Nonlinear mixed-effects models with misspecified random-effects distribution

Reza Drikvandi

Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, UK

Email: r.drikvandi@mmu.ac.uk

Abstract

Nonlinear mixed-effects models are being widely used for the analysis of longitudinal data, especially from pharmaceutical research. They use random effects which are latent and unobservable variables, so the random-effects distribution is subject to misspecification in practice. In this paper, we first study the consequences of misspecifying the random-effects distribution in nonlinear mixed-effects models. Our study is focused on Gauss-Hermite quadrature which is now the routine method for calculation of the marginal likelihood in mixed models. We then present a formal diagnostic test to check the appropriateness of the assumed random-effects distribution in nonlinear mixed-effects models, which is very useful for real data analysis. Our findings show that the estimates of fixed-effects parameters in nonlinear mixed-effects models are generally robust to deviations from normality of the random-effects distribution, but the estimates of variance components are very sensitive to the distributional assumption of random effects. Furthermore, a misspecified random-effects distribution will either overestimate or underestimate the predictions of random effects. We illustrate the results using a real data application from an intensive pharmacokinetic study.

Keywords: *Diagnostic test; Gauss-Hermite quadrature; Longitudinal data; Nonlinear mixed-effects models; Prediction; Random-effects distribution.*

1. Introduction

Nonlinear mixed-effects models are well suited for the analysis of longitudinal data, especially from pharmaceutical research. For example, in pharmacokinetics, often a nonlinear function for drug concentration is achieved over time after administration of a drug and the goal is to characterise drug disposition[1,2]. The term "mixed-effects" refers to the presence of both fixed effects and random effects in the model. Fixed effects are regression parameters constant across

subjects, while random effects are subject-specific random variables incorporated to capture the inter-subject variability.

To fit a nonlinear mixed-effects model, one often needs to assume a distribution for the random effects. Inferences are then based on the marginal likelihood function after integrating out the random effects over their assumed distribution. It is common to assume that the random effects follow a (multivariate) normal distribution. However, the normality assumption of random effects may not always be valid in practice[3,4]. Since the random-effects distribution is crucial in the calculation of the marginal likelihood, it is important to study the impact of misspecifying the random-effects distribution on inferences about parameters and random effects. Note that it is difficult to find out the true distribution of random effects because they are latent and unobservable variables.

Unlike linear mixed-effects models that enjoy a closed-form marginal likelihood[5], the nonlinear mixed-effects models often produce intractable marginal likelihood functions which need to be calculated using approximation methods. Davidian and Giltinan[6] classify approximation methods into two main categories: analytical approximation and numerical approximation. Analytical approximations are based on analytical manipulations to justify approximations to either the marginal likelihood or to the first two moments of the individual marginal distributions. Two commonly used analytical approximations are the first-order expansion (linearisation) and the Laplace's approximation. Numerical approximations directly approximate the integrals in the marginal likelihood by some numerical integration technique. Under the normality assumption of random effects, a routine numerical approximation is Gauss-Hermite quadrature, although it can also be used with non-normal random effects[7].

It has been shown that numerical approximations generally provide more accurate parameter estimates compared to the analytical approximations. In a recent paper Harring and Liu[8] showed, via extensive simulations, that Gauss-Hermite quadrature outperforms both the first-order expansion and the Laplace's approximation in terms of estimation accuracy. Furthermore, in their simulations with two random effects, they observed that Gauss-Hermite quadrature and the first-order expansion do not suffer from convergence issues, while the Laplace's approximation encounters convergence difficulty.

Unlike linear and generalised linear mixed models[9–13], very little is known in the literature about the impact of a misspecified random-effects distribution on inferences in nonlinear mixed-effects models. To the best of our knowledge, only two papers have studied the effects of misspecifying the random-effects distribution in nonlinear mixed-effects models. The first work,

which has been done by Hartford and Davidian [14], is based on two analytical approximations of the marginal likelihood: the first order approximation and the Laplace's approximation. But, as discussed above, these two analytical approximations do not generally provide accurate parameter estimates even if the model is fully correctly specified. Clearly, the actual impact of a misspecified random-effects distribution will be revealed if a more accurate approximation to the marginal likelihood is employed. The second work, which is a recent paper by Harring and Liu [8], mainly aimed to compare several different methods for estimation of the parameters, rather than focus on the impact of misspecifying the random-effects distribution.

After its availability in standard software in the last decade, Gauss-Hermite quadrature has become the routine method for calculation of the marginal likelihood in mixed models especially the nonlinear mixed-effects models. As already discussed, Gauss-Hermite quadrature generally provides reliable parameter estimates. In this paper, we first focus on nonlinear mixed-effects models and study the consequences of misspecifying the random-effects distribution when Gauss-Hermite quadrature is used. Our investigation not only targets the estimates of model parameters but also concerns the prediction of random effects which is a kind of 'individual inference'. The impact on individual inference has not been studied in the context of nonlinear mixed-effects models. We then describe a formal diagnostic test to check the adequacy of the assumed random-effects distribution in nonlinear mixed-effects models, which is very helpful for practical use.

2. Nonlinear mixed-effects models

In this section, we briefly explain the general form of nonlinear mixed-effects models for the analysis of longitudinal data. For a detailed discussion, see Fitzmaurice et al. [15]. Consider a longitudinal study in which N subjects are followed over time. Let $Y_{i1}, \ldots, Y_{in_i}$ be n_i repeated measurements on the ith subject, where Y_{ij} is the outcome for subject i measured at time t_{ij}. For example, Y_{ij} could be the blood pressure measured after administration of a drug. Also, let W_i denote a vector of within-subject covariates for subject i. Likewise, let X_i be a vector of between-subject covariates for subject i that do not change during the study. The nonlinear

mixed-effects model can then be expressed as a two-stage hierarchy as follows:

$$\begin{aligned} &\text{Stage 1: Individual-Level Model} \\ &\qquad Y_{ij} = m(t_{ij}, W_i, \phi_i) + \varepsilon_{ij}, \qquad j = 1, \ldots, n_i, \\ &\text{Stage 2: Population Model} \\ &\qquad \phi_i = h(X_i, \beta, b_i), \qquad\quad i = 1, \ldots, N. \end{aligned} \tag{1}$$

In stage 1, m is a nonlinear function of time t_{ij}, depending on the within-subject covariates W_i and an $r \times 1$ vector of parameters ϕ_i specific to subject i. Also, ε_{ij}'s are independent measurements errors, each of which has a normal distribution with mean 0 and variance σ^2. In stage 2, h is an r-dimensional function that describes the relationship between the elements of ϕ_i and the between-subject covariates X_i in terms of a $p \times 1$ fixed parameter β whose elements are referred to as fixed effects, and a $q \times 1$ vector b_i of random effects representing the inter-subject variability. Because ϕ_i can vary from subject to subject, the random effects b_i are incorporated to capture the inter-subject variability through a hierarchical analysis. The random effects b_i, which are unobservable variables with an unknown distribution, are typically assumed to have a multivariate normal distribution with mean 0 and a covariance matrix D whose elements are known as variance components. Note that, unlike the random effects, the random errors are assumed to be additive.

Let $\theta = \big(\beta, \sigma^2, \text{vech}(D)\big)'$ represent all unknown parameters in the nonlinear mixed-effects model (1). Unless a fully Bayesian approach is followed, the estimates of parameters are usually obtained using the maximum likelihood estimation method. Denoting $Y_i = (Y_{i1}, \ldots, Y_{in_i})'$ and assuming a multivariate normal distribution for the random effects b_i, one can write the marginal log-likelihood function for model (1) as follows

$$l(\theta) = \ln \prod_{i=1}^{N} \int_{R^q} f_i(y_i|b_i)\varphi(b_i)db_i = \ln \prod_{i=1}^{N} \int_{R^q} \left[\prod_{j=1}^{n_i} f_i(y_{ij}|b_i)\right] \varphi(b_i)db_i, \tag{2}$$

where $f_i(y_{ij}|b_i)$ denotes the conditional distribution of Y_{ij} given the random effects b_i, and $\varphi(b_i)$ is the density of the multivariate normal distribution with mean 0 and covariance matrix D. As discussed in the introduction, Gauss-Hermite quadrature is a numerical approximation that can provide an accurate approximation to the marginal log-likelihood (2) in order to facilitate maximisation with respect to θ. Obviously, the random-effects distribution is crucial in the calculation of the marginal log-likelihood function (2) and a misspecified random-effects distribution could lead to biased parameter estimates.

