REVSTAT – Statistical Journal
Volume 16, Number 4, October 2018, 487–514

ON FITTING GENERALIZED LINEAR MIXED EFFECTS MODELS FOR LONGITUDINAL BINARY DATA USING DIFFERENT CORRELATION STRUCTURES

Authors: M. Salomé Cabral
– CEAUL and Departamento de Estatística e Investigação Operacional, Faculdade de Ciências da Universidade de Lisboa, Portugal
salome@fc.ul.pt

M. Helena Gonçalves
– CEAUL and Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade do Algarve, Portugal
mhgoncal@ualg.pt

Received: January 2016 Revised: October 2016 Accepted: October 2016

Abstract:

- The generalized linear mixed effects model (GLMM) approach is widely used to analyze longitudinal binary data when the goal of the study is a subject-specific interpretation because it allows missing values on the response, provided they are missing at random (MAR), and accounts the correlation among the repeated observations of the same subject by the inclusion of random effects in the linear predictor. However, in GLMM it is assumed that the observations of the same subject are independent conditional to the random effects and covariates which may be not true. To overcome this problem [9] extended this model using binary Markov chains as the basic stochastic mechanism. The aim of this paper is to give a statistical assessment of both approaches in terms of properties such as efficiency and coverage probability, as well as, to give some guidelines for the choice of the statistical approach to an applied researcher. Both procedures are described and a simulation study is carried out to compare their performance. An analysis of a longitudinal binary data set illustrates the performance of both procedures in a practical example. The `R` packages `lme4` and `bild` are used.

Key-Words:

- *binary longitudinal data; exact likelihood; random effects; Markov chain; missing data.*

AMS Subject Classification:

- 60J10, 62J99.

1. INTRODUCTION

Longitudinal binary data studies are a powerful design and they have become increasingly popular in a wide range of applications across all disciplines. In these studies repeated observations of a response variable are taken over time on each subject in one or more treatment groups. In such cases the repeated measures of each vector of responses are likely to be correlated and the autocorrelation structure for the repeated data plays a significant role in the estimation of regression parameters. Although longitudinal studies are design to collect data on every subject in the sample at each time of follow-up, many studies have missing data since it is difficult to have complete records of all subjects for a variety of reasons. When longitudinal binary data are incomplete there are important implications for their analysis and one of the main concerns is to distinguish different reasons of missingness. The nature of missing data mechanism has been classified by [16] and [13] as: missing completely at random (MCAR), missing at random (MAR) and non missing at random (NMAR). Another important distinction is whether missing values occur intermittently or as dropouts. When missing values occur as dropouts, an individual is observed only at a certain time and misses all the subsequent observations. When missing values occur as intermittently, an individual may miss some measurement times among a common set of predefined measurement times. To all these situations several methods have been proposed ([4], [5], [1], [6], [14]). A review of this topic is given in [12].

In [12] is argued that methods based on likelihood, such generalized linear mixed effects model [3], usually denoted by GLMM, are recommended when the goal of the study is a subject-specific interpretation and missing values are allowed on the response, provide they are MAR in the standard terminology of [16]. In the GLMM the correlation among the repeated observations of the same subject is account by the inclusion of random effects in the linear predictor and it is assumed that observations to the same subject are independent conditional to the random effects and covariates. Although in GLMM this independence is assumed they may still be correlated. To overcome this problem [9] used a binary Markov chain model to accommodate serial dependence and odds-ratio to measure dependence between successive observations. This methodology is a development of the alternative likelihood-based formulation for a logistic regression presented by [1] which allows: (i) a first order and a second order Markov dependence; (ii) a random intercept term in the linear predictor; (iii) missing values on the response, provided they are MAR. Both approaches, GLMM and generalized linear mixed effects model with binary Markov chain (GLM3C) as the basic stochastic mechanism, are implemented in `R` [18] packages. The goal of this paper is to give information to the practitioners about which of the two procedures, GLMM or GLM3C, is more appropriate to use for their data at hand.

To achieved that goal a simulation study was carried out to compare the two aforementioned approaches in terms of properties such as efficiency and coverage probability. For GLM3C approach the estimates were obtained through the `bild` function of the `R` package `bild`. When the GLMM approach was used the estimates were achieved through the `glmer` function of the `R` package `lme4` [2] as well as the `bild` function of `R` package `bild` [11] with the independence structure selected.

The paper is organized as follows: Section 2 gives a summary of the models used. Section 3 reports a simulation study to assess the performance of the procedures. In Section 4 a real data is used to illustrate the two procedures as well as the key results of the simulation study. Section 5 concludes the paper.

2. MODEL FOR BINARY DATA

Suppose that n independent individuals are observed at times $t = 1, ..., T_i$, which need not be the same for all n individual and, to establish notation, denote by $y_{it} \in \{0, 1\}$ the binary response value at time t from individual i $(i = 1, ..., n)$, and by Y_{it} its generating random variable whose mean value is $\Pr(Y_{it} = 1) = \theta_{it}$. The sequence $(y_{i1}, ..., y_{iT_i})$ will be collectively referred as the i-th individual profile and associated with each observation time and each subject, a set of p covariates is available, denoted by x_{it}.

The logistic regression model which links the covariates and the marginal mean of Y_{it} assumes the form

$$\operatorname{logit} \theta_{it} = x_{it}^{\top} \beta, \tag{2.1}$$

where β is the p-dimensional parameter of interest and $\operatorname{logit} \theta = \log\{\theta/(1-\theta)\}$.

In longitudinal studies the repeated measures of each vector of responses are likely to be correlated. To account for the within-subject association the GLMM uses random effects, b_i, in the linear predictor. The correlation among observations from one subject can be thought of as arising from sharing a set of underlying random effects.

In what follows only the random intercept model is considered.

2.1. Generalized linear mixed effects model

The introduction of random effects can be formulated by adding a $q \times 1$ vector b_i of random effects in (2.1) associated to a $q \times 1$ vector of covariates, z_{it}, (in general a subset of $\mathbf{x}_{it}$). In the random intercept model the vector b_i is reduced to a single ($q = 1$) random effect $b_i \sim N(0, \sigma^2)$ and $z_{it} = 1$ for all $i = 1, ..., n$ and $t = 1, ..., T_i$ leading to

$$\text{logit}\,\Pr(Y_{it} = 1|b_i) = x_{it}^\top \beta + b_i, \qquad (i = 1, ..., n) \tag{2.2}$$

where the b_i's are assumed to be sampled independently from each other and that conditioning on x_{it} and b_i, the Y_{it}'s are independent.

The likelihood inference is based on a sample of n individual profiles that are assumed to be independent from each other. The contribution of the i-th subject to the likelihood of the random intercept model is

$$L_i^R(\beta, \omega) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\mathbb{R}} L_i^F(\beta^{b_i}|b_i) \exp\left(-\frac{b_i^2}{2\,\sigma^2}\right) \mathrm{d}b_i \tag{2.3}$$

where β^{b_i} is a p-vector of parameters like β, but where the first component is now $\beta_0 + b_i$, instead of β_0 and $\omega = \log \sigma^2$. In expression (2.3) the term $L_i^F(\beta^{b_i}|b_i) = \exp\{\ell_i^F(\beta^{b_i}|b_i)\}$ where

$$\ell_i^F(\beta^{b_i}|b_i) = \sum_{t=1}^{T_i} \big[y_{it}\,\text{logit}(\theta_{it}) + \log(1 - \theta_{it})\big].$$

The log-likelihood for the whole sample is given by

$$\ell^R(\beta, \omega) = \sum_{i=1}^{n} \log L_i^R(\beta, \omega). \tag{2.4}$$

The integrals in (2.3) have no analytical solution and appropriate numerical integration methods must be used.

This methodology is implemented in the `R` package `bild` [11] and the integrals in (2.3) are computed using adaptive Gaussian quadrature. Other `R` packages have this procedure implemented and one of the most popular is the `lme4` [2] package that also uses adaptive Gaussian quadrature to compute the integrals in (2.3).

2.2. Generalized linear mixed effects model with Markov chain correlation

Although in GLMM it is assumed that conditioning on x_{it} and b_i, the Y_{it}'s are independent they may still be correlated. To overcome this problem [9] proposed the use of binary by Markov chains to model the serial dependence between successive observations of the same subject. As they note Markov chains provide the simplest stochastic mechanism to introduce serial dependence for discrete random variables. In their approach the serial dependence between successive observation can be regulated (i) by one dependence parameter (first order dependence structure) or (ii) by two dependence parameters the (second order dependence structure). In both cases the odds-ratio is the quantity used to measure dependence between variables. One advantage of odds-ratios as measures of association is that, unlike marginal correlations, they are not constrained by marginal probabilities ([1], [8]). Their approach can be summarized as follows. To simplify notation the subscript i is dropped temporarily.

For the first order dependence structure (MC1), the serial dependence is modeled using $\psi_1 = OR(Y_t, Y_{t-1})$ where

$$OR(Y_t, Y_{t-1}) = \frac{\Pr(Y_{t-1} = Y_t = 1)\ \Pr(Y_{t-1} = Y_t = 0)}{\Pr(Y_{t-1} = 0,\ Y_t = 1)\ \Pr(Y_{t-1} = 1,\ Y_t = 0)} = \frac{p_1/(1-p_1)}{p_0/(1-p_0)}$$

where p_j are the transition probabilities given by

$$p_j = \Pr(Y_t = 1 | Y_{t-1} = j), \quad j = 0, 1; t = 2, ..., T. \tag{2.5}$$

For the second order dependence structure (MC2) is considered the joint distribution of three components of the process at time, (Y_{t-2}, Y_{t-1}, Y_t) and impose the constraints

$$\begin{aligned} OR(Y_{t-1}, Y_{t-2}) &= \psi_1 = OR(Y_t, Y_{t-1}) \\ OR(Y_t, Y_{t-2} | Y_{t-1} = 0) &= \psi_2 = OR(Y_t, Y_{t-2} | Y_{t-1} = 1) \end{aligned}$$

ψ_1 and ψ_2 denote two positive parameters. The transition probabilities are given by

$$p_{hj} = \Pr(Y_t = 1 | Y_{t-2} = h, Y_{t-1} = j), \quad h, j = 0, 1; t = 3, ..., T, \tag{2.6}$$

see [8] for a full account.

