Published in final edited form as:

J Stat Comput Simul. 2015 ; 85(13): 2656–2665. doi:10.1080/00949655.2014.948442.

Approximate testing in two-stage nonlinear mixed models

J.H. Burton^{a,*} and J. Volaufova^b

^aPennington Biomedical Research Center, LSU, Baton Rouge, LA, USA;

^bLSUHSC School of Public Health, New Orleans, LA, USA;

Abstract

We investigate here small sample properties of approximate F -tests about fixed effects parameters in nonlinear mixed models. For estimation of population fixed effects parameters as well as variance components, we apply the two-stage approach. This method is useful and popular when the number of observations per sampling unit is large enough. The approximate F -test is constructed based on large sample approximation to the distribution of nonlinear least squares estimates of subject-specific parameters. We recommend a modified test statistic that takes into consideration approximation to the large sample Fisher information matrix (See [1]). Our main focus is on comparing finite sample properties of broadly used approximate tests (Wald test and likelihood ratio test) and the modified F -test under the null hypothesis, especially accuracy of p -values (See [2]). For that purpose two extensive simulation studies are conducted based on pharmacokinetic models (See [3, 4]).

Keywords

Two-stage nonlinear mixed model; approximate test; accuracy of p -value

1. Introduction

Two-stage or *random coefficients* models can be used to model nonlinear longitudinal data in which responses are observed repeatedly over time. Common goals in this setting involve inference about population mean parameters. Much literature exists describing several estimation techniques (see, e.g., [5–9], and the references therein). Less attention has been given to testing mean parameters in nonlinear models. The large-sample tests often used are the likelihood ratio test (LRT) and the Wald test. It is well documented that rejection rates of these tests under the null hypothesis are greater than the nominal level (see, e.g., [10],[11], and more recently [12] and the references therein), especially when the so called first-order approximation to the marginal likelihood is used. Thus, our main objective is to investigate and compare properties of commonly used tests when constructed from a two-stage estimation method not involving the marginal likelihood. We recommend a slight improvement to the asymptotic test suggested in [7]. In particular, accuracy of p -values [2, 13] will be assessed.

*Corresponding author. Jeffrey.Burton@pbrc.edu.

We use here the notational conventions: for a matrix A , A' denotes the transpose of A , A^{-1} denotes the inverse of A (if it exists) and A^- the generalized inverse (see [14]). For matrices A and B , $A \otimes B$ denotes the Kronecker product of A and B . If A is square, $\text{tr}(A)$ denotes the trace of A .

2. Model and Methods

We investigate a nonlinear random coefficients model defined as a hierarchical model with two levels: the individual-level and the population-level (see [15] and the references therein). For the i -th of the N sampling units, the response variable y is observed at time points $t_{ij}, j = 1, \dots, T_i$. The model function $f(\cdot)$ is characterized by an unknown subject-specific p -vector parameter θ_i and a covariate vector, say x_{ij} . The vector x_{ij} represents a set of fixed, known within-subject covariates, including observation times. The individual level model for the T_i -dimensional response vector Y_i is then defined as

$$Y_i = f(\theta_i, X_i) + \varepsilon_i, \quad i=1, \dots, N. \quad (1)$$

Let t_i denote the vector of time points with entries $t_{ij}, j = 1, \dots, T_i$. The T_i -dimensional error vector ε_i is assumed to be distributed as $N_{T_i}(0, R_i(\eta))$, where the covariance matrix $R_i(\eta)$ may depend on a vector of unknown variance-covariance components η . $f(\theta_i, X_i)$ denotes the T_i -dimensional vector of nonlinear model functions with j -th entry $f(\theta_i, x_{ij})$. In order to obtain estimators of θ_i on an individual level with favorable asymptotic properties, the function $f(\cdot)$ must be twice differentiable with respect to θ_i (see [16–18]). The population level of the nonlinear random coefficients model is defined by

