Simulation Studies on Nonlinear Mixed Effects Models

in Population Pharmacokinetics Studies

北里大学大学院臨床統計部門　竹内正弘　（Masahiro Takeuchi）

南博文　　（Hirofumi Minami）

Division of Biostatistics, Kitasato University Graduate School

1.　Introduction

Drug review procedure and drug development strategies are changing rapidly due to *"The International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use"*, which was initiated in 1990 (ICH, 1990). The ICH seeks for improvement of the efficiency of the development and review processes for promising new drugs by unifying necessary documentation and its associated formats for new drug applications (NDA) to regulatory agencies. In particular, E5 guideline regarding ethnic factors in the acceptability of foreign clinical data has a significant impact on a new drug's development by allowing the extrapolation of foreign clinical data as a part of an NDA submission to the regulatory agency in a new region (ICH, 1998). In the extrapolation of the foreign clinical data to the new regions, a population pharmacokinetics study (PPK) is valuable for the evaluation of pharmacokinetics parameters in order to investigate intrinsic factors among populations. In the population pharmacokinetics studies, the two types of the statistical analyses are commonly employed. We will investigate the statistical properties of the analyses through simulation studies.

2.　Definition and Properties

For each subject i, $1 \leq i \leq N$, $n_i \times 1$ observation vector, y_i, will follow a nonlinear mixed effects model defined as

$$y_i = f(X_i, \beta_i) + \varepsilon_i \quad \beta_i = \beta + b_i \tag{1}$$

where $f(X_i, \beta_i)$ is a nonlinear function of pharmacokinetics model, β_i is an individual $r \times 1$ vector of regression coefficients, X_i is a known $n_i \times t$ design matrix, β is a $r \times 1$ fixed effects parameters, b_i is a $r \times 1$ random effects parameters, ε_i is a $n_i \times 1$ vector of error terms. We assume that b_i is normally distributed with a mean 0, covariance matrix Ψ denoted by, $N(0, \Psi_{r \times r})$, ε_i is normally distributed with mean 0, covariance matrix $\sigma^2 \Lambda_i(\gamma)$ denoted by, $N(0, \sigma^2 \Lambda_i(\gamma)_{n_i \times n_i})$, where γ are unknown parameters. We assume that b_i and ε_i are mutually independent. Our main goal is an estimation of fixed effects parameter β. Therefore we need to know its marginal distribution of individual observations.

Define that $P(y_i \mid b_i)$ and $P(b_i)$ are probability density functions of $y_i \mid b_i$ and b_i. The marginal density function of y_i, $P(y_i)$, can be defined as

$$P(y_i) = \int P(y_i \mid b_i) P(b_i) db_i \tag{2}$$

In contrast to a linear mixed effects model (Laird and Ware, 1982), their expected values of the observed data are nonlinear functions of both the fixed effects and the random effects. In general there is no closed forms existed. Commonly two types of the 1st order Taylor expansion are employed in order to get the closed forms. The first approximation method is the 1st order Taylor expansion around its expected values, 0 (Sheiner and Beal, 1980, 1985, Vonesh and Carter, 1992) and the second approximation method is the 1st order Taylor expansion around its estimated values of the random effects, $\hat{b}_i$ (Lindstrom and Bates, 1990).

The first Taylor expansion around the expected values, 0, of the observed data can be written as

$$y_i = f(X_i, \beta, b_i = 0) - \left[\frac{\partial f}{\partial b_i^T} \bigg|_{b_i = 0} \right] b_i + \varepsilon_i \tag{3}$$

where

$$Z_i = \left[\frac{\partial f}{\partial b_i^T} \bigg|_{b_i = 0} \right]$$

Then the observed data can be approximated as

$$y_i = f(X_i, \beta, b_i = 0) - Z_i b_i + \varepsilon_i \qquad (4)$$

It can be shown that the expected values of the observed data, $E[y_i]$, and the variance,

$Var[y_i]$, can be written as

$$E[y_i] = f(X_i, \beta, b_i = 0) = f(X_i, \beta) \qquad (5)$$

$$Var[y_i] = Z_i \Psi Z_i^T + \sigma^2 A_i(\gamma) \qquad (6)$$

The expected values are only functions of the fixed effects, and the variance are functions of both the fixed effects and the random effects. Two points are noteworthy to mention in this Taylor expansion: (1) this approximation is appropriate when the variance of the random effects, b_i, $Var[b_i]$, is very small so that we can ignore individual variations (Solomon and Cox, 1992) and (2) the expected values of the marginal distribution must be correctly specified for inference of its covariance parameters, γ and σ^2 (Breslow and Clayton, 1993).