The serial dependence for MC2 models is regulated by $\lambda = (\lambda_1, \lambda_2) = (\log \psi_1, \log \psi_2)$, which are assume to be constant across time and subjects. When $\lambda_2 = 0$, the Markov chain reduces to MC1 models and the serial dependence is regulated by λ_1.

The likelihood inference is based on a sample of n individual profiles that are assumed to be independent from each other. The contribution of the i-th subject to the likelihood of the random intercept model is

$$L_i^R(\beta, \lambda, \omega) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\mathbb{R}} L_i^F(\beta^{b_i}, \lambda | b_i) \exp\left(-\frac{b_i^2}{2\,\sigma^2}\right) \mathrm{d}b_i \tag{2.7}$$

where β^{b_i} and ω are defined as in Section 2.1. In expression (2.7) the term $L_i^F(\beta^{b_i}, \lambda | b_i) = \exp\{\ell_i^F(\beta^{b_i}, \lambda | b_i)\}$ is computed, under a serial dependence MC1, from

$$\ell_i^F(\beta, \lambda) = y_1 \operatorname{logit}(\theta_1) + \log(1 - \theta_1) + \sum_{t=2}^{T_i} [y_t \operatorname{logit}(p_j) + \log(1 - p_j)] \tag{2.8}$$

and under a serial dependence MC2 from

$$\begin{aligned} \ell_i^F(\beta, \lambda) \;=\; & \big[y_1 \operatorname{logit}(\theta_1) + \log(1 - \theta_1)\big] + \big[y_2 \operatorname{logit}(p_j) + \log(1 - p_j)\big] \\ & + \sum_{t=3}^{T_i} \Big[y_t \operatorname{logit}(p_{hj}) + \log(1 - p_{hj})\Big] \end{aligned} \tag{2.9}$$

where the three blocks on the right-hand side represent the contribution to the log-likelihood from y_1, y_2, and $(y_3, ..., y_T)$, respectively, where p_{hj} is given by (2.6) and p_j by (2.5). The log-likelihood for the whole sample is given by (2.4). For a full account see [8].

In this approach missing values are allowed on the response, provided they are MAR. If missing data occur at the beginning or at the end of an individual profile, this poses no problems, since this case is equivalent to a designed unbalance in the length profile T_i for that individual. Some restrictions exist for the presence of missing data when they occur in the middle of the profile due to the imposed correlation structure. If MC1 model is considered and if there is a missing value at time point $t - 1$, it is required that there are observations at time points $t - 2$ and t. If MC2 model is considered and if there is a missing value at time point $t - 2$, it is required that there are observations at time points $t - 4, t - 3, t - 1$ and t, except for the two end portions of the observation period, where no restriction is made.

This approach is implemented in the `R` package `bild` ([11]) and, as in the previous approach, the integrals in (2.7) are computed using adaptive Gaussian quadrature.

3. A SIMULATION STUDY

A simulation study was carried out to compare both approaches when simulated data has a serial dependence MC1 or MC2. The model considered in the simulation included a dichotomous treatment, a linear effect time and an interaction between time and treatment and is given by

$$\Pr(Y_{it}=1|t)=\frac{\exp(\beta_0+b_i+\beta_1 t+\beta_2 x_i+\beta_3(x_i\times t))}{1+\exp(\beta_0+b_i+\beta_1 t+\beta_2 x_i+\beta_3(x_i\times t))} \tag{3.1}$$

where $x_i=0$ for half the population and 1 for the remainder. The fixed effect coefficients were set at $\beta_0=-1$, $\beta_1=0.5$, $\beta_2=1$, $\beta_3=1$ and the random effect distribution was simulated with $b_i\sim N(0,\sigma^2)$. In both serial dependence structures several designs were considered to reflect the range of experimental data encountered in practice. The number of subjects was set to either small ($n=20$) or large ($n=50$). The length of profile on each subject was short ($T=7$) or long ($T=13$) and the time points were set for $T=7$ at $t=-1.5,-1,-0.5,0,0.5,1,1.5$ and for $T=13$ at $t=-1.5,-1.25,-1,-0.75,-0.5,-0.25,0,0.25,0.5,0.75,1,1.25,1.5$.

In what concerns the variance of the random effect, σ^2, three values were considered $\sigma^2=0.5,1$ and 2.

1. Under MC1 models on each run were generated T binary correlated data under the i-th subject following a first order serial dependence regulated by λ_1. The values considered for λ_1 were 0.05, 0.25, 0.5, 0.75 and 1.
2. Under MC2 models on each run were generated T binary correlated data under the i-th subject following a second order serial dependence regulated by $\lambda=(\lambda_1,\lambda_2)$. For the pair (λ_1,λ_2) the combinations $(0.05,0.05)$, $(0.25,0.25)$, $(0.5,0.5)$, $(0.75,0.75)$ and $(1,1)$ were considered.

In both cases the whole estimation procedure was repeated for 1000 runs and the sample mean of estimate parameter (Mean), the sample mean of bias (Bias) and the sample mean square error (MSE) were computed, as well as, the coverage probabilities of nominal 95% confidence intervals (CI). For each data set the relative efficiency (RE) of the estimators was computed, as usual, by the ratio of the respective MSE. RE>1 means GLM3C estimator is preferred. The coverage probabilities of nominal 95% confidence intervals were computed as the proportion of simulated intervals that cover the true parameter used to generate the simulated data.

For the GLM3C approach the estimates of the parameters were obtained through the function `bild` in the `R` package `bild` and the dependence structure

was chosen trough the argument `dependence` in the function `bild`, `MC1R` (MC1 with random intercept) `MC2R` (MC2 with random intercept), for details see [10]. When GLMM approach (which ignores the conditional dependence between repeated measures in terms of numerical analysis) was considered the estimates were obtained through the `bild` function with the dependence argument set at `indR` (independence structure with random intercept) as well as through the `glmer` function in the `R` package `lme4`, the results obtained were exactly the same.

The results of simulation are displayed from Figures 1–6 and Tables 1–8 for the time effect (β_1) and group-time interaction effect (β_3) the effects usually of most interest in a longitudinal study. Each table lists the following: Mean, Bias, MSE and coverage probability of nominal 95% confidence intervals for β_1 and β_3 over the 1000 simulations to both approaches (GLMM and GLMC3). The GLM3C approach is denoted by GLM3C-MC1 or GLM3C-MC2 if a serial dependence MC1 or MC2, respectively, is considered. The Figures display the results concerned to $\sigma^2 = 0.5$ and $\sigma^2 = 2$, the two extreme values considered to the variance of the random effect b_i.

Taking into account that the goal of the simulation study is to give a statistical assessment of both approaches the main conclusions to serial dependence MC1 and MC2 are given, respectively, in Sections 3.1 and 3.2.

3.1. Serial dependence MC1

Under a serial dependence MC1 and to β_1 and β_3 parameters the main conclusions, based on the approaches GLM3C-MC1 and GLMM, are:

(**i**) The coverage probabilities of both approaches are similar when the dependence structure established by λ_1 is low (0.05-0.25). When λ_1 increases the GLM3C approach gives coverage probabilities closer to nominal than the GLMM approach (Figures 1–2 and Tables 1–4).

(**ii**) When λ_1 is low (0.05) the efficiency of both approaches is similar with the GLMM estimators more efficient in some configurations. As λ_1 increases the GLM3C estimators becomes more efficient than the GLMM for all the design configurations (Figure 3 and Tables 1–4).

(**iii**) In terms of bias the behavior of both approaches is very similar with a slight decrease of the estimated bias associated with the GLM3C approach when λ_1 increases. The exception is for $\hat{\beta}_1$ when $T = 7$, $n = 50$ and for all σ^2 considered (Tables 1–4).

Table 1: Results of the simulation study under a serial dependence MC1 for $n = 20$, $T = 7$.

		GLM3C-MC1					**GLMM**				
		λ_1					λ_1				
		.05	.25	.50	.75	1	.05	.25	.50	.75	1
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.521	0.520	0.513	0.514	0.513	0.519	0.523	0.525	0.532	0.537
	$\hat{\beta}_3$	1.110	1.108	1.111	1.109	1.082	1.111	1.120	1.137	1.152	1.132
Bias	$\hat{\beta}_1$	0.021	0.020	0.013	0.014	0.013	0.019	0.023	0.025	0.032	0.037
	$\hat{\beta}_3$	0.110	0.108	0.111	0.109	0.082	0.111	0.120	0.137	0.152	0.132
MSE	$\hat{\beta}_1$	0.090	0.094	0.098	0.106	0.112	0.087	0.094	0.100	0.112	0.125
	$\hat{\beta}_3$	0.279	0.289	0.311	0.322	0.325	0.279	0.297	0.330	0.356	0.361
CI	β_1	0.947	0.951	0.947	0.951	0.951	0.950	0.945	0.927	0.928	0.927
	β_3	0.953	0.959	0.952	0.952	0.953	0.957	0.950	0.932	0.934	0.929
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.511	0.522	0.514	0.512	0.509	0.508	0.524	0.527	0.536	0.536
	$\hat{\beta}_3$	1.135	1.074	1.109	1.106	1.048	1.135	1.084	1.141	1.158	1.108
Bias	$\hat{\beta}_1$	0.011	0.022	0.014	0.012	0.009	0.008	0.024	0.027	0.036	0.036
	$\hat{\beta}_3$	0.135	0.074	0.109	0.106	0.048	0.135	0.084	0.141	0.158	0.108
MSE	$\hat{\beta}_1$	0.105	0.100	0.106	0.105	0.131	0.104	0.100	0.111	0.116	0.147
	$\hat{\beta}_3$	0.315	0.281	0.329	0.324	0.327	0.313	0.284	0.354	0.369	0.363
CI	β_1	0.947	0.948	0.950	0.949	0.955	0.945	0.942	0.927	0.928	0.916
	β_3	0.947	0.951	0.951	0.957	0.950	0.941	0.940	0.930	0.938	0.922
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.529	0.527	0.528	0.497	0.500	0.527	0.531	0.539	0.516	0.531
	$\hat{\beta}_3$	1.056	1.060	1.036	1.066	1.045	1.053	1.070	1.064	1.113	1.111
Bias	$\hat{\beta}_1$	0.029	0.027	0.028	−0.003	0.000	0.027	0.031	0.039	0.016	0.031
	$\hat{\beta}_3$	0.055	0.060	0.036	0.066	0.04	0.053	0.070	0.064	0.113	0.111
MSE	$\hat{\beta}_1$	0.119	0.121	0.130	0.115	0.131	0.118	0.121	0.136	0.125	0.149
	$\hat{\beta}_3$	0.308	0.292	0.326	0.304	0.315	0.300	0.294	0.347	0.334	0.3623
CI	β_1	0.955	0.948	0.955	0.944	0.948	0.953	0.949	0.949	0.914	0.897
	β_3	0.948	0.955	0.960	0.960	0.962	0.953	0.960	0.951	0.933	0.927

Table 2: Results of the simulation study under a serial dependence MC1 for $n = 20$, $T = 13$.

