Latent class analysis of persistent disturbing behaviour patients by using longitudinal profiles

DOI: 10.1111/j.1467-9876.2009.00704.x
Handle: http://hdl.handle.net/1942/10912

Latent-class Analysis of Persistent Disturbing Behavior Patients Using Longitudinal Profiles

Liesbeth Bruckers*, Jan Serroyen, Geert Molenberghs,

Herman Slaets, and Willem Goeyvaerts

Center for Statistics, Hasselt University,

Agoralaan, Building D, B3590 Diepenbeek, Belgium.

* *E-mail:* liesbeth.bruckers@uhasselt.be

Abstract

Persistent disturbing behavior (PDB) refers to a chronic condition in highly unstable, therapy-resistant psychiatric patients. Because these patients are difficult to maintain in their natural living environment and even in hospital wards, purposefully designed residential psychiatric facilities need to be established. Therefore, it is important to carefully define and circumscribe the group. Serroyen *et al* (2006), starting from the longitudinal analysis of a score based on data from the Belgian national psychiatric registry, undertook a discriminant analysis to distinguish PDB patients from a control group. These authors also indicated that there is scope for further subdividing the PDB patients into two subgroups, using conventional cluster analysis techniques. In this paper, we employ a variety of novel longitudinal-data based cluster analysis techniques. These are based on either conventional growth models, growth-mixture models, or latent class growth models. Unlike in earlier analyses, where some evidence for two groups was found, there now is an indication of three groups, a finding with high practical and organizational relevance.

Keywords: Growth curves; Growth-mixture models; Latent class growth models; Linear mixed models.

1 Introduction

Mental health care institutions in Belgium are confronted with a group of chronic therapy resistant patients, which is problematic in that neither scientific definitions, theory, nor a legal framework is in place. These patients cannot be treated satisfactorily with current therapies and medication. Their behaviour is disturbing in the sense that living together in their natural environment, or

even in a hospital ward, is extremely difficult. Given that their disease systems are unstable, and that their behavior is persistent over time, intensive 24-hour supervision is required. This condition is referred to as *persistent disturbing behavior* (PDB).

The current Belgian health care system is clearly not accommodating to this group. They are predominantly found in psychiatric hospitals, institutions for intensive specialist care, and psychiatric nursing homes, intended for patients with stabilized chronic psychiatric conditions. The PDB group defies both characterizations, while there are no further alternatives.

Serroyen *et al* (2006) argue that the PDB group raises four important questions. First, how can it be distinguished from related but different groups, such as patients with acute or short-term disturbing behaviour. Second, because a clear definition is emerging only now, the size of the PDB group is unclear. Third, it is conceivable that the PDB group consists of a number of subgroups that can be usefully distinguished. Finally, it is not clear in which residential setting such patients should be accommodated, excluding sheltered living and psychiatric wards of general hospitals.

In the absence of a generally accepted definition of PDB, we will use the following working definition. To be classified as PDB, a patient has to exhibit socially inadequate behaviour, that is persistent and treatment resistant, disruptive for the patient's environment, and confronts the therapeutic team with unrealistic demands. This behaviour can take one or several of many different forms such as multiple forms of aggression (directed at oneself or at others), sexually uninhibited behaviour, agitation, decorum loss, and suicidal behaviour. Note that the lack of a definition implies that no formal classification rules existed prior to this research, which therefore is to be seen as the first attempt to introduce an evidence-based approach.

In 1998, a cross-sectional pilot study was set up in the psychiatric hospitals and the psychiatric nursing homes in the Belgian province of Limburg to (1) estimate the size of the PDB group and (2) explore factors to discriminate between PDB and non-PDB patients (Bruckers *et al* 2000). While a useful exercise, it was ridden with residual issues. First, owing to its cross-sectional nature, the focus is on disturbance rather than on persistence. Second, the above working definition did not exist at the time. Third, the group of patients studied was chosen for comparison with a non-PDB control group rather than for representativeness, even though the wards involved were selected so as to exhibit a typical patient mix. This design implies that, while conclusions regarding differences between PDB and non-PDB patients, and conclusions

pertaining to subgroups within the PDB group can be drawn with confidence, caution has to be used when making inferences about the magnitude of the PDB group. For the latter goal, the study should be seen as being a pilot. Nevertheless, it is important to know whether the group is sufficiently large to warrant specific components of care. Even allowing for some overestimation, it is clear that the group is large enough to render its consideration as a single, monolithic group impractical for organization. This is particularly challenging, given the disturbing character of the disorder, necessitating special small-scale care units.

Serroyen *et al* (2006) addressed this point only partially, by discriminating between PDB and non-PDB patients, using longitudinal data analysis techniques (Verbeke and Molenberghs 2000). By making use of the longitudinal nature of the psychiatric registry data, they studied the persistence aspect of the group, in addition to the disturbance aspect studied by Bruckers *et al* (2000), thus increasing insight. They also employed conventional cluster analysis methods to study the important issue of subgroups within the data and hence in the PDB group.

A major issue with conventional clustering is that it starts from cross-sectional data, thus focusing on similarity *at one point in time*. However, patients exhibiting the same characteristics, the same behaviour at one point in time can still evolve, and diverge, in a multitude of ways.

Therefore, in this paper, cluster analysis is refined by making use of the longitudinal nature of the data. This is based on conventional linear mixed models (Verbeke and Molenberghs 2000), and on so-called growth-mixture and latent class growth models (Nagin 1999, Nagin and Tremblay 2001, Erosheva, Fienberg, and Lafferty 2004, Fieuws, Verbeke, and Brant 2005). In this way, linear mixed modeling and clustering methodology will be usefully blended together to suit our purposes. In the light of earlier comments, cluster analysis is extremely important for breaking up the otherwise undivided group of PDB patients into natural, smaller groups, based on longitudinal data. The existence and description of such refined subgroups comprises relevant information for policy makers, institution managers, and fieldworkers alike. In this context, we also refer to Section 2.2, dedicated to relevant earlier analyses.

The data on which our analyses are based are presented in Section 2. The statistical methodologies (conventional growth models, general growth mixture models and latent class growth models) are reviewed in Section 3. Section 4 presents our findings. These findings are used in Section 5 to formulate a perspective on the patient population with persistent disturbing behaviour.