$$\theta_i = g(a_i, \beta) + Bb_i, \quad i=1, \dots, N, \quad (2)$$

where β is an r -dimensional vector of fixed, unknown population parameters, $\beta \in \Omega \subseteq \mathcal{R}^r$, and b_i is a q -dimensional vector of random effects. The random vectors b_i are assumed to be normally distributed with mean zero and unknown positive definite covariance matrix D . The function $g(\cdot)$ is a (twice differentiable) function of β and the vector a_i is a within-subject design vector. The $p \times q$ -matrix B is a known, full-column-rank matrix indicating which elements of θ_i are random. The random vectors ε_i and b_i are assumed mutually independent.

2.1 Second-Stage Estimation

We assume simplifying conditions under which we pursue our developments. Without loss of generality we assume that $g(a_i, \beta) = A_i\beta$, where $A_i = \mathbf{I} \otimes a_i'$ (see, e.g., [19]). We assume that the entire vector θ_i is random. In (2) we assume that $q = p$ and $B = I_p$, which results in $\theta_i \sim N_p(g(a_i, \beta), D)$ for all i . Lastly, we assume the covariance matrix $R_i(\eta)$ to be $\sigma^2 I_{T_i}$ with $\sigma^2 > 0$ unknown.

In the first stage, NLS estimates of θ_i are obtained by minimizing the objective function:

$$Q(\theta_i) = [y_i - f(\theta_i, X_i)]' [y_i - f(\theta_i, X_i)]. \quad (3)$$

Denote by

$$\tilde{\theta}_i = \underset{\theta_i}{\operatorname{argmin}} Q(\theta_i)$$

the solution to the minimization problem (3). The pooled estimate of σ^2 is obtained as

$$\tilde{\sigma}^2 = \frac{1}{\sum_{i=1}^N T_i} \sum_{i=1}^N Q(\tilde{\theta}_i).$$

The second-stage estimation is based on modeling $\tilde{\theta}_i$ and their asymptotic normality. Let

$$Z_i(\theta_i) = \frac{\partial f(\theta_i, X_i)}{\partial \theta_i'},$$

denote the matrix of first derivatives of the function $f(\theta_i, X_i)$. Among the regularity conditions assumed to hold, we assume that for each given i ,

$$\lim_{T_i \rightarrow \infty} \frac{1}{T_i} Z_i(\theta_i)' Z_i(\theta_i) = \Psi(\theta_i), \quad \text{p.d.},$$

(see, e.g., [16–18]). Then, conditioned on θ_i ,

$$\sqrt{T_i}(\tilde{\theta}_i - \theta_i) \xrightarrow{D} N_p(0, \sigma^2 \Psi(\theta_i)).$$

For large enough T_i ,

$$\tilde{\theta}_i \approx A_i \beta + u_i,$$

where

$$u_i \approx N\left(0, D + \tilde{\sigma}^2 \left(Z_i(\tilde{\theta}_i)' Z_i(\tilde{\theta}_i) \right)^{-1}\right).$$

The parameters β , D can be estimated using a multivariate linear mixed model approach (for more details see [7]).

Alternatively, (see [20] and the references therein), under regularity conditions, for large T_i , conditionally on θ_i ,

$$\sqrt{T_i}(\tilde{\theta}_i - \theta_i) \sim N\left(0, \sigma^2 \left[\frac{1}{T_i} Z_i(\theta_i)' Z_i(\theta_i) \right]^{-1}\right).$$

The marginal moments of $\tilde{\theta}_i$ can be approximated as:

$$E(\tilde{\theta}_i) = E[E(\tilde{\theta}_i|b_i)] \approx E(\theta_i) = A_i\beta,$$

and

$$\begin{aligned} Var(\tilde{\theta}_i) &= Var[E(\tilde{\theta}_i|b_i)] + E[Var(\tilde{\theta}_i|b_i)] \\ &\approx Var(\theta_i) + E\left\{\sigma^2[Z_i(\theta_i)'Z_i(\theta_i)]^{-1}\right\} \quad (4) \\ &\approx D + \sigma^2[Z_i(\beta)'Z_i(\beta)]^{-1} \equiv W_i(\beta, \vartheta). \end{aligned}$$