		GLM3C-MC1					**GLMM**				
		λ_1					λ_1				
		.05	.25	.50	.75	1	.05	.25	.50	.75	1
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.500	0.502	0.497	0.495	0.496	0.501	0.505	0.505	0.508	0.516
	$\hat{\beta}_3$	1.041	1.041	1.045	1.052	1.051	1.041	1.048	1.060	1.078	1.091
Bias	$\hat{\beta}_1$	0.000	0.002	−0.003	−0.005	−0.004	0.001	0.005	0.005	0.008	0.016
	$\hat{\beta}_3$	0.041	0.041	0.045	0.052	0.051	0.041	0.048	0.060	0.078	0.091
MSE	$\hat{\beta}_1$	0.050	0.055	0.054	0.060	0.061	0.050	0.055	0.056	0.063	0.067
	$\hat{\beta}_3$	0.134	0.142	0.147	0.160	0.178	0.133	0.143	0.152	0.173	0.200
CI	β_1	0.958	0.946	0.953	0.947	0.950	0.951	0.932	0.936	0.913	0.916
	β_3	0.960	0.950	0.956	0.957	0.956	0.953	0.944	0.933	0.937	0.910
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.501	0.502	0.500	0.501	0.498	0.501	0.505	0.508	0.515	0.518
	$\hat{\beta}_3$	1.043	1.046	1.044	1.051	1.047	1.044	1.053	1.061	1.077	1.088
Bias	$\hat{\beta}_1$	0.001	0.002	0.000	0.001	−0.002	0.001	0.005	0.008	0.015	0.018
	$\hat{\beta}_3$	0.043	0.046	0.044	0.051	0.047	0.044	0.053	0.061	0.077	0.088
MSE	$\hat{\beta}_1$	0.052	0.056	0.057	0.061	0.071	0.052	0.056	0.059	0.065	0.079
	$\hat{\beta}_3$	0.142	0.150	0.158	0.166	0.184	0.142	0.153	0.166	0.181	0.207
CI	β_1	0.951	0.954	0.962	0.964	0.942	0.944	0.939	0.934	0.924	0.894
	β_3	0.964	0.960	0.960	0.968	0.957	0.959	0.949	0.938	0.934	0.913
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.510	0.504	0.506	0.506	0.508	0.510	0.507	0.513	0.519	0.528
	$\hat{\beta}_3$	1.039	1.044	1.048	1.049	1.042	1.038	1.051	1.064	1.075	1.082
Bias	$\hat{\beta}_1$	0.010	0.004	0.006	0.006	0.008	0.010	0.007	0.013	0.019	0.028
	$\hat{\beta}_3$	0.039	0.044	0.048	0.049	0.042	0.038	0.051	0.064	0.795	0.082
MSE	$\hat{\beta}_1$	0.058	0.059	0.064	0.072	0.080	0.058	0.060	0.066	0.076	0.089
	$\hat{\beta}_3$	0.175	0.180	0.194	0.213	0.225	0.173	0.183	0.202	0.229	0.255
CI	β_1	0.960	0.955	0.952	0.947	0.953	0.962	0.946	0.946	0.923	0.905
	β_3	0.954	0.957	0.946	0.946	0.945	0.960	0.956	0.939	0.921	0.907

Table 3: Results of the simulation study under a serial dependence MC1 for $n = 50$, $T = 7$.

		GLM3C-MC1					**GLMM**				
		λ_1					λ_1				
		.05	.25	.50	.75	1	.05	.25	.50	.75	1
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.492	0.489	0.484	0.483	0.481	0.493	0.495	0.498	0.505	0.513
	$\hat{\beta}_3$	1.068	1.062	1.056	1.048	1.040	1.070	1.074	1.084	1.092	1.102
Bias	$\hat{\beta}_1$	−0.008	−0.011	−0.016	−0.017	−0.019	−0.007	−0.005	−0.002	0.005	0.013
	$\hat{\beta}_3$	0.068	0.062	0.056	0.048	0.040	0.070	0.074	0.084	0.092	0.102
MSE	$\hat{\beta}_1$	0.032	0.033	0.036	0.037	0.039	0.032	0.033	0.038	0.040	0.045
	$\hat{\beta}_3$	0.100	0.102	0.107	0.116	0.115	0.099	0.105	0.115	0.130	0.139
CI	β_1	0.959	0.962	0.960	0.957	0.956	0.956	0.954	0.945	0.938	0.926
	β_3	0.951	0.951	0.951	0.947	0.958	0.944	0.936	0.931	0.916	0.911
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.489	0.495	0.482	0.480	0.477	0.499	0.502	0.496	0.503	0.506
	$\hat{\beta}_3$	1.057	1.045	1.062	1.055	1.021	1.058	1.057	1.091	1.100	1.088
Bias	$\hat{\beta}_1$	−0.011	−0.005	−0.018	−0.020	−0.023	−0.001	0.002	−0.004	0.003	0.006
	$\hat{\beta}_3$	0.057	0.045	0.062	0.055	0.021	0.058	0.057	0.091	0.100	0.088
MSE	$\hat{\beta}_1$	0.032	0.038	0.039	0.040	0.046	0.032	0.039	0.041	0.043	0.052
	$\hat{\beta}_3$	0.103	0.106	0.101	0.106	0.115	0.101	0.109	0.111	0.121	0.136
CI	β_1	0.953	0.948	0.947	0.957	0.955	0.949	0.938	0.938	0.937	0.916
	β_3	0.937	0.941	0.956	0.954	0.950	0.935	0.929	0.939	0.925	0.922
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.496	0.490	0.488	0.486	0.469	0.499	0.497	0.503	0.506	0.499
	$\hat{\beta}_3$	1.028	1.081	1.074	1.026	1.041	1.030	1.093	1.102	1.072	1.108
Bias	$\hat{\beta}_1$	−0.004	−0.010	−0.012	−0.014	−0.031	−0.001	−0.003	0.003	0.006	−0.001
	$\hat{\beta}_3$	0.028	0.081	0.074	0.026	0.041	0.030	0.093	0.102	0.072	0.108
MSE	$\hat{\beta}_1$	0.038	0.040	0.044	0.044	0.051	0.037	0.042	0.046	0.047	0.057
	$\hat{\beta}_3$	0.097	0.115	0.119	0.110	0.124	0.095	0.118	0.129	0.124	0.151
CI	β_1	0.948	0.962	0.953	0.945	0.942	0.938	0.955	0.929	0.926	0.893
	β_3	0.953	0.955	0.952	0.953	0.945	0.946	0.943	0.933	0.927	0.905

Table 4: Results of the simulation study under a serial dependence MC1 for $n = 50$, $T = 13$.

		GLM3C-MC1					**GLMM**				
		λ_1					λ_1				
		.05	.25	.50	.75	1	.05	.25	.50	.75	1
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.497	0.496	0.495	0.494	0.500	0.498	0.500	0.503	0.508	0.521
	$\hat{\beta}_3$	1.043	1.040	1.037	1.035	1.028	1.044	1.048	1.055	1.063	1.067
Bias	$\hat{\beta}_1$	−0.003	−0.004	−0.005	−0.006	0.000	−0.002	0.000	0.003	0.008	0.021
	$\hat{\beta}_3$	0.043	0.040	0.037	0.035	0.028	0.044	0.048	0.055	0.063	0.067
MSE	$\hat{\beta}_1$	0.022	0.023	0.026	0.028	0.029	0.022	0.024	0.027	0.030	0.032
	$\hat{\beta}_3$	0.057	0.060	0.064	0.070	0.074	0.058	0.062	0.068	0.078	0.084
CI	β_1	0.953	0.949	0.945	0.942	0.949	0.949	0.939	0.919	0.905	0.890
	β_3	0.942	0.943	0.943	0.944	0.951	0.937	0.932	0.926	0.910	0.900
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.496	0.496	0.496	0.495	0.493	0.497	0.500	0.504	0.508	0.513
	$\hat{\beta}_3$	1.053	1.050	1.046	1.044	1.040	1.055	1.058	1.064	1.072	1.079
Bias	$\hat{\beta}_1$	−0.004	−0.005	−0.004	−0.005	−0.007	−0.003	0.000	0.004	0.008	0.013
	$\hat{\beta}_3$	0.053	0.050	0.046	0.044	0.040	0.055	0.058	0.064	0.072	0.079
MSE	$\hat{\beta}_1$	0.022	0.025	0.028	0.029	0.030	0.022	0.025	0.029	0.031	0.033
	$\hat{\beta}_3$	0.060	0.064	0.069	0.073	0.078	0.060	0.066	0.074	0.081	0.090
CI	β_1	0.955	0.949	0.949	0.951	0.952	0.951	0.934	0.925	0.914	0.891
	β_3	0.955	0.940	0.943	0.950	0.946	0.948	0.930	0.924	0.912	0.902
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.497	0.496	0.495	0.492	0.494	0.498	0.500	0.504	0.505	0.513
	$\hat{\beta}_3$	1.059	1.058	1.062	1.060	1.055	1.061	1.066	1.080	1.090	1.096
Bias	$\hat{\beta}_1$	−0.003	−0.004	−0.005	−0.008	−0.006	−0.002	0.000	0.004	0.005	0.013
	$\hat{\beta}_3$	0.059	0.058	0.062	0.060	0.055	0.061	0.066	0.080	0.090	0.096
MSE	$\hat{\beta}_1$	0.025	0.027	0.029	0.032	0.035	0.025	0.027	0.030	0.035	0.039
	$\hat{\beta}_3$	0.068	0.070	0.077	0.085	0.088	0.066	0.072	0.082	0.095	0.103
CI	β_1	0.959	0.964	0.954	0.953	0.955	0.950	0.949	0.938	0.927	0.895
	β_3	0.955	0.952	0.953	0.944	0.950	0.951	0.940	0.928	0.909	0.896