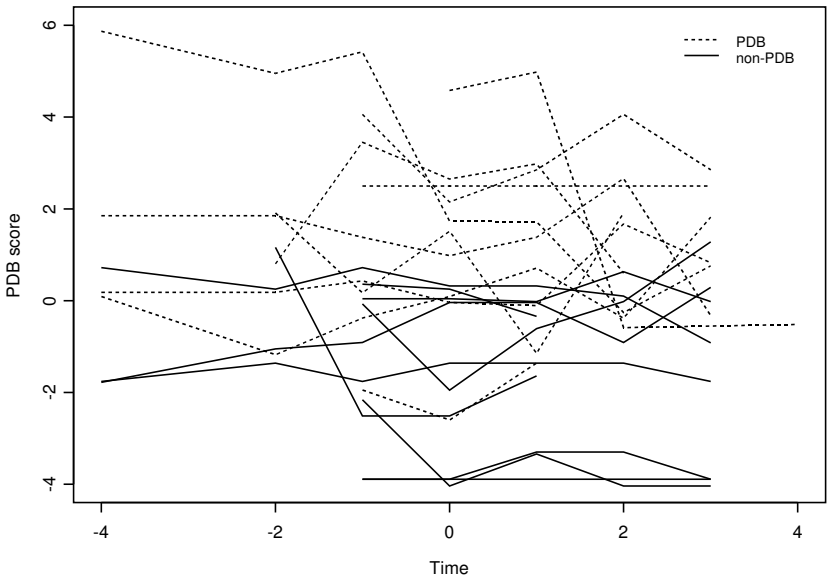

Figure 1: *Random sample of individual PDB score profiles for 10 PDB and 10 non-PDB patients.*

2 The Psychiatric Registry Data and the PDB Score

For every patient admitted to a residential psychiatric care setting in Belgium, data are transferred to a registry, termed *Minimal Psychiatric Data* (MPD), mandatory for psychiatric hospitals since 1996 and psychiatric nursing homes since 1998. The data have a modular structure. The items, relevant for our purposes, are listed in Table 2.

Our data set encompasses 611 patients, among whom 189 are classified as PDB; they are from the province of Limburg. The 189 PDB patients are the basis of inferences on this particular subgroup, while the entire sample of 611 patients will be used in analyses that compare PDB to non-PDB patients. This is a small but very important group of patients with severe pathology, given that organization of care for PDB patients is extremely difficult, time-consuming, and costly.

The variables are mostly of a categorical or ordinal type, although some continuous variables are present as well. The two key continuous variables are the PDB score, expressed as a continuous function on the logit scale, which will be discussed in more detail in Section 2.1, with mean 0.62 (Q1: -0.61; median: 0.34; Q3: 1.69) and age (mean: 47.48 years; Q1: 36 years; median: 46 years; Q3: 60 years).

Figure 1 shows individual profiles of the PDB scores for 20 randomly selected subjects. The

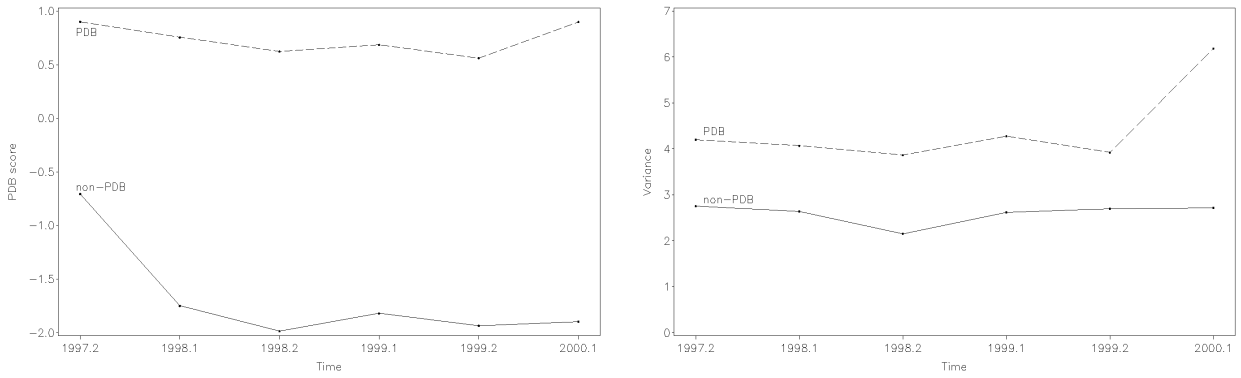

Figure 2: *Mean PDB score over time (lefthand panel) and empirical variance function (righthand panel) for PDB and non-PDB group.*

group-specific average profiles and empirical variance functions are displayed in Figure 2. Table 1 shows the number of measurements available for each of the two sectors (psychiatric hospitals and psychiatric nursing homes), for (a) the subgroup of PDB patients and (b) the number of patients with a PDB score. Note that this starts only at 1996.2 (where the decimal refers to semester), given that 1996 was the year the registration system started and hence it is prudent not to put too much trust in the data for 1996.2. Finally, the table also presents the average PDB scores, together with the standard deviations. PDB patients cannot be discharged from the institution because, owing to the level of disturbance, they are incapable of properly functioning in society. The decreasing number of observations after 1998 can be due to the death of a patient or, in case of the psychiatric hospitals, a transfer to a nursing home. Furthermore, the sudden drop in the available number of measurements at the end of 1999 is caused by the way the MPD had to be exported to the Ministry of Health, and is therefore an extrinsic feature, caused by data collection issues and without clinical basis. For the first semester of 1997, no data are available, owing to the start of the registration system.

2.1 A Cross-sectional and Longitudinal PDB Score

A cross-sectional pilot study was undertaken in both the psychiatric hospitals and the psychiatric nursing homes in the province of Limburg, to obtain a rough estimate of the size of the PDB group and to determine factors that can usefully distinguish between PDB and non-PDB patients. As stated earlier, this pilot study mainly concentrated on the 'disturbing' aspect. Longitudinal data will be introduced, so as to allow for study of the 'persistence' dimension. Arguably, studying

‘

Table 1: *Number of measurements available for the repeated measures profiles of the PDB score, number of patients with PDB score, and average PDB scores (standard deviations), for each of the semesters at which measurements are taken. The decimal index refers to the semester.*

	Year							
Sector	96.2	97.1	97.2	98.1	98.2	99.1	99.2	00.1
Number of patients in the database								
Psychiatric Hospitals	47		112	122	125	120	102	50
Psychiatric Nursing Homes				64	64	63	52	50
Number of patients with PDB score								
Psychiatric Hospitals			107	121	125	119	102	50
Psychiatric Nursing Homes				64	64	62	52	49
Mean PDB scores (standard deviations)								
Psychiatric Hospitals			0.90	0.82	0.60	0.72	0.35	0.58
			(2.05)	(1.79)	(1.80)	(1.83)	(1.50)	(2.14)
Psychiatric Nursing Homes				0.64	0.68	0.62	0.99	1.23
				(2.41)	(2.27)	(2.48)	(2.65)	(2.78)

only one of these two aspects may result in overestimation of the group's size.

Patients were screened by an interdisciplinary team and classified by expert opinion as PDB when the team judged that living together with the patient is hard and that s/he needed continuous supervision. The persistence dimension was approached by restricting attention to patients residing in chronic-patient wards within psychiatric hospitals or in psychiatric nursing homes. Patients residing in one of these wards had in general already had intensive therapy in an acute ward and, in case of a psychiatric nursing home, also a long stay in a chronic ward. The question arises, of course, whether or not the group considered to be PDB in 1998 indeed was chronic in their disturbing behaviour. The fact that these patients are staying in long-stay wards only indicates that we are dealing with chronic disease statuses, not necessarily that the disturbing behaviour is persistent. It is possible that the patient was going through an acute phase of disturbing behaviour, something hard to disentangle based on information localized in time.