3. Gaussian quadrature for calculation of the marginal log-likelihood

In this section we describe Gauss-Hermite quadrature for approximating the marginal log-likelihood function (2). To avoid the numerical complexity with multiple integrals, one might change the variables of integration in (2) to independent standard normally distributed random effects z_i, as suggested by Pinheiro and Bates [16]. This can be done by the Cholesky decomposition Q of the covariance matrix D so that $b_i = Qz_i$. The marginal log-likelihood (2) can then be expressed as

$$l(\theta) = \ln \prod_{i=1}^{N} \int_{-\infty}^{\infty} \phi(z_{iq}) \dots \left\{ \int_{-\infty}^{\infty} \left[\prod_{j=1}^{n_i} f_i(y_{ij}|z_i) \right] \phi(z_{i1}) dz_{i1} \right\} \dots dz_{iq}, \tag{3}$$

where $\phi(\cdot)$ is the univariate standard normal density function. Now, Gaussian quadrature approximates each unidimensional integral in (3) as follows

$$\int_{-\infty}^{\infty} \left[\prod_{j=1}^{n_i} f_i(y_{ij}|z_i) \right] \phi(z_{i1}) dz_{i1} = \sum_{k=1}^{K} w_k \prod_{j=1}^{n_i} f_i(y_{ij}|(a_k, z_{i2}, \dots, z_{iq})), \tag{4}$$

where w_k and a_k are, respectively, the weights and abscissas (quadrature points) of the one-dimensional Gauss-Hermite quadrature rule with K points. The K weights w_k and the K abscissas a_k can be obtained from the tables of Abramowitz and Stegun [17], or can be computed as needed using an algorithm proposed by Golub [18]. Gauss-Hermite quadrature is available in standard software packages like R and SAS. Note that the more quadrature points K are used, the more accurate approximation is achieved. However, the use of a large number of quadrature points can be computationally expensive. Often 10 quadrature points should be adequate to obtain a reliable approximation [19,20].

There are two versions of Gaussian quadrature: adaptive and non-adaptive. The main difference between them is that adaptive Gaussian quadrature centres the quadrature points around the empirical Bayes estimates of random effects, while non-adaptive Gaussian quadrature centres the quadrature points around the expected value of random effects which is 0. Despite the extra burden for calculation of the empirical Bayes estimates, adaptive Gaussian quadrature generally provides more accurate results at lower numbers of quadrature points than its non-adaptive counterpart. It is therefore more common to use adaptive Gaussian quadrature for nonlinear mixed-effects models, though we consider both adaptive and non-adaptive techniques when studying the impact of misspecification of the random-effects distribution.

4. The difficulty with theoretical assessment of bias due to a misspecified random-effects distribution

This section aims to demonstrate that, for nonlinear mixed-effects models, it is difficult to obtain theoretical results on bias of maximum likelihood estimates with a misspecified random-effects distribution. Since we are interested in studying the consequences of misspecifying the random-effects distribution, we assume that the conditional distribution $f_i(y_i|b_i)$ is correctly specified. White [21] gives general theoretical results for misspecified maximum likelihood estimators, which can be used in the context of mixed-effects models by investigating the case in which $\varphi(b_i)$ is incorrectly assumed as the random-effects distribution. Let θ_0 be the true parameter value, which is unknown. Whatever the random-effects distribution is correctly specified or not, the maximum likelihood estimator $\hat{\theta}_{ML}$ converges to θ_* when $N \to \infty$ (i.e., $\hat{\theta}_{ML} \xrightarrow{P} \theta_*$), where θ_* is the minimiser of the Kullback-Leibler information criterion [21]. Equivalently, θ_* satisfies

$$\lim_N E_{f_0}\Big\{ \sum_{i=1}^{N} \frac{\partial}{\partial \theta} \ln f_i(y_i)\Big|_{\theta_*} \Big\} = 0, \tag{5}$$

where $f_i(y_i) = \int_{R^q} f_i(y_i|b_i)\varphi(b_i)db_i$, and moreover the expectation is over the true marginal distribution $f_{0i}(y_i) = \int_{R^q} f_i(y_i|b_i)g_0(b_i)db_i$ with $g_0(b_i)$ as the true random-effects distribution.

Proposition 1. *Under the general regularity conditions, if the assumed random-effects distribution is correctly specified (i.e., $g_0(b_i) = \varphi(b_i)$), then $\theta_* = \theta_0$.*

Proof. First note that the expectation in (5) can be written as

$$\begin{aligned} E_{f_0}\Big\{ \sum_{i=1}^{N} \frac{\partial}{\partial \theta} \ln f_i(y_i)\Big|_{\theta_*} \Big\} &= \sum_{i=1}^{N} \int_{R^{n_i}} \Big[\frac{\partial}{\partial \theta} \ln f_i(y_i)|_{\theta_*} \Big] f_{0i}(y_i)dy_i \\ &= \sum_{i=1}^{N} \int_{R^{n_i}} \Big[\frac{\frac{\partial}{\partial \theta} f_i(y_i)}{f_i(y_i)}|_{\theta_*} \Big] f_{0i}(y_i)dy_i. \end{aligned} \tag{6}$$

Moreover, if $g_0(b_i) = \varphi(b_i)$ then $f_{0i}(y_i) = f_i(y_i)$. Hence, we obtain, from (5) and (6), that $\theta_* = \theta_0$. □

Proposition 1 states that when the random-effects distribution is correctly specified, $\hat{\theta}_{ML}$ is consistent for θ_0. But, if the random-effects distribution is misspecified then $\theta_* - \theta_0$ is the exact asymptotic bias of $\hat{\theta}_{ML}$ for estimating θ. Therefore, for a precise assessment of the bias of $\hat{\theta}_{ML}$ in situations where the random-effects distribution is misspecified, one has to obtain θ_* from (5).

However, it is very difficult to obtain the exact value of θ_* for nonlinear mixed-effects models for two reasons. First, the true random-effects distribution $g_0(b_i)$ is unknown, and second the integral in (6) is analytically intractable for nonlinear mixed-effects models. It is even difficult to find an approximate value of θ_* from (5).

Considering the above theoretical challenge, since $\hat{\theta}_{ML}$ is consistent for θ_* (as $\hat{\theta}_{ML} \xrightarrow{P} \theta_*$), we can use $\hat{\theta}_{ML} - \theta_0$ as an approximation to the actual bias $\theta_* - \theta_0$. It is more convenient to use $(\hat{\theta}_{ML} - \theta_0)/\theta_0$ as the approximate relative bias of the maximum likelihood estimator $\hat{\theta}_{ML}$. In the next section, we conduct simulations to evaluate the approximate relative bias of $\hat{\theta}_{ML}$ in both small and large samples.

5. Simulation study

5.1. Overview

We conducted simulations to assess the impact of misspecifying the random-effects distribution on estimation of parameters and prediction of random effects in nonlinear mixed-effects models. Clearly, a simple model with one random effect is not very helpful in understanding the actual impact of such misspecification. Therefore, for our simulations we considered the following nonlinear mixed-effects model

$$Y_{ij} = \frac{D_i k_{ai} k_{ei}}{C_i(k_{ai} - k_{ei})}\Big[\exp(-k_{ei}t_{ij}) - \exp(-k_{ai}t_{ij})\Big] + \varepsilon_{ij}, \tag{7}$$

which is in fact a one-compartment pharmacokinetic model, with

$$\begin{aligned} C_i &= \exp(\beta_1 + b_{i1}), \\ k_{ai} &= \exp(\beta_2 + b_{i2}), \\ k_{ei} &= \exp(\beta_3 + b_{i3}), \end{aligned}$$

where the response Y_{ij} can be regarded as the drug concentration on subject i at time t_{ij}, D_i is the dose administered to subject i, k_{ai} is the fractional absorption rate constant for subject i, k_{ei} is the fractional elimination rate constant for subject i, and C_i is the clearance for subject i[2].

For each sample size $N = 20, 50, 100, 200, 500$ and with 10 repeated measurements per subject, we generated 500 data sets from the nonlinear mixed-effects model (7). In the simulations, we set $\beta_1 = -3.2$, $\beta_2 = 0.5$, $\beta_3 = -2.4$, and $\sigma^2 = 0.6$. Also, for simplicity in the simulations, we first assumed that the three random effects b_{i1}, b_{i2}, and b_{i3} are uncorrelated. The case of

correlated random effects will be investigated later in this section. We generated each of the three random effects b_{i1}, b_{i2}, and b_{i3} from four distinct distributions: $N(0,1)$, Chi-squared(2), Log-normal$(3,1)$, and F$(1,7)$ where F denotes the Fisher distribution. All the generated random effects were shifted and rescaled such that b_{i1}, b_{i2}, and b_{i3} have zero mean, but with variances equal to $d_1 = 0.1$, $d_2 = 0.3$, and $d_3 = 0.1$ respectively, in accordance with the real data example in Section 6. The normality assumption on random effects is valid only when the random effects are generated from $N(0,1)$, while the three other distributions used for generation of the random effects represent the cases where the random-effects distribution is misspecified since the model is fitted under the normality assumption of random effects.