The second term in (4) was obtained by a first-order Taylor series expansion at $E(\theta_i) = A_i\beta$. The vector ϑ of unknown variance-covariance components consists of unique unknown entries of D as well as σ^2 , i.e., $\vartheta = (\text{vech}(D)', \sigma^2)'$. (See [14]). The vech operator creates from a $n \times n$ symmetric matrix a $n(n+1)/2$ -dimensional column vector by stacking columns of the lower triangular portion of the matrix.

Based on the approximate marginal distribution of $\tilde{\theta}_i$, the second-stage model takes the form

$$\tilde{\theta}_i = A_i\beta + u_i, \quad i=1, \dots, N, \quad (5)$$

where $u_i \approx N(0, W_i(\beta, \vartheta))$. The model in (5) represents a generalized linear mixed model. The unknown model parameters β and ϑ can be estimated by maximum likelihood.

Hypothesis Testing—The hypothesis of interest is of the form:

$$H_0: G'\beta = h_0 \text{ vs. } H_1: G'\beta \neq h_0, \quad (6)$$

where the $r \times c$ -matrix G is a known, full column-rank matrix, $r(G) = c$, $1 \leq c \leq r$, and h_0 is a known vector. The commonly used tests of H_0 are the large-sample likelihood ratio (LRT) and Wald tests. Asymptotically both statistics, under H_0 , follow χ^2 -distributions with c degrees of freedom.

Approximate F -tests can also be constructed as in linear mixed models (see e.g., [21] and the references therein). Consider the statistic

$$F^* = (1/c) (G'\hat{\beta} - h_0)' [\hat{Var}(G'\hat{\beta})]^{-1} (G'\hat{\beta} - h_0). \quad (7)$$

The distribution of F^* in (7) under H_0 is approximated by an F -distribution with c and, say, ν degrees of freedom. The crucial points are the approximation of the estimated covariance matrix $\hat{Var}(G'\hat{\beta})$ of $(G'\hat{\beta})$ and establishing denominator degrees of freedom (ddf) ν .

Using the large-sample assumption of a normal log-likelihood $\ell_i(\beta, \vartheta|\tilde{\theta}_i)$, the approximation of $\hat{Var}(G'\hat{\beta})$ can be based on the approximate Fisher information matrix

(see [20] and the references therein). For the two-stage model, one can approximate the Fisher information matrix for the entire vector of unknown model parameters $\phi = (\beta', \vartheta')$ as:

$$\mathcal{J}(\phi) = - \sum_{i=1}^N E \left[\frac{\partial^2 \ell_i(\phi | \tilde{\theta}_i)}{\partial \phi \partial \phi'} \right] = \begin{pmatrix} J(\phi) & C(\phi) \\ C'(\phi) & B(\phi) \end{pmatrix}, \quad (8)$$

where $J(\phi)$, $C(\phi)$, and $B(\phi)$ denote the blocks of matrix $\mathcal{J}(\phi)$ of appropriate dimensions.

The matrix $C(\phi)$ is in general non-zero due to dependence of covariance matrix of $\tilde{\theta}_i$ on both population fixed effects parameters and variance-covariance parameters.

