilkins, 2006; Wilkins, 2007) or LCM (Roy, 2007) have recently been proposed, but they raise additional estimation problems.

Other important differences between the LCM and PMM concerns the underlying assumption about the missingness process. Indeed, to reach identifiability in PMM, some *a priori* assumptions must be made. Typically, some dropout patterns with few measurements are assumed to follow a common evolution with other dropout patterns. One advantage of PMM is that several sets of constraints may be used to estimate several PMM in the framework of a

sensitivity analysis. When averaging parameter estimates over the pattern to obtain marginal parameters, we need the additional assumption that the pre-dropout model remains valid after the dropout time within each pattern. The LCM avoid *a priori* identifiability constraints as the grouping of dropout patterns is data driven, but they rely on the conditional independence assumption between the time-to-dropout and the responses given the latent classes. That is an missing completely at random (MCAR) assumption within each class. This hypothesis of conditional independence against the alternative of residual dependence on the observed responses may be evaluated using the posterior classification, but it is obviously impossible to assess residual dependence on the unobserved response.

To conclude, we recommend the use of PMM rather than LCM to tackle MNAR data because the parameters are more simply interpreted and are easily implemented with existing software. However, to understand differences in parameter estimates between PMM and MAR analysis, the association between dropout patterns and covariates needs to be described. When data encompass many observed patterns or patterns with few measurements per subject, several pattern mixture analyses should be performed with different constraints on the parameters to ensure identifiability. If LCM are preferred in such a case, they require complementary analyses to characterize the classes and to avoid misinterpretation of the parameters. However, the two approaches share a common drawback compared to the selection models because the parameters of the marginal distribution of the response variable cannot be easily obtained. Despite this drawback, PMM are useful to perform sensitivity analyses for incomplete longitudinal data, whereas LCM should be used with caution in this context, particularly with heterogeneous data. However, as illustrated in the Paquid data set, the latent class models remain very interesting to explore heterogeneity in data by highlighting several profiles of evolution and by giving additional insight into the impact of covariates.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of mental disorders, edition III Revised (DSM IIIR)*. Washington DC: American Psychiatric Association, 1987.

Bauer DJ, Curran PJ. Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychological methods* 2003; **8**: 338-63.

Benton A. *Manuel pour l'application du test de rétention visuelle. Applications cliniques et expérimentales.* Centre de psychologie appliquée. Paris 1965.

Beunckens C, Molenberghs G, Verbeke G, Mallinckrodt C. A latent-class mixture model for incomplete longitudinal gaussian data. *Biometrics* 2007; published online: 30-Jun-2007. doi: 10.1111/j.1541-0420.2007.00837.x

Commenges D, Letenneur L, Joly P, Alioum A, Dartigues JF. Modelling age-specific risk: application to dementia. *Statistics in medicine* 1998; **17**(17): 1973-1988.

Diggle PJ, Kenward MG. Informative dropouts in longitudinal data analysis (with discussion). *Applied statistics* 1994; **43**: 49-93.

Elias MF, Elias PK, D'Agostino RB, Silbershatz H, Wolf PA. Role of age, education, and gender on cognitive performance in the Framingham Heart Study: community-based norms. *Experimental aging research* 1997; **23**(3): 201-35.

Folstein MF, Folstein SE, Mc Hugh PR. Mini-Mental State. A practical method for grading the cognitive state of patients for the clinicians. *Journal of psychiatric Research* 1975; **12**: 189- 198.

Isaacs B, Kennie AT. The Set test as an aid to the detection of dementia in old people. *The British journal of psychiatry: the journal of mental science* 1973; **123**(575): 467-70.

Jacqmin-Gadda H, Fabrigoule C, Commenges D, Dartigues JF. A 5-year longitudinal study of the Mini Mental State Examination in normal aging. *American journal of epidemiology* 1997; **145**: 498-506.

