# The analysis of multivariate longitudinal data: A review

**Geert Verbeke**[1,2,*], **Steffen Fieuws**[1], and **Geert Molenberghs**[2,1]
**Marie Davidian**
Department of Statistics, North Carolina State University, Raleigh, NC, U.S.A.

[1]Interuniversity Institute for Biostatistics and statistical Bioinformatics, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium

[2]Universiteit Hasselt, B-3590 Diepenbeek, Belgium

## Abstract

Longitudinal experiments often involve multiple outcomes measured repeatedly within a set of study participants. While many questions can be answered by modeling the various outcomes separately, some questions can only be answered in a joint analysis of all of them. In this paper, we will present a review of the many approaches proposed in the statistical literature. Four main model families will be presented, discussed and compared. Focus will be on presenting advantages and disadvantages of the different models rather than on the mathematical or computational details.

## Keywords

Mixed models; Random effects; Shared parameters; Marginal models; Conditional models; Latent variables

## 1 Introduction

In many scientific applications, one often needs to analyze data resulting from experiments in which outcomes have been measured repeatedly on a set of units, leading to so-called repeated measurements. Examples are hearing thresholds measured on both ears of a set of subjects, birthweights of all litter members in a toxicologic animal experiment, or weekly blood pressure measurements in a group of treated patients. The last example is different from the first two in the sense that the time dimension puts a strict, and scientifically relevant, ordering on the obtained measurements within the units. Indeed, the key purpose of the experiment is to study the evolution of the blood pressure over time, and how that evolution depends on subject-specific characteristics such as treatment, age, gender, body mass index, etc. The resulting data are referred to as longitudinal data. Obviously, a correct statistical analysis of such data should account for the clustered nature of the data, i.e., allow that measurements within subjects can be correlated, while observations from different subjects are independent. Therefore, classical (generalized) linear regression models are not applicable. An additional complication arises from the fact that such data sets are often highly unbalanced, i.e., the number of available measurements per unit, and the time points at which the measurements were taken, are often very different across units.

A variety of models has been proposed in the statistical literature, during the last few decades (see e.g., Diggle[1] et al., Verbeke and Molenberghs[2], Molenberghs and Verbeke[3]).

---

*Corresponding author: Geert Verbeke. geert.verbeke@med.kuleuven.be.

Attention focused on the analysis of univariate longitudinal data in which a single outcome is analyzed. In practice, one is often confronted with multiple outcomes, all measured repeatedly over time, possibly a different number of times and/or at different time points. Those outcomes may be similar or of disparate types (continuous/discrete), and a variety of scientific questions may be of interest. In toxicological studies, interest may focus on the relationship between dose of a toxic agent and several outcomes reflecting possible deleterious effects of the agent. For example, birth weight, a continuous measure, as well as some binary indicator of malformation, may be recorded on each fetus in a teratogenicity study. Another example occurs in HIV studies, where measures of immunological and virological status, such as CD4 T-cell count and viral RNA copy number ("viral load"), are collected longitudinally on each participant, and interest may be in studying the relation between the evolutions of both outcomes. Finally, consider a study seeking to elucidate how hearing ability changes during aging based on longitudinal measurements of hearing thresholds at various frequencies, potentially measured separately for the left and right ear, respectively. Of particular interest might be to evaluate whether or not the rate of loss of hearing ability is the same at different frequencies.

It should be emphasized that the availability of multivariate longitudinal data does not necessarily require the construction of a joint model for all outcomes simultaneously. In some cases, univariate longitudinal models for each outcome separately may answer all research questions. In other examples, such as the above ones, a joint modeling strategy is inevitable to answer these questions, because interest is in assessing the relation between some covariate and all outcomes simultaneous, in studying how the association between the various outcomes evolves over time, or in investigating the association between the evolutions of all outcomes.

A number of approaches to joint modeling of multivariate longitudinal data have been proposed in the statistical literature, the main differences between which are similar to those existing between the many techniques available for the analysis of univariate longitudinal data. They originate from different modeling traditions, their construction can be motivated by different arguments, and they also may differ in a number of formal characteristics, such as the structure of the data (balanced or unbalanced), the scale of the observed outcomes (continuous, ordinal, binary), or the way the association between and across outcomes is modeled (with or without latent variables). Especially the latter aspect is important given that the use of latent variables allows for more flexible data structures but usually also has important implications with respect to the interpretation of the various model parameters. To focus ideas, consider an experiment as the hearing test described by Brant and Fozard[4], and Pearson[5] et al., and analyzed by Morrell and Brant[6], Fieuws and Verbeke[7], and Fieuws, Verbeke, and Molenberghs[8]. Hearing threshold sound pressure levels (dB) are determined longitudinally for 603 male volunteers in the Baltimore Longitudinal Study on Aging (BLSA, Shock[9] et al.), at 11 different frequencies and for the left and the right ear separately. The number of visits per subject ranges between 1 and 15, over a median follow-up time of 6.9 years, and visits are unequally spaced. Several scientific questions are of interest. For example, one may be interested in studying whether the loss of hearing ability is the same across frequencies or investigating the association between subject-specific evolutions at different frequencies. These questions require jointly modeling all outcomes. When modeling such (high-)dimensional longitudinal data, one option is to use one or more latent variables for the outcome dimension, i.e., to reduce the dimensionality of the multivariate vector of outcomes. It is then assumed that the observed outcomes are measuring one or more underlying concepts characterizing "hearing ability." An alternative option is to use latent variables for the time dimension, i.e., to assume that the repeated measurements of a particular outcome (one frequency at either left or right side) are reflecting a latent evolution for that outcome. Four families of models can be distinguished

based on whether or not latent variables are assumed for the time dimension and/or for the outcome dimension. These families will be successively described in the Sections 2 to 5, and their merits and disadvantages will be discussed.

In Section 2, models for the joint distribution of observed outcomes will be described. They do not rely on the assumption that latent structures can be used to explain the association between repeated measurements of a single outcome, nor the association between the various outcomes measured a particular point in time. Section 3 presents models to jointly analyze latent evolutions over time. Such models are particularly useful in contexts where the observations are not taken at fixed time points for all subjects. In balanced contexts, techniques such as factor analysis or principal components analysis can be used to reduce the dimensionality of the response vector at each occasion, after which the obtained factors or principal components are analyzed longitudinally. Obviously, a key disadvantage is that longitudinal trends are then no longer in terms of the observed outcomes. Longitudinal trends in latent constructs are studied instead. Examples of such models will be discussed in Section 4. Finally, in Section 5 ideas from Sections 3 and 4 will be combined into models for latent evolutions of latent variables. All models imply specific assumptions about (i) the association between the longitudinal measurements within an outcome, (ii) the association between the various outcomes taken at the same time point, and (iii) the association between outcomes at different points in time. The choice for a specific association structure can be driven by the data, by the research question, or by the chosen estimation procedure. This will be illustrated in Section 6, in the context of the hearing data introduced before. Because highly unbalanced data frequently occur in practice, we will be particularly interested in assessing the suitability of the various methods to handle such data. Finally, some concluding remarks will be given in Section 7. While some features of the models discussed here are equally well applicable for the analysis of time series data, we will not include the literature on multivariate time series in our review. Relevant references in that area can be found in Molenaar[10], and Jorgensen[11] et al.

For the remainder of this paper, let $Y_1$, …, $Y_m$ denote the $m$ vectors containing the longitudinal measurements for the $m$ outcomes and let $Y_k(t)$, $k = 1, …, m$ denote the measurement for the $k$th outcome taken at time point $t$. Further, let $Y(t)$ denote the vector of all outcomes measured at time point $t$. Note that, with unbalanced data, not all outcomes may be available at all time points implying that the dimension of $Y(t)$ can be less than $m$. Finally, let $Y$ be the vector of all measurements in $Y_1, …, Y_m$, and interest will be in specifying a model for the entire vector $Y$. Whenever possible, ideas will be explained in the context of $m = 2$ outcomes $Y_1$ and $Y_2$, but extension to higher dimensions will be straightforward. Also, in all models discussed in the following sections, dependence on covariates will be suppressed from notation, and no additional index for subject will be used.