5.2. The impact on estimation

First, using adaptive Gaussian quadrature we fitted the nonlinear mixed-effects model (7) to each of the generated data sets under the normality assumption of b_{i1}, b_{i2}, and b_{i3}. For each fitted model, we calculated the maximum likelihood estimator $\hat{\theta}_{ML}$ for each of the model parameters $\beta_1, \beta_2, \beta_3, \sigma^2, d_1, d_2, d_3$. Because we had 500 replications for each simulation setting which resulted in 500 estimates for each parameter, we defined the simulated relative bias of $\hat{\theta}_{ML}$ as follows

$$RB_{sim} = \frac{\hat{\theta}^*_{ML} - \theta_0}{\theta_0}, \tag{8}$$

in which $\hat{\theta}^*_{ML}$ is the mean of maximum likelihood estimates obtained from the 500 replications. Then, for each simulation setting, we computed the simulated relative bias of $\hat{\theta}_{ML}$ according to (8). The simulations results are presented in Table 1. The results show that the estimates of fixed-effects parameters and residual variance (reported as CV) are quite robust to deviations from normality of the random-effects distribution since the bias is very small, even for the small sample sizes. This result is in line with the results for generalised linear mixed models[10]. But, the estimates of variance components show a relatively large bias which does not tend to improve with the sample size. This is in agreement with the results found for generalised linear mixed models[12]. Also from the results in Table 1 we can see that, when the random-effects distribution is correctly specified, the estimates of parameters tend to have smaller relative bias as the sample size N increases, however the sample size does not help a model with misspecified random-effects distribution. A justification for this could be that the random effects are latent and unobservable variables and the data itself may not contain much information about their actual distribution.

Next, we used non-adaptive Gaussian quadrature to fit the nonlinear mixed-effects model

Table 1: The simulated relative bias multiplied by 100 (and associated standard error multiplied by 100) of the maximum likelihood estimates of parameters in the nonlinear mixed-effects model (7) using adaptive Gaussian quadrature. The model was fitted under the normality assumption of random effects, whereas the random effects were generated from four true random-effects distributions: $N(0,1)$, Chi-squared(2), Log-normal$(3,1)$, and F$(1,7)$.

True distribution	Parameter	$N=20$	$N=50$	$N=100$	$N=200$	$N=500$
	β_1	0.01 (8.1)	−0.21 (4.8)	−0.07 (3.6)	−0.05 (2.5)	0.09 (1.6)
	β_2	2.36 (12.9)	−1.68 (8.1)	0.20 (6.2)	0.83 (4.6)	−0.46 (2.5)
	β_3	0.63 (9.2)	−0.03 (5.2)	0.15 (4.4)	0.10 (2.7)	0.03 (1.9)
Normal$(0,1)$	CV	−2.13 (7.2)	1.82 (4.2)	−2.10 (3.0)	−1.79 (2.3)	−1.58 (1.4)
	d_1	−5.63 (3.6)	−4.46 (2.3)	−2.10 (1.9)	−0.03 (1.4)	−1.48 (0.7)
	d_2	−13.11 (11.3)	−3.19 (8.3)	2.77 (5.7)	0.48 (3.9)	0.32 (1.7)
	d_3	−3.65 (4.6)	3.64 (4.0)	4.76 (3.1)	2.15 (2.5)	3.85 (1.6)
	β_1	0.85 (6.8)	1.37 (5.3)	1.55 (3.3)	1.63 (2.1)	1.62 (1.4)
	β_2	−4.88 (13.3)	−2.34 (8.5)	−1.51 (5.9)	−1.46 (3.8)	−2.30 (2.4)
	β_3	0.97 (8.7)	2.61 (6.2)	2.58 (4.6)	2.47 (2.7)	2.47 (1.8)
Chi-squared(2)	CV	3.08 (7.2)	2.88 (4.8)	3.13 (3.5)	3.60 (2.3)	2.75 (2.2)
	d_1	−47.84 (4.0)	−55.92 (3.0)	−61.19 (1.9)	−63.28 (1.1)	−63.32 (7.8)
	d_2	−23.06 (16.2)	−18.20 (11.9)	−20.64 (5.8)	−19.89 (4.9)	−19.47 (2.6)
	d_3	31.50 (12.2)	61.59 (10.2)	81.77 (8.7)	88.98 (5.9)	86.75 (3.8)
	β_1	0.71 (7.3)	1.12 (4.4)	1.52 (2.9)	1.56 (2.2)	1.53 (1.3)
	β_2	−3.85 (10.8)	−0.99 (7.8)	−3.44 (5.6)	−3.39 (3.2)	−3.13 (2.7)
	β_3	1.40 (9.1)	1.83 (5.4)	2.32 (4.0)	2.24 (3.3)	2.38 (2.0)
Log-normal$(3,1)$	CV	3.43 (7.9)	3.05 (5.1)	3.96 (3.1)	3.72 (2.2)	3.09 (2.5)
	d_1	−56.21 (5.1)	−67.68 (2.6)	−73.22 (1.3)	−74.28 (1.0)	−74.14 (0.6)
	d_2	−37.94 (18.3)	−38.05 (7.4)	−42.02 (5.8)	−42.49 (3.6)	−41.45 (2.7)
	d_3	18.93 (11.6)	43.09 (9.5)	57.36 (8.2)	74.39 (5.5)	61.72 (4.9)
	β_1	0.94 (6.2)	1.20 (3.6)	1.40 (3.1)	1.48 (2.0)	1.55 (1.1)
	β_2	−4.52 (9.4)	−2.39 (6.8)	−3.34 (4.7)	−2.55 (3.3)	−2.22 (2.0)
	β_3	1.84 (9.0)	1.77 (6.2)	2.14 (3.9)	2.23 (3.1)	2.43 (1.8)
F$(1,7)$	CV	4.18 (7.9)	3.51 (4.6)	3.86 (3.5)	3.80 (2.4)	3.77 (2.6)
	d_1	−69.75 (4.1)	−80.45 (1.7)	−80.75 (1.4)	−82.24 (1.0)	−84.18 (0.4)
	d_2	−54.96 (11.4)	−46.03 (9.0)	−48.92 (5.7)	−48.70 (4.1)	−48.64 (2.3)
	d_3	19.14 (15.0)	49.95 (11.7)	59.56 (9.8)	46.79 (5.3)	60.29 (3.5)

(7) to the generated data sets. For this case, we repeated the above calculations to obtain the simulated relative bias of $\hat{\theta}_{ML}$ using (8). The simulations results, which are reported in Table 2, are very similar to those from the adaptive Gaussian quadrature, and the only difference is that the estimates of variance components show little bias under the correct random-effects distribution. This is probably due to the fact that non-adaptive Gaussian quadrature centres the quadrature points around the expected value of 0 instead of the empirical Bayes estimates of random effects.

To understand the role of correlated random effects and the number of repeated measurements on parameter estimation under a misspecified random-effects distribution, we conducted a similar simulation study for the nonlinear mixed-effects model (7), where we used the number of repeated

Table 2: The simulated relative bias multiplied by 100 (and associated standard error multiplied by 100) of the maximum likelihood estimates of parameters in the nonlinear mixed-effects model (7) using non-adaptive Gaussian quadrature. The model was fitted under the normality assumption of random effects, whereas the random effects were generated from four true random-effects distributions: $N(0,1)$, Chi-squared(2), Log-normal$(3,1)$, and F$(1,7)$.

True distribution	Parameter	$N=20$	$N=50$	$N=100$	$N=200$	$N=500$
	β_1	−0.16 (8.2)	0.23 (6.7)	0.08 (5.4)	0.01 (3.7)	0.17 (2.9)
	β_2	−2.23 (11.7)	0.18 (11.9)	1.85 (7.3)	0.05 (5.5)	0.13 (4.1)
	β_3	0.23 (8.3)	−1.12 (8.2)	−1.02 (6.7)	−0.55 (3.8)	−0.70 (3.0)
Normal$(0,1)$	CV	−5.20 (6.4)	7.11 (4.6)	11.99 (6.1)	5.02 (2.9)	6.45 (3.6)
	d_1	−0.65 (3.3)	−6.06 (2.8)	0.67 (3.3)	0.73 (1.6)	1.07 (1.5)
	d_2	8.46 (11.3)	−6.24 (8.0)	−21.05 (4.6)	−10.42 (4.1)	−11.72 (3.2)
	d_3	−7.53 (4.1)	−31.16 (2.4)	−28.21 (2.5)	−19.84 (1.7)	−21.43 (1.1)
	β_1	1.13 (8.1)	1.29 (4.9)	1.37 (4.8)	1.65 (3.2)	1.74 (2.8)
	β_2	1.32 (12.7)	−3.28 (9.0)	−6.80 (7.5)	−3.95 (6.5)	−1.28 (4.7)
	β_3	1.36 (7.9)	0.08 (8.5)	0.77 (6.1)	0.52 (5.7)	0.20 (9.3)
Chi-squared(2)	CV	−5.21 (6.6)	7.33 (4.5)	10.72 (5.9)	12.14 (6.3)	13.76 (7.4)
	d_1	−14.81 (5.8)	−21.73 (5.3)	−31.48 (5.5)	−48.14 (3.6)	−47.59 (4.1)
	d_2	−14.70 (10.9)	−21.82 (9.6)	−25.92 (7.9)	−33.78 (5.0)	−38.87 (3.7)
	d_3	−16.13 (4.8)	0.05 (4.9)	8.45 (3.7)	17.23 (5.1)	38.97 (7.6)
	β_1	−1.59 (5.2)	0.86 (6.5)	1.13 (4.8)	0.97 (4.3)	1.93 (2.9)
	β_2	4.12 (13.0)	−7.46 (8.2)	−7.99 (8.9)	−7.97 (7.2)	4.10 (7.5)
	β_3	0.69 (6.9)	−0.38 (8.6)	−0.05 (11.3)	−1.42 (13.2)	−2.52 (10.7)
Log-normal$(3,1)$	CV	−4.95 (6.3)	7.21 (6.1)	9.66 (5.9)	10.79 (6.8)	13.81 (7.2)
	d_1	−37.90 (5.4)	−32.74 (6.7)	−61.97 (2.9)	−68.05 (2.7)	−62.75 (0.7)
	d_2	−15.94 (2.0)	−42.36 (8.1)	−45.29 (7.0)	−46.41 (6.5)	−45.61 (4.4)
	d_3	−24.14 (5.2)	6.72 (8.8)	31.17 (9.7)	72.53 (6.6)	57.66 (11.4)
	β_1	0.69 (7.5)	1.13 (5.4)	0.86 (4.0)	0.80 (3.8)	1.06 (2.6)
	β_2	−6.68 (12.3)	−3.32 (8.1)	−0.24 (7.7)	−12.71 (5.9)	−4.65 (5.8)
	β_3	0.36 (10.8)	0.42 (8.7)	−0.90 (11.9)	−1.22 (9.8)	−5.09 (13.8)
F$(1,7)$	CV	−4.09 (5.9)	6.70 (4.7)	9.51 (5.4)	9.94 (6.5)	12.87 (7.1)
	d_1	−36.07 (9.3)	−56.04 (3.8)	−60.66 (3.9)	−83.04 (10.8)	−85.21 (0.6)
	d_2	−50.74 (10.5)	−50.96 (7.1)	−44.26 (7.2)	−59.14 (3.2)	−50.49 (5.9)
	d_3	−0.15 (8.0)	−0.98 (8.2)	53.19 (10.2)	17.45 (5.9)	63.37 (16.8)

