T E C H N I C A L
R E P O R T

0527

EVALUATION OF THE PAIRWISE APPROACH FOR FITTING JOINT LINEAR MIXED MODELS: A SIMULATION STUDY

FIEUWS, S. and G. VERBEKE

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Evaluation of the Pairwise Approach for Fitting Joint Linear Mixed Models: A Simulation Study

Steffen Fieuws and Geert Verbeke

Biostatistical Centre, Katholieke Universiteit Leuven, Leuven, Belgium

Summary

A mixed model is a flexible tool for joint modelling purposes, especially when the gathered data are unbalanced. However, computational problems due to the dimension of the joint covariance matrix of the random effects arise as soon as the number of outcomes and/or the number of used random effects per outcome increases. We propose a pairwise approach in which all possible bivariate models are fitted, and where inference follows from pseudo-likelihood arguments. The approach is applicable for linear, generalised linear and nonlinear mixed models, or for combinations of these. This paper evaluates the performance of the pairwise approach for joint linear mixed models using a set of simulation studies.

Key words: Joint Modelling, Multivariate Longitudinal Profiles, Mixed Models, Pseudo-likelihood, Correlated Curves

---

[1]Correspondence to: Steffen Fieuws, Biostatistical Centre, Katholieke Universiteit Leuven, Kapucijnenvoer 35, B-3000 Leuven, Belgium. E-mail : steffen.fieuws@med.kuleuven.ac.be

1. Introduction

Most models proposed in the literature for the analysis of longitudinal data, restrict attention to the analysis of one single outcome variable, measured repeatedly over time (analysis of univariate longitudinal data). Multivariate longitudinal data arise when, instead of a single outcome, a set of different outcomes on the same unit is measured repeatedly over time. As an example of multivariate longitudinal data, consider the situation where different indices of physical and mental health are measured repeatedly over time. In many situations joint modelling of the multivariate longitudinal profiles is needed or has additional advantages over the separate analyses of the different outcomes (Fieuws and Verbeke, 2004).

A flexible approach is to model the different outcomes jointly by using random effects models. In a joint-modelling approach using mixed models, random-effects are assumed for each outcome process, and by imposing a joint multivariate distribution on the random effects, the different processes are associated. This approach has many advantages and is applicable in a wide variety of situations. First, the data can be highly unbalanced. For example, it is not necessary that all outcomes are measured at the same timepoints. Moreover, the approach is applicable in situations where linear, nonlinear or generalised linear mixed models are used to describe the evolution of the individual outcome processes. Also, models can be constructed joining different types of mixed models. For example, a generalised linear mixed model for a discrete outcome combined with a nonlinear mixed model for a continuous outcome. Examples of joint model using the random-effects approach can be found in Buyse et al. 2000, Burzykowski et al. 2001), Thum

(1997), Gueorguieva (2001a), Chakraborty et al. (2003), MacCallum et al. (1997), Thiébaut et al. (2002) and Shah et al. (1997). All these examples refer to situations where the number of different outcomes is low, leading to random vectors of a relatively low dimension. The focus of this paper is to propose a method which allows this random-effects approach for much higher dimensions and to evaluate this method using simulation studies.

This paper is organised as follows. Section 2 introduces the joint modelling approach using mixed models and describes the dimensionality problem. Section 3 details the pairwise fitting approach as a solution to this problem, Section 4 presents the results of a set of simulation studies and Section 5 contains a discussion.

2. Joint Modelling : Random-Effects Approach

2.1 *Univariate Mixed Models*

Let m be the number of different outcomes to be modelled jointly. It will be assumed that each of the m longitudinally measured outcomes can be modelled using a mixed model. More specifically, for one outcome, let y_{ij} denote the jth measurement available for the ith subject, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$, and let $\boldsymbol{y_i}$ denote the vector of all measurements for the ith subject, i.e., $\boldsymbol{y_i}' = (y_{i1}, \ldots, y_{in_i})$. Our general model assumes that $\boldsymbol{y_i}$ (possibly appropriately transformed) satisfies

$$\boldsymbol{y_i}|\boldsymbol{b_i} \ \sim \ F_i(\boldsymbol{\psi}, \boldsymbol{b_i}), \tag{1}$$

i.e., conditional on $\boldsymbol{b_i}$, $\boldsymbol{y_i}$ follows a pre-specified distribution F_i, possibly depending on covariates, and parameterized through a vector $\boldsymbol{\psi}$ of unknown parameters, common to all subjects. Further, $\boldsymbol{b_i}$ is a q-dimensional normal vector of subject-specific parameters with mean zero and covariance matrix

D. Often conditional independence is assumed, implying that the components y_{ij} in $\boldsymbol{y_i}$ are independent, conditionally on $\boldsymbol{b_i}$. The distribution function F_i in (1) then becomes a product over the n_i independent elements in $\boldsymbol{y_i}$. In general, unless a fully Bayesian approach is followed, inference is based on the marginal distribution of $\boldsymbol{y_i}$ with density

$$f_i(\boldsymbol{y_i}) \quad = \quad \int f_i(\boldsymbol{y_i}|\boldsymbol{b_i})g(\boldsymbol{b_i})d\boldsymbol{b_i}, \tag{2}$$

in which $f_i(\boldsymbol{y_i}|\boldsymbol{b_i})$ and $g(\boldsymbol{b_i})$ denote the conditional density of $\boldsymbol{y_i}$ given $\boldsymbol{b_i}$, and the density of $\boldsymbol{b_i}$, respectively. Estimation and inference of the parameters in $f_i(\boldsymbol{y_i})$ is based on maximum likelihood principles, assuming independence across subjects. Special cases are linear and nonlinear mixed models for continuous data, and generalised linear mixed models for discrete data. Linear mixed models assume that the n_i-dimensional vector $\boldsymbol{y_i}$ satisfies

$$\boldsymbol{y_i}|\boldsymbol{b_i} \quad \sim \quad N(X_i\boldsymbol{\beta} + Z_i\boldsymbol{b_i}, \Sigma_i), \tag{3}$$

where X_i and Z_i are $(n_i \times k)$ and $(n_i \times q)$ dimensional matrices of known covariates, $\boldsymbol{\beta}$ is a k-dimensional vector of regression parameters, called fixed effects, and Σ_i is a $(n_i \times n_i)$ covariance matrix which depends on i only through its dimension n_i, i.e. the set of unknown parameters in Σ_i will not depend upon i. The conditional independence assumption then reduces to $\Sigma_i = \sigma^2 I_{n_i}$. The latter restriction is not a prerequisite for joint mixed models, e.g. a serial correlation structure could be allowed. Marginally, $\boldsymbol{y_i}$ follows a normal distribution with mean $X_i\boldsymbol{\beta}$ and covariance matrix $V_i = Z_iDZ_i' + \Sigma_i$. Non-linear mixed models can be considered as an extension of model (3) in the sense that they replace the linear predictor $X_i\boldsymbol{\beta} + Z_i\boldsymbol{b_i}$ by $h(X_i, Z_i, \boldsymbol{\beta}, \boldsymbol{b_i})$ for some known 'link' function h. The generalized linear mixed model assumes that conditionally on random effects $\boldsymbol{b_i}$, the elements y_{ij} of $\boldsymbol{y_i}$ are independent, with density function of the form

$$f_i(y_{ij}|\boldsymbol{b_i}) \quad = \quad \exp\left[(y_{ij}\eta_{ij} - a(\eta_{ij}))/\phi + c(y_{ij}, \phi)\right],$$

with mean $\mathrm{E}(y_{ij}|\boldsymbol{b_i}) = a'(\eta_{ij}) = \mu_{ij}(\boldsymbol{b_i})$ and variance $\mathrm{Var}(y_{ij}|\boldsymbol{b_i}) = \phi a''(\eta_{ij})$, and with $h(\boldsymbol{\mu_i}(\boldsymbol{b_i})) = X_i\boldsymbol{\beta} + Z_i\boldsymbol{b_i}$ for a known link function h. For generalized linear and the nonlinear mixed models, the integral in (2) cannot be calculated analytically such that numerical approximations are required.