## 2 Models for evolutions in measured outcomes

A first approach attempts to specify directly the joint density $f(y)$ of $Y$ (see, e.g., Galecki[12]; see also Molenberghs and Verbeke[3] Section 24.1 for an overview). Such models may or may not result from formulating models for the various elements in $Y$ by conditioning on other elements in $Y$, resulting in marginal and conditional models, respectively. In Section 2.1, an overview will be given of models that do not require any conditioning, while conditional models will be presented in Section 2.2.

### 2.1 Marginal models

Specification of a marginal model for $Y$ will require making assumptions about the marginal association among the longitudinally-measured elements within each of the vectors $Y_k$, but also must include assumptions on the nature of the association between elements of any two

vectors $Y_k$ and $Y_{k'}$, $k \neq k'$. Especially when the $Y_k$ are of different types (e.g., continuous-discrete) and/or in the case of (highly) unbalanced data, this becomes cumbersome.

When all outcomes are Gaussian a multivariate linear regression model can sometimes be used. When the data set is fully balanced, the covariance matrix $V$ of $Y$ can have an unstructured form, but would contain $mn(mn + 1)/2$ covariance parameters, where $n$ is the number of time points at which measurements have been taken. To reduce the number of covariance parameters, a more parsimonious structure is sometimes used, such as a Kronecker product of the covariance matrix for, respectively, the $m$ outcomes and the $n$ time points (Galecki[12], O'Brien and Fitzmaurice[13]). Even if unstructured $m \times m$ and $n \times n$ covariance matrices are used for the $m$ outcomes and $n$ time points, respectively, an important reduction in the number of covariance parameters can be obtained, when compared to the full unstructured $mn \times mn$ covariance matrix. Some further structure, such as a first-order autoregressive covariance for the repeated measurements of each outcome, can lead to an even more parsimonious covariance model. Note however that the Kronecker product model implies that the cross-correlations, i.e., correlations between distinct outcomes at various points in time, are products of the marginal correlations specified for the $m$ outcomes and the $n$ time points, which may be too restrictive to be realistic. Another marginal modeling approach for multivariate linear regression models, applicable for irregularly timed observations, has been proposed by Carey and Rosner[14] who assume that the intra-outcome and inter-outcome correlations over time follow a dampened autoregressive correlation structure of the form $\text{Corr}\{Y_k(s), Y_k(t)\} = \exp\{\alpha(|s - t|^{\theta_k})\}$ and $\text{Corr}\{Y_k(s), Y_{k'}(t)\} = \exp\{\alpha(|s - t| + 1)^{\theta_{kk'}}\}$, respectively. The model implies that the variances as well as the correlations between the outcomes remain constant over time, which again may not be realistic in many applications.

Also, when the data are discrete, likelihood-based marginal models can be formulated (see for example Molenberghs and Verbeke[3]), but then are very difficult to implement, unlike in the Gaussian case, unless $nm$ is sufficiently small. In addition, full specification of the distribution of $Y$ not only requires specification of all first- and second-order moments, but of all higher-order moments as well. In the case of binary balanced data, $2^{nm} - 1$ multinomial probabilities need to be modeled and it is not always clear how sensitive inferences for the parameters of interest are with respect to misspecification of some of these multinomial probabilities. Examples of likelihood-based marginal models designed for a longitudinal context do exist, but only for small $nm$, see for example, Daskalakis, Laird, and Murphy[15], or Molenberghs and Lesaffre[16].

A class of models which is very flexible for jointly analyzing outcomes of mixed types is the so-called copula model (Sklar[17], Nelsen[18]). For $m$ outcomes $Y_1, \ldots, Y_m$ with univariate distribution functions $F_1(y_1), \ldots, F_m(y_m)$, a copula model is defined by an $m$-dimensional cumulative distribution function $C(u_1, \ldots, u_m)$ with uniform marginals. The multivariate distribution function $F(y_1, \ldots, y_m) = C(F_1(y_1), \ldots, F_m(y_m))$ then has the prespecified marginals. From this perspective, a copula can be viewed as an association function and the model enjoys the benefit of separating the formulation of the marginals from the specification of the association between the outcomes. For fixed marginals, many different multivariate models can be obtained from considering different copula functions $C(u_1, \ldots, u_m)$. For example, the model of Molenberghs and Lesaffre[16] for ordinal outcomes is an example of the so-called Plackett copula (Plackett[19], Mardia[20]). While the construction of copulas is mathematically elegant, parameter estimation is often not evident, especially in high-dimensional situations. Furthermore, application in longitudinal contexts is often not straightforward due to unbalanced data structures often encountered in practice. To our knowledge, very limited applications of copulas for the analysis of multivariate longitudinal outcomes have been reported. One example can be found in Lambert and Vandenhende[21]

who jointly analysed three outcomes measured longitudinally at 12 pre-specified fixed time points.

Specification of the full joint distribution for discrete data can be avoided by using a generalized estimating equations (GEE) approach. Even when the within-subject associations are (partially) misspecified, valid inferences can be obtained for the regression parameters relating the mean to the set of covariates (Liang and Zeger[22], Prentice[23]). To capture the association, correlations as well as odds ratios can be used. Carey, Zeger and Diggle[24] modeled binary data using odds ratios leading to so-called alternating logistic regression which combines a marginal logistic model for the mean with a conditional logistic model for the association parameters. Their method has the advantage of avoiding the computational burden associated with second order GEE (Zhao and Prentice[25]) which combines estimating equations for the mean with estimating equations for odds ratios, but assumes a parsimonious model for the $\binom{nm}{2}$ pairwise odds ratios, i.e., $m \binom{n}{2}$ intraoutcome

odds ratios, $n \binom{m}{2}$ inter-outcome odds ratios, and $\binom{m}{2} \left( n^2 - n \right)$ cross odds ratios. Ways to specify parsimonious models have been presented by O'Brien and Fitzmaurice[13].

Ten Have and Morabia[26] constructed a bivariate longitudinal model for binary outcomes by combining a logit model for each outcome through a log odds ratio model for the association at a specific point in time. This marginal model addressed the association at each time point. The longitudinal association was modeled using random effects, using ideas similar to those we will present in Section 3. A major advantage of this model is that it can easily handle unbalanced data.

A model for a binary outcome and a continuous outcome, both measured longitudinally, has been proposed by Rochon[27] who combined two GEE models for the two outcomes, using an autoregressive-type working correlation matrix for the intra- and inter-outcome dependence over time. Building on earlier work for continuous outcomes (Gray and Brookmeyer[28]), Gray and Brookmeyer[29] proposed a marginal model for outcomes of different types, in a context where interest is in inference for a treatment effect. They assumed that the treatment group and the control group follow the same time trajectory, but at a different rate. The association was treated as nuisance and modeled using measures of association which depend on the type of outcomes to be analyzed (correlations, odds ratios, … ). The advantage of the approach is that the treatment effect can easily be compared across the various outcomes, irrespective of the metric of the outcome, and that a common treatment effect can be estimated.

The main advantage of all these models is that they allow for direct inferences for marginal characteristics of the outcomes, such as average evolutions and associations. This is also reflected by the symmetric treatment of the two outcomes, which is in strong contrast to several of the other approaches which we will discuss next.

## 2.2 Conditional models

One way to avoid direct specification of a joint distribution for $Y$ is to model a subject's measurement on a given outcome at a particular time point, conditional on all other $mn - 1$ measurements. Geys, Molenberghs, and Ryan[30] applied this in a non-longitudinal context with binary data. In a longitudinal context, it is often considered natural only to condition on the past, which can be done through so-called transition models. Transition models for univariate discrete longitudinal data (Diggle[1] et al.) consider the time course as a sequence of states and transition probabilities to be in a specific state at a particular point in time depends on the state at the previous time point(s) and possibly on a set of covariates, but

extensions to multivariate longitudinal binary data are possible (see, e.g., Zeng and Cook[31], Liang and Zeger[32]). These extensions differ in the way the cross associations are modeled.