measurements $n_i = 5, 10, 20$ and fixed $N = 30$. Here the random effects were generated from two multivariate distributions: a multivariate normal distribution and a multivariate log-normal distribution. The simulation results, presented in Table 3, suggest that increasing the number of repeated measurements would help the estimates of the fixed-effects parameters, however it does not improve the estimates of variance components when the random-effects distribution is misspecified. Furthermore, similar to the previous results in Table 1 and Table 2, when the random effects are correlated the estimates of variance components show high relative bias under a misspecified random-effects distribution.

Table 3: The simulated relative bias multiplied by 100 (and associated standard error multiplied by 100) of the maximum likelihood estimates of parameters in the nonlinear mixed-effects model (7) with $N = 30$ and with the number of repeated measurements $n_i = 5, 10, 20$, using adaptive Gaussian quadrature for the two multivariate random-effects distributions considered.

True distribution	Parameter	$n_i = 5$	$n_i = 10$	$n_i = 20$
Multivariate normal	β_1	−0.69 (7.9)	−0.11 (6.2)	−0.06 (3.2)
	β_2	−6.60 (10.1)	−3.01 (8.7)	0.52 (4.1)
	β_3	−1.19 (8.2)	0.17 (5.9)	−0.17 (3.3)
	CV	−2.48 (7.6)	2.91 (5.4)	2.95 (3.5)
	d_{11}	−7.31 (6.1)	−5.56 (4.6)	−3.60 (2.8)
	d_{12}	8.08 (5.9)	−4.25 (4.3)	−2.46 (3.0)
	d_{22}	10.10 (6.2)	−5.12 (4.8)	−3.91 (3.1)
	d_{13}	−8.19 (5.3)	−5.15 (4.6)	2.52 (2.9)
	d_{23}	−9.28 (4.9)	−6.24 (3.9)	1.03 (3.1)
	d_{33}	−6.23 (5.9)	−4.40 (4.3)	−2.10 (2.9)
Multivariate log-normal	β_1	−1.22 (8.3)	−1.20 (6.2)	0.73 (4.0)
	β_2	−12.57 (9.8)	−2.36 (8.5)	−1.14 (3.7)
	β_3	−5.09 (10.2)	1.38 (6.3)	−1.15 (3.4)
	CV	4.16 (8.1)	3.02 (6.2)	2.80 (3.6)
	d_{11}	−49.11 (7.9)	−46.23 (5.5)	−42.11 (4.2)
	d_{12}	−21.24 (8.0)	−18.16 (5.3)	−16.31 (3.7)
	d_{22}	−36.77 (7.7)	−35.13 (5.3)	−34.85 (4.0)
	d_{13}	−20.06 (6.1)	−16.35 (5.0)	−17.11 (3.5)
	d_{23}	23.01 (5.8)	24.19 (4.8)	19.85 (2.9)
	d_{33}	−31.16 (7.3)	−29.98 (5.0)	−28.79 (3.3)

5.3. The impact on prediction

We then investigated how misspecification of the random-effects distribution affects the prediction of random effects in the nonlinear mixed-effects model (7). Our approach to assessing the impact on predictions is based on the comparison of the predicted values of random effects obtained under the true distribution with those obtained under a misspecified random-effects distribution. This approach makes sense because the true model is believed to produce correct predictions, especially when the sample size is large enough ($N \geq 200$), and our simulation results in Table 1 and Table 2 confirm that the model fitting is generally accurate under the true random-effects distribution. Recall that in our simulations $N(0, 1)$ represents the case where the random-effects distribution is correctly specified, while Chi-squared(2), Log-normal(3, 1), and F(1, 7) are used for the cases where the random-effects distribution is misspecified as the model is fitted under the normality assumption of random effects. Figure 1 shows the predictions of random effects for two sample sizes of 200 and 500 obtained for these four distributions using adaptive Gaussian quadrature. It can be seen that the misspecified random-effects distributions tend to overestimate the predictions (see Figure 1(c,d,e,f)) or underestimate the predictions (see Figure 1(a,b)). In fact, the predictions

of random effects are influenced with the shape of the random-effects distribution. A similar behaviour is observed for the predictions obtained using non-adaptive Gaussian quadrature (see Figure 2).

6. A diagnostic test for checking the random-effects distribution

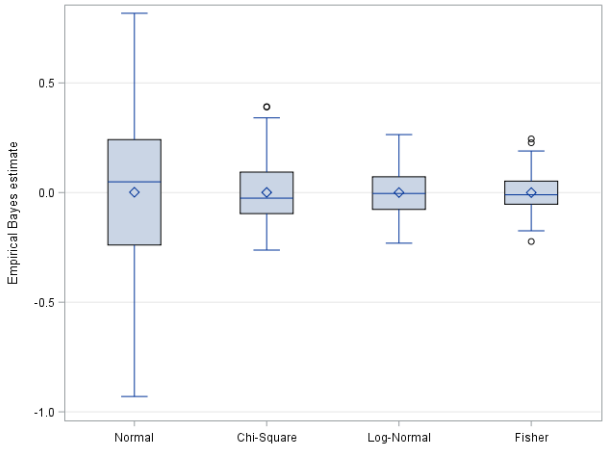
In the previous sections we studied the impact of misspecifying the random-effects distribution on estimation and prediction. We now describe a formal diagnostic test to verify whether an assumed random-effects distribution is correctly specified or not.

Let G be the assumed random-effects distribution, which is typically a multivariate normal distribution (i.e., $G = \varphi$). To check the appropriateness of the assumed random-effects distribution G, Verbeke and Molenberghs [22] suggested to use the so-called gradient function given by

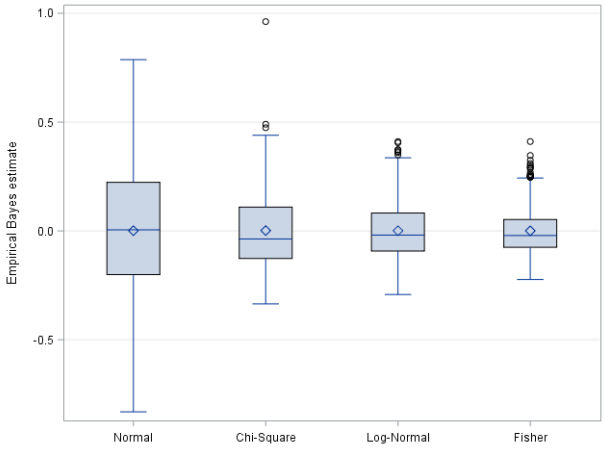
$$\Delta(G, b) = \frac{1}{N} \sum_{i=1}^{N} \frac{f_i(y_i|b)}{f_i(y_i|G)}, \qquad b \in R^q, \tag{9}$$

where $f_i(y_i|b)$ and $f_i(y_i|G)$ are, respectively, the conditional (given random-effect point b) and marginal distributions of Y_i. They proved that if the random-effects distribution G is correctly specified, then $\Delta(G, b) \leq 1$ for all $b \in R^q$, and furthermore $\Delta(G, b) = 1$ for all b in the support of G. Therefore, deviations of the gradient function from 1 in the support points of G indicate inadequacy of G. As an informal approach, they suggested to plot the gradient function versus points b in the support of G, and if the gradient plot is close to 1 then the adequacy of G is confirmed. Note that the gradient can be interpreted as an average of likelihood ratios, each ratio measuring how much more likely Y_i is to be observed for subject i if the corresponding random effect b_i equals b rather than being sampled from the distribution G. It can be seen that the calculation of the gradient function is easy because it only requires the calculation of the marginal and conditional distributions for all N subjects.