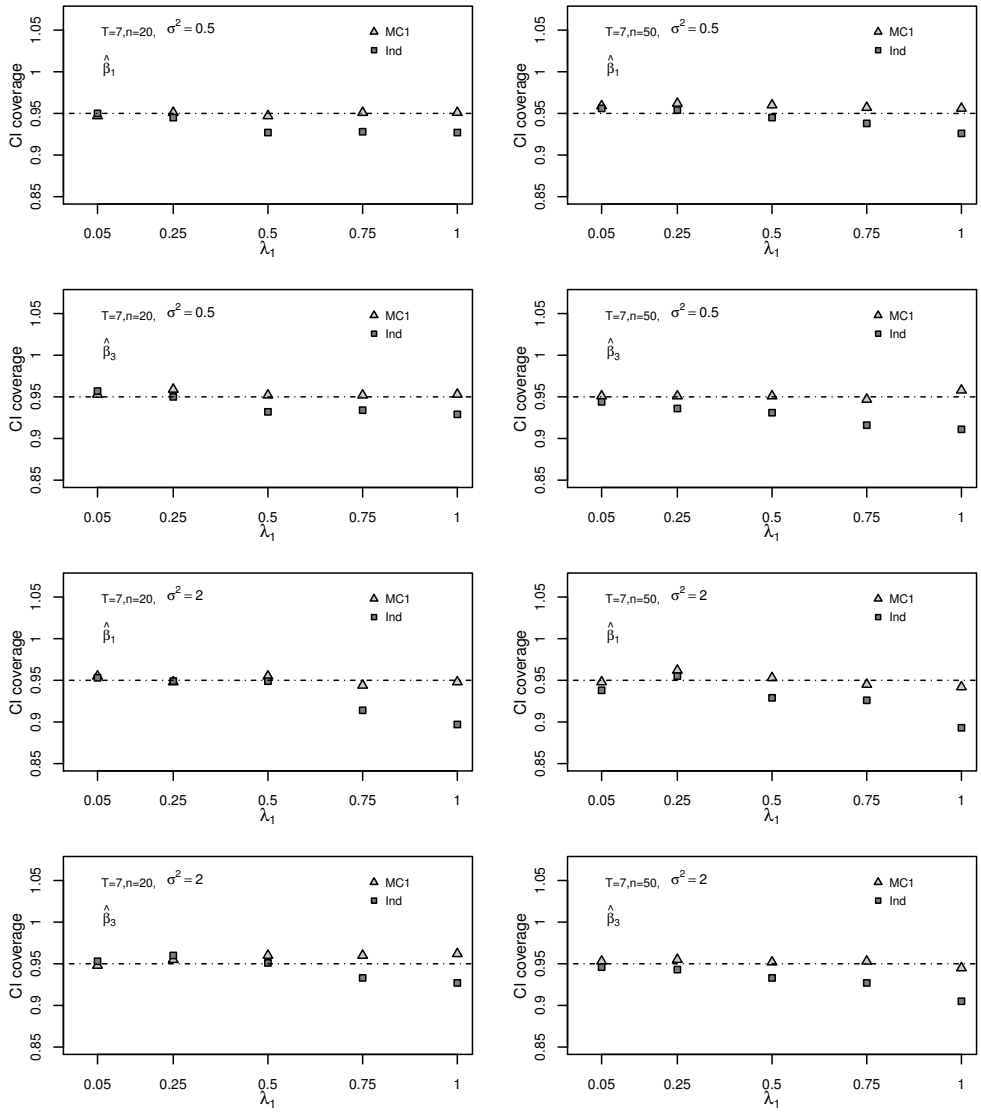

Figure 1: Coverage probabilities of nominal 95% confidence intervals (CI coverage) for β_1 ($\hat{\beta}_1$) and β_3 ($\hat{\beta}_3$) based on 1000 runs using GLM3C and GLMM estimation procedures (T = length of profile on each subject, n = number of subjects, σ^2 = variance of the random effect). Data set on each run has a serial dependence MC1 regulated by λ_1. Coding for estimation procedures: MC1 (GLM3C) and Ind (GLMM).

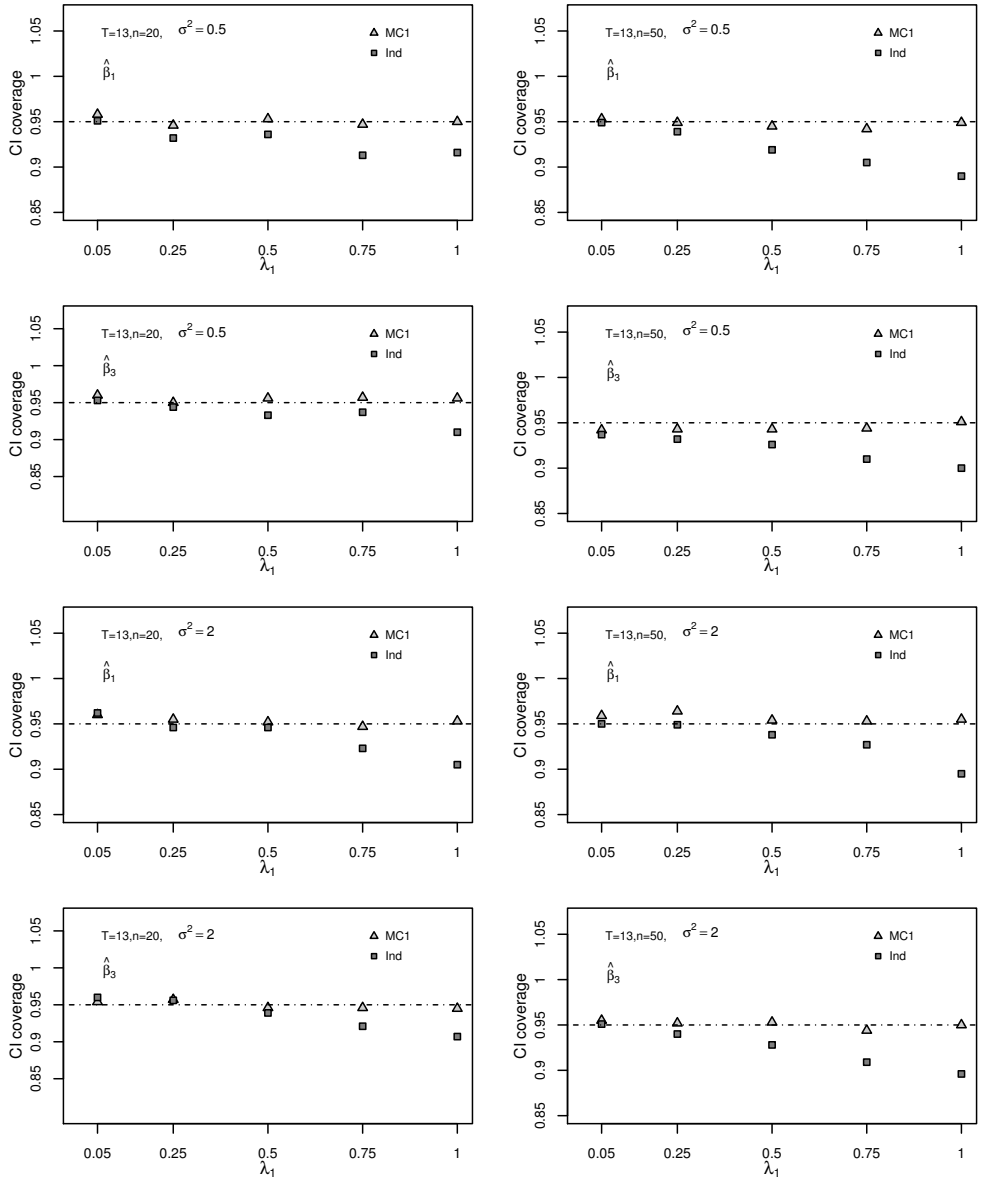

Figure 2: Coverage probabilities of nominal 95% confidence intervals (CI coverage) for β_1 ($\hat{\beta}_1$) and β_3 ($\hat{\beta}_3$) based on 1000 runs using GLM3C and GLMM estimation procedures (T = length of profile on each subject, n = number of subjects, σ^2 = variance of the random effect). Data set on each run has a serial dependence MC1 regulated by λ_1. Coding for estimation procedures: MC1 (GLM3C) and Ind (GLMM).