2.2 *Joint Mixed Model*

The joint model assumes a mixed model for each outcome, and these univariate models are combined through specification of a joint multivariate distribution for all random effects. Obviously, the joint model can be considered as a new mixed model of the form (1), but with a random-effects vector $\boldsymbol{b_i}$ of a higher dimension. Let $\boldsymbol{\Theta^*}$ the vector containing all parameters (fixed effects parameters as well as covariance parameters), then $l_i(\boldsymbol{Y}_{1i}, \boldsymbol{Y}_{2i}, \ldots, Y_{mi}|\boldsymbol{\Theta^*})$ refers to the loglikelihood contribution of subject i to the full joint mixed model. Strictly speaking, standard software for linear, non-linear or generalised linear mixed models can be used to obtain parameter estimates for this joint mixed model. Examples based on SAS can be found in Thiébaut et al. (2002b). However, computational problems will arise as the dimension of the random-effects vector $\boldsymbol{b_i}$ in the joint model increases, even in the case of linar mixed models where the marginal density in (2) of $\boldsymbol{y_i}$ can be calculated analytically. The approach presented in Section 3 solves this by fitting joint models to all pairs of outcomes separately.

3. Pairwise Modelling Approach

To resolve the computational complexity of high-dimensional joint random effect models, the dimensionality of the problem needs to be reduced. This is obtained by fitting in a first step all pairwise bivariate models separately, instead of maximising the likelihood of the full joint model presented in the previous section. Assuming the full joint model is correct, all possible pairwise models are correct. In a maximum likelihood framework, each pairwise model yields estimates, with classical optimal asymptotic properties, for a part in $\mathbf{\Theta^*}$. For some elements in $\mathbf{\Theta^*}$ multiple ML estimates will be obtained. Therefore, in a second step, these estimates will be combined to obtain one single estimate for each parameter in $\mathbf{\Theta^*}$ of the full joint model.

3.1 *Fitting Pairwise Models*

Instead of maximising the loglikelihood of the joint mixed model, loglikelihoods of the following form will be maximised separately

$$\sum_{i=1}^{N} l_{rsi}(\boldsymbol{Y}_{ri}, \boldsymbol{Y}_{si}|\mathbf{\Theta}_{r,s}), \tag{4}$$

$r = 1, \ldots, m-1, s = r+1, \ldots, m$. The total number of subjects is indicated by N. $\mathbf{\Theta}_{r,s}$ represents the vector of all parameters in the bivariate joint mixed model corresponding to the specific pair (r, s). To simplify the notation in the remainder, expression (4) can be rewritten as $\sum_{i=1}^{N} l_{pi}(\mathbf{\Theta}_p)$ with $p = 1, \ldots, P$ and $P = m(m-1)/2$ representing the total number of possible pairs. Finally, let $\mathbf{\Theta}$ then be the stacked vector combining all pair-specific parameter vectors $\mathbf{\Theta}_p$. Estimates for the elements in $\mathbf{\Theta}$ are obtained by maximising each of the P likelihoods separately.

It is important to note that the parameter vectors $\mathbf{\Theta}$ and $\mathbf{\Theta^*}$ are not equivalent. Indeed, some parameters in $\mathbf{\Theta^*}$ will have a single counterpart in $\mathbf{\Theta}$, e.g. the covariance between random effects from two different outcomes. Other elements in $\mathbf{\Theta^*}$ will have multiple counterparts in $\mathbf{\Theta}$, e.g. the covariance between random effects from the same outcome. In the latter case a single estimate is obtained by averaging all corresponding pair-specific ML estimates in $\widehat{\mathbf{\Theta}}$. By definition, this linear combination of ML estimates shares the same nice asymptotic proporties of its constituting elements. Calculation of the precision of this average is adressed in the next subsection.

3.2 *Inference for* $\mathbf{\Theta}$

Standard errors of the so-obtained estimates clearly cannot be obtained from averaging standard errors or variances. Indeed, the variability amongst the pair-specific estimates needs to be taken into account. Furthermore, two pair-specific estimates corresponding to two pairwise models with a common outcome are based on overlapping information and hence correlated. This correlation should also be accounted for in the sampling variability of the combined estimates in $\widehat{\mathbf{\Theta}}^*$. Borrowing ideas from the pseudo-likelihood framework, first a covariance matrix for the elements in $\widehat{\mathbf{\Theta}}$ will be constructed. The idea behind pseudo-likelihood estimation (Besag, 1975) is to replace the joint likelihood by a suitable product of marginal or conditional densities, where this product is easier to evaluate than the original likelihood. Examples of pseudo-likelihood estimation can be found in Arnold and Strauss (1991), Geys et al. (1997) and Renard et al. (2004). Although in the pairwise approach a set of likelihoods is maximised separately, the approach fits within the pseudo-likelihood framework. Indeed, fitting all possible pairwise models is equivalent to maximising a pseudo-likelihood function of the following form

$$\begin{array}{rcl} pl(\mathbf{\Theta}) & = & l(\boldsymbol{Y}_1, \boldsymbol{Y}_2|\mathbf{\Theta}_{1,2}) + l(\boldsymbol{Y}_1, \boldsymbol{Y}_3|\mathbf{\Theta}_{1,3}) + \ldots + l(\boldsymbol{Y}_{m-1}, \boldsymbol{Y}_m|\mathbf{\Theta}_{m-1,m}) \\ & = & \sum_{p=1}^{P} l_p(\mathbf{\Theta}_p). \end{array} \tag{5}$$

Note however that, in the classical examples of pseudo-likelihood estimation, the same parameter is present in the different parts of the pseudo-likelihood function whereas in (5) the set of parameters in $\mathbf{\Theta}_p$ is considered pair-specific (subscript p at this stage, only at a later stage the estimates will be combined). This separate parametrisation is needed to be able to maximise the different parts in expression (5) separately. Since the pairwise approach fits within the pseudo-likelihood framework, an asymptotic multivariate normal distribution for $\widehat{\mathbf{\Theta}}$ can be derived. Asymptotic normality of the pseudo-likelihood estimator in the single parameter case and in the vector valued parameter case is shown in Arnold and Strauss (1991), and in Geys (1999), respectively. The asymptotic multivariate normal distribution for $\widehat{\mathbf{\Theta}}$ is given by

$$\sqrt{N}(\widehat{\mathbf{\Theta}} - \mathbf{\Theta}) \sim MVN(\mathbf{0}, J^{-1}KJ^{-1}) \tag{6}$$

where J is a block-diagonal matrix with diagonal blocks J_{pp}, and where K is a symmetric matrix containing blocks K_{pq}, given by

$$J_{pp} = \frac{1}{N}\sum_{i=1}^{N} E\left(\frac{\partial^2 l_{pi}}{\partial\boldsymbol{\theta}_p\partial\boldsymbol{\theta}'_p}\right), \quad K_{pq} = \frac{1}{N}\sum_{i=1}^{N} E\left(\frac{\partial l_{pi}}{\partial\boldsymbol{\theta}_p}\frac{\partial l_{qi}}{\partial\boldsymbol{\theta}'_q}\right), \quad p, q = 1, \ldots, P$$

Estimates are obtained by dropping the expectations and replacing the unknown parameters by their estimates.

3.3 *Combining Information: Inference for* $\mathbf{\Theta^*}$

In a final step, estimates for the parameters in $\mathbf{\Theta^*}$ can be calculated, as suggested before, by taking averages over all pairs. This is obtained by $\widehat{\mathbf{\Theta}}^* = A\widehat{\mathbf{\Theta}}$ with $\widehat{\mathbf{\Theta}}^*$ following a multivariate normal distribution with mean $\mathbf{\Theta^*}$ and covariance matrix $A\Sigma(\widehat{\mathbf{\Theta}})A'$. A is a matrix containing the appropriate coefficients to calculate the averages and $\Sigma(\widehat{\mathbf{\Theta}})$ equals the covariance matrix for $\widehat{\mathbf{\Theta}}$ obtained by expression (6).

4. Simulation Study

Four simulation studies were conducted to evaluate the performance of the estimators from the pairwise approach. In a first study, the performance of the pairwise approach will be evaluated in a setting where the full joint model is trivariate. The aim of this study is (1) to confirm the expected unbiasedness of the estimates (based on theoretical arguments), (2) to verify whether the pairwise approach yields valid standard errors (meaning that the standard errors reflect the sampling variability), (3) to verify whether the standard errors are robust agains model misspecification and (4) to study the possibility of efficiency loss in the pairwise approach. In a second simulation study the parameter recovery has been verified in a higher dimensional setting. A third and fourth simulation study have been performed to study into detail the efficiency loss observed in study 1. The latter two studies are based on an artificial dataset, whereas the first two simulation studies are based on a real dataset containing longitudinal pure-tone hearing tresholds.