To keep the burden on the fieldworkers as low as possible, it was decided to include a sample of wards and to rely, as far as possible, on existing information, rather than initiating further data

collection. In November 1998, a number of wards were screened for PDB behaviour; information was supplemented with relevant MPD items. Based on these, a discriminant function was developed, giving the probability of dealing with a PDB patient, based on discriminatory MPD items. When this probability exceeds a threshold value we classify the patient as PDB. The function turned out to have good discriminative power. The screening status and the classification status agree for about 80 % of the screened patients. Of course, one has to be careful when interpreting these results, because use of a training sample may lead to overly optimistic results. To estimate the size of the PDB group, this function was applied to the registration data of all patients residing in a psychiatric institution on December 31, 1998. More details of this study are reported in Bruckers *et al* (2000).

The functional form of the discriminant function, as obtained from logistic regression, for the patients admitted in a psychiatric hospital is:

$$\begin{aligned} \mathsf{PDB}_{ij} \;=\; & -4.81 + 1.73 \cdot \mathsf{Aggr.A}_{ij} + 0.62 \cdot \mathsf{Aggr.P}_{ij} + 0.33 \cdot \mathsf{Suicid}_{ij} + 0.47 \cdot \mathsf{Appear}_{ij} \\ & +0.40 \cdot \mathsf{Respect}_{ij} - 0.03 \cdot \mathsf{Age}_{i} + 1.81 \cdot \mathsf{Sex}_{i} - 1.50 \cdot \mathsf{DDAC}_{i} \\ & +0.56 \cdot \mathsf{Schizo}_{i} - 2.32 \cdot \mathsf{Resid}_{i}. \end{aligned} \tag{1}$$

The index i refers to patient, whereas j is the measurement occasion (semester) within a patient. The predictive covariates have the following meaning: 'Aggr.A' stands for aggression towards oneself (auto-aggression), 'Aggr.P' for aggression against other people, 'Suicid' for suicide danger, 'Appear' for appearance, 'Respect' for respect for others, 'Age' for age in 1998, 'Sex' is a binary indicator for a patient's sex, 'DDAC' for the diagnostic class Delirium, Dementia, Amnestic and Cognitive disorders, 'Schizo' for the diagnostic class schizophrenia, and 'Resid' for the residual diagnostic class (so-called *V-codes*, a technical term for conditions that are the focus of clinical attention without being considered disorders). That sex and age, for example, are included in the score might be met with surprise. Such concern would be warranted when a purely behavioural approach is envisaged. However, the goal here is to put forward rules that classify a patient as belonging to the PDB group as accurately as possible. To complicate matters, in some of the analyses, such as the one reported here, the aim is to do this at a single point in time. Of course, then only essentially 'disturbance'-oriented information is available, rather than reliable longitudinal 'persistence' information. We believe that this offers an extra motivation to employ background covariates.

For the psychiatric nursing home patients, the functional form of the discriminant function is:

$$\begin{aligned} \text{PDB}_{ij} &= -6.39 + 1.26 \cdot \text{Aggr.A}_{ij} + 1.15 \cdot \text{Aggr.O}_{ij} + 0.65 \cdot \text{Asoc}_{ij} + 1.21 \cdot \text{Separ}_{ij} \\ &\quad +0.70 \cdot \text{Social}_{ij} + 0.59 \cdot \text{Respect}_{ij} - 0.85 \cdot \text{Retar}_{i}, \end{aligned} \tag{2}$$

with the same abbreviations as in (1) and in addition 'Aggr.O' standing for aggression against objects, 'Asoc' for anti-social attitude, 'Separ' for need for separation or isolation, 'Social' for socially unacceptable behaviour, and 'Retar' for the diagnostic class mental retardation.

The *mental signs and symptoms*, i.e., 'Aggr.A,' 'Aggr.P,' 'Aggr.O,' and 'Asoc,' are direct indications for whether or not a patient's disturbing behaviour contributed to his/her admission or extension of stay. Indirect indications for the degree of disturbance are captured by the items referring to preventive suicidal supervision, 'Suicid,' and the need for separation, 'Separ.' For each item, whether or not the activity was performed at least once was scored. 'Respect' and 'Social' belong to a set of *patient functioning* items that describe the interaction between the patient and his/her immediate surroundings, as well as the relationship between the patients and their co-residents. The diagnostic classes 'DDAC,' 'Schizo,' 'Retar,' and 'Vcode' are constructed according to a consensus document, designed by the Limburg collaborative network in psychiatry, SPIL, summarizing the diagnostic DSM-IV codes in 11 classes (SPIL-RPL 1997, Munson 2001).

The MPD items contributing significantly to the logistic-regression-based classification are presented in Table 2. The ROC c statistic (Agresti 2002), which quantifies the quality of fit, equals 0.85 for the psychiatric hospitals and 0.88 for the psychiatric nursing homes. Bruckers *et al* (2000) observed that the sensitivity and specificity for the psychiatric hospitals (psychiatric nursing homes) were 77.2% (71.9%) and 78.7% (85.5%), respectively. These were attained for a cutoff value of 0.40 (0.28) for the hospitals (nursing homes).

An important conclusion from Bruckers *et al* (2000) was that, following such a discriminant rule, 35.5% of the patient population in a psychiatric hospital might belong to the PDB group, with a similar figure (32.1%) for the psychiatric nursing homes. The corresponding 95% confidence intervals are [198;242] and [100;135]. Of course, these findings have to be taken with some caution. First, as stated in the introduction, the data used for analysis constituted a pilot sample of PDB patients and controls, 611 in total, intended to build the classification rule from. Hence, its use lies in the ability to compare the two groups, rather than in being representative for a larger population. Second, and more important, the discriminant function focuses on the disturbance

Table 2: *Logistic regression analysis results, separately for psychiatric hospitals and psychiatric nursing homes. Odds ratios [95% confidence intervals] are reported.*

Item	Hospitals	Nursing homes
Psychiatric signs and symptoms:		
Auto-aggressive actions	5.62 [1.93; 16.42]	3.52 [1.61; 7.72]
Aggression against people	1.87 [1.21; 2.88]	
Aggression against objects		3.17 [1.26; 7.99]
Anti-social attitude		1.92 [1.05; 3.50]
Intensified supervision:		
Suicide danger	1.39 [1.19; 1.63]	
Separation/isolation		3.34 [1.12; 9.95]
Patient Functioning:		
Appearance	1.60 [1.18; 2.17]	
Respect for others	1.49 [1.03; 2.16]	1.81 [1.16; 2.82]
Socially unacceptable behaviour		2.02 [1.27; 3.19]
Age	0.97 [0.95; 0.99]	
Sex	6.10 [2.89; 12.90]	
Primary diagnosis at admission:		
Mental Retardation		0.43 [0.19; 0.97]
DDAC[1]	0.22 [0.06; 0.87]	
Schizophrenia	1.75 [0.95; 3.21]	
V-codes	0.10 [0.01; 0.92]	

[1] Delirium, Dementia, Amnestic and Cognitive disorders

aspect, neglecting the persistence. However, as stated in Section 2, the items constituting the discriminant function have been recorded in the MPD registry twice a year, from the second semester in 1996 to 2000. When applying discriminants (1) and (2) to the MPD registry at the other measurement occasions, repeated measurements of the PDB score, a continuous score on the logit scale, are obtained. It is these repeated measures that will be further scrutinized.