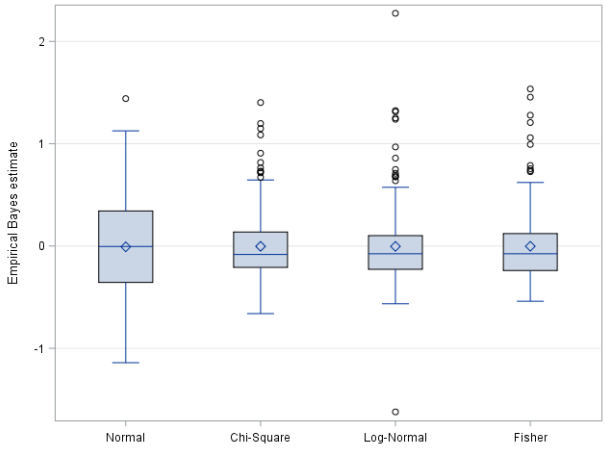
Based on the gradient function (9), we recently developed a formal diagnostic test for the random-effects distribution (see Drikvandi et al. [4]). Suppose that the null hypothesis H_0 says the random-effects distribution G is correctly specified and the alternative hypothesis H_1 says otherwise. Having considered all deviations of the gradient function from 1, we constructed a



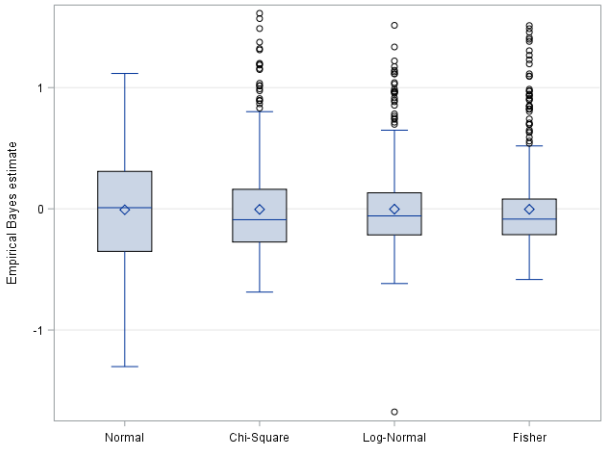
(a) Prediction of b_{i1} when $N = 200$.



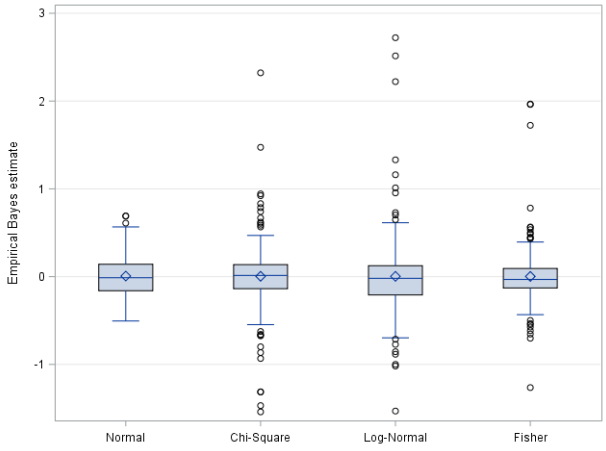
(b) Prediction of b_{i1} when $N = 500$.



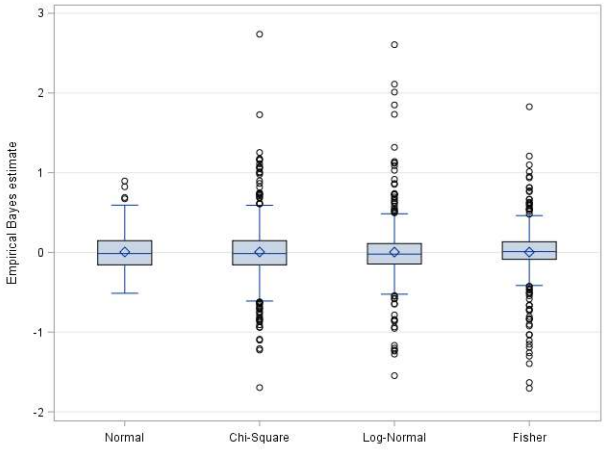
(c) Prediction of b_{i2} when $N = 200$.



(d) Prediction of b_{i2} when $N = 500$.

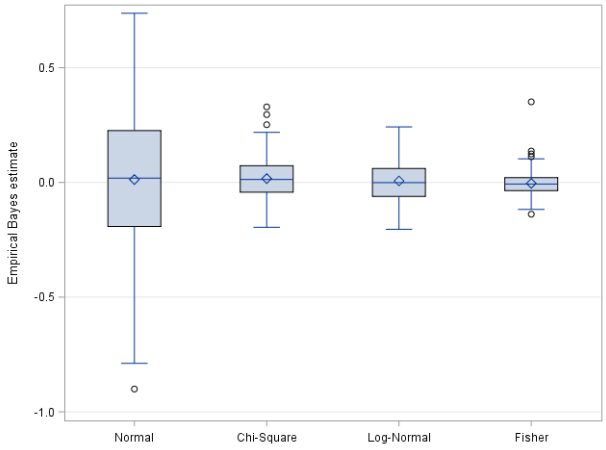


(e) Prediction of b_{i3} when $N = 200$.

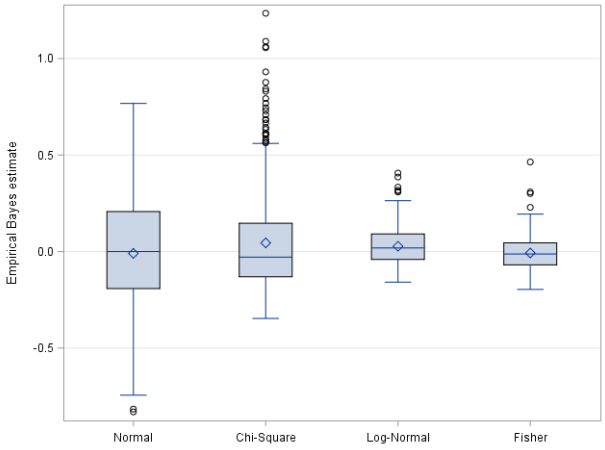


(f) Prediction of b_{i3} when $N = 500$.

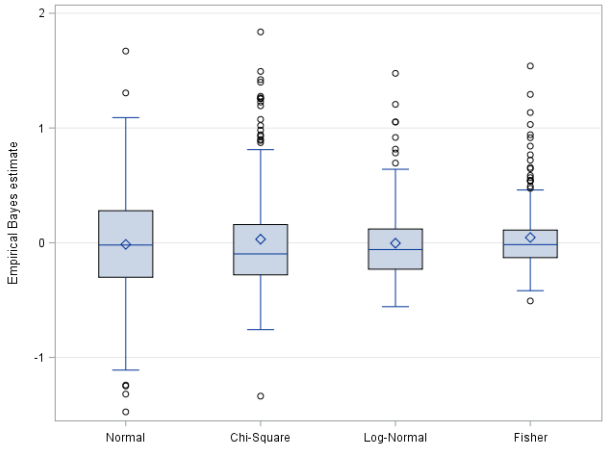
Figure 1: The prediction of random effects in the nonlinear mixed-effects model (7) using adaptive Gaussian quadrature. The model was fitted under the normality assumption of random effects, whereas the random effects were generated from the four true random-effects distributions considered.



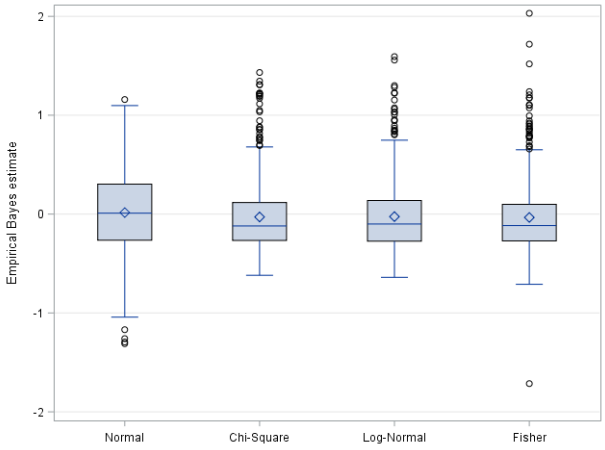
(a) Prediction of b_{i1} when $N = 200$.



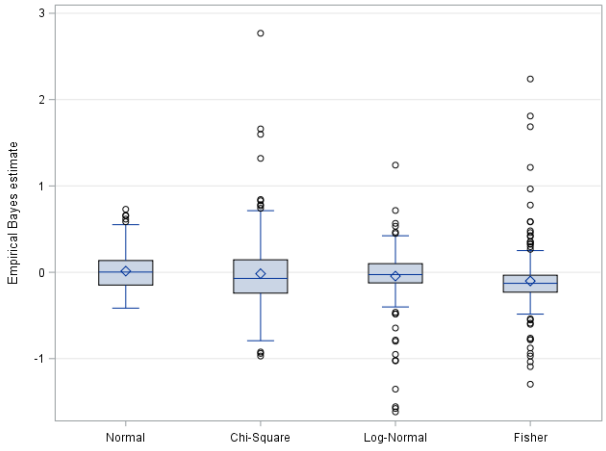
(b) Prediction of b_{i1} when $N = 500$.