4.1 *Pure-Tone Hearing Thresholds*

In a hearing test, hearing threshold sound pressure levels (dB) are determined at different frequencies to evaluate the hearing performance of a subject. A hearing threshold is the lowest signal intensity a subject can detect at a specific frequency. In this study, hearing thresholds measured at eleven different frequencies (125Hz, 250Hz, 500Hz, 750Hz, 1000Hz, 1500Hz, 2000Hz, 3000Hz, 4000Hz, 6000Hz and 8000Hz), obtained on 603 male participants from the Baltimore Longitudinal Study of Aging (BLSA, Shock et al. 1984), are considered. Hearing thresholds are measured at the left as well as at the right ear, leading to 22 outcomes measured repeatedly over time. The number of visits per subject varies from 1 to 15 (a median follow-up time of 6.9 years). Visits are unequally spaced. The age at first visit of the participants ranges from 17.2 to 87 years (with a median age at first visit of 50.2 years). Analyses of the hearing data collected in the BLSA study can be found in Brant and Fozard (1990), Morrell and Brant (1991) and Pearson et al. (1995). Ear- and frequency specific profiles (for all 22 outcomes) for one randomly chosen subject are shown in Figure 1.

[Figure 1 about here.]

4.2 *Study 1*

In the first study, three outcomes were selected (500 Hz, 1000 Hz and 2000 Hz, all taken at the right ear) from the example on the hearing thresholds. Verbeke and Molenberghs (2000) proposed the following linear mixed model to analyse the evolution of the hearing threshold for a single frequency. Let $Y_i(t)$ denote the hearing threshold at some frequency for a subject i taken at time t, the model is specified as

$$\begin{aligned} Y_i(t) &= (\beta_1 + \beta_2 Age_i + \beta_3 Age_i^2 + a_i) \\ &\quad +(\beta_4 + \beta_5 Age_i + b_i)t \\ &\quad +\beta_6 Visit1(t) + \varepsilon_i(t) \end{aligned} \tag{7}$$

in which t is time expressed in years from entry in the study and Age_i equals the age of subject i at the time of entry in the study. Since there is evidence for the presence of a learning effect from the first to the subsequent visits, a time-varying covariate $Visit1$ has been added. This covariate is defined to be one at the first measurement and zero for all other visits. Finally, the a_i are random intercepts, the b_i are the random slopes for time, and the ε_i represent the usual error components, independent of the random effects. The vector $(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)'$ of fixed effects describes the average evolution of the hearing threshold and the vector $(a_i, b_i)'$ of random effects describes how the profile of the ith subject deviates from the average profile. The classical normality assumptions apply for all random terms in this model.

Considering the 3 outcomes, let $Y_{1i}(t)$, $Y_{2i}(t)$, $Y_{3i}(t)$ denote the hearing thresholds of the 3 frequencies for a subject i taken at time t. Each of the 3 responses is described using the linear mixed-effects model (7). More specifically,

$$\begin{cases} Y_{1i}(t) = \mu_1(t) + a_{1i} + b_{1i}t + \varepsilon_{1i}(t) \\ Y_{2i}(t) = \mu_2(t) + a_{2i} + b_{2i}t + \varepsilon_{2i}(t) \\ Y_{3i}(t) = \mu_3(t) + a_{3i} + b_{3i}t + \varepsilon_{3i}(t) \end{cases}$$

where $\mu_1(t)$, $\mu_2(t)$, $\mu_3(t)$ refer to the average evolutions:

$$\begin{cases} \mu_1(t) &= \beta_{11} + \beta_{21} Age_i + \beta_{31} Age_i^2 + (\beta_{41} + \beta_{51} Age_i)t + \beta_{61} Visit1(t) \\ \mu_2(t) &= \beta_{12} + \beta_{22} Age_i + \beta_{32} Age_i^2 + (\beta_{42} + \beta_{52} Age_i)t + \beta_{62} Visit1(t) \\ \mu_3(t) &= \beta_{13} + \beta_{23} Age_i + \beta_{33} Age_i^2 + (\beta_{43} + \beta_{53} Age_i)t + \beta_{63} Visit1(t) \end{cases}$$

with the second subscript for each β parameter referring to outcomes 1, 2 or 3. The 3 outcome trajectories are tied together through a joint distribution for the random effects

$$\left(\; a_{1i}, a_{2i}, a_{3i}, b_{1i}, b_{2i}, b_{3i} \; \right)' \sim N\left(\mathbf{0}, \boldsymbol{D}\right)$$

where $\boldsymbol{D}$ is the 6×6 covariance matrix of the random effects. The error components,

$$\left(\; \varepsilon_{1i}(t), \varepsilon_{2i}(t), \varepsilon_{3i}(t) \; \right)' \sim N\left(\mathbf{0}, R\right) \quad \text{for all } t,$$

are assumed independent of the random effects. Note that the error components are assumed independent across timepoints.

Two different settings of the above joint linear mixed model are considered. In a first setting (scenario A), all fixed and random effects were considered to be outcome-specific, and unstructured 6×6 and 3×3 covariance matrices were used for the random effects and the error components respectively. In a second settting (Scenario B), the model where the data are drawn from is simplified by considering the interaction between the linear evolution and age at first entry, as well as the random linear evolution to be common to the three outcomes. This implies that the following restrictions apply under scenario B:

$$\begin{cases} \beta_{51} &= \beta_{52} &= \beta_{53} \\ b_{i1} &= b_{i2} &= b_{i3} \end{cases}$$

Both models will be denoted as full trivariate mixed models in the remainder. In both settings, 1000 data sets were simulated from the full trivariate model with parameters obtained from the hearing data. Each time, the full trivariate model was fitted, and the pairwise approach was applied as well. In scenario A, two versions of the model were used: The correct model and an incorrect model. The latter model does not contain random slopes and the error components are assumed to be uncorrelated. In the full trivariate approach, model-based as well as robust standard errors are calculated. The latter are the asymptotically consistent 'sandwich' estimators described in Liang and Zeger (1986) and in Diggle, Heagerty, Liang and Zeger (1994).

4.2.1 Bias Tables 1 and 2 present results obtained under scenario A, and tables 3 and 4 show results obtained under scenario B. Comparison of the parameter-specific biases obtained with the pairwise and the trivariate approach clearly indicates that the pairwise approach yields unbiased estimates, irrespective of whether parameters are shared or not by a set of outcomes. Note that this result was expected from theoretical arguments presented in section 3.1.

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

[Table 4 about here.]

Additionally, to express the degree of agreement between the results from the pairwise and the trivariate approach, intra-class correlations have been calculated for each parameter. In a perfectly balanced dataset with all fixed and random effects parameters outcome-specific, it can be expected from theory on balanced growth curve models (Lange and Laird, 1989; Reinsel, 1982) that there will be a perfect agreement between estimates obtained under the trivariate model and estimates obtained under the pairwise models. However, this perfect agreement does not hold anymore if there are parameters common to a set of outcomes and/or if the dataset is unbalanced. In scenario A, the ICC's vary from 0.981 to 0.999 for the 27 covariance parameters and from 0.993 to 0.999 for the 18 fixed effects parameters. In scenario B, the ICC's vary from 0.974 to 0.998 for the 16 covariance parameters and from 0.977 to 0.999 for the 16 fixed effects parameters.

4.2.2 Valid Standard Errors Tables 1 and 3 contain also information on the sampling variability ($\hat{\sigma}_\beta$) of the 1000 estimates of each fixed effect. Comparing this sampling variability with the mean of the standard errors ($\overline{se}$) clearly indicates that the pairwise method yields standard errors which reflect the true sampling variability.

4.2.3 Robustness Table 5 presents the results for the fixed effects obtained under the incorrect model of scenario A (the model which does not contain random slopes and assumes that the error components are uncorrelated). The results of the incorrect model allow to assess the robustness (to model misspecification) of the standard errors obtained under the pairwise approach. Columns 2 and 3 give an indication of the sampling variability of the 1000 estimates under the pairwise and the trivariate approach respectively. Column 4 gives the mean of the standard errors obtained under the pairwise approach. Columns 5 and 6 give the mean of the model-based and the empirically corrected standard errors obtained under the trivariate approach respectively. The ratio's presented in columns 7, 8 and 9 indicate if the standard errors obtained under the pairwise approach, the model-based standard errors of the trivariate approach, and the empirically corrected standard errors of the trivariate approach, reflect the true sampling variability. As could be expected (Diggle, Heagerty, Liang and Zeger, 2002), this ratio fluctuates for all parameters around 1 (column 9) if empirically corrected standard errors are used. This is not the case for the model-based standard errors. Indeed, these standard errors are no longer valid. This is especially true for the parameters involving the evolution over time, where the ratio's range from 0.66 to 0.77. Contrary to this, the standard errors of the pairwise approach are valid. Column 7 indicates that for all parameters the ratio is close to 1, ranging from 0.969 to 1.016.