Blocks on the main diagonal of the inverse of (8) are used to obtain approximate large sample covariance matrices of estimates $\hat{\beta}$ and $\hat{\vartheta}$ (see, e.g., [22, 23], or [1]).

$$Var_S(\hat{\beta}) \approx J(\phi)^{-1} + J(\phi)^{-1} C(\phi) Var_S(\hat{\vartheta}) C'(\beta, \vartheta) J(\phi)^{-1}. \quad (9)$$

$$Var_S(\hat{\vartheta}) \approx [B(\phi) - C'(\phi) J(\phi)^{-1} C(\phi)]^{-1} \quad (10)$$

Here the subscript S stands for the Schur complement based approximation. The resulting expression for the approximate large sample covariance matrix of $\hat{\beta}$ is the sum of two matrices. The first is the negative expected value of the matrix of second partial derivatives of the log-likelihood with respect to β (see also (8)). The second matrix is a function of the large sample approximate covariance matrix of estimates of ϑ , $Var_S(\hat{\vartheta})$. An expression for the estimated approximation to the covariance matrix of $\hat{\beta}$ is obtained by plugging the approximate maximum likelihood estimates of β and ϑ into (9). We get

$$\hat{Var}_S(\hat{\beta}) \approx J(\hat{\beta}, \hat{\vartheta})^{-1} + J(\hat{\beta}, \hat{\vartheta})^{-1} C(\hat{\beta}, \hat{\vartheta}) \hat{Var}_S(\hat{\vartheta}) C'(\hat{\beta}, \hat{\vartheta}) J(\hat{\beta}, \hat{\vartheta})^{-1}. \quad (11)$$

Alternatively, we may ignore the dependence of the covariance matrix of $\tilde{\theta}_i$ on β in model (5) (see, e.g., [24, 25]). We obtain approximate model

$$\tilde{\theta}_i = A_i \beta + u_i^*, \quad u_i^* \sim N(0, D), \quad i=1, \dots, N, \quad (12)$$

from which we get

$$Var_A(\hat{\beta}) \approx \left[\sum_{i=1}^N A_i' D^{-1} A_i \right]^{-1}, \quad (13)$$

where subscript A is for “alternative”. To get an estimate of the approximated large sample covariance matrix of $\hat{\beta}$, we plug in into (13) the maximum likelihood estimate $\hat{D}$ of D from model (12).

We can construct LRT, Wald test, and approximate F -tests using the two approximations to the estimated covariance matrix of $\hat{\beta}$, (9) and (13), respectively. From these, we may construct several approximate F -tests using the statistic (7) combined with different choice of ddf . The Fai-Cornelius-Satterthwaite type approximation [26] to ν is frequently used (see, e.g., [11]). A simpler approach is using fixed degrees of freedom $\nu = N - q$, where q is the number of random effects in the model (see, e.g., [27]). The following section will detail simulations carried out to compare properties of tests described here.

3. Simulations

One-Compartment Model for IV Dose

Pharmacokinetic (PK) studies are carried out in order to investigate how a drug is processed in the body. Researchers often want to compare pharmacokinetic profiles between subjects from different populations following drug administration. That is, investigators want to know if different populations process the drug differently. Compartmental methods, in which specified models are used to estimate mean plasma drug concentrations over time, may be used. In the first simulation study, a one-compartment PK model for an intravenous (IV) drug dose is used.

PK models of this type are often used for illustration of nonlinear mixed model methodology (see, e.g., [3]). The model has the following form:

$$Y_{ij} = \frac{d_i}{V_i} e^{-\left(\frac{Cl_i}{V_i}\right)t_{ij}} + \varepsilon_{ij}, \quad i=1, \dots, N, \quad j=1, \dots, T_i. \quad (14)$$

Here, Y_{ij} is the observed plasma concentration for subject i at time t_{ij} , d_i is the administered dose, V_i and Cl_i are unknown subject-specific PK parameters representing volume of distribution and clearance rate respectively, and ε_{ij} are the mutually independent random error terms, $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$.

We are interested in testing the null hypothesis that the parameters for volume and clearance for two populations are equal. The vector composed of the individual PK parameters is denoted $(Cl_i, V_i)' = (e^{\theta_{1i}}, e^{\theta_{2i}})'$. The elements θ_{1i} and θ_{2i} are exponentiated to ensure positive values. Then

$$\theta_{1i} = \beta_1 + b_{1i}; \quad \theta_{2i} = \beta_2 + b_{2i} \quad (15)$$

for subjects from the first population and

$$\theta_{1i} = \beta_1 + \delta_1 + b_{1i}; \quad \theta_{2i} = \beta_2 + \delta_2 + b_{2i} \quad (16)$$

for subjects from the second population. The null hypothesis $H_0 : \delta_1 = \delta_2 = 0$ is investigated.