2.2 Previous Clustering Attempts

Classical clustering techniques (Johnson and Wichern 2003, Krzanowksi 1988) applied to the PDB group of 1998, indicated that three subgroups can be defined based on the patient's sex and age and, to a lesser extent, the diagnosis at admission (Bruckers *et al* 2000).

Serroyen *et al* (2006), using a K-means method, retained two PDB clusters, separating in terms of ordinal variables that measure the pathological behaviour: mobility, recognition of persons, notion of time, initiative, social, respect, and conflicts. A cluster with high values for these pathological behaviour items was separated from a cluster with lower values. Furthermore, the mean PDB score for the first cluster (0.78) is higher than for the second cluster (0.47). Notwithstanding this finding, both clusters exhibit residual heterogeneity, as is clear from the standard deviations, which take values 1.80 and 2.11, respectively. We will therefore consider more elaborate clustering methods that start from the longitudinal nature of the data, to study the evidence for the existence of finer clusters.

3 Statistical Methodology

Growth models were applied to the PDB scores to identify individual differences in the change in PDB score and to understand the process of change itself. The longitudinal models discussed by Serroyen *et al* (2006) aim at analyzing change too, where intra-individual change and inter-individual differences in such changes were considered. The data were, however, treated as if collected from a single population. This assumption of homogeneity in the growth parameters is often unrealistic. Finite mixture models (McLachlan and Basford 1988, McLachlan and Peel 2000) partition the population in an unknown number of latent classes or subpopulations, with class membership determined by specific parameters. Muthén and Shedden (1999) and Muthén and Muthén (1998–2007) proposed a general growth mixture modeling (GGMM) framework.

An important question is what sample sizes are required to detect a certain number of clusters. It is also a difficult question since the answer depends on: (1) true separation of clusters; (2) number of repeated measures per subject (in our case this is not constant); (3) the presence of incompletely observed profiles; (4) the particular clustering technique chosen. For these reasons, and also given the pilot status of our study, we believe that caution should be used when interpreting the results. They motivate consideration of not one but several clustering methods, with different methodological foundations.

In Section 3.1 a conventional growth modeling framework will be sketched, whereas Section 3.2 is dedicated to a general description of the GGMM framework.

3.1 Conventional Growth Models

Growth models are used to describe the individual and average trajectory of a measurement over time (Verbeke and Molenberghs, 2000) and are able to capture individual differences in evolution by introducing patient-specific parameters, or effects, such as random intercepts and random slopes. In a conventional model, also termed a mixed model, the random effects are assumed to follow a continuous, usually normal, distribution. Even when normality is violated, it has been shown that the linear mixed model, unlike its generalized counterparts, is relatively robust in its inferences (Lesaffre and Verbeke 1998, Litière *et al* 2007).

To formulate a growth model for the repeated PDB scores, let us denote by Y_{ij} the PDB score for patient i at occasion j, by t_{ij} a time-related variable, such as the calendar time at which a measurement is taken, and by x_{ij} a vector of time varying or time-invariant covariates. If all covariates are time-independent, one can write x_i instead. Such models are also termed multilevel models (Raudenbush and Bryk 1992, Goldstein 1995). In the multi-level spirit, one example of a model for the PDB scores would take the form:

$$\text{Level 1} \quad : \quad Y_{ij} = \alpha_{0i} + \alpha_{1i} t_{ij} + \varepsilon_{ij}, \tag{3}$$

$$\text{Level 2} \quad : \quad \alpha_{0i} = \beta_{00} + \beta_{01} x_i + r_{0i}, \tag{4}$$

$$\alpha_{1i} = \beta_{10} + \beta_{11} x_i + r_{1i}. \tag{5}$$

Here, α_{0i} and α_{1i} are random intercepts and slopes, respectively, varying across individuals. The residuals ε, r_0 and r_1 are assumed to be zero-mean normally distributed. Level 2 residuals r_0 and r_1 are possibly correlated but uncorrelated with ε_{ij}. The time-specific residuals $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})'$ have an unstructured covariance matrix, Θ say. This allows for residual correlation across time. The r residuals have covariance matrix Ψ.

In the model formulated above, the random effects α_0 and α_1 are latent variables, called growth factors, that capture heterogeneity between patients. The model allows for individual differences in the evolution over time because the growth intercept, α_{0i}, and slope, α_{1i}, vary across individuals, producing individual-specific time courses. Nevertheless, the random effects are assumed to be generated from a single normal distribution. This implies that all patients are drawn from a single population, characterized by one set of parameters. This assumption is revisited in GGMM.

3.2 General Growth Mixture Modeling

Growth Mixture Modeling (GMM) relaxes the assumption of a single population by allowing for different, unknown classes of individuals to vary around different mean growth curves. To this end, the conventional growth model (3)–(5) is extended to a general growth mixture model for K latent trajectory classes, where in class k ($k = 1, \ldots, K$):

$$\alpha_{0i} = \beta_{00}^{k} + \boldsymbol{\beta}_{01}^{k'}\boldsymbol{x}_i + r_{0i}, \quad (6)$$

$$\alpha_{1i} = \beta_{10}^{k} + \boldsymbol{\beta}_{11}^{k'}\boldsymbol{x}_i + r_{1i}. \quad (7)$$

Here, the $\boldsymbol{\beta}_{\cdot\cdot}^{k}$ parameters vary across classes, resulting in different types of trajectories. The $\boldsymbol{\beta}_{01}$ and $\boldsymbol{\beta}_{11}$ parameters allow for variation across classes in how a covariate influences the growth factors. Class-specific covariance matrices Ψ_k for the r terms as well as class-specific covariance matrices Θ_k for ε are allowed. In the absence of covariates, the β_{00}^{k} and β_{10}^{k} are interpretable as the mean growth factors, i.e., k mean intercepts and k mean slopes.

3.3 Latent Class Growth Analysis

Nagin (1999) and Nagin and Land (1993) consider a particular type of growth mixture model, by assuming that the risk factors of patient i and his/her repeated measurements over time are independent, conditional on group membership. This assumption puts constraints on the variances and covariances of the growth factors, in the sense that the variance components governing (r_{0i}, r_{1i}) are set equal to zero. Individuals within a class are treated as homogenous relative to their development. This form of GMM is referred to as *latent class growth analysis* (LCGA). The SAS procedure TRAJ, (Jones, Nagin, and Roeder 2001) can be used to estimate the LCGA model parameters. The procedure uses maximum likelihood with the iterative procedure based on a quasi Newton procedure. Prediction of class membership uses the mode of the posterior probabilities.