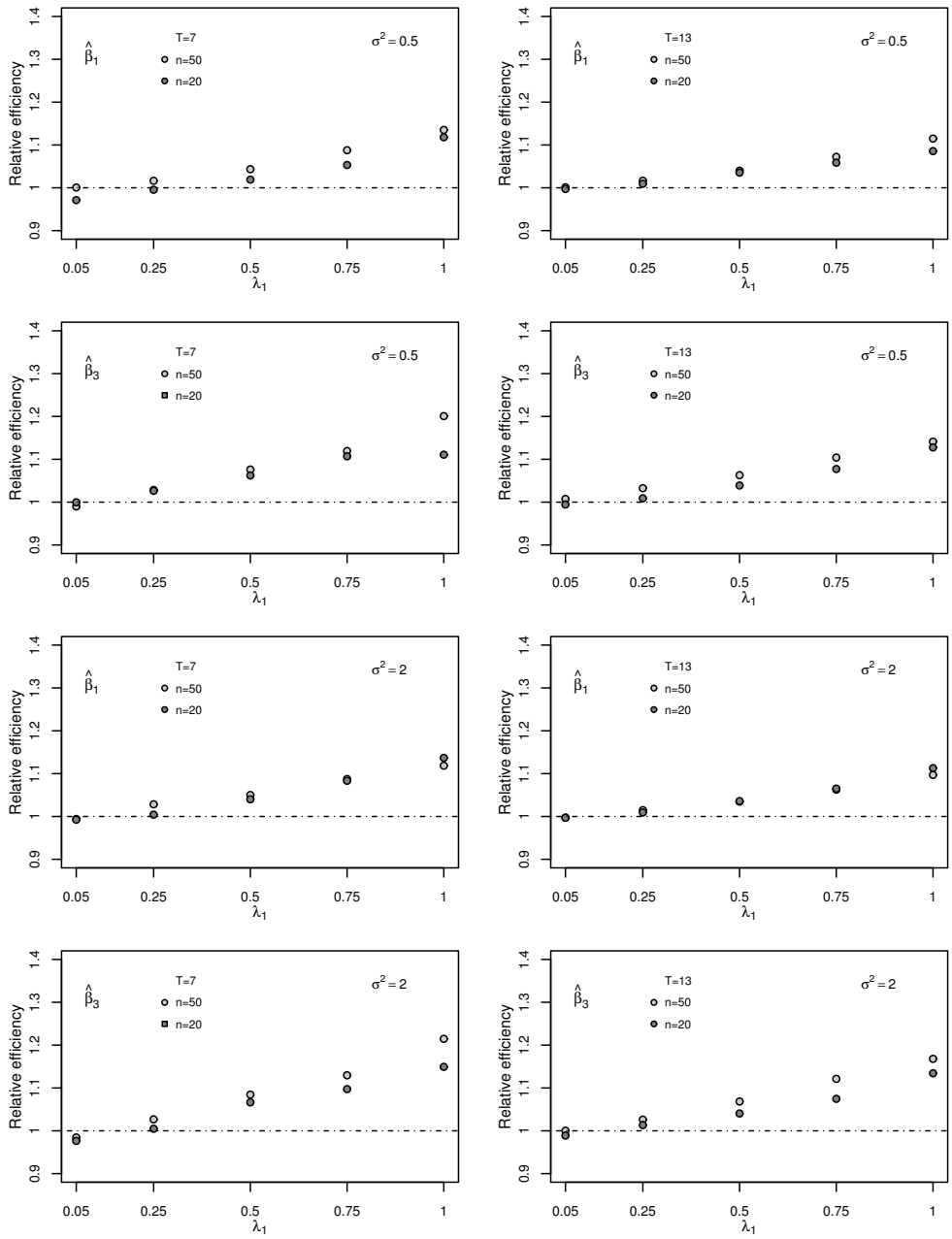

Figure 3: Relative efficiency (see text for definition) of $\hat{\beta}_1$ and $\hat{\beta}_3$ based on 1000 runs using GLM3C and GLMM estimation procedures (T = length of profile on each subject, n = number of subjects, σ^2 = variance of the random effect). Data set on each run has a serial dependence MC1 regulated by λ_1.

3.2. Serial dependence MC2

Under a serial dependence MC2 and to β_1 and β_3 parameters the main conclusions, based on the approaches GLM3C-MC2 and GLMM, are:

(**i**) In both approaches the coverage probabilities are similar when the dependence structure established by (λ_1, λ_2) is low, (0.05,0.05) and (0.25,0.25). The exception occurs when$(\lambda_1, \lambda_2) = (0.05,0.05)$ for $\hat{\beta}_1$ and $\hat{\beta}_3$ when $T = 13$, $n = 20$ and $\sigma^2 = 2$, where the coverage probabilities of GLMM approach are closer to nominal. When (λ_1, λ_2) is greater than (0.25,0.25) and for all the design configurations the coverage probabilities are closer to nominal for the GLM3C approach than for the GLMM approach. This is so much better applied as the length of the profile of each subject increases as well as the dependence structure established by (λ_1, λ_2) (Figures 4–5 and Tables 5–8).

(**ii**) The results of simulation show that when (λ_1, λ_2) is greater than (0.25,0.25) the GLM3C estimators are more efficient than the GLMM estimators and this is so much better applied as (λ_1, λ_2) increases. For values of (λ_1, λ_2) equal to (0.25,0.25) the efficiency of both approaches is similar. When (λ_1, λ_2) is equal to (0.05,0.05) the GLMM estimators are more efficient than the GLM3C in some design configurations (Figure 6 and Tables 5–8).

(**iii**) The estimate bias for the GLMM approach becomes greater than the associated with the GLM3C as (λ_1, λ_2) increases and for all the designs configurations except for $\hat{\beta}_3$ when $T = 13$, $n = 20$ and $\sigma^2 = 2$ (Tables 5–8).

Table 5: Results of the simulation study under a serial dependence MC2 for $n = 20$, $T = 7$.

		GLM3C-MC2					**GLMM**				
		(λ_1, λ_2)					(λ_1, λ_2)				
		$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$	$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.522	0.514	0.531	0.537	0.546	0.518	0.518	0.550	0.576	0.605
	$\hat{\beta}_3$	1.066	1.130	1.114	1.114	1.123	1.050	1.139	1.152	1.197	1.259
Bias	$\hat{\beta}_1$	0.021	0.014	0.031	0.037	0.046	0.018	0.018	0.050	0.076	0.105
	$\hat{\beta}_3$	0.066	0.130	0.114	0.114	0.123	0.050	0.139	0.152	0.197	0.259
MSE	$\hat{\beta}_1$	0.091	0.106	0.115	0.127	0.140	0.085	0.104	0.123	0.147	0.177
	$\hat{\beta}_3$	0.253	0.318	0.338	0.350	0.372	0.233	0.316	0.358	0.420	0.504
CI	β_1	0.944	0.944	0.946	0.937	0.944	0.946	0.947	0.928	0.905	0.893
	β_3	0.963	0.941	0.946	0.954	0.961	0.974	0.928	0.925	0.914	0.913
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.521	0.526	0.514	0.529	0.548	0.515	0.530	0.539	0.571	0.616
	$\hat{\beta}_3$	1.066	1.067	1.130	1.105	1.112	1.060	1.080	1.177	1.198	1.254
Bias	$\hat{\beta}_1$	0.021	0.026	0.014	0.029	0.048	0.015	0.030	0.039	0.071	0.116
	$\hat{\beta}_3$	0.066	0.067	0.130	0.105	0.112	0.060	0.080	0.177	0.198	0.254
MSE	$\hat{\beta}_1$	0.098	0.108	0.114	0.128	0.144	0.092	0.104	0.123	0.147	0.183
	$\hat{\beta}_3$	0.267	0.276	0.364	0.413	0.372	0.263	0.278	0.391	0.482	0.487
CI	β_1	0.944	0.943	0.936	0.938	0.939	0.942	0.944	0.919	0.910	0.892
	β_3	0.955	0.955	0.949	0.936	0.946	0.958	0.946	0.920	0.897	0.904
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.507	0.525	0.524	0.496	0.522	0.505	0.536	0.558	0.550	0.601
	$\hat{\beta}_3$	1.054	1.006	0.991	1.112	1.103	1.051	1.033	1.062	1.233	1.272
Bias	$\hat{\beta}_1$	0.007	0.025	0.024	−0.004	0.022	0.005	0.036	0.058	0.050	0.101
	$\hat{\beta}_3$	0.054	0.006	−0.009	0.112	0.103	0.051	0.033	0.062	0.233	0.272
MSE	$\hat{\beta}_1$	0.099	0.098	0.096	0.124	0.132	0.094	0.100	0.109	0.149	0.173
	$\hat{\beta}_3$	0.269	0.266	0.282	0.371	0.385	0.260	0.276	0.307	0.478	0.542
CI	β_1	0.941	0.951	0.956	0.941	0.947	0.944	0.939	0.931	0.902	0.902
	β_3	0.954	0.956	0.952	0.945	0.963	0.957	0.945	0.924	0.904	0.906

Table 6: Results of the simulation study under a serial dependence MC2 for $n = 20$, $T = 13$.

		GLM3C-MC2					**GLMM**				
		(λ_1, λ_2)					(λ_1, λ_2)				
		$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$	$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.501	0.497	0.501	0.507	0.538	0.501	0.503	0.521	0.540	0.594
	$\hat{\beta}_3$	1.042	1.051	1.064	1.070	1.041	1.044	1.068	1.103	1.141	1.141
Bias	$\hat{\beta}_1$	0.001	−0.003	0.001	0.007	0.038	0.001	0.003	0.021	0.040	0.094
	$\hat{\beta}_3$	0.042	0.051	0.064	0.070	0.041	0.044	0.068	0.103	0.141	0.141
MSE	$\hat{\beta}_1$	0.051	0.057	0.064	0.073	0.081	0.051	0.058	0.071	0.086	0.112
	$\hat{\beta}_3$	0.136	0.155	0.177	0.198	0.210	0.134	0.160	0.199	0.242	0.275
CI	β_1	0.948	0.951	0.947	0.950	0.953	0.946	0.942	0.915	0.876	0.842
	β_3	0.956	0.948	0.956	0.950	0.961	0.955	0.941	0.922	0.895	0.865
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.506	0.505	0.535	0.532	0.516	0.507	0.512	0.556	0.569	0.567
	$\hat{\beta}_3$	1.025	1.044	0.944	0.996	1.087	1.030	1.060	0.991	1.07	1.199
Bias	$\hat{\beta}_1$	0.006	0.005	0.035	0.032	0.016	0.007	0.012	0.056	0.069	0.067
	$\hat{\beta}_3$	0.025	0.044	−0.056	−0.004	0.087	0.030	0.060	−0.009	0.077	0.199
MSE	$\hat{\beta}_1$	0.054	0.058	0.071	0.076	0.087	0.053	0.060	0.077	0.091	0.114
	$\hat{\beta}_3$	0.145	0.157	0.156	0.196	0.242	0.144	0.164	0.169	0.231	0.324
CI	β_1	0.942	0.953	0.930	0.939	0.951	0.946	0.938	0.891	0.868	0.847
	β_3	0.936	0.959	0.950	0.948	0.948	0.937	0.948	0.919	0.891	0.854
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.536	0.510	0.549	0.549	0.553	0.537	0.516	0.571	0.585	0.608
	$\hat{\beta}_3$	0.981	1.048	0.894	0.882	0.883	0.980	1.063	0.947	0.980	1.024
Bias	$\hat{\beta}_1$	0.036	0.010	0.049	0.049	0.053	0.037	0.016	0.071	0.085	0.108
	$\hat{\beta}_3$	−0.019	0.048	−0.106	−0.118	−0.117	−0.020	0.063	−0.053	−0.020	0.024
MSE	$\hat{\beta}_1$	0.068	0.066	0.072	0.078	0.089	0.067	0.068	0.080	0.094	0.119
	$\hat{\beta}_3$	0.164	0.194	0.168	0.183	0.216	0.159	0.199	0.177	0.205	0.262
CI	β_1	0.939	0.939	0.937	0.948	0.946	0.948	0.928	0.900	0.870	0.835
	β_3	0.933	0.949	0.939	0.948	0.943	0.946	0.946	0.923	0.912	0.880