While conditional models have the advantage of reducing the modeling task to the specification of a model for each of the outcomes separately, they also have a number of severe shortcomings, for two longitudinal outcomes, but even more so for larger numbers of outcomes. To illustrate the major ones, let us consider the situation where interest is in jointly modeling two outcomes $Y_1$ and $Y_2$, both measured longitudinally. The joint density can be factorized as

$$f(y_1, y_2) = f(y_1|y_2) f(y_2) = f(y_2|y_1) f(y_1). \quad (1)$$

Specification of a conditional model such as $f(y_1|y_2)$ requires very careful reflection about plausible associations between $Y_1$ and $Y_2$, where the latter plays the role of a time-varying covariate, and different choices can lead to very different, sometimes completely opposite, results and conclusions (see, e.g., Diggle[1] et al., Chapter 12; see also literature on time-dependent treatment and confounding, e.g., Zhang[33] et al.). Another drawback of conditional models is that they do not directly lead to marginal inferences. Suppose scientific interest would be in a comparison of the rate of longitudinal change in both average outcomes. The first factorization in (1) directly allows for inferences about the marginal evolution of $Y_2$, but the marginal expectation of $Y_1$ requires computation of

$$E(Y_1) = E[E(Y_1|Y_2)] = \int \left[ \int y_1 f(y_1|y_2) \, dy_1 \right] f(y_2) \, dy_2,$$

which, depending on the actual models, may be far from straightforward. One way to circumvent this would be to fit both factorizations in (1) but specification of both models in a compatible way often requires direct specification of the joint density $f(y_1, y_2)$, thus involving the problems discussed in Section 2.1. Furthermore, the marginal mean of $Y_1$ is not, in general, of the same form as the original conditional mean. For example, a logistic regression model for the conditional mean of $Y_1$ given $Y_2$ does not marginalize to a logistic regression for the marginal mean of $Y_1$. Hence, marginalization, even when computationally straightforward or feasible, is not always useful or helpful.

In some situations, the asymmetric treatment of the outcomes is very unappealing. In a clinical trial, for example, none of the factorizations in (1) will be of interest due to the conditioning on a post-randomization outcome which may (partially) attenuate the treatment effect on the other. Finally, with (many) more than two outcomes, many possible factorizations are possible, all potentially leading to different results. Hence, conditional models are often not the preferred choice for the analysis of high-dimensional multivariate longitudinal data.

## 3 Models for associations between latent evolutions

A very flexible class of models, often used for the analysis of univariate longitudinal data, is the family of mixed models. They assume that the observations represent realizations of a latent subject-specific trajectory which can be modeled parsimoniously using a relatively small number of subject-specific parameters. Since subjects are believed to be randomly sampled from some population, the subject-specific parameters are assumed to be random as well, and are therefore referred to as random effects. The association between the repeated measures is then modeled through the assumption that the random effects are shared by the observations made for that outcome. Such models have the major advantage that they do not assume balancedness, allowing for different numbers of observations per subject and/or

measurements of different subjects taken at different time points. Various mixed model families, such as linear mixed, generalised linear mixed, and nonlinear mixed models have been introduced in the statistical literature. See, e.g., Laird and Ware[34], Breslow and Clayton[35], Davidian and Giltinan[36], Verbeke and Molenberghs[2], Molenberghs and Verbeke[3]. This class of models is also known under a number of other names, including multilevel models (Goldstein[37]), hierarchical linear models (Bryk and Raudenbush[38]) and variance components models (Longford[39]). The choice of the type of mixed model depends on the type of outcome (continuous, ordinal, categorical) and on the functional form of the relation between the outcome and the covariates in the model (linear, generalised linear, nonlinear). Due to the flexibility of the models and the widespread availability of commercial software to fit them, mixed models have developed into a very popular tool for analysing longitudinal data in many areas.

The idea of using random effects to account for the correlation between measurements within a subject can also be exploited to construct joint models for multivariate longitudinal outcomes. More specifically, it will be assumed that, conditionally on the random vector $b_k$, $Y_k$ follows a distribution with density $f(y_k|b_k)$, possibly depending on additional population-specific parameters $\theta_k$, suppressed from notation. Some models assume all $b_k$ to be identical, leading to so-called shared parameter models. Other models allow the different outcomes $Y_k$ to be modeled with separate but correlated random vectors $b_k$, resulting in so-called random-effects models. McCulloch[40] used the same ideas to jointly model multiple outcomes of mixed types, although not restricted to the longitudinal context. Both families will be discussed and compared in the Sections 3.1 and 3.2, respectively. Note that the same two approaches are sometimes used in the analysis of correlated time-to-event outcomes, leading to so-called shared frailty models and correlated frailty models (see, e.g., Duchateau and Janssen[41]).

## 3.1 Shared parameter models

Let $b$ denote the vector of random effects shared by all outcomes $Y_k$, i.e., $b_1 = \ldots = b_m = b$, with density $f(b)$, possibly depending on a vector of unknown population parameters. Often, $b$ is assumed normally distributed, but alternatives are possible (see Chapter 2 in Fitzmaurice[42] et al. for a general discussion). Under the assumption of conditional independence, i.e., assuming that the outcomes $Y_k$ are mutually independent, conditionally on $b$, the joint marginal density is given by

$$f(y) = \int f(y_1, \ldots, y_m | b) \, f(b) \, db = \int \prod_{j=1}^{m} f(y_j | b) \, f(b) \, db. \quad (2)$$

The assumption that all $Y_k$ are conditionally independent given $b$ reflects the belief that a common set $b$ of underlying characteristics of the individual governs both outcome processes. This assumption, while convenient particularly if the outcomes $Y_k$ are of a different type, can be relaxed in some situations (see, e.g., Fieuws and Verbeke[7]).

One of the main advantages of shared parameter models is that the outcomes $Y_k$ do not have to be of the same type. Linear mixed models for continuous outcomes can be combined with logistic mixed models for binary ones and/or Poisson mixed models for counts. Moreover, the parameters in the joint model have the same interpretation as in each of the univariate models. Finally, extending the model to more than two outcomes is straightforward, and, because dimensionality of the integration in (2) does not increase, does not entail any additional computational burden.

A key disadvantage of shared parameter models is that they often involve very strong, sometimes unrealistic, assumptions about the association between the outcomes. As an example, suppose that the $Y_1$ and $Y_2$ be well-described by random-intercept models given by

$$
\begin{aligned}
Y_1(t) &= \beta_1 + b + \beta_2 t + e_1(t), \\
Y_2(t) &= \beta_3 + \gamma b + \beta_4 t + e_2(t),
\end{aligned} \quad (3)
$$

respectively, with $b$ a normally distributed mean-zero random effect common to both models and with $e_1(t)$ and $e_2(t)$ mean-zero normally distributed error process, independent of $b$ and of each other for all $t$. Furthermore, denote $\text{Var}\{e_1(t)\}$ by $\sigma_1^2$ and $\text{Var}\{e_2(t)\}$ by $\sigma_2^2$, constant for all $t$. The parameter $\gamma$ is used to scale the shared random effect $b$ in the model for $Y_2(t)$. It directly follows from model (3) that the marginal cross-correlations between elements from both outcomes are given by

$$
\text{Corr}\{Y_1(s), Y_2(t)\} = \sqrt{\text{Corr}\{Y_1(s), Y_1(t)\}} \sqrt{\text{Corr}\{Y_2(s), Y_2(t)\}}, \qquad \forall s \neq t, \quad (4)
$$

implying that the correlation structures of the individual outcomes dictate the association between pairs of measurements from different outcomes. For example, the model would not allow that $Y_1$ and $Y_2$ to be independent if repeated measures of $Y_1$ and $Y_2$ are strongly correlated. Similar restrictions hold in more general models that involve shared random effect vectors beyond just a scalar random intercept. As we will describe next, these restrictions can be relaxed by allowing for different but correlated random effects for the various outcomes.