3.4 Growth Mixture Modeling Estimation

Generally, estimation of the GGMM parameters can be done conveniently via maximum likelihood, coupled with the EM algorithm (Dempster, Laird, and Rubin 1977). The class memberships c_i are considered missing data in the EM algorithm. It is convenient to replace the categorical c_i by a set of dummies c_{ik}, which are 1 if $c_i = k$ and 0 otherwise. The observed data log-likelihood is

a mixture:

$$\ell = \sum_{i=1}^{n} \log \left\{ \sum_{k=1}^{K} P(c_{ik} = 1|\boldsymbol{x}_i)\phi(\boldsymbol{y}_i|c_{ik} = 1, \boldsymbol{x}_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \right\},$$

where $\boldsymbol{y}_i$ is the vector of repeated measures for subject i and $\boldsymbol{x}_i$ is the vector assembling the covariate information. Further, $\phi(\cdot)$ represents a normal density with mean $\boldsymbol{\mu}_i$ and variance-covariance matrix $\boldsymbol{\Sigma}_i$. Note that a mixture of normals is a flexible device, which can generate, for example, unimodal and multimodal densities, both skewed and symmetric. The model parameters were estimated using the MPlus software (Muthén and Muthén 1998–2007).

3.5 Modelling Strategy

The following model fitting strategy was employed for the PDB data, in line with the proposal of Muthén (2001). First, the data were explored by plotting individual observed and fitted curves. Second, a curve was fitted and plotted for each person. It is difficult to see clusters in this set of individual curves.

A conventional, one-class growth model was fitted to obtain some initial insight into the growth factor variation. Second, groups were identified through latent-class membership. LCGA was carried out for differing numbers of classes. Subsequently, a series of unconditional growth mixture models with two to four classes were considered. Various constraints for the growth factor variances and residual variances of the PDB scores were investigated. The variances of the growth factors were set equal to zero in Model I, which corresponds to the approach of Nagin and Tremblay (2001), they were constrained to be equal across classes in Model II, and they were allowed to range freely in Model III. The residual variances were constrained to be constant across classes but allowed to change over time, Model A, to be constant over time but with a different variance per class, Model B, and finally the residual variances were left unconstrained in Model C.

Selection of a final model was based on the Bayesian Information Criterion value (Schwarz 1978), $BIC = -2\ln(L) + p\ln(n)$, where p is the number of model parameters and n is the sample size. Additionally, and by way of sensitivity analysis, we also consider Akaike's Information Criterion (Akaike 1974), $AIC = -2\ln(L) + 2p$, supplemented with the likelihood ratio test proposed by Lo, Mendell, and Rubin (2001), designed to compare a given model with a model containing one class less; the associated p-value represents the evidence in favour of the simpler model. A parametric

bootstrapped likelihood ratio test (McLachlan and Peel 2000) could also be considered, but this method is rather time consuming.

4 Analysis of the Psychiatric Registry Data

4.1 Psychiatric Hospitals

Model fitting procedures for the single-class mixture model resulted in a log-likelihood value of -1040.503 and a BIC of 2099.420. The estimated slope is -0.0852 (p=0.012) and indicates that PDB-score decreased over the 6 assessments. The estimated intercept variance of 2.365 (p=0.0114) for the PDB-score, and the variance of the latent slope score of 0.050 (p <0.0001), indicate that substantial variation exists among the PDB patients, mainly in initial status, at time zero, but also in the rate of change over time. The significance of the random effects were investigated by means of likelihood ratio statistics with asymptotic null distributions a mixture of two chi-squared distributions. In graphical displays, time zero will correspond to 1997.2.

Table 3 displays model fit results of unconditional growth mixture models with two to four classes. Models I, II and III, defining the growth factor variances and models A, B, and C specifying the residual variances, are defined in Section 3.5. We show the value of the log-likelihood, the sample size adjusted BIC, the number of parameters in the variance-covariance matrix, and the entropy (Ramaswamy *et al* 1993). Lower observed BIC values are indicative of an improved fit, owing to MPlus' sign reversal of the said quantity. The entropy is a summary measure for the classification accuracy, based on model-based probabilities. Higher values indicate better classification.

Based on the information criteria, the three-class linear Model IIIB was selected as the optimal model. This choice is a compromise between goodness-of-fit on the one hand and the desire to select a model that is not overly complex on the other, bearing in mind that more elaborate models might be less than optimal for prediction and classification purposes. This three-class mixture model resulted in a log-likelihood value of -975.78, a BIC of 1985.03 and an entropy estimate of 0.708. Parameter estimates and standard errors for all fixed effects and all variance components are reported in Table 4. Five percent (n=6) of the patients were allocated to the first class, 33% (n=41) to the second class and 63% (n=79) to the third class. Based on the intercept and slope factor, these classes were labeled: (1) 'Low group', (2) 'High but improving', and (3) 'Middle group'.

Table 3: *Models for the psychiatric hospitals: log likelihood, sample size adjusted BIC, number of parameters in the variance-covariance matrix, and the entropy are reported: log likelihood ℓ, sample size adjusted BIC and AIC, the entropy, the Lo-Mendell-Rubin likelihood ratio test (p value), and the number of parameters in the variance-covariance matrix, and are reported. Growth factor variances are: I, equal to zero; II, equal across classes; III, unconstrained. Residual variances are: A, constant across classes; B, constant over time; C, unconstrained.*

Classes		ℓ	BIC	AIC	entropy	LMR-LRT (p)	#par
				Model I			
2	A	-1162.75	2343.92	2347.51	0.810	173.54 (0.1594)	11
	B	-1161.81	2335.34	2337.62	0.794	190.04 (0.4205)	7
	C	-1150.11	2328.67	2334.22	0.801	206.06 (0.3584)	17
3	A	-1103.13	2229.70	2234.26	0.879	111.56 (0.4513)	14
	B	-1082.05	2182.50	2186.09	0.868	151.68 (0.1280)	11
	C	-1073.26	2190.05	2198.53	0.870	150.24 (0.2031)	26
4	A	-1063.46	2155.38	2160.92	0.893	74.23 (0.0647)	17
	B	-1051.25	2127.62	2132.51	0.901	58.56 (0.0020)	15
	C	-1038.43	2135.44	2146.85	0.882	68.11 (0.1745)	35
				Model II			
2	A	-1032.62	2088.38	2092.94	0.933	15.03 (0.2139)	14
	B	-992.31	2001.35	2004.61	0.687	102.13 (0.0000)	10
	C	-983.00	1999.48	2006.00	0.719	112.42 (0.0001)	20
3	A	-1026.95	2082.36	2087.91	0.950	10.33 (0.3051)	17
	B	-986.95	1997.34	2001.91	0.758	10.18 (0.0050)	14
	C	-970.01	1988.57	1998.02	0.702	25.40 (0.3117)	29
4	A	-1023.33	2080.13	2086.65	0.951	6.96 (0.0800)	20
	B	-982.37	1994.88	2000.76	0.736	8.71 (0.3831)	18
	C	-954.10	1971.81	1984.20	0.788	31.11 (0.2241)	38
				Model III			
2	A	-1025.28	2079.02	2084.57	0.478	29.43 (0.0198)	17
	B	-988.37	1998.50	2002.74	0.664	111.97 (0.0000)	13
	C	-981.07	2000.64	2008.14	0.688	116.86 (0.0809)	23
3	A	-1018.21	2074.91	2082.41	0.395	13.68 (0.3559)	23
	B	-975.78	1985.03	1991.55	0.708	24.46 (0.0903)	20
	C	-964.87	1988.33	1999.74	0.694	31.85 (0.3329)	35
4	A	-1014.57	2077.69	2087.15	0.528	7.81 (0.7622)	29
	B	-971.96	1989.11	1985.91	0.806	19.08 (0.0846)	27
	C	-946.01	1970.70	1986.02	0.720	28.73 (0.1646)	47