[Table 5 about here.]

4.2.4 Efficiency The last column in Tables 1, 2, 3 and 4 gives for each parameter the relative efficiency, expressed as the ratio of the mean squared errors of the estimates obtained with the trivariate and the pairwise approach. An RE lower than 1 thus indicates presence of efficiency loss. When applying the pairwise approach on a joint mixed model where all parameters are outcome-specific (scenario A), there is no clear indication of efficiency loss. However, when some parameters are common to the three outcomes (scenario B), some efficiency loss is clearly present. Figure 2 presents a histogram of the relative efficiencies of all parameters, comparing scenario A and scenario B. In scenario B, of the 32 parameters, 19 have a relative efficiency higher than 0.99. The lowest observed relative efficiency equals 0.931. A possible explanation is that maximisation in the pairwise approach is less restricted than in the full trivariate model. E.g., for the interaction between the slope and age at entry (β_5), the constraint in the trivariate model is $\beta_{51} = \beta_{52} = \beta_{53}$, whereas in the pairwise approach the less stringent constraints $\beta_{51} = \beta_{52}$, $\beta_{51} = \beta_{53}$ and $\beta_{52} = \beta_{53}$ apply respectively in each of the three pairwise models. Obviously, the less restricted the parameter space over which the maximisation takes place, the more variability will be observed. However, the efficiency loss for β_5 under scenario B seems rather low (RE=0.989), whereas other parameters, not subject to restrictions, show more efficiency loss. The third and fourth simulation study will explore the efficiency loss further.

[Figure 2 about here.]

4.3 *Study 2*

A second simulation study has been used to study in detail the parameter recovery in a higher-dimensional situation. To this purpose, 11 outcomes were selected (all hearing tresholds taken at the right ear) from the example on the hearing thresholds. Again, two scenario's were considered. In scenario A, all fixed and random effects were considered to be outcome-specific, and unstructured 22×22 and 11×11 covariance matrices were used for the random effects and the error components respectively. In scenario B the model where the data are drawn from is simplified by considering the linear evolution, the interaction between the linear evolution and age at entry, and the random linear evolution to be common to respectively (1) outcomes 125, 250, 500 and 750 Hz, (2) outcomes 1000, 1500, 2000 and 3000 Hz and (3) outcomes 4000, 6000 and 8000 Hz. In both scenario's, 250 datasets were simulated from a full multivariate model (parameters were obtained by applying the pairwise approach on the real data), and the pairwise approach has been applied. Note that in this study no information is available on efficiency since the full multivariate models in scenario A and B, with respectively 22 and 14 random effects, can not be fitted directly.

To explore the parameter recovery, for each parameter the bias (the difference between the mean of the 250 samples and the truth) has been expressed as a proportion of the standard deviation of the 250 samples. In scenario A, this ratio varied from -0.147 to 0.138 (with mean 0.001) for all 66 fixed effects parameters, and from -0.157 to 0.117 (with mean 0.011) for all $253+66 = 319$ covariance parameters. In scenario B, this ratio varied from -0.125 to 0.108 (with mean 0.01) for all 50 fixed effects parameters, and from -0.141 to 0.128 (with mean -0.018) for all $105+66 = 171$ covariance parameters. A histogram of these ratio's can be found in figure 3. As expected, the pairwise approach clearly yields unbiased estimates in a high-dimensional setting, irrespective some parameters are shared or not shared by a set of outcomes.

[Figure 3 about here.]

4.4 *Study 3*

The first simulation study indicated that if some parameters are common to a set of outcomes, efficiency loss can be expected when using the pairwise approach. As a possible explanation, we referred to differences between the pairwise and the trivariate approach in the restrictions put on the parameter space. To verify this hypothesis in a very simple setting, the following simulation has been performed. This simulation is not based on the hearing data, but on an artificial dataset, which is a balanced one. This to ensure that the results are purely attributable to the presence of common parameters and not to the unbalanced structure of the data. 100 subjects are equally divided over a control and a treatment group. The following model has been used to sample from:

$$Y_{ijk} = \alpha_k + b_{ik} + \beta_k T_i + \varepsilon_{ijk} \tag{8}$$

where Y_{ijk} denotes the jth response of subject i on outcome k, with $i = 1, \ldots, 100$, $j = 1, \ldots, 5$ and $k = 1, \ldots, 3$. T is a binary indicator taking value 1 for the subjects in the treatment group. The following set of restrictions applies:

$$\begin{cases} \alpha_1 = \alpha_2 = \alpha_3 \\ \beta_1 = \beta_2 = \beta_3 \\ b_{i1} = b_{i2} = b_{i3} \end{cases}$$

The parameters common to the three outcomes will be denoted with α, β and b_i. The random intercepts b_i follow a zero-mean normal distribution with variance δ^2. The only outcome-specific parameters are the variances of the error components, which are uncorrelated between outcomes and denoted with respectively σ_1^2, σ_2^2 and σ_3^2. 500 datasets are sampled and for each dataset model (8) is fitted using the full trivariate approach as well as the pairwise approach. For all parameters the relative efficiency has been calculated as:

$$RE = \frac{var(\hat{\theta}_{triv})}{var(\hat{\theta}_{pair})}$$

where $var(\hat{\theta}_{triv})$ and $var(\hat{\theta}_{pair})$ denote the sampling variability of the 500 estimates observed under the full trivariate and the pairwise approach respectively.

Table 6 compares the sampling variabily for the estimates obtained under the pairwise and the trivariate approach. Also, the observed sampling variability in each pair is given. As can be seen, the pair-specific sampling variability of each parameter is slightly higher than the sampling variability under the trivariate approach. However, due to the lack of perfect correlation between the estimates obtained in different pairs, the sampling variability of the mean will be reduced compared to the sampling variability of the pair-specific estimates. For the fixed effects parameters and the variability of the random intercept, the sampling variability of the mean of the pair-specific estimates is comparable to the sampling variability under the trivariate approach. Hence, no efficiency loss is present for these parameters, although they are common to the three outcomes. Only for the outcome-specific variances of the error components, a small efficiency loss seems present. This study clearly suggest that the pairwise approach does not automatically suffers from efficiency loss for parameters which are shared by a set of outcomes.

[Table 6 about here.]

4.5 *Study 4*

A fourth simulation study has been set up to explore further situations where efficiency loss can occur. A similar model as in study 3 has been used. However, nine instead of three different outcomes are involved. The model used to sample from is:

$$Y_{ijk} = \alpha_k + b_{ik} + \beta_k T_i + \varepsilon_{ijk} \quad (9)$$

where Y_{ijk} denotes the jth response of subject i on outcome k, with $i = 1, \ldots, 100$, $j = 1, \ldots, 5$ and $k = 1, \ldots, 9$. A common fixed and random intercept are considered for successive trios of outcomes, such that $\alpha_1 = \alpha_2 = \alpha_3$, $\alpha_4 = \alpha_5 = \alpha_6$, $\alpha_7 = \alpha_8 = \alpha_9$, $b_{i1} = b_{i2} = b_{i3}$, $b_{i4} = b_{i5} = b_{i6}$ and $b_{i7} = b_{i8} = b_{i9}$. The outcome-common fixed intercepts will be denoted with α_{123}, α_{456} and α_{678}. The compound symmetric covariance matrix of the 3 random effects, having equal variance, is denoted by D. R denotes the covariance matrix of the error components ε_{ijk} which are uncorrelated, also having equal variance. T is a binary indicator taking value 1 for the 50 subjects in the treatment group. 500 datasets are sampled and for each dataset model (9) is fitted using the full multivariate approach as well as the pairwise approach. Seven versions of this model are fitted, differing in the restrictions put on the β_k in the model. In the seven scenario's, the β's are assumed to be common to respectively the first three outcomes, the first four outcomes, the first five outcomes, the first six outcomes, the first seven outcomes, the first eight outcomes, or to all nine outcomes:

$$\begin{cases} \text{Scenario 1} & : \quad \beta_1 = \beta_2 = \beta_3 \\ \text{Scenario 2} & : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 \\ \text{Scenario 3} & : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 \\ \text{Scenario 4} & : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 \\ \text{Scenario 5} & : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 \\ \text{Scenario 6} & : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 \\ \text{Scenario 7} & : \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 \end{cases}$$

[Figure 4 about here.]