Table 7: Results of the simulation study under a serial dependence MC2 for $n = 50$, $T = 7$.

		GLM3C-MC2					**GLMM**				
		(λ_1, λ_2)					(λ_1, λ_2)				
		$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$	$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.506	0.495	0.496	0.513	0.532	0.506	0.509	0.528	0.567	0.612
	$\hat{\beta}_3$	1.018	1.070	1.077	1.052	1.045	1.017	1.098	1.142	1.163	1.210
Bias	$\hat{\beta}_1$	0.006	−0.005	−0.003	0.013	0.032	0.006	0.009	0.028	0.067	0.112
	$\hat{\beta}_3$	0.018	0.070	0.077	0.052	0.045	0.017	0.098	0.142	0.163	0.210
MSE	$\hat{\beta}_1$	0.037	0.035	0.039	0.043	0.049	0.037	0.036	0.043	0.055	0.074
	$\hat{\beta}_3$	0.101	0.111	0.123	0.115	0.124	0.099	0.117	0.146	0.157	0.196
CI	β_1	0.953	0.956	0.951	0.946	0.950	0.947	0.946	0.933	0.907	0.874
	β_3	0.947	0.952	0.941	0.953	0.957	0.943	0.924	0.912	0.909	0.881
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.503	0.512	0.520	0.517	0.520	0.505	0.527	0.554	0.572	0.598
	$\hat{\beta}_3$	1.049	1.045	1.044	1.063	1.070	1.050	1.072	1.114	1.176	1.237
Bias	$\hat{\beta}_1$	0.003	0.012	0.020	0.017	0.020	0.005	0.027	0.054	0.072	0.098
	$\hat{\beta}_3$	0.049	0.045	0.044	0.063	0.070	0.050	0.072	0.114	0.176	0.237
MSE	$\hat{\beta}_1$	0.036	0.040	0.044	0.043	0.049	0.036	0.042	0.051	0.057	0.070
	$\hat{\beta}_3$	0.101	0.099	0.123	0.133	0.151	0.098	0.103	0.141	0.174	0.229
CI	β_1	0.952	0.943	0.947	0.961	0.949	0.950	0.926	0.922	0.910	0.885
	β_3	0.955	0.948	0.938	0.944	0.937	0.945	0.932	0.914	0.897	0.866
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.509	0.506	0.498	0.510	0.524	0.511	0.522	0.535	0.568	0.609
	$\hat{\beta}_3$	1.062	1.047	1.065	1.049	1.038	1.063	1.076	1.144	1.171	1.215
Bias	$\hat{\beta}_1$	0.009	0.006	−0.002	0.010	0.024	0.011	0.022	0.035	0.068	0.109
	$\hat{\beta}_3$	0.062	0.047	0.065	0.049	0.038	0.063	0.076	0.144	0.171	0.215
MSE	$\hat{\beta}_1$	0.040	0.039	0.042	0.046	0.049	0.040	0.041	0.048	0.059	0.075
	$\hat{\beta}_3$	0.105	0.103	0.117	0.129	0.140	0.102	0.107	0.145	0.170	0.215
CI	β_1	0.949	0.941	0.938	0.955	0.966	0.943	0.929	0.917	0.909	0.883
	β_3	0.951	0.953	0.951	0.953	0.951	0.947	0.939	0.909	0.902	0.873

Table 8: Results of the simulation study under a serial dependence MC2 for $n = 50$, $T = 13$.

		GLM3C-MC2					**GLMM**				
		(λ_1, λ_2)					(λ_1, λ_2)				
		$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$	$(.05, .05)$	$(.25, .25)$	$(.5, .5)$	$(.75, .75)$	$(1, 1)$
		$\sigma^2 = 0.5$									
Mean	$\hat{\beta}_1$	0.497	0.499	0.503	0.506	0.510	0.498	0.507	0.523	0.541	0.564
	$\hat{\beta}_3$	1.042	1.044	1.041	1.038	1.044	1.045	1.062	1.083	1.113	1.160
Bias	$\hat{\beta}_1$	−0.003	−0.001	0.003	0.006	0.010	−0.001	0.007	0.023	0.041	0.064
	$\hat{\beta}_3$	0.042	0.044	0.041	0.038	0.044	0.045	0.062	0.083	0.113	0.160
MSE	$\hat{\beta}_1$	0.022	0.024	0.028	0.030	0.035	0.022	0.025	0.031	0.038	0.049
	$\hat{\beta}_3$	0.058	0.065	0.073	0.079	0.089	0.059	0.069	0.085	0.104	0.138
CI	β_1	0.950	0.953	0.952	0.949	0.949	0.953	0.934	0.901	0.884	0.836
	β_3	0.944	0.945	0.943	0.938	0.948	0.941	0.919	0.900	0.867	0.818
		$\sigma^2 = 1$									
Mean	$\hat{\beta}_1$	0.498	0.506	0.506	0.499	0.507	0.500	0.516	0.527	0.536	0.562
	$\hat{\beta}_3$	1.035	1.025	1.027	1.056	1.044	1.038	1.043	1.069	1.129	1.160
Bias	$\hat{\beta}_1$	−0.002	0.006	0.006	−0.001	0.007	0.000	0.016	0.027	0.036	0.062
	$\hat{\beta}_3$	0.035	0.025	0.027	0.056	0.044	0.038	0.043	0.069	0.129	0.160
MSE	$\hat{\beta}_1$	0.022	0.025	0.028	0.028	0.033	0.022	0.026	0.031	0.034	0.045
	$\hat{\beta}_3$	0.051	0.060	0.068	0.071	0.083	0.052	0.063	0.078	0.094	0.128
CI	β_1	0.950	0.941	0.943	0.952	0.954	0.940	0.922	0.901	0.883	0.856
	β_3	0.954	0.953	0.944	0.956	0.942	0.953	0.926	0.889	0.874	0.828
		$\sigma^2 = 2$									
Mean	$\hat{\beta}_1$	0.492	0.500	0.503	0.503	0.513	0.493	0.509	0.524	0.539	0.568
	$\hat{\beta}_3$	1.048	1.036	1.038	1.041	1.037	1.051	1.054	1.079	1.114	1.148
Bias	$\hat{\beta}_1$	−0.008	0.000	0.003	0.003	0.013	−0.007	0.009	0.024	0.039	0.068
	$\hat{\beta}_3$	0.048	0.036	0.038	0.041	0.037	0.051	0.054	0.079	0.114	0.148
MSE	$\hat{\beta}_1$	0.024	.030	0.033	0.038	0.042	0.024	0.031	0.037	0.045	0.055
	$\hat{\beta}_3$	0.060	0.071	0.079	0.088	0.097	0.061	0.076	0.091	0.114	0.143
CI	β_1	0.952	0.936	0.939	0.943	0.948	0.951	0.928	0.900	0.876	0.848
	β_3	0.944	0.951	0.956	0.959	0.950	0.942	0.935	0.915	0.880	0.844