The other Taylor expansion around the estimated random values, $\hat{b}_i$, of the observed values can be defined as

$$y_i = f(X_i, \beta, b_i = \hat{b}_i) - \left[\frac{\partial f}{\partial b_i^T} \bigg|_{b_i = \hat{b}_i} \right] (b_i - \hat{b}_i) + \varepsilon_i \qquad (7)$$

where

$$Z^*_i = \left[\frac{\partial f}{\partial b_i^T} \bigg|_{b_i = \hat{b}_i} \right]$$

Then the expected values and its covariance can be written as follows:

$$E[y_i] = f(X_i, \beta, b_i = \hat{b}_i) - Z^*_i \hat{b}_i \qquad (8)$$

$$Var[y_i] = Z^*_i \Psi Z^{*T}_i + \sigma^2 A_i(\gamma) \qquad (9)$$

In contrast to the results derived from the Taylor expansion around the expected values, 0, both the expected values and its covariance are function of the estimated random value, $\hat{b}_i$. These results will influence an asymptotic distribution of estimators, $\hat{\beta}$, of the fixed effects. The likelihood function derived from the Taylor expansion around the estimated random values, $\hat{b}_i$, can be derived by Laplace

approximation (Wolfinger, 1993, Vonesh, 1996). It can be shown that the required condition for consistency can be given as

$$\hat{\beta} - \beta = O_p\left[\max\left\{\frac{1}{\sqrt{N}}, \quad \frac{1}{\min(n_i)}\right\}\right] \tag{10}$$

(Vonesh, 1996). The condition indicates that an increasing rate of a sample size and of observations per subject must be constant in order to obtain the consistency of $\hat{\beta}$.

It is well known that population pharmacokinetics analysis is suitable to data consisting of a few data points per subject from many individuals. We can apply the first order Taylor expansion around the expected values, 0, when we can ignore the variability among individuals, i.e., $Var[b_i]$. In that case we require a large number of individuals, N, with few data points per subject, n_i, for the consistency of $\hat{\beta}$. On the other hand, the main purpose of population pharmacokinetics analysis is to estimate population parameters adjusting for variability among subjects, $Var[b_i]$. The first order Taylor expansion around the estimated random effects, $\hat{b}_i$, is appropriate in this case. This method requires a large number of observations per individual, which is contradict to the merit of population pharmacokinetics analysis requiring few observations per subject from many individuals for the consistency of the population parameters. We investigate behaviors of the estimated population parameters, $\hat{\beta}$, influenced by either a total sample size, N, or observations per subject, n_i, depending on a degree of $Var[b_i]$ through simulation studies. Finally we consider an appropriate study design in clinical trial settings.

3. Simulation Studies

Plasma concentrations are simulated from the following 1-compartment pharmacokinetics model defined by equation (11) after multiple dosing:

$$C_{m_{ij}} = \left\{ A * \left[\frac{1-\exp(-m*k_{ei}*\tau)}{1-\exp(-k_{ei}*\tau)}\right] * \exp(-k_{ei}*t_{ij}) - A * \left[\frac{1-\exp(-m*k_{ai}*\tau)}{1-\exp(-k_{ai}*\tau)}\right] * \exp(-k_{ai}*t_{ij})\right\} * \exp(e_{ij})$$

$$A = \frac{k_{ai} * F_i * D_i}{VD_i * (k_{ai} - k_{ei})},$$

$$k_{ei} = \frac{CL_i}{VD_i}, CL_i = \mu_{CL} * \exp(Z_{CL_i}), VD_i = \mu_{VD} * \exp(Z_{VD_i}), k_{ai} = \mu_{k_a} * \exp(Z_{k_{ai}})$$

The observations, C_{mij}, are simulated plasma concentrations of subject i at time t_{ij}, after multiple m-dosing on the log-scale to provide additive residual error. The times at each multiple-dosing are fixed as [0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 8, 10, 12, 24]. τ is fixed as 24. The observations on each individual are randomly selected from a uniform distribution.

The fixed effects population parameters of interest are k_a, VD and CL. The parameter, k_a, is an absorption rate constant, VD is a total distribution volume constant and CL is a total body clearance rate constant. The individual-specific constant, k_{ai}, VD_i and CL_i, can be defined as lognormal random variables. The normal random variables, Z_{CL_i}, Z_{VD_i} and $Z_{k_{ai}}$, which define CL_i, VD_i and k_{ai}, are assumed to be independent with mean zero and variances, CV_{CL}, CV_{VD} and CV_{ka}, respectively. The residual errors e_{ij} are assumed to be lognormally distributed and be independent of Z_{CL_i}, Z_{VD_i} and Z_{kai} with mean zero and variance CV_{CP}. Therefore, the variance parameters of interest are CV_{CL}, CV_{VD}, CV_{ka} and CV_{CP}.