Figure 4 gives the relative efficiencies for the fixed intercept shared by trios of outcomes and for the outcome-common treatment effect. A substantial increase in efficiency loss is observed for the outcome-common intercepts α_{123} and α_{456} under the second scenario. In this scenario, the restriction $\beta_1 = \beta_2 = \beta_3 = \beta_4$ applies, involving the trio of outcomes 1, 2, 3 as well as the trio of outcomes 4, 5, 6. The same phenomenon is present in scenario 5 where the restriction on the treatment effect involves the last trio of outcomes (7, 8, 9), yielding a similar increase in efficiency loss as has been observed for the two other trios in scenario 2. Also for the common treatment effect a reduction in efficiency is observed under scenario 2 and scenario 5, the situation where the restriction on the treatment effect involves a new trio of outcomes.

[Figure 5 about here.]

Figure 5 gives the relative efficiencies for the outcome-specific treatment effects. Irrespective the scenario, there is always a mild efficiency loss for the outcome-specific effect (relative efficiencies fluctuating around 0.9). However, additional decreases in efficiency for the pairwise approach are observed in scenarios 2, 3, 5 and 6.

To help understanding the pattern of efficiency losses for the fixed effects parameters, Table 7 distinguishes the scenarios by the answers on the following two questions: (1) Do all outcomes of a trio share the β-parameter? (2) Is the β-parameter specific for all outcomes of a trio?. Only in scenario's 1, 4 and 7 the answer is yes on one of both questions for all trios. These are the scenario's where the smallest amount of efficiency loss is observed.

[Table 7 about here.]

Figure 6 presents the results for all variance components. For none of the variance components, the scenario has an influence on the observed efficiency loss. A plausible explanation is that the scenarios only varied with respect to restrictions put on the fixed effects part of the model. Except for the covariances of the random effects, some efficiency loss is present for most variance components.

[Figure 6 about here.]

5. Discussion

This paper evaluates a joint-modelling approach designed to model high-dimensional multivariate longitudinal data. The approach is based on fitting bivariate mixed models for all pairs of outcomes. As long as each bivariate mixed model can be fitted, estimates can be obtained for the full multivariate mixed model. Obviously, this approach is advantageous whenever fitting the full multivariate mixed model is not possible or too time-consuming. Different simulation studies were used to assess the performance of the pairwise approach. The simulation studies indicated that the pairwise approach yields unbiased estimates with robust standard errors reflecting the true sampling variability. It was shown that also in a high-dimensional settings the pairwise approach recovered all parameters correctly. Efficiency loss can be present when some parameters are shared by a set of outcomes. However, the presence of shared parameters is a necessary, but not a sufficient condition to observe efficiency loss for all parameters. There is also no indication that the efficiency loss will increase as the number of outcomes sharing a parameter increases.

Acknowledgments

Both authors acknowledge support from the Interuniversity Attraction Poles Program P5/24 - Belgian State - Federal Office for Scientific, Technical and Cultural Affairs. They also wish to thank Larry Brant for providing the hearing data from the Baltimore Study of Aging.

References

Arnold, B.C. and Strauss, D. (1991), “Pseudolikelihood estimation: Some examples”, *Sankyhya, Series B*, **53**, 233–243.

Besag, J. (1975), “Statistical analysis of non-lattice data”, *The Statistician*, **24**(3), 179–195.

Brant, L.J. and Fozard, J.L. (1990), “Age changes in pure-tone hearing thresholds in a longitudinal study of normal human aging”, *Journal* *of the Acoustical Society of America*, **88**, 813–820.

Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H., and Renard, D. (2001), "Validation of surrogate endpoints in multiple randomized clinical trials with failure time end points", *Applied Statistics*, **50**, 405–422.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., and Geys, H. (2000), "The validation of surrogate endpoints in meta-analyses of randomized experiments", *Biostatistics*, **1**, 49–67.

Chakraborty, H., Helms, R.W., Sen, P.K., and Cohen, M.S. (2003), "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*, **22**, 1457–1464.

Diggle, P.J., Heagerty, P., Liang, K.Y., and Zeger, S.L. (2002), *Analysis of longitudinal data*, Clarendon Press, Oxford.

Fieuws, S. and Verbeke, G. (2004), "Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach", *Statistics in Medicine*, **23**, 3093–3104.

Geys, H., Molenberghs, G., and Ryan, L. (1997), "Pseudo-likelihood inference for clustered binary data", *Communications in Statistics: Theory and methods*, **26**(11), 2743–2767.

Gueorguieva, R. (2001), "A multivariate generalized linear mixed model for joint modelling of clustered outcomes in the exponential family", *Statistical Modelling*, **1**(3), 177–193.

Lange, N. and Laird, N.M. (1992), "The effect of covariance structure on variance estimation in balanced growth-curve models with random parameters", *Journal of the American Statistical Association*, **87**, 615–632.

Liang, K.Y. and Zeger, S.L. (1986), "Longitudinal data analysis using generalized linear models", *Biometrika*, **73**, 13–22.

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

Morrell, C.H. and Brant, L.J. (1991), "Modelling hearing thresholds in the elderly", *Statistics in Medicine*, **10**, 1453–1464.

Pearson, J.D., Morrell, C.H., Gordon-Salant, S., Brant, L.J., Metter, E.J., Klein, L.L., and Fozard, J.L. (1995), "Gender differences in a longitudinal study of age-associated hearing loss", *Journal of the Acoustical Society of America*, **97**, 1196–1205.

Reinsel, G. (1982), "Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure", *Journal of the American Statistical Association*, **77**, 190 – 195.

Renard, D., G.Molenberghs, and Geys, H. (2004), "A pairwise likelihood approach to estimation in multilevel probit models", *Computational statistics and data analysis*, **44**, 649–667.

Shah, A., Laird, N., and Schoenfeld, D. (1997), "A random-effects model for multiple characteristics with possibly missing data", *Journal of the American Statistical Association*, **92**, 775–779.

Shock, N.W., Greullich, R.C., Andres, R., Arenberg, D., Costa, P.T., Lakatta, E.G., and Tobin, J.D. (1984), "Normal human aging : The Baltimore Longitudinal Study of Aging", *National Institutes of Health publication 84-2450*.

Thiébaut, R., Jacqmin-Gadda, H., Chêne, G., Leport, C., and Commenges, D. (2002), "Bivariate linear mixed models using SAS PROC MIXED", *Computer Methods and Programs in Biomedicine*, **69**, 249–256.

Thum, Y.M. (1997), "Hierarchical linear models for multivariate outcomes", *Journal of Educational and Behavioral Statistics*, **22**(1), 77–108.

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

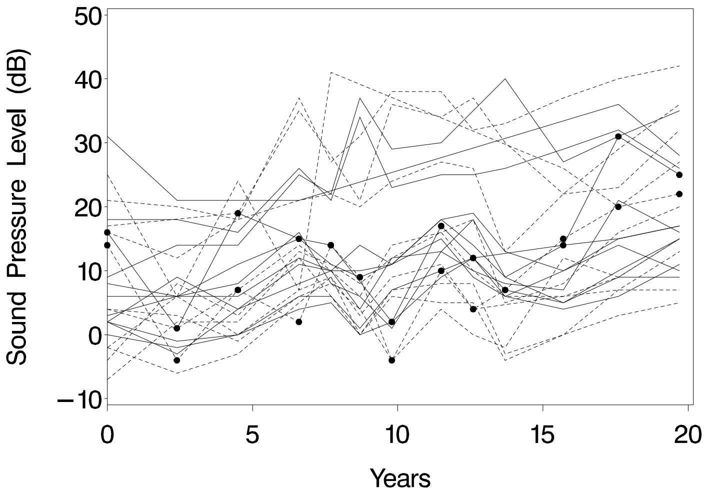

Figure 1. *Ear- and frequency specific profiles of hearing thresholds of 1 randomly selected subject. Dotted and solid lines represent the frequency specific profiles at the left ear and right ear respectively. Two profiles are marked with filled circles. These represent the hearing thresholds at both ears for frequency 250 Hz.*