## 3.2 Random-effects models

As discussed in Section 3.1, shared parameter models enjoy many desirable properties such as (1) allowing for different types of outcomes, (2) allowing for (highly) unbalanced data, and (3) parameter interpretation identical to univariate models. Only, the rigid constraint (4) often makes them unrealistic to be applicable in practice. This can be solved by allowing for different random effects $b_k$ for the outcomes $Y_k$. The association between the outcomes is then generated by allowing the random effects themselves to be correlated. Let $b = (b_1, \ldots, b_m)$ be the vector of all random effects and let $f(b)$ denote the assumed density (often multivariate normal). Under the conditional independence assumption, the joint marginal density is now given by

$$
f(y) = \int f(y_1 \ldots, y_m | b) f(b) \, db = \int \ldots \int \prod_{j=1}^{m} f(y_j | b_j) f(b_1, \ldots, b_m) \, db_1 \ldots db_m. \quad (5)
$$

To illustrate that this model implies less strict assumptions about the associations between the outcomes, let us re-consider the random-intercepts models from (3), but modified to involve separate random intercepts for both processes, i.e.,

$$
\begin{aligned}
Y_1(t) &= \beta_1 + b_1 + \beta_2 t + e_1(t), \\
Y_2(t) &= \beta_3 + b_2 + \beta_4 t + e_2(t),
\end{aligned} \quad (6)
$$

with similar assumptions as before, and assuming that $b_1$ and $b_2$ jointly follow a bivariate zero-mean normal distribution. Models (3) and (separate intercepts) assume the same correlation structures for each of the outcomes separately, but the marginal cross-correlations between elements from both outcomes are now given by

$$
\text{Corr}\{Y_1(s), Y_2(t)\} = \text{Corr}(b_1, b_2) \sqrt{\text{Corr}\{Y_1(s), Y_1(t)\}} \sqrt{\text{Corr}\{Y_2(s), Y_2(t)\}}, \qquad \forall s \neq t,
$$

showing that correlation between the outcome-specific random effects generates a between-process association, but also that the restriction imposed by the shared parameter model (3) is relaxed in the sense that the correlation structures of the individual outcomes no longer dictate the association between pairs of measurements from different outcomes. For example, the model would now perfectly allow $Y_1$ and $Y_2$ to be independent even if repeated measures of $Y_1$ and $Y_2$ would be strongly correlated. Finally, note that the shared-parameter model (3) can be obtained as a special case of (6) by specifying perfect correlation between $b_1$ and $b_2$.

Many authors have proposed the use of random-effects models for multivariate repeated measures data. Reinsel[43] already introduced the multivariate linear mixed model for continuous outcomes, but estimation was restricted to balanced observed data. Beckett, Tancredi and Wilson[44] analyzed 4 continuous outcomes in an observational data study on elderly people with a total of 8 random effects. An example in a bivariate highly unbalanced setting with continuous data can be found in Chakraborty[45] et al. in the context of HIV data. Other examples of multivariate linear mixed models can be found in MacCallum[46] et al. and Matsuyama and Ohashi[47]. Still within the context of linear models, Shah, Laird, and Schoenfeld[48], Sy, Taylor, and Cumberland[49], Heitjan and Sharma[50], and Fieuws and Verbeke[7] relaxed the conditional independence assumption by allowing the error components of the $m$ outcomes to be correlated. Examples in the context of multivariate nonlinear or generalized linear mixed models are far less common. Ribaudo and Thompson[51] compared two treatments for lung-cancer using longitudinal measurements of 6 binary quality-of-life outcome measures QOL, by combining 6 random-intercept logistic regression models. Agresti[52] developed a multivariate extension of the Rasch model for a non-longitudinal context. An example of the use of random-effects models for multivariate longitudinal data of a mixed type can be found in, e.g., Fieuws[53] et al. who combined linear, non-linear, and generalized linear mixed models to predict renal graft failure in renal transplant patients. Another application where such models are widely used is in the joint analysis of longitudinal pharmacokinetic and pharmacodynamic outcomes (e.g., Davidian and Giltinan[36] Chapter 9). Finally, Blozis[54,55] modeled continuous multivariate growth curve data using non-linear random-effects where fixed effects and random effects entered the models in a non-linear and a linear way, respectively.

All multivariate random-effects models discussed so far assumed that the vector $b$ of random effects is multivariate normally distributed. Thum[56] replaced this by a multivariate $t$-distribution, while Nagin and Land[57], Nagin[58], and Nagin and Tremblay[59] assumed a multinomial distribution to identify different classes of subjects with respect to their evolution over time. This class of models is sometimes also referred to as latent-class growth analysis (Muthén[60]). Molenberghs[61] et al. and Vangeneugden[62] et al. extended generalized linear mixed models with conjugate random effects to simultaneously model longitudinal association and overdispersion, and Njeru Njagi[63] et al. extended this to the multivariate setting.

Finally, a further extension is obtained if no specific parametric functional form for the latent process underlying the longitudinally observed outcomes is assumed anymore. The resulting models borrow ideas from factor analysis to model each outcome vector $Y_k$ as $\Gamma_k b_k + e_k$, for some matrix $\Gamma_k$ with columns containing basis functions representing aspects of change which are not fixed a priori. Examples in the multivariate longitudinal context can be found in, e.g., Stoolmiller[64], MacCallum[46] et al., Willett and Keiley[65], and Ferrer and McArdle[66].

Obviously, (5) still fits within the general framework of mixed models allowing model fitting and inference to be based on likelihood theory for mixed models, and many models

can be fitted using commercially available software packages such as the SAS procedures MIXED for linear models, GLIMMIX for generalized linear models, and NLMIXED for nonlinear models. We refer to Thiébaut[67] et al. for an example in the context of linear mixed models. While most procedures require random effects to be normally distributed, appropriate reformulation of the model sometimes allows fitting models with other random-effects distributions, see, e.g., Liu and Yu[68] and Nelson[69] et al. for examples with the SAS procedure NLMIXED. Note that the general idea of joining separate mixed models by allowing their model-specific random effects to be correlated is applicable irrespective of the number of outcomes involved. The main disadvantage is that the dimensionality of the total vector of random effects in the resulting multivariate model for all outcomes grows with the number of outcome variables, often leading to computational problems in the evaluation or maximization of the marginal density (5) when the number of outcomes exceeds two, and/or when some of the outcomes are best described by a generalized linear, or a non-linear mixed model. For multivariate random-effects models with normal random effects, Fieuws and Verbeke[7] noted that all parameters in the joint model can be estimated from fitting all bivariate models implied by the multivariate model (see also, e.g., Molenberghs and Verbeke[3], Chapter 25; Fieuws, Verbeke and Molenberghs[8]). More specifically, for all $\binom{m}{2}$ pairs $(Y_s, Y_t)$, $1 \leq s < t \leq m$, the joint model based on

$$f(y_s, y_t) = \int \int f(y_s|b_s) f(y_t|b_t) \, db_s db_t \quad (7)$$

is fitted using maximum likelihood estimation and estimators for the parameters in the joint multivariate model (5) are obtained by simply averaging over the results from the $\binom{m}{2}$ pairwise model fits. Obviously, the resulting estimators do not maximize the likelihood corresponding to (5), hence inferences do not follow from standard maximum likelihood theory. Instead, Fieuws and Verbeke[7] have shown that pseudo-likelihood theory can be used to derive the asymptotic distribution of the so-obtained estimators. We refer to Arnold and Strauss[70], Geys, Molenberghs, and Ryan[71], and Molenberghs and Verbeke[3] (Chapter 9) for more details about pseudo-likelihood theory. Replacing the log-likelihood of the original model by a sum of implied marginal or conditional log-likelihoods is also sometimes referred to as composite likelihood (Lindsay[72]). Heagerty and Lele[73] and Curriero and Lele[74] applied composite likelihood theory in the context of spatial data, another instance of high-dimensional correlated data. The main advantage of the pairwise model fitting approach is that it can be applied irrespective of the number of outcomes involved in the joint model, and irrespective of the type of mixed models (linear, generalized linear, nonlinear) that are combined into the joint model, provided that all pairwise models (7) can be fitted. The method has been illustrated in the analysis of a 22-dimensional vector of continuous data, a 7-dimensional vector of binary data, and 4-dimensional vector of mixed outcomes, see Fieuws and Verbeke[7] and Fieuws[53,75] et al., respectively.