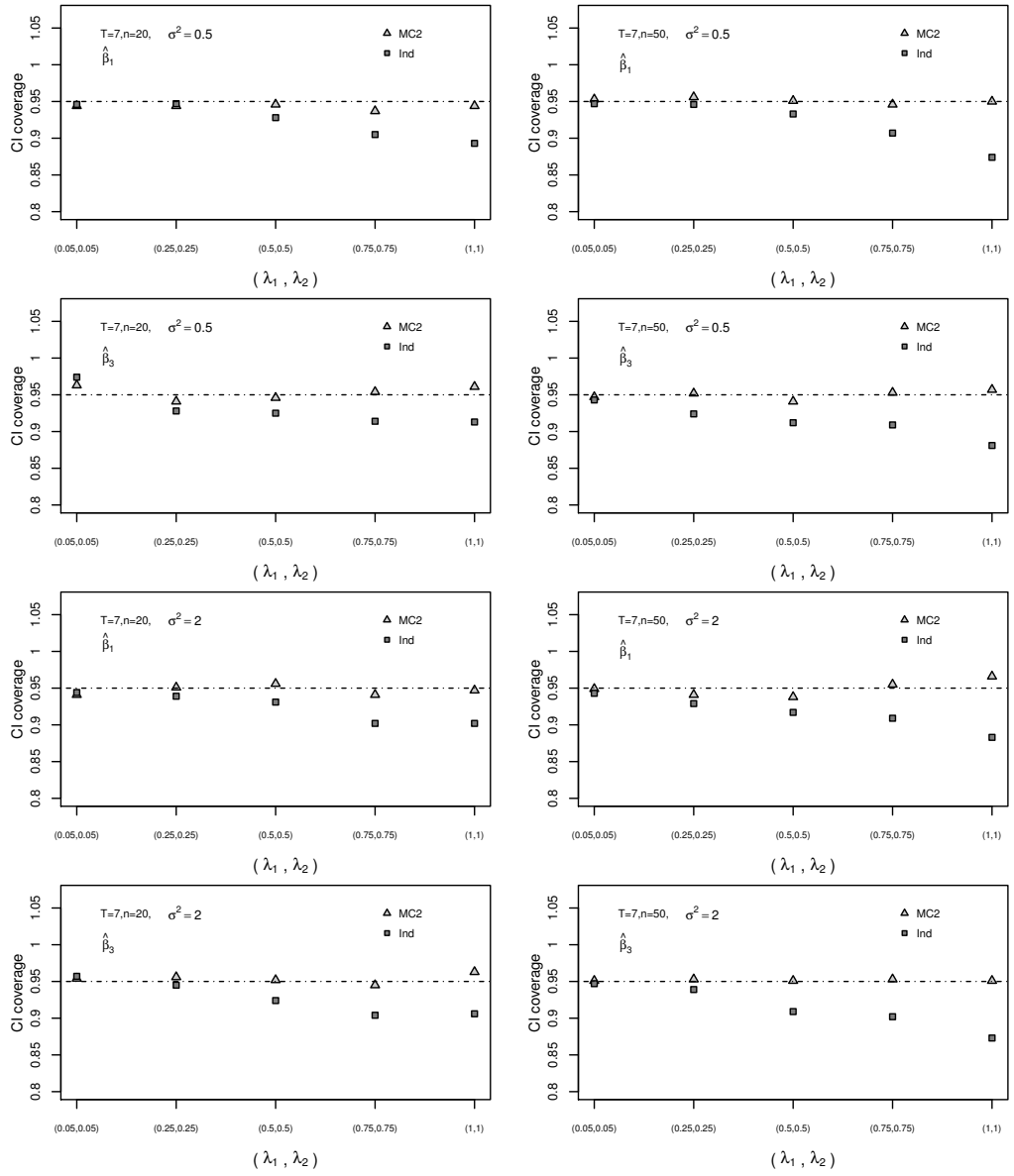

Figure 4: Coverage probabilities of nominal 95% confidence intervals (CI coverage) for β_1 ($\hat{\beta}_1$) and β_3 ($\hat{\beta}_3$) based on 1000 runs using GLM3C and GLMM estimation procedures (T = length of profile on each subject, n = number of subjects, σ^2 = variance of the random effect). Data set on each run has a serial dependence MC2 regulated by $\lambda = (\lambda_1, \lambda_2)$. Coding for estimation procedures: MC2 (GLM3C) and Ind (GLMM).

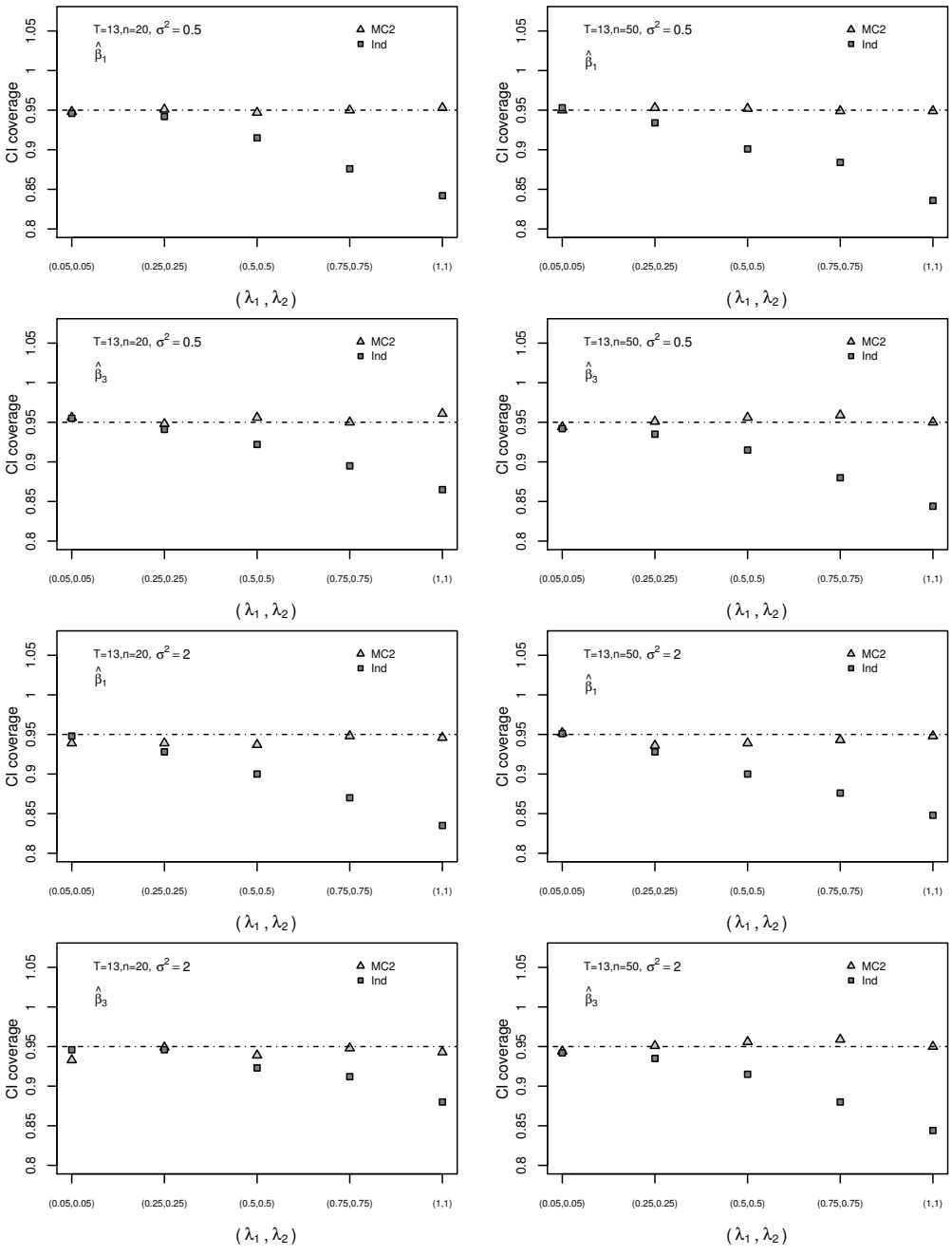

Figure 5: Coverage probabilities of nominal 95% confidence intervals (CI coverage) for β_1 ($\hat{\beta}_1$) and β_3 ($\hat{\beta}_3$) based on 1000 runs using GLM3C and GLMM estimation procedures (T = length of profile on each subject, n = number of subjects, σ^2 = variance of the random effect). Data set on each run has a serial dependence MC2 regulated by $\lambda = (\lambda_1, \lambda_2)$. Coding for estimation procedures: MC2 (GLM3C) and Ind (GLMM).

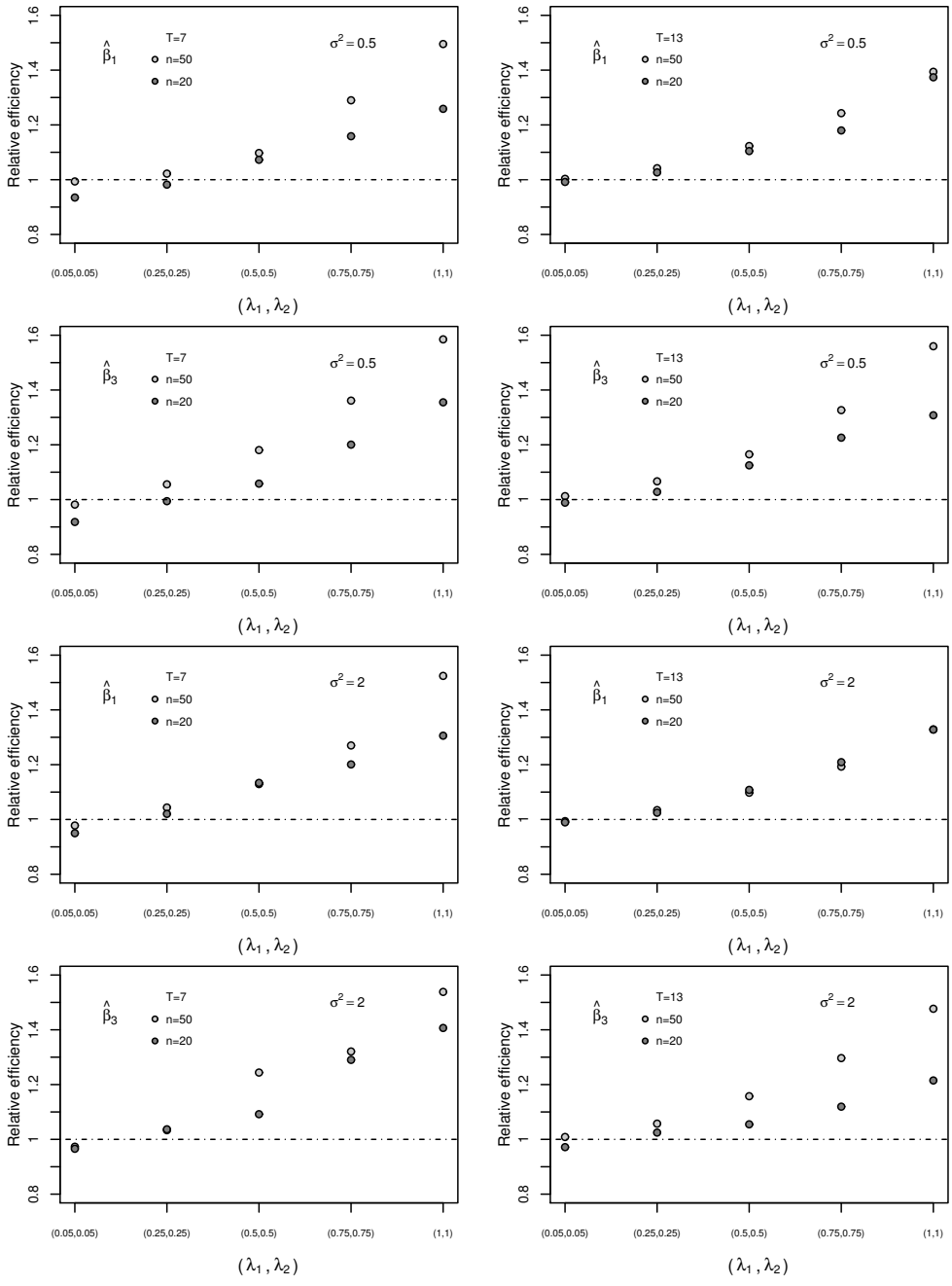

Figure 6: Relative efficiency (see text for definition) of $\hat{\beta}_1$ and $\hat{\beta}_3$ based on 1000 runs using GLM3C and GLMM estimation procedures (T = length of profile on each subject, n = number of subjects, σ^2 = variance of the random effect). Data set on each run has a serial dependence MC2 regulated by $\lambda = (\lambda_1, \lambda_2)$.