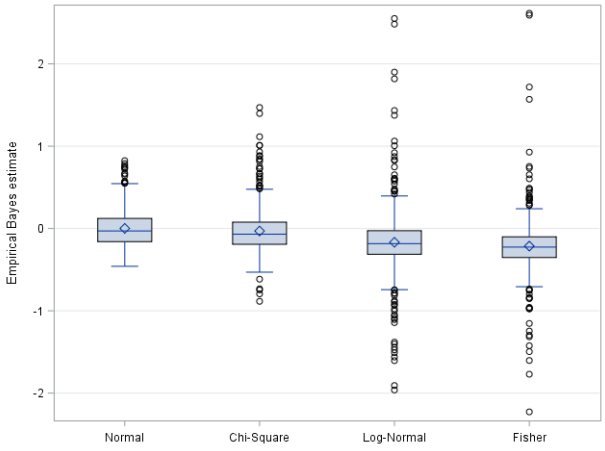
(c) Prediction of b_{i2} when $N = 200$.



(d) Prediction of b_{i2} when $N = 500$.



(e) Prediction of b_{i3} when $N = 200$.



(f) Prediction of b_{i3} when $N = 500$.

Figure 2: The prediction of random effects in the nonlinear mixed-effects model (7) using non-adaptive Gaussian quadrature. The model was fitted under the normality assumption of random effects, whereas the random effects were generated from the four true random-effects distributions considered.

test statistic for testing H_0 versus H_1 as follows

$$T = \int_{R^q} \left(\hat{\Delta}(\hat{G}, b) - 1\right)^2 d\hat{G}(b), \tag{10}$$

where $\hat{G}$ is the estimated random-effects distribution obtained by replacing the covariance matrix D by its maximum likelihood estimate, and $\hat{\Delta}$ denotes the estimated gradient function based on $\hat{G}$ obtained simply by replacing the unknown parameters in $f_i(y_i|b)$ and $f_i(y_i|\hat{G})$ by their maximum likelihood estimates. If T deviates much from 0, we can reject H_0 implying that the assumed random-effects distribution G is not appropriate for random effects.

The asymptotic distribution of T is given in Theorem 1 of Drikvandi et al. [4], which is essentially a weighted sum of independent chi-squared distributions each with one degree of freedom. However, the asymptotic distribution should only be used when the sample size N is sufficiently large. For small-sample situations, we also proposed a parametric bootstrap procedure to obtain the finite-sample distribution of the test statistic T in (10). The key step in our bootstrap procedure, in order to obtain a bootstrap sample, is to first generate random effects b_i^s, $i = 1, \ldots, N$, from $\hat{G}$ and then generate a bootstrap sample Y_i^s, $i = 1, \ldots, N$, from $\hat{f}_i(y_i|b_i^s)$. We use 200 bootstrap samples to conduct the bootstrap test. Below we illustrate how the bootstrap test based on (10) can be performed.

Implementation of the bootstrap test

The bootstrap test can be carried out by the following steps:

1. Generate K, say 1000, random-effect points b_k from $\hat{G}$.
2. Compute the gradient function (9) and its squared deviation from 1 for each b_k.
3. Calculate the test statistic T being the average of the K squared deviations obtained in step 2 (which is the Monte Carlo approximation of T), and denote it by T_{obs}.
4. For each bootstrap step s, $s = 1, \ldots, 200$, repeat the following two steps:
   i. First generate random effects b_i^s, $i = 1, \ldots, N$, from $\hat{G}$ and then generate a bootstrap sample Y_i^s, $i = 1, \ldots, N$, from $\hat{f}_i(y_i|b_i^s)$.
   ii. Calculate the test statistic T for the bootstrap sample obtained in step i and denote it by T^s.
5. If the proportion of T^s exceeding T_{obs} is less than 0.05, then reject H_0, indicating that the assumed random-effects distribution G is not appropriate for random effects.

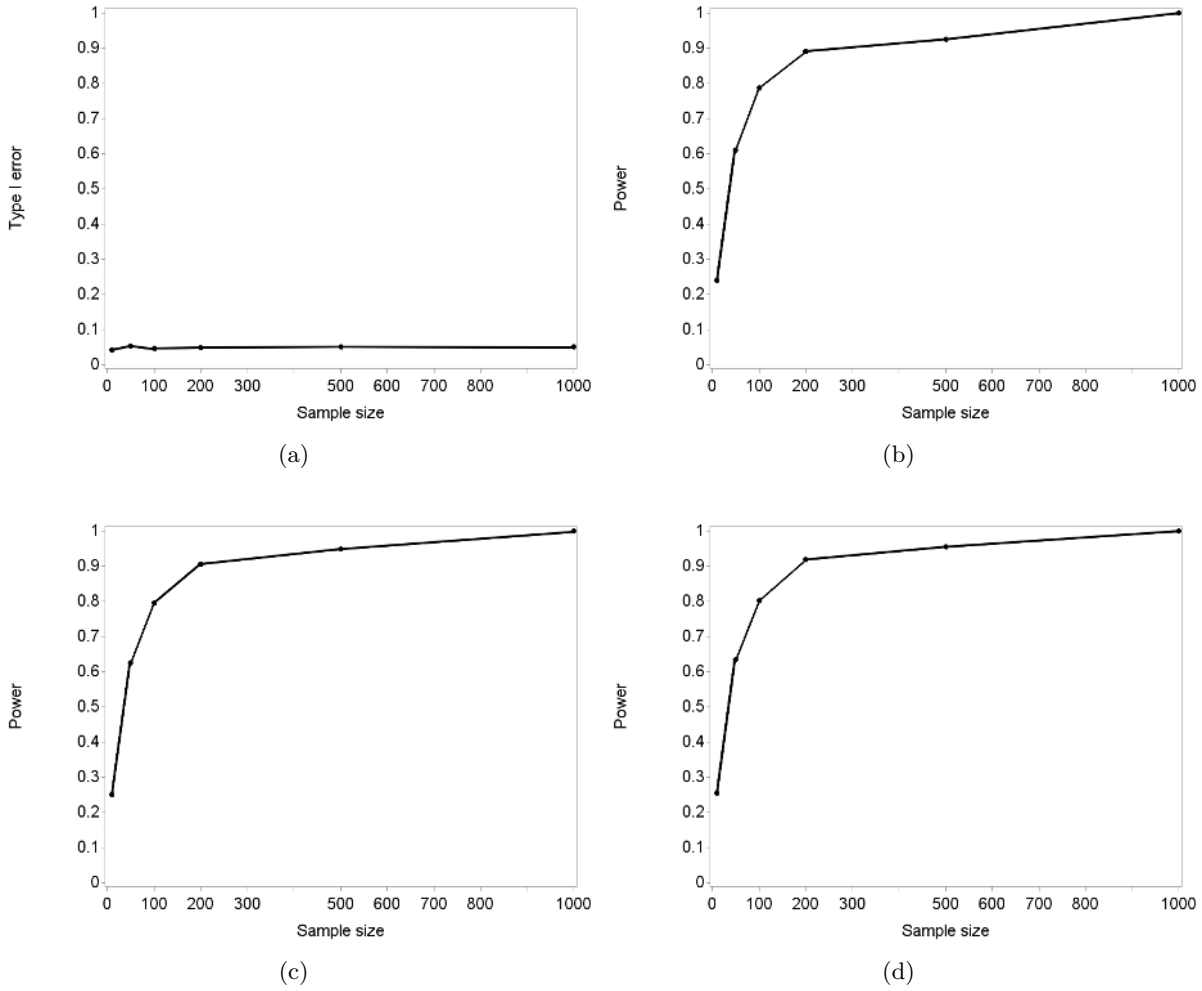


Figure 3: Type I error and power of the bootstrap test at the significance level 0.05 for detecting the misspecification of random-effects distribution in the nonlinear mixed-effects model (7) using adaptive Gaussian quadrature, with the four true random-effects distributions considered: (a) Normal$(0, 1)$, (b) Chi-squared(2), (c) Log-normal$(3, 1)$, and (d) F$(1, 7)$. Note that the Type I error is for the case of Normal$(0, 1)$ where the random-effects distribution is correctly specified.

We evaluated the empirical performance of this bootstrap test for a general class of mixed-effects models in Drikvandi et al.[4]. Here we calculated the power and Type I error of the permutation test at the 5% significance level for the nonlinear mixed-effects model (7), under the same simulation setting as in Section 5 and with the additional sample size of $N = 1000$. The results, presented in Figure 3, indicate that the bootstrap test has a Type I error close to the nominal level 0.05 and further it shows a high power for detecting the misspecification of random-effects distribution, especially when the sample size N is sufficiently large (e.g., $N \geq 100$). We prepared a SAS code for implementation of the above test which is available online on the journal website. We will apply this diagnostic test to our real data application in the next section.