## 4 Models for evolutions of latent variables

When (many) more than two longitudinally measured outcomes need to be analyzed jointly, most approaches described earlier are no longer feasible, involve numerical difficulties, or are based on extremely strong, often unrealistic, assumptions about the association structure between the various outcomes in the multivariate response vector. Therefore, a number of methods have been proposed based on dimension reduction. The general idea is to use a factor-analytic, or principal-component type, analysis to first reduce the dimensionality of the response vector and to use standard longitudinal models for the analysis of the principal factors. A typical example for continuous data is the model by Oort[76], who uses $r$ latent factors to reduce the dimensionality of $Y(t)$ at any point $t$ in time:

$$Y(t) = \tau(t) + \Lambda(t)\xi(t) + e(t), \quad (8)$$

for $\tau(t)$ a vector of $m$ intercepts at time $t$, $\Lambda(t)$ a $m \times r$ matrix of (common) factor loadings, $\xi(t) \sim N[\kappa(t), \Phi(t)]$ a random $r$-dimensional vector of scores on the latent variables, and with $e(t)$ a $m$-dimensional vector of classical error components, normally distributed with mean zero and covariance $\Sigma(t)$, and independent of $\xi(t)$. Combining the information over all time points yields

$$Y = \tau + \Lambda\xi + e, \quad (9)$$

where $Y$, $\tau$, $\xi$, and $e$ are stacked vectors and $\Lambda$ is block-diagonal with blocks $\Lambda(t)$. Note that model (9) is over-specified and additional restrictions are needed for model identification. In our longitudinal context, the interpretation of the latent factors should remain the same, implying the restrictions that all $\Lambda(t)$ and all $\tau(t)$ are equal for all time points $t$. The resulting model is the so-called longitudinal three-mode model (Oort[77]). The mean evolution over time of the observed outcomes is reflected by the changes in $\kappa(t)$. Special cases of the model imply further restrictions on how the factor scores $\xi(t)$ evolve over time, in a way which is common to all subjects, but which may vary between outcomes. Examples can be found in Oort[76] and Sivo[78], and a more detailed discussion of this class of models can be found in Fieuws and Verbeke[79].

As an alternative to the above factor analytic (FA) approach, a principal component (PC) type of analysis can be used as well. The main difference is that, the FA model in (9) contains distributional assumptions with respect to the between-subject variability, while the different modes are considered fixed in the PC approach. The PC model is therefore more exploratory. Examples of PC models can be found in Kiers and ten Berge[80], Timmerman and Kiers[81,82], and Kiers, ten Berge, and Bro[83] which differ primarily in the identifying restrictions imposed on the models.

In a non-continuous context, Ilk and Daniels[84] used one latent variable for $m$ binary outcomes, but the evolution over time was modeled using a Markov structure for the observed outcomes, whereas Oort[76] and Sivo[78] modeled the within-subject dependence over time on the latent level. Liu and Hedeker[85] introduced a model for multivariate longitudinal ordinal data in a psychometric context. The ordinal outcomes are believed to be discretized versions of underlying continuous outcomes a joint model for which is obtained using an FA model with one latent factor.

In (8), the random scores on the latent variables $\xi(t)$ were assumed normally distributed. Other distributions have been proposed as well. For example, Reboussin and Reboussin[86] used a multinomial distribution leading to a so-called latent class model at each time point $t$, and the evolution of $\xi(t)$ over time was modeled using a first-order stationary model, hereby assuming that the current state only depends on the previous one, and this association remains constant over time. Similarly to constraining the $\Lambda(t)$ in (8) to be constant over time, the probability of class membership conditional on the outcome is time-invariant to ensure that the meaning of the latent variable does not change.

## 5 Models for latent evolutions of latent variables

In Section 3, latent structures were used for the time-dimension of each longitudinal outcome. In Section 4, latent variables were used for the outcome-dimension at each time-point. Both approaches can also be combined, assuming latent trajectories, not for the observed outcomes but for the few latent variables which summarize the information in the multivariate outcome vector. Such models are often referred to as second-order latent

growth models, given that they model change in a latent variable, as opposed to the so-called first-order latent growth models described in Section 3 which model change in observed variables. Second-order latent growth models have been introduced by McArdle[87] and Duncan and Duncan[88]. More recently, the link with models for mean and covariance structure analysis has been recognized, facilitating parameter estimation (see, e.g., Sayer and Cumsille[89], Hancock, Kuo and Lawrence[90], Muthén[60]). The basic idea is to build upon model (8) by putting additional structure on the way the vector $\xi(t)$ of factor scores changes over time. Oort[76] assumed a factor analysis model with similar structure as the model (8) for the original outcome variables, while Roy and Lin[91] used a linear mixed model in a context where only one latent factor was used in (8), the advantage being that covariates can be incorporated in the model and that model equally well applies for balanced as well as unbalanced data sets. On the other hand, the model of Oort[76] offers more flexibility with respect to the number of latent variables that can be used in (8), the underlying evolution for the latent variable, and the assumed association structure, but these advantages come at the expense of the need for balanced data structures. A more detailed comparison of both models can be found in Fieuws and Verbeke[79]. A very extensive and didactically oriented overview of various linear models for latent constructs is given by Chan[92]. Extensions towards non-linear models for latent variables can be found in Blozis[93] and Harring[94]. All these examples deal with continuous outcomes only. To our knowledge, no extensions have been formulated which would allow analysis of discrete outcomes or outcomes of mixed type.

## 6 Illustration: Modeling loss of hearing ability

In the previous sections, a variety of models for the joint analysis of multivariate longitudinal data has been presented. To illustrate how to select an appropriate modeling strategy in the context of a particular application, we re-consider the hearing data example mentioned in Section 1, and primarily focus on the model choice rather than the results and the subject-matter insights provided by the statistical analysis. Hearing threshold sound pressure levels (dB) are determined at eleven different frequencies (125Hz, 250Hz, 500Hz, 750Hz, 1000Hz, 1500Hz, 2000Hz, 3000Hz, 4000Hz, 6000Hz and 8000Hz), for both ears, on 603 male volunteers. The $11 \times 2 = 22$ outcomes are measured repeatedly over time, with up to 15 measurements per subject. Because the data set is observational, measurements were not taken at equidistant time points. A hearing threshold is the lowest signal intensity a subject can detect at a specific frequency. As such, increasing longitudinal trends are expected and a number of subject-matter research questions about these trends have been formulated:

**Q1** Does the average rate of change depend on the age at which the subject enters the study? If so, hearing loss would be different at different ages.

**Q2** Is the relation between the average rate of change and age (if any) the same across frequencies ? If not, this would indicate selective hearing loss, i.e., that hearing loss is different for different types of sounds.

**Q3** Is a longitudinal trend observed for a subject at one particular frequency strongly associated with the trend for that subject at a different frequency? Very strong associations are expected for the high frequencies and weaker associations for the lower ones.

Note that, while question Q1 can be answered by fitting a model for each outcome separately, addressing the questions Q2 and Q3 requires a joint model for all outcomes. A number of considerations are to be made when an appropriate analysis technique is to be selected. These should involve the data structure, the nature of the outcomes, and the

research questions to be answered. For the hearing data, the following considerations were made:

C1     All 22 outcomes are continuous. Strictly speaking, this allows for the use of a marginal model (Section 2.1) based on the multivariate normal distribution. However, the large number of observations per subject calls for a parsimonious covariance structure, while the unbalanced nature of the data, with observations taken at irregularly spaced time points which are even different for all subjects, seriously reduces the number of realistic covariance structures. Note that this is already the case when one single outcome is to be analyzed.

C2     All questions Q1-Q3 are in terms of the original hearing thresholds, implying that they cannot be directly answered by a conditional approach (Section 2.2), nor by any of the methods based on dimension reduction (Sections 4 or 5).

C3     Questions Q1 and Q2 are with respect to average trends in the population, while question Q3 is with respect to subject-specific trends. Linear mixed models have the advantage that fixed effects have population-average interpretations while random effects provide information about the evolution of individual study participants. An additional advantage is that a parsimonious but flexible marginal covariance structure is implied (see C1).

C4     The high dimensionality complicates model building and model checking. The models discussed in Section 3 have the advantage that model building can be done for each outcome separately after which the multivariate model is easily obtained by joining the various univariate models.

C5     If linear mixed models are used for all outcomes, question Q3 can be reformulated in terms of the association between random (subject-specific) effects. In order not to assume a priori that random effects show perfect correlation, shared parameter models (Section 3.1) should be avoided.