To generate data under the null hypothesis, the values of the population parameters from [3] were used. For clearance and volume, $\beta_1 = \log(1.25)$ and $\beta_2 = \log(10)$, respectively. For the covariance matrix of the random effects b_i ,

$$D = \begin{pmatrix} 0.086 & 0.052 \\ 0.052 & 0.086 \end{pmatrix}.$$

Two values, $\sigma^2 = 1$ and $\sigma^2 = 2$, were investigated. Dose $d_i \equiv d = 100$ for all subjects.

Samples were generated with an equal number of observations per subject with the vector of time points being $t_i = (0.25, 0.5, 1, 2, 4, 8, 12, 24)'$. For sample sizes, three options $(n_1, n_2) = (10, 8), (20, 25)$, and $(50, 35)$ were used. For each of the 6 possible model configurations, 5000 replications were generated.

Using generated data, estimates of unknown parameters β and $\vartheta = (\lambda', \sigma^2)'$ were obtained using the two-stage method. For all calculations, SAS software versions 9.2 and 9.3 were used. For NLS estimates, SAS PROC NLIN was applied with the Gauss-Newton optimization method used to minimize the least squares objective function. Initial values for the optimization were chosen close to the true values, from which data were generated. Convergence was achieved for all 5000 replications under all model configurations. For the second stage, PROC IML was used to calculate ML estimates as well as the Wald, F_S^* , and LRT statistics. PROC MIXED was used to obtain the F_A^* test statistic.

The p -values for LRT and Wald tests were calculated from the $\chi^2(2)$ distribution. For F_S^* and F_A^* , the p -values were obtained from an F -distribution with numerator degrees of freedom $r(G) = 2$. In Table 1, a list of test statistics and derived ddf s are shown.

Accuracy of p-Values

Denote by P the random variable “ p -value” and by p its observed value. For a test statistic with known probability distribution, P follows uniform distribution ($UNI(0, 1)$) under the null hypothesis. This is not generally the case if the null distribution of the test statistic (under H_0) is approximated. Hence it is reasonable to investigate how close the distribution of P is to $UNI(0, 1)$; in other words, the “accuracy” of the p -value. The p -value is accurate if, under H_0 , $P_0(P \leq p) = p$ (i.e., if its cumulative distribution function (CDF) is the identity function) (see, e.g., [2]). Accuracy of p -values can be investigated as a surrogate for accuracy of the probability distribution of the test statistic under H_0 . To assess accuracy, the empirical CDF of P is compared to the CDF of the $UNI(0, 1)$ distribution. The accuracy of the p -value is closely related to “validity” of the p -value (see, e.g., [28]). A p -value is valid if, under H_0 , $P_0(P \leq p) \leq p$.

Based on a p -value P that is not accurate, the probability of Type I error is inaccurate. For the p -value that is not valid, one takes the risk of higher rejection rate of a true null hypothesis. That is, it is more likely that a difference will be inferred that does not exist. For this reason, an inaccurate p -value is less “harmful” for inference if it is valid.

Plots of p -values under each model configuration for the two smaller sample size configurations are shown in Figure 1. The vertical axis represents the observed p -values and the horizontal axis represents their empirical CDF. Axes only range from 0 to 0.10 because p -values outside of this range are rarely of interest; in a typical setting, the significance level is usually below 0.10.

In the accuracy plots, $Satt$ denotes the approximate F -test constructed from F_S^* with Fai-Cornelius-Satterthwaite type ddf . $Wolf$ denotes the test constructed in the same manner but using $ddf = \nu = N - q$. Lastly, A denotes the approximate F -test constructed using F_A^* and Fai-Cornelius-Satterthwaite type ddf .