Table 4: *Summary of the three-class model, selected for the PDB data. Parameter estimates, standard errors and t value for the class-specific intercepts and slopes are shown.*

Effect	**Estimate**	**Standard Error**	t**-value**
Fixed effects			
Intercepts			
Class 1	-0.913	0.758	-1.025
Class 2	1.558	0.296	5.256
Class 3	0.372	0.160	2.325
Slopes			
Class 1	0.064	0.009	6.929
Class 2	-0.271	0.092	-2.944
Class 3	0.021	0.030	0.705
Random effects			
Variance of intercepts			
Class 1	3.507	1.857	
Class 2	2.552	0.698	
Class 3	1.155	0.279	
Variance of slopes			
Class 1	-0.006	0.002	
Class 2	0.132	0.064	
Class 3	0.014	0.010	
Covariance of intercept, slope			
Class 1	0.038	0.018	
Class 2	-0.120	0.164	
Class 3	0.007	0.040	
Residual Variance			
Class 1	0.110	0.045	
Class 2	1.849	0.234	
Class 3	0.376	0.052	

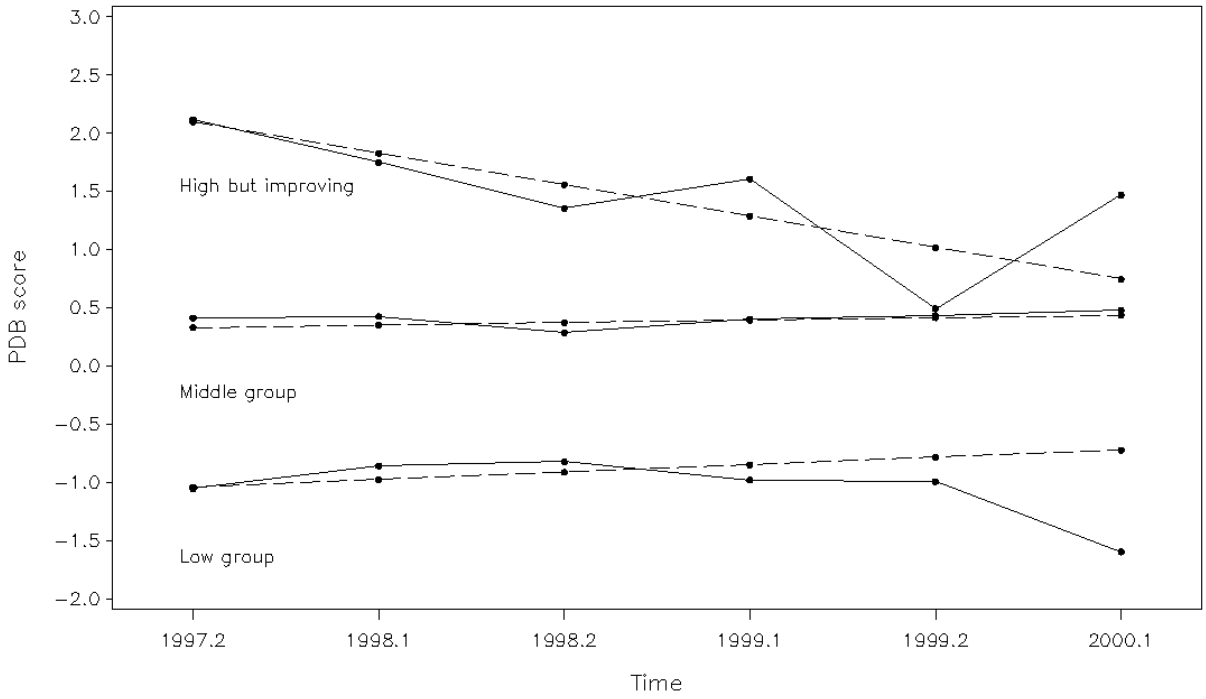

Figure 3: *Observed (full lines) and predicted (dashed lines) trajectories for the 3-class model (psychiatric hospitals).*

Figure 3 displays the predicted trajectories for the three classes. The linear trends appear to describe the data well.

Table 5 shows the posterior class membership probabilities for the three-class model. High diagonal and low off-diagonal values indicate good classification. The highest probability for 'correct' classification, in the sense of agreement between the latent class membership and status in terms of average class probability, occurred for class 2, the patients with high PDB-scores. Patients of this class had 91% chance to be assigned to class 2. The highest misclassification probability occurred for class 3, i.e., for the stable patients: patients classified in this group had 0.15 probability to belong to class 1. Classes 1 and 3 are the most difficult to distinguish.

More than 50% of the PDB patients belong to class 3, with an average PDB score of 0.37. About one in three patients has high PDB scores, averaging around 1.56 at time 0. It is sensible to conclude that the behaviour of this group is more disturbing than the behaviour of the other groups. That there is no evolution to the PDB scores over time for class 3 does not mean that the behaviour itself is constant. The only conclusion that can be drawn, is that the behaviour remains disturbing in the same degree. The type of disturbing behaviour can however change

Table 5: *Agreement between the classification probabilities based on the average class probabilities and latent class membership, for the three-class model displayed in Table 4.*

	Latent class		
Average class prob.	**1**	**2**	**3**
1	0.840	0.011	0.149
2	0.000	0.914	0.086
3	0.052	0.105	0.842

over time.

The variances of the intercepts are relatively large; indicating that even within a class there is still heterogeneity. The significance of the random effects was investigated with likelihood ratio statistic; with asymptotic null distributions a mixture of chi-squared distributions. The variance of the intercept is significant for all three classes; the variance of the slope is highly significant for classes 1 and 2, but not for class 3.

When studying the sample variances, weighted by the estimated class probabilities we find that the variances are not constant over time. For classes 1 and 2, the variance is smallest around times 2 and 3. This is when the pilot study was performed. The repeated PDB scores are constructed based on a discriminant function that was built using the data of the pilot study in 1998. It is therefore important that the analysis can accommodate non-constant variances, as fortunately is the case.

4.2 Psychiatric Nursing Homes

Model fitting procedures for the single-class mixture model resulted in a log-likelihood value of -586.768 and a BIC of 1183.651. PDB patients residing in a psychiatric nursing home, have constant PDB scores over time. The estimated slope of 0.085 is not significant ($p = 0.1905$). The large intercept variance of 4.0334 ($p < 0.0001$) for the PDB-score shows that variation exists among PDB patients in psychiatric nursing homes in terms of their (average) PDB-score. The estimate of the slope variance equals zero, indicating that, under a hierarchical interpretation of the model, the random slope can be removed.