Kenward MG. Selection models for repeated measurements with non-random dropout: an illustration of sensitivity. *Statistics in medicine* 1998; **17**: 2723-2732.

Le Carret N, Lafont S, Mayo W, Fabrigoule C. The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve. *Developmental neuropsychology* 2003; **23**(3): 317-37.

Letenneur L, Commenges D, Dartigues JF, Barberger-Gateau P. Incidence of

dementia and Alzheimer's disease in elderly community residents of southwestern France. *International journal of epidemiology* 1994; **23**: 1256-61.

Letenneur L, Proust-Lima C, Le Gouge A, Dartigues JF, Barberger-Gateau P. Flavonoid intake and cognitive decline over a 10 year period. *American journal of epidemiology* 2007; **165**: 1364-1371.

Lin H, Turnbull BW, Mc Culloch CE, Slate EH. Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* 2002; **97**: 53-65.

Lin H, McCulloch CE, Rosenheck RA. Latent pattern mixture models for informative intermittent missing data in longitudinal studies. *Biometrics* 2004; **60**(2): 295-305.

Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. Wiley: New York, 1987.

Little RJA. Pattern-Mixture Models for Multivariate Incomplete Data. *Journal of the American Statistical Association* 1993; **88**: 125-134.

Little RJA. Modeling the Drop-Out Mechanism in Repeated-Measures Studies. *Journal of the American Statistical Association* 1995; **90**: 1112-1121.

Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. *SIAM journal of applied mathematics* 1963; **11**: 431-41.

Molenberghs G, Michiels B, Kenward M, Diggle P. Missing data mechanisms and pattern-mixture models. *Statistica Neerlandica* 1998; **52**: 153-161.

Proust C, Jacqmin-Gadda H, Taylor J, Ganiayre J, Commenges D. A nonlinear model with latent process for cognitive evolution using multivariate longitudinal data. *Biometrics* 2006; **62**: 1014-1024.

Proust-Lima C, Joly P, Dartigues JF, Jacqmin-Gadda H. Joint modelling of multivariate longitudinal outcomes and time-to-event: a nonlinear latent class approach. Submitted for publication.

Redner RA, Walker HF. Mixture densities, maximum likelihood and the EM

algorithm. *SIAM review. Society for Industrial and Applied Mathematics* 1984; **26**: 195-239.

Roy J. Modeling longitudinal data with nonignorable dropouts using a latent dropout class model. *Biometrics* 2003; **59**(4): 829-36.

Roy J, Daniels MJ. A General Class of Pattern Mixture Models for Nonignorable Dropout with Many Possible Dropout Times. *Biometrics* 2007; in press.

Schwartz G. Estimating the dimension of a model. *The annals of statistics* 1978; **6**: 461-464.

Thijs H, Molenberghs G, Michiels B, Verbeke G, Curran D. Strategies to fit pattern-mixture models. *Biostatistics* 2002; **3**: 245-265.

Verbeke G, Molenberghs G. *Linear mixed models for longitudinal data*. Springer Series in Statistics: New York, 2000.

Verbeke G, Molenberghs G, Thijs H, Lesaffre E, Kenward MG. Sensitivity analysis for nonrandom dropout: a local influence approach. *Biometrics* 2001; **57**(1): 7-14.

Wiederholt WC, Cahn D, Butters NM, Salmon DP, Kritz-Silverstein D, Barrett-Connor E. Effects of age, gender and education on selected neuropsychological tests in an elderly community cohort. *Journal of the American Geriatrics Society* 1993; **41**(6): 639-47.

Wilkins KJ , Fitzmaurice GM . A Hybrid Model for Nonignorable Dropout in Longitudinal Binary Responses *Biometrics* 2006; **62** (1): 168-176.

Wilkins KJ , Fitzmaurice GM . A marginalized pattern-mixture model for longitudinal binary data when nonresponse depends on unobserved responses *Biostatistics* 2007;**8** (2): 297-305.