Combining all considerations C1-C5, Fieuws and Verbeke[7] concluded that a random effects model (Section 3.2) is optimally suited to answer questions Q1-Q3. Using similar notation as in (6), they assumed that each outcome $Y_k$, $k = 1, \ldots, 22$ can be appropriately modeled using a linear mixed model of the form

$$Y_k(t) = \begin{array}{l} \left(\beta_{k,1} + \beta_{k,2} Age_i + \beta_{k,3} Age^2 + a_k\right) \\ + \left(\beta_{k,4} + \beta_{k,5} Age + b_k\right) t + \beta_{k,6} L(t) + e_k(t), \end{array} \quad (10)$$

in which $t$ is time expressed in years since entry in the study and $Age$ equals the age of the subject at the first measurement. The binary time-varying covariate $L(t)$ represents a learning effect from the first to subsequent visits. Finally, the random vectors $(a_k, b_k)'$ contain subject-specific intercepts and slopes, while $e_k(t)$ reflects how the actual observation $Y_k(t)$ for the $k$th outcome at time point $t$ deviates from the model-based prediction. Model (10) is the result of a model-building exercise that was conducted for each outcome separately. The multivariate model for all outcomes $(Y_1, \ldots, Y_{22})'$ is obtained by assuming all random effects to jointly follow a 44-dimensional multivariate normal distribution with zero mean and general $44 \times 44$-dimensional covariance matrix. Although the resulting model is still a linear mixed model, the model cannot be fitted using standard software, due to the high dimensionality of the random effects. Instead, the pairwise model fitting approach summarized in Section 3.2 was used and the following answers to questions Q1-Q3 were obtained:

**A1** All parameters $\beta_{k,5}$ significantly differ from zero ($p < 0.05$), except for the corresponding parameter in the model for the hearing thresholds of 250Hz for the left ear.

**A2** Addressing Q2 can be done for both ears simultaneously, or for each ear separately if systematic differences between both sides were to be expected. Separate testing would require testing the null-hypotheses $H_0 : \beta_{1,5} = \beta_{2,5} = \ldots = \beta_{11,5}$ for the right side, and $H_0 : \beta_{12,5} = \beta_{13,5} = \ldots = \beta_{22,5}$ for the left side. This can be done using asymptotic Wald-type tests yielding highly significant results for the right side ($\chi^2_{10} = 110.9$, $p < 0.0001$), as well as for the left side ($\chi^2_{10} = 90.4$, $p < 0.0001$), with more severe loss of hearing ability at higher frequencies.

**A3** The fitted covariance matrix for the random slopes ($b_1, \ldots, b_{22}$)$'$ indicates that the hearing loss for the high frequencies is very highly correlated, while this is far less the case for small frequencies. Furthermore, a principal components analysis based on the correlation matrix yields two principal components representing 69.3% and 52.4%, for the left and right ear outcomes respectively. This provides additional evidence that a shared parameter model (Section 3.1) in which the same random effects are shared by all outcomes is not realistic for the data set at hand.

More detailed results of the analyses have been reported in Fieuws and Verbeke[7] and Fieuws, Verbeke and Molenberghs[8].

## 7 Concluding remarks

In most longitudinal experiments, the number of outcomes measured repeatedly in the participating subjects exceeds one. Often, subject-matter research questions can be answered by analysing all outcomes separately. However, whenever interest is a comparison of longitudinal trends between outcomes, or interest is in the association between the outcomes and how that association evolves over time, joint analysis of all outcomes is required. While focus has been primarily on the analysis of one longitudinal outcome, extensions towards multivariate settings have been proposed during the last decade. Recently, Bandyopadhyay, Ganguli, and Chatterjee[95] discussed a number of possible approaches for the joint analysis of eleven continuous lung function outcomes measured longitudinally on a set of 73 dogs. The aim of our paper was to give a more general overview of the various models and model families, not restricted to the context of one particular experiment and/or one particular type of outcomes, with a discussion of their relative advantages and disadvantages. As discussed by Verbeke and Davidian[96], and Verbeke, Molenberghs, and Rizopoulos[97], many of the ideas presented here can equally well be applied in other contexts such as the joint analysis of one longitudinally measured outcome and one or more time-to-event outcomes.

The construction of a joint model for multivariate longitudinal data usually involves a trade-off between increased computational complexity on the one hand and gain in information on the other hand. As a result, most joint models proposed in the literature are limited to the joint modeling of a relatively small number of outcomes, although counterexamples exist in each of the model families discussed. The ultimate choice will depend on the research questions, the data structure (balanced/unbalanced), the desire to model observed outcomes rather than latent constructs, the dimension of the problem, the nature of the outcomes, etc. Research questions may be in terms of (a comparison of ) average trends or in terms of associations between the (evolutions of) outcomes. Measurements may be taken at fixed time points for all subjects and all outcomes, or may be taken at arbitrary time points and/or different time points for different outcomes. In some cases, observations are believed to be repeated measurements of some underlying latent construct which is the objective of the

inferences while in other cases interests are primarily in the outcomes themselves. Finally, a (very) large number of outcomes, or outcomes of mixed distributional types seriously limits the possible choices for the analysis. Section 6 illustrates how such considerations are needed in the selection of a model appropriate for a particular data set at hand and for a particular research question.

A very versatile model is the random-effects model presented in Section 3.2. First, it is not restricted to balanced settings, which makes the model particularly useful for the analysis of observational data such as the Baltimore Longitudinal Study on Aging, mentioned in Section 1 and discussed in Section 6. Second, since linear, generalised, and nonlinear mixed models can be combined, outcomes can be of different types. Third, the model is constructed as a combination of univariate models, allowing for model building for each outcome separately. Fourth, the interpretation of the parameters in the model is the same as their interpretation in their univariate counterparts, which is in strong contrast to, e.g., conditional models. Finally the structure of the model does not impose any restrictions on the dimensionality of the multivariate outcome vector. The only potential restriction is a computational one, as the dimension of the random-effects distribution increases with the number of outcomes. When all outcomes can be described with linear mixed models, the resulting multivariate model is again a linear mixed model and the integration in (5) can be solved analytically. Model fitting can then be performed using software for mean/covariance models in the multivariate context and the dimension can be relatively large. When some outcomes require a generalised linear or nonlinear mixed model, integration in (5) cannot be done analytically anymore and approximation methods are needed, implying severe restrictions on the number of outcomes that can be incorporated in the outcome vector in order for those approximations to be sufficiently accurate. The pairwise model fitting approach of Fieuws and Verbeke[7], summarized in Section 3.2, then offers a very convenient solution as long as none of the pairwise models is too complex to be fitted using the available standard approximation techniques.

When the association structure is not of any interest, it can be considered a nuisance, and valid inferences for the fixed effects of interest are often still possible, using GEE-type techniques. However, when the focus of the analysis includes certain aspects of the association structure, the construction of the joint model becomes more complex as it implies making assumptions about the within-outcome, the between-outcome, and the cross-outcome association, and inferences of interest can be very sensitive with respect to the assumptions made. In the context of the hearing data example introduced in Section 1, Fieuws and Verbeke[98] illustrated that the association of the evolution of two outcomes and the evolution of the association between two outcomes are two different aspects of the association structure and that both highly depend on assumptions made about seemingly unrelated components in the model.

Often, the choice for a specific type of model is guided by characteristics of the specific problem such as the structure of the data or the measurement scale of the considered outcomes. Also, the background of the researcher will influence this choice. We have shown, however, that various models can severely differ in the assumptions they make and the research questions they answer. While it seems natural that the model choice should predominantly be guided by the subject-matter research question(s) to be answered, this is often hampered by the fact that the various models stem from different research traditions.

## Acknowledgments

# References

1. Diggle, P.; Heagerty, P.; Liang, K.; Zeger, S. Analysis of longitudinal data. Clarendon Press; Oxford: 2002.

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

3. Molenberghs, G.; Verbeke, G. Models for discrete longitudinal data. Springer Series in Statistics; Springer, New-York: 2005.