Table 6 displays the results of fitting unconditional growth mixture models with two to four

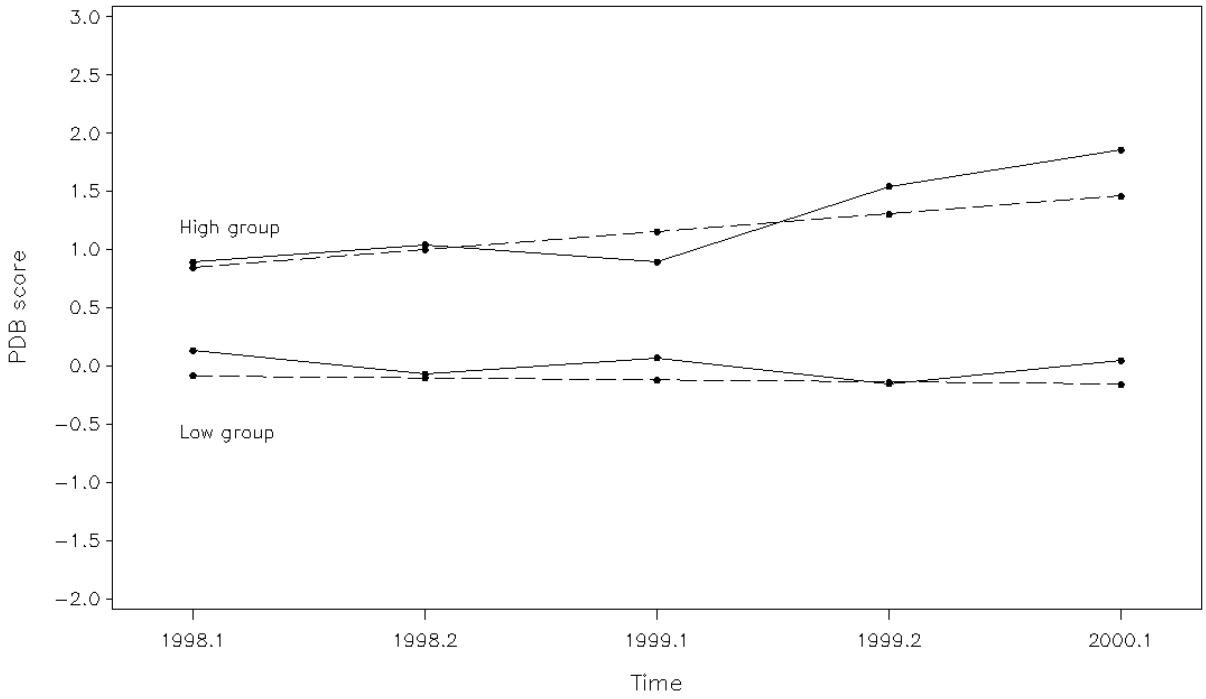

Figure 4: *Observed (full lines) and predicted (dashed lines) trajectories for the 2-class model (psychiatric nursing homes).*

classes. Models I, II, and III, defining the growth factor variances and models A, B, and C specifying the residual variances, are as in Section 3.5.

The two-class Model IIIC is the preferred choice, using logic similar to the one employed in the case of psychiatric hospitals. This two-class mixture model resulted in a log likelihood value of -541.656, a BIC of 1104.55, and an entropy estimate of 0.835. Figure 4 displays the observed and predicted trajectories of the classes. The estimate of the mean PDB-score equals -0.10 in class 1 and 1.00 in class 2. Two thirds of patients belong to class 2. The variances of the intercepts shows that within a class patients still differ. The heterogeneity is largest in class 2. The probability for correct classification equals 0.91 for class 1 and 0.98 for class 2, indicating that the groups are well separated.

5 Conclusion

Persistent disturbing behaviour (PDB) is a highly disruptive condition. Proper treatment and organization of care pose important challenges. So far, it had not been properly defined, let alone

Table 6: *Models for the psychiatric nursing homes: log likelihood, sample size adjusted BIC, number of parameters in the variance-covariance matrix, and the entropy are reported: log likelihood ℓ, sample size adjusted BIC and AIC, the entropy, the Lo-Mendell-Rubin likelihood ratio test (p value), and the number of parameters in the variance-covariance matrix, and are reported. Growth factor variances are: I, equal to zero; II, equal across classes; III, unconstrained. Residual variances are: A, constant across classes; B, constant over time; C, unconstrained. Empty entries result from convergence failure.*

Classes		ℓ	BIC	AIC	entropy	LMR-LRT (p)	#par
				Model I			
2	A	-622.76	1255.64	1265.53	0.839	99.37 (0.5228)	10
	B	-619.02	1245.13	1252.05	0.796	110.89 (0.0779)	7
	C	-615.02	1245.22	1260.04	0.825	119.19 (0.0912)	15
3	A	-599.77	1212.68	1225.53	0.876	42.59 (0.0959)	13
	B	-584.05	1179.22	1190.09	0.900	65.99 (0.0531)	11
	C	-571.72	1166.71	1189.45	0.945	84.07 (0.0460)	23
4	A	-585.62	1186.91	1202.73	0.899	26.67 (0.1752)	16
	B	-542.17	1099.51	1114.33	0.926	79.01 (0.0639)	15
	C	-544.11	1119.58	1150.22	0.911	53.62 (0.0432)	31
				Model II			
2	A	-576.45	1166.05	1178.90	0.976	19.11 (0.0141)	13
	B	-558.30	1126.71	1136.36	0.741	75.05 (0.0003)	10
	C	-551.18	1120.56	1138.36	0.717	69.10 (0.0023)	18
3	A	-575.24	1166.65	1182.47	0.979	2.25 (0.1838)	16
	B	-553.16	1120.47	1134.31	0.835	9.70 (0.0858)	14
	C	-541.65	1109.61	1135.31	0.700	18.49 (0.3939)	26
4	A	-573.61	1166.44	1185.22	0.895	3.01 (0.7151)	19
	B	-551.67	1121.54	1139.34	0.818	2.81 (0.3170)	18
	C						
				Model III			
2	A	-574.88	1165.95	1181.76	0.494	22.86 (0.0485)	16
	B	-557.53	1128.22	1141.07	0.718	78.40 (0.0009)	13
	C	-541.66	1104.55	1125.31	0.835	88.30 (0.0900)	21
3	A	-565.08	1152.42	1174.17	0.578	10.90 (0.2472)	22
	B	-545.58	1111.39	1131.16	0.846	23.11 (0.0403)	20
	C	-516.92	1066.22	1097.85	0.899	48.40 (0.0295)	32
4	A						
	B	-536.71	1100.74	1127.43	0.905	20.17 (0.5365)	27
	C	-499.18	1041.85		0.816		43

Table 7: *Summary of the two-class model, selected for the PDB data. Parameter estimates, standard errors and t value for the class-specific intercepts and slopes are shown.*