One-Compartment Model for Oral Dose

In the second simulation study, data were generated from a one-compartment PK model, in which the plasma concentration is modeled following oral administration of a drug:

$$Y_{ij} = \frac{d_i K_i k_{a_i}}{Cl_i (k_{a_i} - K_i)} \left[e^{-K_i t_{ij}} - e^{-k_{a_i} t_{ij}} \right] + \varepsilon_{ij}, i=1, \dots, N, j=1, \dots, T_i.$$

Here, Y_{ij} is the plasma concentration for subject i at time t_{ij} ; d is the dose, K_i is elimination rate, k_{a_i} is absorption rate, and Cl_i is clearance for the i -th subject. The mutually independent errors $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. In order to ensure positive values for the subject-specific parameters, similarly to (15) and (16), define $K_i = e^{\theta_{1i}}$, $k_{a_i} = e^{\theta_{2i}}$, and $Cl_i = e^{\theta_{3i}}$. Then $\theta_{ji} = \beta_j + b_{ji}$, $j = 1, 2, 3$, for subjects from the first population and $\theta_{ji} = \beta_j + \delta_j + b_{ji}$, $j = 1, 2, 3$, for subjects from the second population. Thus, the null hypothesis takes the form $H_0 : \delta_1 = \delta_2 = \delta_3 = 0$.

For data generation under H_0 , we used population parameter values from a simulation study described in [4]. For fixed effects parameters, $\beta_1 = -2$, $\beta_2 = 0.5$, and $\beta_3 = -3$ were used. For covariance matrix D of random effects b_i , a diagonal matrix with main-diagonal values $(0.01, 0.05, 0.03)'$ was set. The parameter σ^2 was set to either $\sigma^2 = 0.25$ or $\sigma^2 = 0.75$. Dose ($d_i \equiv d = 5$) was the same for all subjects.

Data were generated with equal number of observations per subject, using the same vector of time points used in the IV model. The same sample size configurations were used and the same methodology was utilized to calculate estimates of unknown parameters; construct LR, Wald, and approximate F -tests; and obtain p -values.

4. Results and Discussion

Concluding from Figure 1, it appears that p -values from approximate F -test based on F_A^* are the most accurate within the smallest sample size setting (for both values of σ^2). The p -values from other approximate F -tests are still valid, while those from the LRT and Wald test are not. As sample size increases, the LRT becomes more accurate while accuracy of approximate F -tests remains the same. For larger sample sizes, accuracy of the LRT and F_A^* are similar. When comparing these two for the moderate sample size setting, F_A^* is

preferable. For the largest sample size setting, the accuracy is difficult to differentiate (graphs not shown). In addition, it appears that the accuracy of tests is not sensitive with respect to changes in nuisance parameter σ^2 .

In Figure 2, all tests appear slightly less valid than in the first PK model. Under the smallest sample size setting, only the approximate F -tests *Wolf* and *Satt* are valid, with our suggested test *Satt* being the more accurate of the two under both variance parameter configurations. As sample size increases, all tests become more accurate. Only tests *Wolf* and *Satt* are valid over the entire range of p -values presented. These results are the same regardless of the value of σ^2 , again indicating that these tests in this setting are not affected by changes in σ^2 .

Under both models, we have also investigated unequal time points settings. The results about accuracy and validity of p -values were very similar to those presented for the equal time points setting and, thus, are not included here.

The purpose of our research was not only to investigate properties of tests, but also to suggest improvement to existing methodologies. The improvement within the two-stage approach is in construction of the test statistic using the Schur complement of the large sample information matrix. Additionally, application of the Fai-Cornelius-Satterthwaite degrees of freedom improved accuracy of the test; however, the Wolfinger-type degrees of freedom also performed acceptably. We make two important conclusions based on our limited simulations. First, all investigated tests appear to have low sensitivity to changes in nuisance parameter σ^2 . Secondly, the test F_S^* was *less* accurate (and invalid) under the three-parameter PK model while tests *Wolf* and *Satt* were *more* accurate under the more complex model (although both were valid under both models).