4. Brant L, Fozard J. Age changes in pure-tone hearing thresholds in a longitudinal study of normal human aging. Journal of the Acoustical Society of America. 1990; 88:813–820. [PubMed: 2212307]

5. Pearson J, Morrell C, Gordon-Salant S, Brant L, Metter E, Klein L, Fozard J. Gender differences in a longitudinal study of age-associated hearing loss. Journal of the Acoustical Society of America. 1995; 97:1196–1205. [PubMed: 7876442]

6. Morrell C, Brant L. Modelling hearing thresholds in the elderly. Statistics in Medicine. 1991; 10:1453–1464. [PubMed: 1925173]

7. Fieuws S, Verbeke G. Pairwise fitting of mixed models for the joint modelling of multivariate longitudinal profiles. Biometrics. 2006; 62(2):424–431. [PubMed: 16918906]

8. Fieuws S, Verbeke G, Molenberghs G. Random-effects models for multivariate repeated measures. Statistical Methods in Medical Research. 2007; 16(4):387–398. [PubMed: 17656450]

9. Shock N, Greullich R, Andres R, Arenberg D, Costa P, Lakatta E, Tobin J. Normal human aging: The Baltimore Longitudinal Study of Aging. National Institutes of Health publication. 1984:84–2450.

10. Molenaar P. A dynamic factor model for the analysis of multivariate time series. Psychometrika. 1985; 50:181–202.

11. Jørgensen B, Lundbye-Christensen S, Song P, Xue-Kun, Sun L. State-space models for multivariate longitudinal data of mixed types. The Canadian Journal of Statistics. 1996; 24:385–402.

12. Galecki A. General class of covariance structures for two or more repeated factors in longitudinal data analysis. Communications in Statistics-Theory and Methods. 1994; 23:3105–3119.

13. O'Brien L, Fitzmaurice G. Analysis of longitudinal multiple-source binary data using generalized estimating equations. Applied Statistics. 2004; 53:177–193.

14. Carey V, Rosner B. Analysis of longitudinally observed irregularly timed multivariate outcomes: Regression with focus on cross-component correlation. Statistics in Medicine. 2001; 20:21–30. [PubMed: 11135345]

15. Daskalakis C, Laird N, Murphy J. Regression analysis of multiple-source longitudinal outcomes: A stirling county depression study. American Journal of Epidemiology. 2002:88–94. [PubMed: 11772789]

16. Molenberghs G, Lesa re E. Marginal modelling of correlated ordinal data using a multivariate plackett distribution. Journal of the American Statistical Association. 1994; 89:633–644.

17. Sklar A. Fonctions de répartition à n dimensions et leur marges. Publications de l'Institut de Statistique de l'Université de Paris. 1959; 8:229–231.

18. Nelsen, R. An introduction to copulas. No. 139 in Lecture Notes in Statistics. Springer-Verlag; New-York: 1998.

19. Plackett R. A class of bivariate distributions. Journal of the American Statistical Association. 1965; 60:516–522.

20. Mardia, K. Families of Bivariate Distributions. Gri n; London: 1970.

21. Lambert P, Vandenhende F. A copula-based model for multivariate non-normal longitudinal data: Analysis of a dose titration safety study on a new antidepressant. Statistics in Medicine. 2002; 21:3197–3217. [PubMed: 12375299]

22. Liang K, Zeger S. Longitudinal data analysis using generalized linear models. Biometrika. 1986; 73:13–22.

23. Prentice R. Correlated binary regression with covariates specific to each binary observation. Biometrics. 1988; 44:1033–1048. [PubMed: 3233244]

24. Carey V, Zeger S, Diggle P. Modelling multivariate binary data with alternating logistic regressions. Biometrika. 1993; 80:517–526.

25. Zhao L, Prentice R. Correlated binary regression using a quadratic exponential model. Biometrika. 1990; 77:642–648.

26. Ten Have T, Morabia A. Mixed effects models with bivariate and univariate association parameters for longitudinal bivariate binary response data. Biometrics. 1999; 55:85–93. [PubMed: 11318182]

27. Rochon J. Analyzing bivariate repeated measures for discrete and continuous outcome variables. Biometrics. 1996; 52:740–750. [PubMed: 8672710]

28. Gray S, Brookmeyer R. Estimating a treatment effect from multidimensional longitudinal data. Biometrics. 1998; 54:976–988. [PubMed: 9750246]

29. Gray S, Brookmeyer R. Multidimensional longitudinal data: estimating a treatment effect from continuous, discrete or time-to-event response variables. Journal of the American Statistical Association. 2000; 95:396–406.

30. Geys H, Molenberghs G, Ryan L. Pseudolikelihood modeling of multivariate outcomes in developmental studies. Journal of the American Statistical Association. 1999; 94:734–745.

31. Zeng, L.; Cook, R. Tech. rep. University of Water-loo; 2004. Transition models for multivariate longitudinal binary data; p. 2004-038.Working Paper

32. Liang K, Zeger S. A class of logistic regression models for multivariate binary time series. Journal of the American Statistical Association. 1989; 84:447–451.

33. Zhang M, Tsiatis A, Davidian M, Pieper K, Maha ey K. Inference on treatment effects from a randomized clinical trial in the presence of premature treatment discontinuation: The SYNERGY trial. Biostatistics. 2011; 12:258–269. [PubMed: 20797983]

34. Laird N, Ware J. Random-effects models for longitudinal data. Biometrics. 1982; 38:963–974. [PubMed: 7168798]

35. Breslow N, Clayton D. Approximate inference in generalized linear mixed models. Journal of the American Statistical Association. 1993; 88:9–25.

36. Davidian, M.; Giltinan, D. Nonlinear models for repeated measurement data. Chapman & Hall; 1995.

37. Goldstein, H. Multilevel Statistical Models. Wiley; New York: 1995.

38. Bryk, A.; Raudenbush, S. Hierarchical linear models in social and behavioral research: Applications and data analysis methods. Sage; Newbury Park, CA: 1992.

39. Longford N. A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. Biometrika. 1987; 74:817–827.

40. McCulloch C. Joint modelling of mixed outcome types using latent variables. Statistical Methods in Medical Research. 2008; 17:53–73. [PubMed: 17855745]

41. Duchateau, L.; Janssen, P. The frailty model. Springer-Verlag; New-York: 2008.

42. Fitzmaurice, G.; Davidian, M.; Verbeke, G.; Molenberghs, G. Longitudinal data analysis. Handbooks of Modern Statistical Methods; Chapman & Hall/CRC: 2009.

43. Reinsel G. Estimation and prediction in a multivariate random effects generalized linear model. Journal of the American Statistical Association. 1984; 79:406–414.

44. Beckett L, Tancredi D, Wilson R. Multivariate longitudinal models for complex change processes. Statistics in Medicine. 2004; 23:231–239. [PubMed: 14716725]

45. Chakraborty H, Helms R, Sen P, Cohen M. Estimating correlation by using a general linear mixed model: Evaluation of the relationship between the concentration of HIV-1 RNA in blood and semen. Statistics in Medicine. 2003; 22:1457–1464. [PubMed: 12704609]

46. MacCallum R, Kim C, Malarkey W, Kiecolt-Glaser J. Studying multivariate change using multilevel models and latent curve models. Multivariate Behavioral Research. 1997; 32:215–253.

47. Matsuyama Y, Ohashi Y. Mixed models for bivariate response repeated measures data using Gibbs sampling. Statistics in Medicine. 1997; 16:1587–1601. [PubMed: 9257414]

48. Shah A, Laird N, Schoenfeld D. A random-effects model for multiple characteristics with possibly missing data. Journal of the American Statistical Association. 1997; 92:775–779.

49. Sy J, Taylor J, Cumberland W. A stochastic model for the analysis of bivariate longitudinal aids data. Biometrics. 1997; 53:542–555. [PubMed: 9192450]

50. Heitjan D, Sharma D. Modelling repeated-series longitudinal data. Statistics in Medicine. 1997; 16:347–355. [PubMed: 9044525]

51. Ribaudo H, Thompson S. The analysis of repeated multivariate binary quality of life data: A hierarchical model approach. Statistical Methods in Medical Research. 2002; 11:69–83. [PubMed: 11923995]

52. Agresti A. A model for repeated measurements of a multivariate binary response. Journal of the American Statistical Association. 1997; 92:315–321.