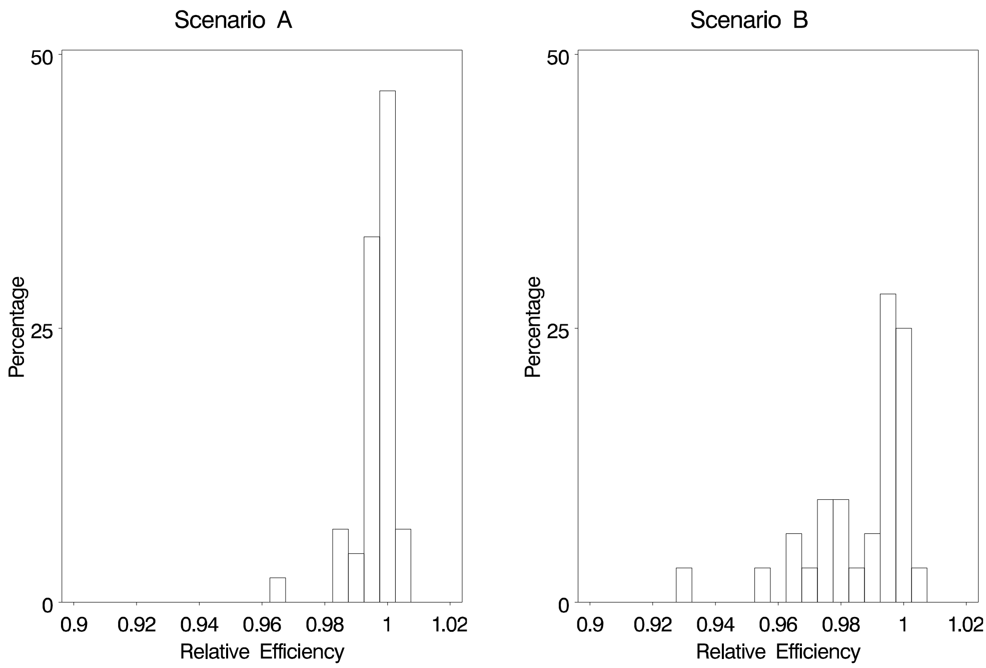

Figure 2. *Histogram of the relative efficiencies for the 45 parameters in Scenario A, and for the 32 parameters in scenario B.*

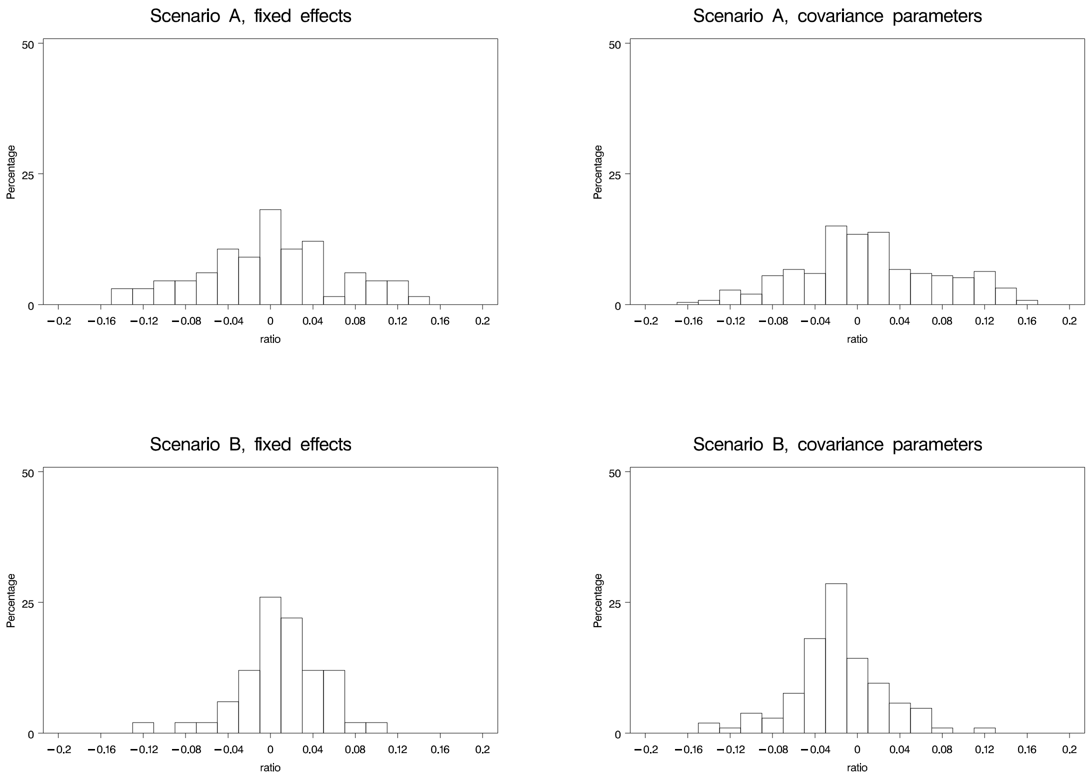

Figure 3. *Evaluation of parameter recovery in a high-dimensional setting: histogram of the bias for the 66 fixed effects under scenario A (upper left panel), the 50 fixed effects under scenario B (lower left panel), the 319 covariance parameters under scenario A (upper right panel) and the 171 covariance parameters under scenario B (lower right panel). The biases are expressed as a proportion of the standard deviation of the estimates.*

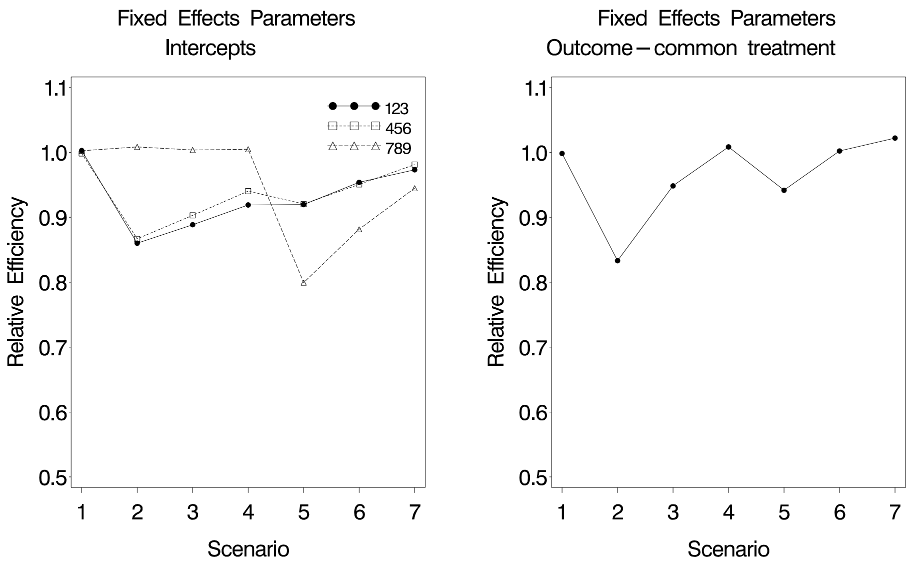

Figure 4. *Relative efficiencies of fixed effects parameters as a function of the seven different scenario's. The left panel shows the results for the outcome-common intercepts* α_{123}*,* α_{456} *and* α_{678} *and the figure at the right the results for the common treatment effect*

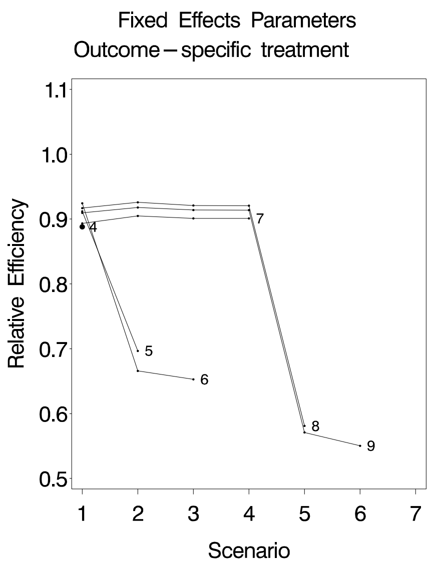

Figure 5. *Relative efficiencies of outcome-specific fixed effects parameters as a function of the seven different scenario's. The numbers in the figure denote the results for the outcome-specific treatment effect β_k, with $k = 4, \ldots, 7$.*