4. ANALYSIS OF CONTRACEPTING WOMEN DATA

The key results of the simulation study are illustrated, using data from a longitudinal clinical trial of contracepting women given in [7]. In this trial, and following their description, women received an injection of either 100 mg or 150 mg of depot-medroxyprogesterone acetate (DMPTA) on the day of randomization and three additional injections at 90-day intervals. There was a final follow-up visit 90 days after the fourth injection. The outcome of interest is a binary response indicating whether the i^{th} woman experienced amenorrhea ($Y_{ij} = 1$) in the j^{th} four successive three-month intervals, or not ($Y_{ij} = 0$). A feature of this clinical trial is that there was substantial dropout (17% dropped out after receiving one injection of DMPA, 13% dropped out after receiving only two injections, and 7% dropped out after receiving three injections).

The mixed effects logistic model proposed by [7]

$$\begin{aligned}\text{logit } [E(Y_{ij}|b_i)] \;=\; & \beta_0 + \beta_1 time_{ij} + \beta_2 time_{ij}^2 + \beta_3 (time_{ij} \times dose_i) \\ & + \beta_4 (time_{ij}^2 \times dose_i) + b_i\end{aligned}$$

with $j = 1, ..., 4$, $time = 1, 2, 3, 4$ and *dose* a binary variable taking the value 1 if the i-th woman is randomized to 150mg of DMPA and 0 otherwise, was fitted to data with different dependence structures:

(**i**) Independence (Model I).

(**ii**) Serial dependence MC1 (Model II).

(**iii**) Serial dependence MC2 (Model III).

Models I correspond to the GLMM approach (model fitted by [7]), Models II and III correspond to the GLM3C approach. The analysis of all models was performed using the `bild` function of the `R` package `bild` with the dependence argument sated to `indR` to Model I, `MC1R` and `MC2R`, respectively to Models II and III.

Tables 9 and 10 display the results of fitting the different models to data. Table 9 reports the log-likelihood, the change in deviance with corresponding p-values. The estimated values of the parameters, as well as their standard errors, t-ratio and corresponding p-values are given in Table 10.

The first step of the analysis is to choose the appropriate serial dependence to account correlation between successive observation of the same subject. The results of Table 10 show, among other things, that the estimates of λ_1 and λ_2 in MC2 model (Model III) as well as the estimate of λ_1 in MC1 model (Model II) point strongly to a first order serial dependence. The change of deviance between this two models, compared with the χ_1^2 reference distribution, produces a p-value = 0.3467 (Table 9) confirming that there is no significant difference between this

two models at 5% level. To explore further this point Model II was compare to Model I, which assume independence between successive observation of the same subject. The change of deviance between this two models, compared with the χ^2_1 reference distribution, produces a p-value= 0.0001 (Table 9) and so Model II with a serial dependence MC1 is significantly preferable to Model I.

Table 9: Log-likelihood and change in deviance between models.

Model	**LogL**	**Δ D**	**p-value**
I	−1937.54		
II	−1930.108	14.866	0.0001
II	−1930.108		
III	−1929.665	0.885	0.3467

The results displayed in Table 10 also show that the model with the appropriate serial dependence (MC1 model–Model II) produce smaller standard errors, as point out in the simulation study, as well as a decrease in the value of the estimative of σ^2 face a more complex serial dependence model as remarked by Pinheiro and Bates (2000) ([17]).

Table 10: Parameters estimates, Standard errors, t-ratio and p-value for models I, II and III.

Model	**Parameter**	**Estimate**	**SE**	**t-ratio**	**p-value**
I	β_0	−3.799	0.305	−12.471	0.0000
	β_1	1.131	0.268	4.221	0.0000
	β_2	−0.042	0.055	−0.763	0.4457
	β_3	0.562	0.192	2.932	0.0034
	β_4	−0.109	0.050	−2.206	0.0274
	σ^2	5.030			
II	β_0	−3.443	0.304	−11.328	0.0000
	β_1	1.033	0.247	4.188	0.0000
	β_2	−0.039	0.050	−0.781	0.4346
	β_3	0.522	0.177	2.943	0.0033
	β_4	0.105	0.177	2.943	0.0234
	λ_1	0.744	0.226	3.293	0.0009
	σ^2	3.598			
III	β_0	−3.384	0.397	−8.524	0.0000
	β_1	1.014	0.253	4.004	0.0000
	β_2	−0.038	0.049	−0.781	0.4349
	β_3	0.516	0.178	2.904	0.0037
	β_4	−0.105	0.046	−2.262	0.0237
	λ_1	0.820	0.397	2.068	0.0387
	λ_2	0.092	0.398	0.230	0.8178
	σ^2	3.376			

The model fitted by [7] corresponds to Model I. The estimated values of the parameters, as well as, their standard errors reported in Table 10 are in close agreement to those obtained by [7]. To fit Model I [7] used the `PROC NLMIXED` procedure in `SAS` and the estimation was based on 50-point adaptive Gaussian quadrature.

5. FINAL REMARKS

This paper is concerned with the asses of performance of the GLM3C and GLMM approaches both implemented in `R` package `bild` for the analysis of longitudinal binary data. The GLM3C approach seems to be preferable to GLMM in the situations considered by checking that its performance is so much better the higher the serial correlation between observations of the same subject, regardless of the number of subjects involved in the study, the length of their profile or the variance of the random effect. In spite of the use of the adaptive Gaussian quadrature method the users may be aware that this method needs careful handling to ensure converge even in simple random-effects models for categorical outcome data as referred in [15].

The results pointed out in the simulation study are illustrated in the example analyzed where a MC1 model was need to account dependence between successive observations of the same subject. The program codes for analysing the data set are available under request from the authors.

Finally, the `R` package `bild` allows the practitioners to choose the serial dependence adequate to use for their data at hand.

ACKNOWLEDGMENTS

This work has been partially funded by FCT – Fundação Nacional para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2013. The authors thank the Referee, the Associate Editor and the Editor-in-Chief for their comments which have led to significant improvement of this paper.

REFERENCES

[1] Azzalini, A. (1994). Logistic Regression for Autocorrelated Data with Application to Repeated Measures, *Biometrika*, **81**(4), 767–775.

[2] Bates, D.; Maechler, M. and Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes *R package* version 0.999375-42.
URL: `http://CRAN.R-project.org/package=lme4`

[3] Breslow, N.E. and Clayton, D.G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association*, **88**, 9–25.

[4] Fitzmaurice, G.M. and Laird, N.M. (1993). A likelihood-based method for analysing longitudinal binary responses, *Biometrika*, **80**, 141–151.

[5] Fitzmaurice, G.M.; Laird, N.M. and Rotnitzky, A.G. (1993). Regression models for discrete longitudinal responses, *Statistical Science*, **8**(3), 141–151.

[6] Fitzmaurice, G.M.; Laird, N.M. and Lipsitz, S.R. (1994). Analyzing incomplete longitudinal binary responses: a likelihood-based approach, *Biometrics*, **50**, 601–612.

[7] Fitzmaurice, G.M.; Laird, N.M. and Ware, J.H. (2004). *Applied Longitudinal Analysis*, John Wiley & Sons, New York.

[8] Gonçalves, M.H. (2002). *Likelihood methods for discrete longitudinal data*, PhD thesis, University of Lisbon.

[9] Gonçalves, M.H. and Azzalini, A. (2008). Using Markov chains for marginal modelling of binary longitudinal data in an exact likelihood approach, *Metron*, **LXVI**(1), 157–181.

[10] Gonçalves, M.H.; Cabral, M.S. and Azzalini, A. (2012). The R package bild for the analysis of binary longitudinal data, *Journal of statistical software*, **46**(9), 1–17.

[11] Gonçalves, M.H.; Cabral, M.S. and Azzalini, A. (2013). bild: A Package for Binary Longitudinal Data, Version 1.1-5.
URL: `http://CRAN.R-project.org/package=bild`

[12] Jansen, I.; Beunckens, C.; Molenberghs, G.; Verbeke, G. and Mallinckrodt, C. (2006). Analyzing incomplete discrete longitudinal clinical trials data, *Statistical Science*, **21**(1), 52–69.

[13] Laird, N. (1988). Missing data in longitudinal studies, *Statistics in Medicine*, **7**, 505–315.

[14] Lipsitz, S.R.; Molenberghs, G.; Fitzmaurice, G.M. and Ibrahim, J. (2000). GEE with Gaussian estimation of the correlations when data are incomplete, *Biometrics*, **56**, 528–536.

[15] Lesaffre, E. and Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random-effects model: an example, *Applied Statistics*, **50**, 325–335.

[16] Little, R.J.A. and Rubin, D.B. (1987). *Statistical Analysis with Missing Data*, John Wiley & Sons, New York.

[17] Pinheiro, J.C. and Bates, D.M. (2004). *Mixed effects models in S and S-Plus*, Springer-Verlag, New York.

[18] R DevelopmentCore Team (2013). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: `http://www.R-project.org`