With consideration to future work in this area, it may prove interesting to investigate properties of tests constructed from two-stage estimates with some model assumptions relaxed. In the research described here, we assume the within-subjects covariance matrix $R_i(\eta) = \sigma^2 I_{T_i}$. This is done for simplicity such that ordinary NLS methodology may be used for estimation of subject-specific parameters. To extend the results, a more general within-subjects covariance matrix (e.g., allowing for unequal variances and/or non-zero covariances) could be assumed. Estimation of individual level parameters could be achieved using an estimated generalized least squares approach or an extended least squares approach, in which fixed effects and variance-covariance parameters are jointly estimated (see [7] and the references therein).

Acknowledgments

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health which funds the Louisiana Clinical and Translational Science Center under Grant 1 U54 GM104940.

References

- [1]. Volafova, J.; Burton, JH. Oral presentation at: LINSTAT 2012/21st IWMS. Bedlewo, Poland: 2012. Note on hypothesis testing in mixed models.
- [2]. Volafova J, LaMotte L. Comparison of approximate tests of fixed effects in linear repeated measures design models with covariates. *Tatra Mountains*. 2008; 39:17–25.

- [3]. Hartford A, Davidian M. Consequences of misspecifying assumptions in nonlinear mixed effects models. *Comput Stat and Data An.* 2000; 34:139–164.
- [4]. Pinheiro J, Bates D. Approximations to the log-likelihood function in the non-linear mixed-effects model. *J Comput Graph Stat.* 1995; 4(1):12–35.
- [5]. Beal SL, Sheiner LB. Estimating population kinetics. *Crit Rev Biomed Eng.* 1981; 8:195–222. [PubMed: 6754254]
- [6]. Lindstrom MJ, Bates DM. Nonlinear mixed effects models for repeated measures data. *Biometrics.* 1990; 46:673–687. [PubMed: 2242409]
- [7]. Vonesh, E.; Chinchilli, V. *Linear and nonlinear models for the analysis of repeated measurements.* Marcel Dekker, Inc.; 1997.
- [8]. Wolfinger R. Laplace's approximation for nonlinear mixed models. *Biometrika.* 1993; 80:791–795.
- [9]. Dubois A, Lavielle M, Gsteiger S, Pigeolet E, Mentré F. Model-based analyses of bioequivalence crossover trials using the stochastic approximation expectation maximisation algorithm. *Stat Med.* 2010; 30:2582–2600. [PubMed: 21793036]
- [10]. Bertrand J, Comets E, Mentré F. Comparison of model-based tests and selection strategies to detect genetic polymorphisms influencing pharmacokinetic parameters. *J Biopharm Stat.* 2008; 18:1084–1102. [PubMed: 18991109]
- [11]. Bertrand J, Comets E, Chenel M, Mentré F. Some alternatives to asymptotic tests for the analysis of pharmacogenetic data using nonlinear mixed effects models. *Biometrics.* 2012; 68:146–155. [PubMed: 22049987]
- [12]. Volafova J. Approximate small sample tests in nonlinear mixed models. *Commun Stat - Simul C.* 2014 In print.
- [13]. Volafova J, LaMotte LR. A simulation comparison of approximate tests for fixed effects in random coefficients growth curve models. *Commun Stat - Simul C.* 2013; 42(2):344–359.
- [14]. Harville, DA. *Matrix algebra from a statistician's perspective.* Springer; 1997.
- [15]. Davidian M, Giltinan D. Nonlinear models for repeated measurement data: An overview and update. *J Agr Biol and Envir St.* 2003; 8(4):387–419.
- [16]. Jennrich R. Asymptotic properties of non-linear least squares estimators. *Ann Math Stat.* 1969; 40(2):633–643.
- [17]. Wu C. Asymptotic theory of nonlinear least squares estimation. *Ann Stat.* 1981; 9(3):501–513.
- [18]. Amemiya, T. Chapter 6, Non-linear regression models. In: Griliches, Z.; Intriligator, MD., editors. *Handbook of Econometrics.* Edition 1. Elsevier Science; 1983. p. 333–389.
- [19]. Vonesh E, Carter R. Efficient inference for random-coefficient growth curve models with unbalanced data. *Biometrics.* 1987; 43(3):617–628. [PubMed: 3663819]
- [20]. Burton, JH. Improved hypothesis tests of linear combinations of fixed effects parameters in nonlinear random coefficients models [dissertation]. Louisiana State University Health Sciences Center; New Orleans (LA): 2013.
- [21]. Kenward M, Roger J. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics.* 1997; 53(3):983–997. [PubMed: 9333350]
- [22]. Retout S, Mentré F, Bruno R. Fisher information matrix for nonlinear mixed effects models: Evaluation and application for optimal design of Enoxaparin population pharmacokinetics. *Stat Med.* 2002; 21:2623–2639. [PubMed: 12228881]
- [23]. Retout S, Mentré F. Further developments of the fisher information matrix in nonlinear mixed effects models with evaluation in population pharmacokinetics. *J Biopharm Stat.* 2003; 13:209–227. [PubMed: 12729390]
- [24]. Mielke, T.; Schwabe, R. Some considerations on the fisher information in nonlinear mixed effects models. In: Giovagnoli, A.; Atkinson, AC.; Torsney, B.; May, C., editors. *mODa 9 - Advances in Model-Oriented Design and Analysis; Proceedings of the 9th International Workshop in Model-Oriented Design and Analysis; Bertinoro, Italy. 2010 Jun 14-18; Springer, Physica Verlag Heidelberg; 2010. p. 129-136.*