53. Fieuws S, Verbeke G, Maes B, Renterghem YV. Predicting renal graft failure using multivariate longitudinal profiles. Biostatistics. 2008; 9:419–431. [PubMed: 18056686]

54. Blozis S. Structured latent curve models for the study of change in multivariate repeated measures. Psychological Methods. 2004; 9:334–353. [PubMed: 15355152]

55. Blozis S. On fitting nonlinear latent curve models to multiple variables measured longitudinally. Structural Equation Modeling. 2007; 14:179–201.

56. Thum Y. Hierarchical linear models for multivariate outcomes. Journal of Educational and Behavioral Statistics. 1997; 22(1):77–108.

57. Nagin D, Land K. Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed poisson model. Criminology. 1993; 31:327–362.

58. Nagin D. Analysing developmental trajectories: Semi-parametric, group-based approach. Psychological Methods. 1999; 4:139–177.

59. Nagin D, Tremblay R. Analyzing developmental trajectories of distinct but related behaviors: A group-based method. Psychological Methods. 2001; 6:18–34. [PubMed: 11285809]

60. Muthén B, Beyond SEM. General latent variable modeling. Behaviormetrika. 2002; 29:81–117.

61. Molenberghs G, Verbeke G, Demétrio C, Viera A. A family of generalized linear models for repeated measures with normal and conjugate random effects. Statistical Science. 2010; 25:325–347.

62. Vangeneugden T, Molenberghs G, Verbeke G, Demétrio C. Marginal correlation from an extended random-effects model for repeated and overdispersed counts. Journal of Applied Statistics. 2010; 38:215–232.

63. Njeru Njagi E, Molenberghs G, Verbeke G, Kenward M, Dendale P, Willekens K. A flexible joint-modelling framework for longitudinal and time-to-event data with overdispersion. submitted. 2011

64. Stoolmiller M. Antisocial behavior, delinquent peer association and un-supervised wandering for boys:growth and change from childhood to early adolescence. Multivariate Behavioral Research. 1994; 29:263–288.

65. Willett, J.; Keily, M. Using covariance structure analysis to model change over time. In: Tinsley, HEA.; Brown, SD., editors. Handbook of applied multivariate statistics and mathematical modelling. Vol. 23. Academic Press; San Diego: 2000. p. 665-669.

66. Ferrer E, McArdle JJ. Alternative structural models for multivariate longitudinal data. Structural Equation Modeling. 2003; 10:493–524.

67. Thiébaut R, Jacqmin-Gadda H, Chêne G, Leport C, Commenges D. Bivariate linear mixed models using SAS PROC MIXED. Computer Methods and Programs in Biomedicine. 2002; 69:249–256. [PubMed: 12204452]

68. Liu L, Yu Z. A likelihood reformulation method in non-normal random effects models. Statistics in Medicine. 2007; 27:3105–3124. [PubMed: 18038445]

69. Nelson K, Lipsitz S, Fitzmaurice G, Ibrahim J, Parzen M, Strawderman R. Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. Journal of Computational and Graphical Statistics. 2006; 15:39–57.

70. Arnold B, Strauss D. Pseudo-likelihood estimation: some examples. Sankhya: The Indian Journal of Statistics - Series B. 1991; 53:233–243.

71. Geys H, Molenberghs G, Ryan L. Pseudo-likelihood inference for clustered binary data. Communications in Statistics: Theory and Methods. 1997; 26:2743–2767.

72. Lindsay B. Composite likelihood methods. Contemporary Mathematics. 1988; 80:221–239.

73. Heagerty P, Lele S. A composite likelihood approach to binary spatial data. Journal of the American Statistical Association. 1998; 93:1099–1111.

74. Curriero F, Lele S. A composite likelihood approach to semivariogram estimation. Journal of Agricultural, Biological, and Environmental Statistics. 1999; 4:9–28.

75. Fieuws S, Verbeke G, Boen F, Delecluse C. High-dimensional multivariate mixed models for binary questionnaire data. Applied Statistics. 2006; 55(4):1–12.

76. Oort F. Three-mode models for multivariate longitudinal data. British journal of mathematical and statistical psychology. 2001; 54:49–78. [PubMed: 11393902]

77. Oort F. Stochastic three-mode models for mean and covariance structures. British journal of mathematical and statistical psychology. 1999; 52:243–272.

78. Sivo S. Multiple indicator stationary time series models. Structural Equation Modeling. 2001; 8:599–612.

79. Fieuws, F.; Verbeke, G. Joint models for high-dimensional longitudinal data. In: Fitzmaurice, G.; Davidian, M.; Verbeke, G.; Molenberghs, G., editors. Longitudinal data analysis. Vol. 16. Handbooks of Modern Statistical Methods; Chapman & Hall/CRC: 2009. p. 367-391.

80. Kiers H, ten Berge J. Hierarchical relations between methods for simultaneous component analysis and a technique for rotation to a simple simultaneous structure. British Journal of Mathematical and Statistical Psychology. 1994; 47:109–126.

81. Timmerman M, Kiers H. Four simultaneous component models for the analysis of multivariate time series from more than one subject to model intraindividual and interindividual differences. Psychometrika. 2003; 1:105–121.

82. Timmerman M, Kiers H. Three-way component analysis with smooth-ness constraints. Computational Statistics and Data Analysis. 2002; 40:447–470.

83. Kiers H, ten Berge J, Bro R. PARAFAC2-Part 1: A direct fitting algorithm for the PARAFAC2 model. Journal of Chemometrics. 1999; 13:275–294.

84. Ilk O, Daniels M. Marginalised transition random effects models for multivariate longitudinal binary data. Canadian Journal of Statistics. 2007; 35:105–123.

85. Liu L, Hedeker D. A mixed-effects regression model for longitudinal multivariate ordinal data. Biometrics. 2005; 3 doi: 10.1111 / j.1541-0420.2005.00408.x.

86. Reboussin B, Reboussin D. Latent transition modeling of progression of health-risk behavior. Multivariate Behavioral Research. 1998; 33:457–478.

87. McArdle, J. Dynamic but structural equation modeling of repeated measures data. In: Nesselroade, J.; Catell, R., editors. Handbook of multivariate experimental psychology. American Psychological Association; Washington DC: 1988. p. 561-614.

88. Duncan S, Duncan T. Multivariate latent growth curve analysis of adolescent substance abuse. Structural Equation Modeling. 1996; 3:323–347.

89. Sayer, A.; Cumsille, P. Second-order latent growth models. In: Collins, L.; Sayer, A., editors. New Methods for the Analysis of Change. Vol. 6. American Psychological Association; Washington DC: 2001.

90. Hancock G, Kuo W, Lawrence F. An illustration of second-order latent growth models. Structural Equation Modeling. 2001; 8:470–489.

91. Roy J, Lin X. Latent variable models for longitudinal data with multiplecontinuous outcomes. Biometrics. 2000; 56:1047–1054. [PubMed: 11129460]

92. Chan D. The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). Organizational Research Methods. 1998; 1:421–483.

93. Blozis, S. A second-order structured latent curve model for longitudinal data. In: van Montfort, K.; Oud, H.; Satorra, A., editors. Longitudinal models in the behavioural and related sciences. Lawrence Erlbaum Associates; 2006. p. 189-214.

94. Harring J. A nonlinear mixed effects model for latent variables. Journal of Educational and Behavioural Statistics. 2009; 34:293–318.

95. Bandyopadhyay S, Ganguli B, Chatterjee A. A review of multivariate longitudinal data analysis. Statistical Methods in Medical Research. 2011; 20:299–330. [PubMed: 20212072]

96. Verbeke, G.; Davidian, M. Joint models for longitudinal data: Introduction and overview. In: Fitzmaurice, G.; Davidian, M.; Verbeke, G.; Molenberghs, G., editors. Longitudinal data analysis. Vol. 13. Handbooks of Modern Statistical Methods; Chapman & Hall/CRC: 2009. p. 319-326.

97. Verbeke, G.; Molenberghs, G.; Rizopoulos, D. Random effects models for longitudinal data. In: Montfort, K.; HO, H.; Satorra, A., editors. Longitudinal research with latent variables. Vol. 2. Springer-Verlag; New York: 2010. p. 37-96.

98. Fieuws S, Verbeke G. Joint modelling of multivariate longitudinal profiles: Pitfalls of the random-effects approach. Statistics in Medicine. 2004; 23:3093–3104. [PubMed: 15449333]