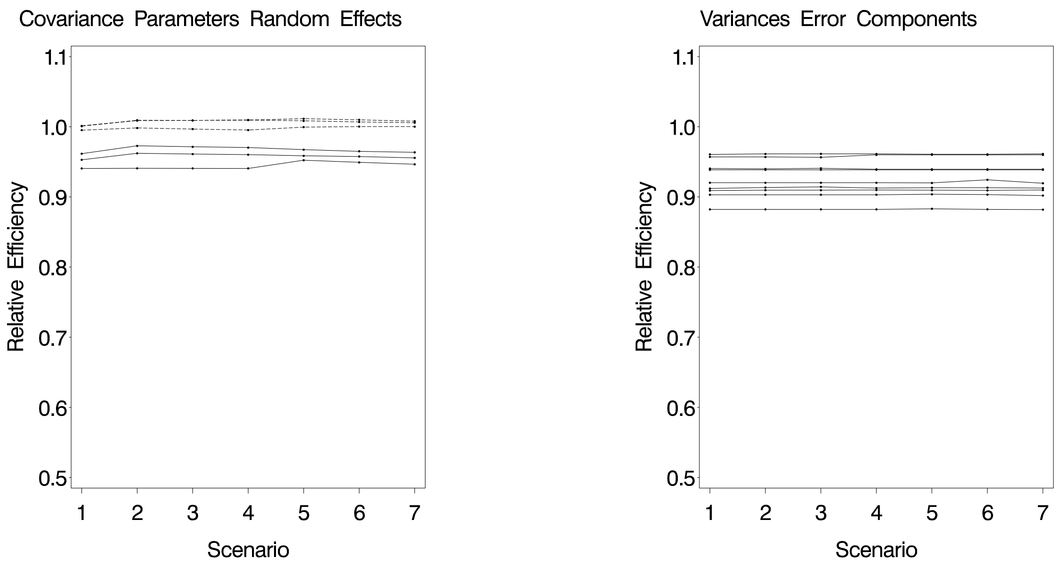

Figure 6. *Relative efficiencies of covariance parameters as a function of the seven different scenario's. The left panel shows the results for the elements in D, the covariance matrix of the random effects, with the solid lines and dotted lines representing the results for the variances and covariances of the random intercepts respectively. The right panel shows the results for the variances of the error components.*

Table 1

Scenario A. Results for fixed effects β_{jk}'s, with $j = 1, \ldots, 6$ (cfr. expression 7), $k = 1, \ldots, 3$, the latter referring to respectively 500 Hz, 1000 Hz and 2000 Hz. True values are given, the bias under the pairwise ($Bias_P$) and the trivariate approach ($Bias_T$), the standard deviation of the 1000 estimates ($\hat{\sigma}_\beta$) the mean of the standard errors obtained under the pairwise approach ($\overline{se}$), the root mean squared error under the pairwise ($RMSE_P$) and the trivariate approach ($RMSE_T$) and the relative efficiency (RE) of the pairwise approach.

$^{(*)}$ parameters multiplied with 10, $^{(+)}$ parameters multiplied with 1000

	truth	$Bias_P$	$Bias_T$	$\hat{\sigma}_\beta$	$\overline{se}$	$RMSE_P$	$RMSE_T$	*RE*
β_{11}	7.092	-0.051	-0.055	2.442	2.389	2.442	2.441	0.999
β_{12}	2.135	-0.032	-0.034	2.805	2.873	2.803	2.798	0.996
β_{13}	3.772	0.027	0.025	3.712	3.679	3.710	3.704	0.996
$\beta_{21}{}^{(*)}$	-2.827	0.022	0.024	1.040	1.007	1.039	1.039	1.000
$\beta_{22}{}^{(*)}$	-1.296	0.036	0.036	1.187	1.215	1.187	1.185	0.996
$\beta_{23}{}^{(*)}$	-2.203	0.019	0.020	1.557	1.556	1.556	1.554	0.997
$\beta_{33}{}^{(+)}$	4.570	-0.023	-0.025	0.998	0.963	1.000	1.000	1.000
$\beta_{32}{}^{(+)}$	3.696	-0.047	-0.047	1.139	1.162	1.140	1.140	1.000
$\beta_{33}{}^{(+)}$	6.029	-0.032	-0.033	1.488	1.489	1.483	1.483	1.000
$\beta_{41}{}^{(*)}$	-2.948	0.022	0.024	0.801	0.785	0.801	0.795	0.986
$\beta_{42}{}^{(*)}$	-5.907	-0.020	-0.020	0.848	0.842	0.847	0.846	0.998
$\beta_{43}{}^{(*)}$	-9.965	-0.066	-0.065	1.081	1.099	1.082	1.083	1.002
$\beta_{51}{}^{(+)}$	13.750	-0.007	-0.012	1.634	1.598	1.633	1.623	0.987
$\beta_{52}{}^{(+)}$	19.370	0.060	0.061	1.687	1.707	1.673	1.673	1.000
$\beta_{53}{}^{(+)}$	29.050	0.126	0.124	2.167	2.210	2.168	2.168	1.000
β_{61}	1.387	0.010	0.010	0.306	0.306	0.306	0.306	1.002
β_{62}	-0.735	0.001	0.001	0.291	0.297	0.291	0.290	0.997

Table 2

Scenario A. Results for covariance parameters with $d_{11}, d_{21}, \ldots, d_{66}$ the elements of the covariance matrix of the random effects D and $r_{11}, r_{21}, \ldots, r_{33}$ the elements of the covariance matrix of the error components R. True values are given, the bias under the pairwise ($Bias_P$) and the trivariate approach ($Bias_T$), the root mean squared error under the pairwise ($RMSE_P$) and the trivariate approach ($RMSE_T$) and the relative efficiency (RE) of the pairwise approach.

	truth	$Bias_P$	$Bias_T$	$RMSE_P$	$RMSE_T$	RE
d11	42.166	-0.051	-0.047	3.137	3.132	0.997
d21	0.401	0.013	0.012	0.175	0.175	0.997
d22	0.063	-0.002	-0.002	0.015	0.015	0.996
d31	42.627	-0.061	-0.060	3.328	3.327	0.999
d32	0.683	0.011	0.011	0.211	0.211	1.001
d33	68.629	-0.021	-0.024	4.683	4.677	0.998
d41	0.341	0.008	0.008	0.199	0.198	0.996
d42	0.051	-0.001	-0.001	0.013	0.013	0.986
d43	0.447	0.013	0.013	0.245	0.245	0.994
d44	0.108	-0.001	-0.001	0.019	0.019	0.997
d51	39.073	-0.101	-0.104	3.888	3.871	0.991
d52	0.537	0.006	0.002	0.279	0.274	0.967
d53	63.407	-0.035	-0.037	5.066	5.074	1.003
d54	0.837	0.000	0.001	0.302	0.302	0.998
d55	114.820	-0.124	-0.132	7.480	7.472	0.998
d61	0.234	0.004	0.003	0.260	0.260	1.003
d62	0.035	-0.001	-0.001	0.017	0.016	0.991
d63	0.248	0.010	0.010	0.309	0.309	1.001
d64	0.082	-0.001	-0.001	0.021	0.021	0.999
d65	0.511	0.001	0.002	0.392	0.392	1.003
d66	0.225	-0.001	-0.001	0.033	0.033	1.002
r11	26.586	0.033	0.031	0.820	0.819	0.998
r21	11.672	0.026	0.025	0.614	0.612	0.994
r22	24.187	0.012	0.012	0.746	0.746	1.000
r31	10.745	0.012	0.014	0.717	0.714	0.992
r32	17.710	0.007	0.008	0.736	0.735	0.997
r33	34.316	-0.034	-0.032	1.057	1.054	0.993