- [25]. Mielke, T. Approximation of the fisher information and design in nonlinear mixed effects models [unpublished preprint]. Available from: http://www.math.unimagdeburg.de/institute/imst/ag_schwabe/preprints/2011_22.pdf
- [26]. Fai A, Cornelius P. Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. *J Stat Comput Sim.* 1996; 54:363–378.
- [27]. Wolfinger, RD. Fitting nonlinear mixed models with the new NLMIXED procedure [technical report 287]. SAS Institute; Cary (NC): 2000.
- [28]. Casella, G.; Berger, RL. *Statistical Inference*. Second edition. Duxbury: 2002.

Figure 1.

Accuracy of p -values for LRT, Wald test, and approximate F -tests: A -type, $Satt$ -type, and $Wolf$ -type for equal time points. Calculations based on data generated from one-compartment pharmacokinetic model for single IV bolus under various sample size and variance parameter configurations. (a)-(b): $\sigma^2 = 1$; (c)-(d): $\sigma^2 = 2$.

Figure 2.

Accuracy of p -values for LRT, Wald test, and approximate F -tests: A -type, $Satt$ -type, and $Wolf$ -type for equal time points. Calculations based on data generated from one-compartment pharmacokinetic model for single oral bolus under various sample size and variance parameter configurations. (a)-(b): $\sigma^2 = 0.25$; (c)-(d): $\sigma^2 = 0.75$.

Table 1

Descriptions of tests of fixed effects used in simulations.

Cov. Approx.	Statistic	df	Distribution
N/A	LRT	c	χ^2
$\hat{Var}_S(\hat{\beta})$	Wald($= c F_S^*$)	c	χ^2
	F_S^* (Wolfinger)	$N - q$	$F(c, N - q)$
	F_S^* (Satterth)	ν_{2sat}	$F(c, \nu_{2sat})$
$\hat{Var}_S(\hat{\beta})$	F_A^* (Satterth)	ν_{sat}	$F(c, \nu_{sat})$