Effect	**Estimate**	**Standard Error**	t**-value**
Fixed effects *Intercepts*			
Class 1	-0.103	0.446	-0.231
Class 2	0.999	0.372	2.689
Slopes			
Class 1	-0.019	0.011	-1.638
Class 2	0.153	0.086	1.774
Random effects			
Variance of intercepts			
Class 1	2.794	0.608	
Class 2	4.375	1.053	
Variance of slopes			
Class 1	-0.029	0.020	
Class 2	-0.068	0.107	
Covariance of intercept, slope			
Class 1	-0.005	0.012	
Class 2	-0.089	0.188	
Residual Variance			
Class 1			
time -1	0.438	0.198	
time 0	-0.042	0.025	
time 1	0.582	0.340	
time 2	0.467	0.271	
time 3	0.561	0.399	
Class 2			
time -1	2.961	1.093	
time 0	1.167	0.428	
time 1	2.687	0.853	
time 2	5.751	2.022	
time 3	4.368	1.947	

circumscribed and characterized. Previous analyses, based on a pilot study, indicated that the group of PDB patients is likely larger than generally believed, complicating fieldwork organization because, additionally, it is desirable that PDB wards are small. The analyses presented in this article provide some basis for grouping patients into organizational units based on the degree and evolution of their condition. This does not mean that they might be able to function socially together, but rather that they will be receiving similar types and intensities of care. This will be advantageous for the care givers involved. Our analyses have been based on a pilot study, encompassing 189 PDB and 422 control patients, deemed representative. Data from the psychiatric registry are supplemented with PDB-related information, collected for the purposes of the pilot study. The status of the data may cause concern but, while on the one hand one ought to be careful when using the data for estimating the size of the PDB group, it is very useful to: (1) distinguish between PDB and non-PDB patients; and (2) to explore whether or not it is sensible to assume the presence of subgroups within the PDB group.

Our analyses, based on general growth mixture modeling, lead to two important conclusions. First, meaningful, plausible groups may well exist, in spite of previous findings that were less optimistic (Serroyen *et al* 2006). While previous analyses indicated, at best, the presence of two group, we reached plausible evidence for three groups, categorized as high, medium, and low, in terms of PDB-score profiles. Second, the GGMM analyses clearly show that there is a lot of variability, even within a group of patients whose behaviour is experienced as disturbing by the care team. The group with extreme disturbing behaviour is about one third, 35%, of the entire group. Setting up specialized wards for this group could also lead to better living circumstances for the remaining patients at the wards.

For the entire PDB group in psychiatric hospitals, a linear decrease in the average PDB scores was reported previously. A repeated-measurements analyses indicated that patient-specific characteristics are important and that some patients have intrinsically high values, while others have low values (Serroyen *et al* 2006). These findings stemmed from a conventional growth model. Juxtaposing the results of the conventional growth model and the results of the GMM, we discern that part of the heterogeneity in the PDB population is explainable by it being a mixture of classes, which differ not only in their mean values but also in their evolution. Most patients have moderately stable PDB-scores.

For the PDB patients in psychiatric nursing homes, the GMM analyses reveal two classes. The

distinction between the groups is essentially the average level of the score. Thus, the condition does not worsen or improve. This is not surprising, as we are dealing with a chronic, therapy-resistant group of patients. With current knowledge of therapy and medication, the behaviour of these patients cannot be improved. At the same time, the absence of worsening underscores the chronic nature of the group, which reaches and gets trapped in its worst condition.

Once again, our findings are based on a pilot study and would require verification. To this end, currently a follow-up study is ongoing in the mental health institutions in Limburg. The selection of the patients to be screened is being done centrally to obtain a random sample of long stay patients. A checklist with detailed criteria was developed. Only when all criteria are met is the patient considered to be of the PDB type. This checklist will lead to better insight into the differentiating characteristics of a chronic non-PDB patient and a PDB patient.

Acknowledgment

The authors gratefully acknowledge financial support from the Interuniversity Attraction Pole Research Network P6/03 of the Belgian Government (Belgian Science Policy).

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Bruckers, L., Molenberghs, G., Poncelet, J., Brouns, K., Cuypers, W., Slaets, H., and Vanheyst, I. (2000). Identificatie en inschatting van de omvang van de groep patiënten met persisterend storend gedrag. *Acta Hospitalia*, **40**, 21–30.

Bryk, A.S. and Raudenbush, S.W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park: Sage Publications.

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

Erosheva, E., Fienberg, S., and Lafferty, J. (2004). Mixed-membership models of scientific publications. *Proceedings of the National Academy of Sciences*, **101S**, 5220–5227.

Fieuws, S., Verbeke, G., and Brant, L.J. (2005). Classification of longitudinal profiles using nonlinear mixed-effects models. *Submitted for publication*.

Goldstein, H. (1995). *Multilevel Statistical Models*. Kendall's Libary of Statistics 3. London: Arnold.

Johnson, R. A. & Wichern, D. W. (2003). *Applied Multivariate Statistical Analysis.* 5th ed. Englewood Cliffs: Prentice-Hall.

Jones, B.L., Nagin, D.S., and Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods and Research*, **29**, 374–393.

Krzanowski, W.J. (1988). *Principles of Multivariate Analysis.* Oxford: Clarendon Press.

Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, **54**, 570–582.

Litière, S., Alonso, A., Molenberghs, G., and Geys, H. (2007) The impact of a misspecified random-effects distribution on the estimation and performance of inferential procedures in generalized linear mixed models. *Statistics in Medicine*, **00**, 000–000.

Lo, Y., Mendell, N.R., and Rubin, D.B. (2001). Testing the number of components in a normal mixture. *Biometrika*, **88**, 767–778.

McLachlan, G.J. and Basford, K.E. (1988). *Mixture models. Inference and Applications to Clustering*. New York: Marcel Dekker.

McLachlan, G.J. and Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons.

Munson, C.E. (2001). *The Mental Health Diagnostic Desk Reference.* New York: Haworth Press.

Muthén, L.K. and Muthén, B.O. (1998–2007). *Mplus User's Guide*. Fourth edition. Los Angeles, CA: Muthén and Muthén.

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In: Collins, L.M. and Sayer, A. (eds.), *New Methods for the Analysis of Change* (pp. 291–322). Washington, D.C.: APA.

Muthén, B. and Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM-algorithm, *Biometrics*, **55**, 463–469.

Nagin, D.S. (1999). Analyzing developmental trajectories: a semiparametric, group-bases approach. *Psychological Methods*, **4**, 139–157.

Nagin, D.S. and Land, K.C. (1993). Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed Poission model. *Criminology*, **31**, 327-362.

Nagin, D.S. and Tremblay, R.E. (2001). Analyzing developmental trajectories of distinct but related behaviours: a group-based method. *Psychological Methods*, **6**, 18–34.

Ramaswamy, V., DeSarbo, W., Reibstein, D., and Robinson, W. (1993). An empirical pooling approach for estimating marketing elasticities with PIMS data. *Marketing Science*, **12**, 103–124.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.

Serroyen, J., Bruckers, L., Rogiers, G., and Molenberghs, G. (2006) Characterizing persistent disturbing behaviour using longitudinal and multivariate techniques. *Submitted for publication*.

Verbeke, G. and Molenberghs, G. (2000). *Linear mixed models for longitudinal data*, Springer Series in Statistics, Springer-Verlag, New-York.