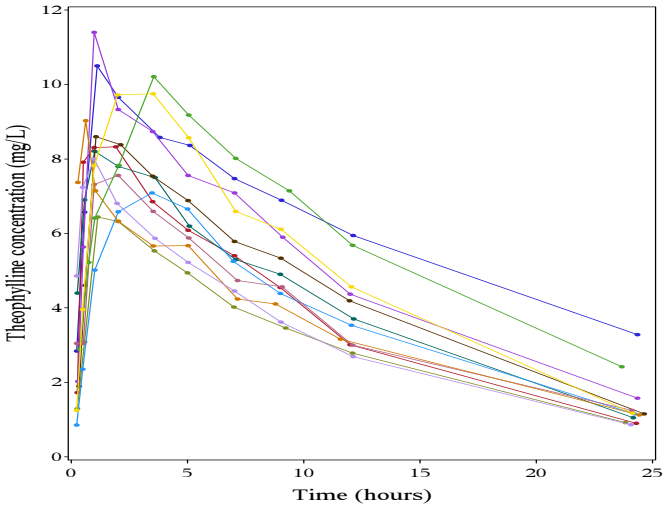


Figure 4: Theophylline data: individual profiles for 12 subjects.

7. Real data application

Theophylline is a well-known anti-asthmatic agent, administered orally[23,24]. In a pharmacokinetic study, 12 subjects were given the same oral dose (mg/kg) of theophylline, and blood samples were taken at several times following administration were assayed for theophylline concentration[24]. The main objective was to gain insight into within-subject pharmacokinetic processes of absorption, distribution, and elimination governing concentrations of drug achieved. The individual profiles, presented in Figure 4, show that the theophylline concentrations have a similar shape for all subjects, but peak concentration achieved, rise, and decay vary significantly across the subjects. These differences are due to the inter-subject variability in the underlying pharmacokinetic processes, understanding of which is critical for developing dosing guidelines.

To characterise these processes formally, we consider the one-compartment model (7) with first-order absorption and elimination, as also suggested by Davidian and Giltinan[24]. The inter-subject variability in the pharmacokinetic processes is accounted for by the subject-specific random effects (b_{i1}, b_{i2}, b_{i3}), which are assumed to have mean 0 and covariance matrix given by $D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{12} & d_{22} & d_{23} \\ d_{13} & d_{23} & d_{33} \end{bmatrix}$. We fit the one-compartment model (7) to the theophylline data, by assuming two different multivariate distributions for the random effects (b_{i1}, b_{i2}, b_{i3}): a multivariate normal distribution and a multivariate log-normal distribution. The maximum likelihood estimates of parameters along with their associated standard errors obtained under each distribution are reported in Table 4. It can be seen that the estimates of fixed-effects parameters are similar

Table 4: Theophylline data: the maximum likelihood estimates of parameters and associated standard errors obtained from fitting the one-compartment model (7), once assuming a multivariate normal distribution for the random effects, once assuming a multivariate log-normal distribution for the random effects. Note that the test statistics and p-values are reported only for the multivariate normal distribution.

	Multivariate normal			Multivariate log-normal
Parameter	Estimate (s.e.)	t-value	p-value	Estimate (s.e.)
Fixed effects:				
β_1	-3.277 (0.046)	-70.64	< 0.0001	-3.053 (0.188)
β_2	0.537 (0.063)	8.52	< 0.0001	0.701 (0.380)
β_3	-2.454 (0.064)	-38.38	< 0.0001	-2.395 (0.079)
Residual variance:				
σ^2	0.624 (0.083)	7.55	< 0.0001	0.542 (0.080)
Variance components:				
d_{11}	0.057 (0.022)	2.56	0.0308	0.631 (0.461)
d_{12}	-0.012 (0.018)	-0.67	.5188	-0.018 (0.027)
d_{22}	0.264 (0.054)	4.92	0.0008	2.215 (1.852)
d_{13}	0.030 (0.020)	1.52	0.1636	0.039 (0.025)
d_{23}	-0.025 (0.017)	-1.47	0.1743	-0.043 (0.034)
d_{33}	0.035 (0.017)	2.05	0.0702	0.090 (0.056)
-2 log-likelihood		341.7		358.6

under the two assumed random-effects distributions, but the estimates of variance components are quite different. These results are consistent with our simulation findings in Section 5. Note that the test statistics and p-values are reported only for the multivariate normal distribution, and one should be cautious about the p-values for variance components because their associated tests require testing on the boundary of the parameter space[25].

Since the true random-effects distribution is unknown, it is not clear which parameter estimates in Table 4 are correct, though the multivariate normal distribution provides a larger marginal likelihood. To check this formally, we apply the diagnostic bootstrap test in Section 6 to see whether a multivariate normal distribution is appropriate for the random effects. The bootstrap test with 200 bootstrap samples and with 1000 Monte Carlo integration nodes produces a test statistic of 3.91, giving a p-value of 0.15. This suggests that a multivariate normal distribution is appropriate for the random effects b_{i1}, b_{i2} and b_{i3}.

Pharmacokineticists are more interested in the estimates of pharmacokinetic parameters C_i, k_{ai}, and k_{ai}. Figure 5 shows that the estimates of pharmacokinetic parameters for the theophylline data are roughly similar between the two different random-effects distributions. It is because the estimates of pharmacokinetic parameters here are more affected by the estimates of fixed-effects parameters rather than the predictions of random effects.

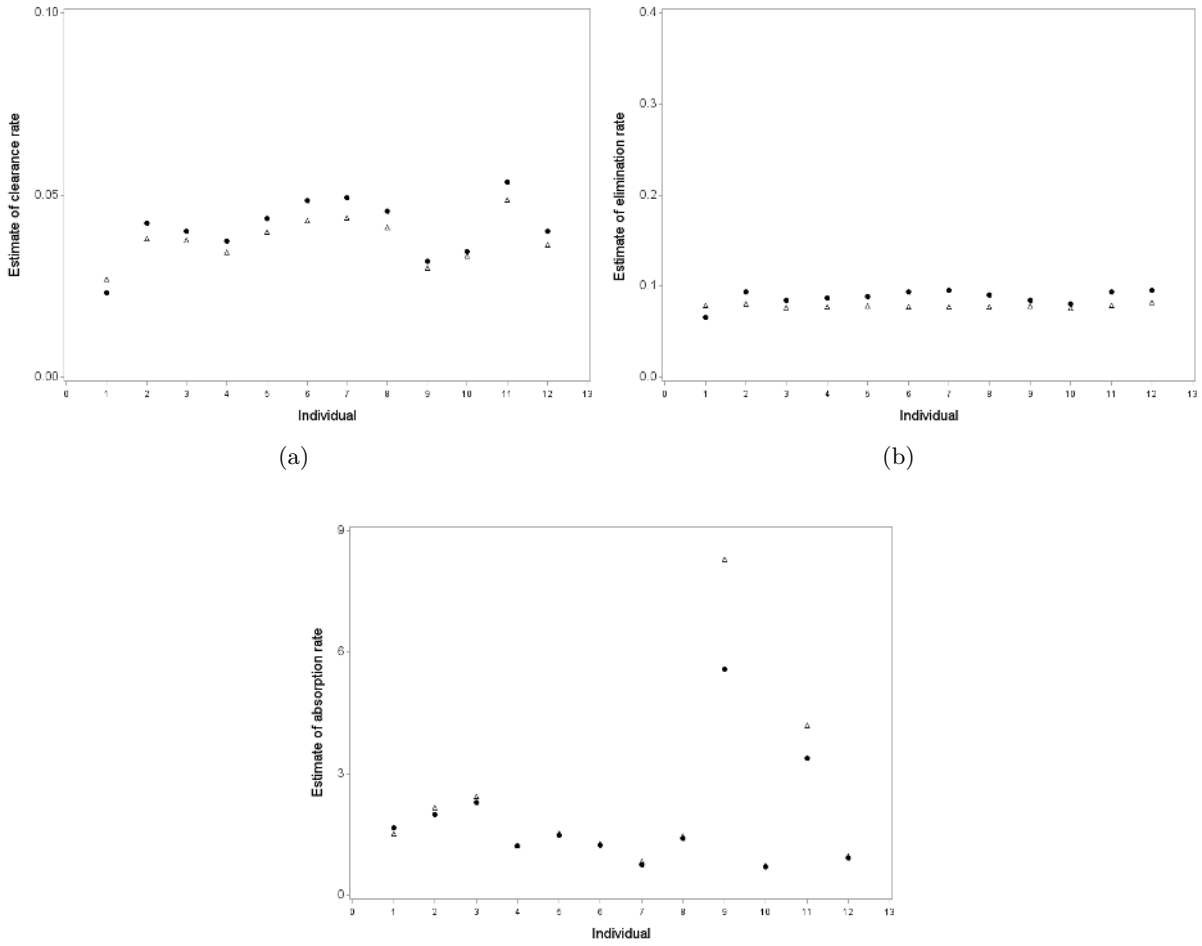


Figure 5: Theophylline data: the estimates of pharmacokinetic parameters for the 12 subjects. In each plot, the filled dot symbol is used to show the estimates obtained under the multivariate normal distribution for random effects, while the triangle symbol is used to show the estimates obtained the multivariate log-normal distribution for random effects.