Table 3

Scenario B. Results for fixed effects β_{jk}'s, with $j = 1, \ldots, 6$ (cfr. expression 7), $k = 1, \ldots, 3$, the latter referring to respectively 500 Hz, 1000 Hz and 2000 Hz. Note that β_5 has no additional subscript since this parameter is common to the three frequencies. True values are given, the bias under the pairwise ($Bias_P$) and the trivariate approach ($Bias_T$), the standard deviation of the 1000 estimates ($\hat{\sigma}_\beta$) the mean of the standard errors obtained under the pairwise approach ($\overline{se}$), the root mean squared error under the pairwise ($RMSE_P$) and the trivariate approach ($RMSE_T$) and the relative efficiency (RE) of the pairwise approach.
$^{(*)}$ parameters multiplied with 10, $^{(+)}$ parameters multiplied with 1000

	truth	$Bias_P$	$Bias_T$	$\hat{\sigma}_\beta$	$\overline{se}$	$RMSE_P$	$RMSE_T$	*RE*
β_{11}	8.437	-0.068	-0.072	2.370	2.416	2.370	2.367	0.998
β_{12}	1.851	-0.062	-0.050	2.816	2.872	2.815	2.807	0.994
β_{13}	0.726	-0.048	-0.038	3.661	3.708	3.659	3.659	1.000
$\beta_{22}{}^{(*)}$	-3.258	0.023	0.025	1.008	1.022	1.008	1.007	0.998
$\beta_{22}{}^{(*)}$	-1.194	0.018	0.014	1.197	1.216	1.197	1.194	0.995
$\beta_{23}{}^{(*)}$	-1.265	0.011	0.006	1.545	1.571	1.544	1.544	1.000
$\beta_{33}{}^{(+)}$	4.830	-0.021	-0.022	0.971	0.979	0.971	0.970	0.998
$\beta_{32}{}^{(+)}$	3.630	-0.015	-0.011	1.152	1.164	1.151	1.149	0.996
$\beta_{33}{}^{(+)}$	5.480	-0.003	0.001	1.477	1.505	1.476	1.476	0.999
$\beta_{41}{}^{(*)}$	-4.950	0.016	0.020	0.680	0.729	0.679	0.681	1.003
$\beta_{42}{}^{(*)}$	-5.361	0.033	0.029	0.715	0.727	0.715	0.690	0.931
$\beta_{43}{}^{(*)}$	-4.873	0.030	0.029	0.726	0.744	0.726	0.717	0.976
$\beta_5{}^{(+)}$	18.12	-0.029	-0.029	1.383	1.438	1.383	1.375	0.989
β_{61}	1.483	0.004	0.004	0.314	0.309	0.314	0.313	0.996
β_{62}	-0.768	0.008	0.007	0.312	0.298	0.312	0.311	0.995
β_{63}	-1.936	0.015	0.015	0.377	0.372	0.377	0.378	1.001

Table 4

Scenario B. Results for covariance parameters with $d_{11}, d_{21}, \ldots, d_{44}$ the elements of the covariance matrix of the random effects D and $r_{11}, r_{21}, \ldots, r_{33}$ the elements of the covariance matrix of the error components R. True values are given, the bias under the pairwise ($Bias_P$) and the trivariate approach ($Bias_T$), the root mean squared error under the pairwise ($RMSE_P$) and the trivariate approach ($RMSE_T$) and the relative efficiency (RE) of the pairwise approach.

	truth	$Bias_P$	$Bias_T$	$RMSE_P$	$RMSE_T$	*RE*
d11	44.333	-0.051	-0.050	3.159	3.143	0.990
d21	43.074	-0.115	-0.110	3.316	3.306	0.994
d22	67.865	-0.159	-0.148	4.495	4.465	0.987
d31	36.543	-0.094	-0.113	3.771	3.730	0.978
d32	62.046	-0.228	-0.192	4.841	4.769	0.971
d33	115.93	-0.358	-0.345	7.201	7.177	0.993
d41	0.250	0.009	0.009	0.167	0.165	0.977
d42	0.591	0.010	0.008	0.198	0.194	0.956
d43	0.737	0.009	0.008	0.256	0.252	0.966
d44	0.065	0.000	0.000	0.012	0.012	0.978
r11	26.792	-0.002	-0.001	0.803	0.801	0.994
r21	11.235	0.025	0.024	0.612	0.606	0.982
r22	25.041	0.017	0.015	0.784	0.782	0.995
r31	9.638	0.008	0.010	0.734	0.721	0.964
r32	17.856	0.004	0.000	0.780	0.770	0.975
r33	38.846	0.044	0.041	1.126	1.126	1.000

Table 5

Robustness of Standard Errors (Scenario A, Incorrect Model). Results for fixed effects β_{jk}'s, with $j = 1, \ldots, 6$ (cfr. expression 7), $k = 1, \ldots, 3$, the latter referring to respectively 500 Hz, 1000 Hz and 2000 Hz.
Given are the standard deviation of the 1000 estimates of the pairwise ($\hat{\sigma}_\beta(P)$) and the trivariate ($\hat{\sigma}_\beta(T)$) approach, the mean of the standard errors obtained under the pairwise approach ($\overline{se}(P)$), mean of the model-based standard errors (($\overline{se}(T)$) and mean of the empirically corrected standard errors ($\overline{se}(T_e)$)of the trivariate approach, and the respective ratio's of the mean and the standard error of the 1000 estimates ($ratio_P$, $ratio_T$, $ratio_{T_e}$).
$^{(*)}$ *parameters multiplied with 10,* $^{(+)}$ *parameters multiplied with 1000*

	$\hat{\sigma}_\beta(P)$	$\hat{\sigma}_\beta(T)$	$\overline{se}(P)$	$\overline{se}(T)$	$\overline{se}(T_e)$	$ratio_P$	$ratio_T$	$ratio_{T_e}$
β_{11}	2.459	2.459	2.421	2.493	2.421	0.984	1.014	0.985
β_{12}	2.874	2.880	2.920	3.010	2.923	1.016	1.045	1.015
β_{13}	3.746	3.746	3.746	3.865	3.748	1.000	1.032	1.001
$\beta_{22}{}^{(*)}$	1.046	1.046	1.022	1.051	1.021	0.977	1.005	0.977
$\beta_{22}{}^{(*)}$	1.220	1.222	1.236	1.272	1.237	1.013	1.040	1.012
$\beta_{23}{}^{(*)}$	1.578	1.578	1.585	1.633	1.586	1.004	1.035	1.005
$\beta_{33}{}^{(+)}$	1.004	1.003	0.977	1.003	0.976	0.973	1.000	0.973
$\beta_{32}{}^{(+)}$	1.172	1.174	1.183	1.215	1.184	1.009	1.035	1.009
$\beta_{33}{}^{(+)}$	1.512	1.512	1.518	1.561	1.519	1.004	1.033	1.005
$\beta_{41}{}^{(*)}$	0.842	0.836	0.817	0.644	0.813	0.970	0.770	0.972
$\beta_{42}{}^{(*)}$	0.910	0.912	0.895	0.628	0.898	0.984	0.689	0.985
$\beta_{43}{}^{(*)}$	1.188	1.192	1.198	0.791	1.202	1.009	0.663	1.009
$\beta_{51}{}^{(+)}$	1.719	1.709	1.665	1.318	1.657	0.969	0.771	0.970
$\beta_{52}{}^{(+)}$	1.823	1.829	1.824	1.290	1.830	1.001	0.705	1.001
$\beta_{53}{}^{(+)}$	2.440	2.449	2.429	1.616	2.437	0.995	0.660	0.995
β_{61}	0.312	0.312	0.310	0.308	0.310	0.992	0.987	0.993
β_{62}	0.303	0.305	0.304	0.302	0.306	1.004	0.989	1.003
β_{63}	0.370	0.371	0.373	0.376	0.374	1.007	1.013	1.008

Table 6

Study 3: Sampling variability of parameter estimates obtained under the trivariate approach and the pairwise approach. For the pairwise approach, the sampling variability of pair-specific estimates is given (pair 1, pair 2, pair 3), as well as the sampling variability of the mean over the pairs. Relative efficiency of the pairwise approach is denoted with RE

	Pairwise Approach				Trivariate	RE
	Pair 1	Pair 2	Pair 3	Mean		
α	0.346	0.340	0.333	0.334	0.334	1.000
β	0.678	0.669	0.660	0.658	0.658	1.000
δ^2	5.185	5.306	5.301	5.063	5.069	1.001
σ_1^2	0.406	0.390	-	0.389	0.387	0.995
σ_2^2	0.375	-	0.375	0.368	0.358	0.973
σ_3^2	-	0.390	0.387	0.380	0.376	0.990

Table 7

Study 4: Characterisation of the seven scenarios. 'C' indicates that the treatment effect is common to all outcomes of a trio and 'S' indicates that the treatment effect is specific for all outcomes of a trio. Empty cells refer to situations where none of both possibilities apply.

Scenario	Trio 1	Trio 2	Trio 3
1	C	S	S
2	C		S
3	C		S
4	C	C	S
5	C	C	
6	C	C	
7	C	C	C