8. Conclusions and discussion

Since random effects are latent and unobservable variables, it is difficult to find out their true underlying distribution. Consequently, in practice, the random-effects distribution can be subject to misspecification. We presented a formal diagnostic test to check the appropriateness of the assumed random-effects distribution in nonlinear mixed-effects models. Such a diagnostic tool is very useful for practical use.

In the paper, we focused on Gauss-Hermite quadrature which is now the default method in standard software (**nlmer** function in *R* and **PROC NLMIXED** in *SAS*), and generally provides

reliable approximation to the marginal likelihood in nonlinear mixed-effects models. This enabled us to obtain fair results on consequences of misspecifying the random-effects distribution. The theory suggests that the maximum likelihood estimates of parameters in nonlinear mixed-effects models with a misspecified random-effects distribution are not asymptotically unbiased. The asymptotic bias is equal to $\theta_* - \theta_0$ where θ_* is the minimiser of the Kullback-Leibler information criterion and θ_0 is the true parameter value. Because calculation of θ_* is difficult, we used the maximum likelihood estimator $\hat{\theta}_{ML}$ as a consistent estimator of θ_* for evaluating the bias in our simulations.

The main findings/conclusions of our simulation studies and real data analysis are summarised below. We should emphasise that these results essentially apply to the model and range of scenarios considered in the simulations and real data application.

- The maximum likelihood estimates of fixed-effects parameters are robust to the normality assumption of random effects.
- The maximum likelihood estimate of residual variance is also robust to the normality assumption of random effects.
- The maximum likelihood estimates of variance components show substantial bias when the random-effects distribution is misspecified.
- Predictions of random effects are highly affected with a misspecified random-effects distribution and, in fact, they are either overestimated or underestimated.
- The sample size N and the number of repeated measurements n_i do not help much with the relative bias of the maximum likelihood estimates of variance components when the random-effects distribution is misspecified. A possible reason is that the random-effects distribution is a mixing distribution, and mainly the marginal distribution will benefit from increasing the sample size.
- All the above conclusions hold for both adaptive and non-adaptive Gaussian quadrature. The only difference between the results from these two methods was that the estimates of variance components obtained using non-adaptive Gaussian quadrature show little bias under the correct random-effects distribution. This is due to the fact that non-adaptive Gaussian quadrature centres the quadrature points around the expected value of 0 instead of the empirical Bayes estimates of random effects.

In our simulations, the variances of the random effects ($d_1 = 0.1$, $d_2 = 0.3$, and $d_3 = 0.1$) were very small and we chose them based on the case study. In another simulation study (not reported

in the paper) with the same settings but with larger variance components, we found similar results for the fixed-effects parameters, but the impact of the mispecification on the estimates of variance components and predictions of random effects was more substantial when the variance components are larger, that is, the relative bias of the estimates of variance components tends to be higher and the predictions of random effects tend to get farther from the true predictions.

The fact that the estimates of fixed-effects parameters (often the main parameters of interest) are robust to departures from normality of the random-effects distribution should not be a reason to depreciate the distributional assumptions on random effects because, as our simulations revealed, the estimates of variance components and predictions of random effects are highly affected with such misspecification. Clearly, variance components are crucial in calculation of the standard errors of the fixed-effects parameters estimates, hence wrong estimates of them could affect the confidence intervals and hypothesis tests regarding the fixed-effects parameters (see[26,27]). Moreover, individual-specific inferences could be misleading when the predictions of random effects are inaccurate.

Last but not least, the more random effects, the more challenging their distribution would be. Therefore, it is important to avoid any unnecessary random effects in the model. This can be done, for example, via a test for zero random effects. Classical tests such as the likelihood ratio and score tests cannot be easily applied to this testing problem because it requires testing on the boundary of parameter space. Bootstrap and permutation tests have been suggested for testing random effects in linear and generalised linear mixed models[25,28,29], however little is known on testing random effects in nonlinear mixed-effects models. Developing such tests for random effects in nonlinear mixed-effects models will be very helpful.

9. Data availability statement

The data that support the findings of this paper are available from the author upon request.

References

[1] Dubois, A., Gsteiger, S., Balser, S., Pigeolet, E., Steimer, J. L., Pillai, G., and Mentré, F. Pharmacokinetic similarity of biologics: Analysis using nonlinear mixed-effects modeling. *Clinical Pharmacology & Therapeutics*, 91:234–242, 2012.

[2] Drikvandi, R. Nonlinear mixed-effects models for pharmacokinetic data analysis: assessment

of the random-effects distribution. *Journal of Pharmacokinetics and Pharmacodynamics*, 44: 223–232, 2017.

[3] Efendi, A., Drikvandi, R., Verbeke, G., and Molenberghs, G. A goodness-of-fit test for the random-effects distribution in mixed models. *Statistical Methods in Medical Research*, 26: 970–983, 2017.

[4] Drikvandi, R., Verbeke, G., and Molenberghs, G. Diagnosing misspecification of the random-effects distribution in mixed models. *Biometrics*, 73:63–71, 2017.

[5] Verbeke, G. and Molenberghs, G. *Linear mixed models for longitudinal data*. New York: Springer Verlag, 2009.

[6] Davidian, M. and Giltinan, D. M. Some general estimation methods for nonlinear mixed-effects model. *Journal of Biopharmaceutical Statistics*, 3:23–55, 1993.

[7] Liu, L. and Yu, Z. A likelihood reformulation method in non-normal random effects models. *Statistics in Medicine*, 27:3105–3124, 2008.

[8] Harring, J. R. and Liu, J. A comparison of estimation methods for nonlinear mixed-effects models under model misspecification and data sparseness: A simulation study. *Journal of Modern Applied Statistical Methods*, 15:539–569, 2016.

[9] Verbeke, G. and Lesaffre, E. The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data. *Computational Statistics & Data Analysis*, 23: 541–556, 1997.

[10] Heagerty, P. J. and Kurland, B. F. Misspecified maximum likelihood estimates and generalised linear mixed models. *Biometrika*, 88:973–985, 2001.

[11] Agresti, A., Caffo, B., and Ohman-Strickland, P. Examples in which misspecification of a random effects distribution reduces efficiency, and possible remedies. *Computational Statistics & Data Analysis*, 47:639–653, 2004.

[12] Litière, S., Alonso, A., and Molenberghs, G. The impact of a misspecified random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models. *Statistics in Medicine*, 27:3125–3144, 2008.

[13] Alonso, A., Litière, S., and Laenen, A. A note on the indeterminacy of the random-effects distribution in hierarchical models. *The American Statistician*, 64:318–324, 2010.

[14] Hartford, A. and Davidian, M. Consequences of misspecifying assumptions in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, 34:139–164, 2000.

[15] Fitzmaurice, G., Davidian, M., Verbeke, G., and Molenberghs, G. *Longitudinal data analysis: a handbook of modern statistical methods*. New York: Chapmann & Hall/CRC, 2008.

[16] Pinheiro, J. C. and Bates, D. M. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of computational and Graphical Statistics*, 4:12–35, 1995.

[17] Abramowitz, M. and Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Table*. New York: Dover, 1964.

[18] Golub, G. H. Some modified matrix eigenvalue problems. *Siam Review*, 15:318–334, 1973.

[19] Hedeker, D. and Gibbons, R. D. MIXOR: a computer program for mixed-effects ordinal regression analysis. *Computer Methods and Programs in Biomedicine*, 49:157–176, 1996.

[20] Lesaffre, E. and Spiessens, B. On the effect of the number of quadrature points in a logistic random effects model: an example. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 50:325–335, 2001.

[21] White, H. Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–25, 1982.

[22] Verbeke, G. and Molenberghs, G. The gradient function as an exploratory goodness-of-fit assessment of the random-effects distribution in mixed models. *Biostatistics*, 14:477–490, 2013.

[23] Boekmann, A., Sheiner, L., and Beal, S. Nonmem user's guide, part v: introductory guide. *University of California, San Francisco*, 1992.

[24] Davidian, M. and Giltinan, D. M. Nonlinear models for repeated measurement data: an overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, 8: 387–419, 2003.

[25] Drikvandi, R., Verbeke, G., Khodadadi, A., and PartoviNia, V. Testing multiple variance components in linear mixed-effects models. *Biostatistics*, 14:144–159, 2013.

[26] Litière, S., Alonso, A., and Molenberghs, G. Type I and type II error under random-effects misspecification in generalized linear mixed models. *Biometrics*, 63:1038–1044, 2007.

[27] Woods, C. M. Likelihood-ratio DIF testing: Effects of nonnormality. *Applied Psychological Measurement*, 32:511–526, 2008.

[28] Sinha, S. Bootstrap tests for variance components in generalized linear mixed models. *The Canadian Journal of Statistics*, 37:219–234, 2009.

[29] Drikvandi, R., Khodadadi, A., and Verbeke, G. Testing variance components in balanced linear growth curve models. *Journal of Applied Statistics*, 39:563–572, 2012.