We focus on the results derived from the 1st order Taylor expansion around $\hat{b}_i$ in this paper. The results derived from the 1st order Taylor expansion around the expected values, 0, can be found in Minami, 2002. Table 3-1 shows the parameter values for each simulation study on 15 scenarios. We perform simulations up to "100" to obtain optimal parameter estimates.

Table 3-1: Simulation Data Sets

Data set No.	m	N	n_i	CL	VD	k_a	CV_{CL}, CV_{VD} and CV_{ka}	CV_{CP}
1	28	25	2	0.693	4.0	1.386	0.50	0.20
2	28	25	3	0.693	4.0	1.386	0.50	0.20
3	28	25	4	0.693	4.0	1.386	0.50	0.20
4	28	25	5	0.693	4.0	1.386	0.50	0.20
5	28	25	6	0.693	4.0	1.386	0.50	0.20
6	28	50	2	0.693	4.0	1.386	0.50	0.20
7	28	50	3	0.693	4.0	1.386	0.50	0.20
8	28	50	4	0.693	4.0	1.386	0.50	0.20
9	28	50	5	0.693	4.0	1.386	0.50	0.20
10	28	50	6	0.693	4.0	1.386	0.50	0.20
11	28	100	2	0.693	4.0	1.386	0.50	0.20
12	28	100	3	0.693	4.0	1.386	0.50	0.20
13	28	100	4	0.693	4.0	1.386	0.50	0.20
14	28	100	5	0.693	4.0	1.386	0.50	0.20
15	28	100	6	0.693	4.0	1.386	0.50	0.20

Equation (12) defines bias (%) for evaluation of the estimated fixed parameters from the simulated data sets.

$$Bias\,(\%) = \frac{\hat{\beta} - \beta_0}{\beta_0} \times 100\,(\%) \tag{12}$$

$\hat{\beta}$: Parameter estimate of fixed effects

β_0 : True values of fixed effects

Table 3-2 and Figure 4-1 show results of the bias(%) of CL from 100 simulations derived from the 1st order Taylor expansion around $\hat{b}_i$. In Figure 3-1, the symbol (■) denotes a mean of the estimated parameters, $\hat{\beta}$. The symbol ($\triangle$) of a top or a bottom denotes a maximum or a minimum value of the bias(%) of CL from 100 simulations, and the symbol ($\times$) of a bottom or a top denotes 1% or 99% quantile, respectively. The bottom and top edges of the box plots denote 25% and 75% quantiles. The line in an interior of the box denotes a median (50% quantile).

Table 3-2: Results from the 1st order Taylor Expansion around $\hat{b}_i$

No.	N	n_i	Mean	S.D.	Min	Max	Quantiles				
							10%	25%	50%	75%	90%
1	25	2	-1.15	5.96	-12.69	14.20	-8.56	-5.43	-1.14	2.71	6.34
2	25	3	-0.49	5.36	-15.96	9.71	-8.10	-4.00	-0.36	3.76	6.49
3	25	4	-0.05	4.91	-12.03	11.30	-7.00	-3.24	0.15	3.73	6.08
4	25	5	-0.46	4.53	-10.76	10.37	-6.62	-3.73	-0.35	3.14	4.87
5	25	6	-0.30	4.37	-9.90	9.26	-6.20	-4.07	-0.09	2.98	4.66
6	50	2	-0.59	4.28	-13.91	8.75	-6.17	-3.65	-0.82	2.92	4.90
7	50	3	-0.45	3.39	-8.46	6.88	-5.07	-2.60	-0.34	2.24	4.12
8	50	4	-0.32	3.19	-8.24	5.96	-4.56	-2.76	0.05	2.07	4.17
9	50	5	-0.37	3.25	-7.93	7.19	-4.14	-2.83	-0.62	1.98	3.87
10	50	6	-0.27	3.14	-8.10	6.63	-4.08	-2.45	-0.59	2.15	3.99
11	100	2	-0.05	3.13	-9.12	7.40	-3.91	-2.11	-0.26	1.93	4.19
12	100	3	-0.13	2.58	-6.80	5.07	-3.41	-2.16	0.21	1.54	3.37
13	100	4	-0.21	2.34	-6.54	4.85	-3.58	-1.74	0.10	1.34	2.79
14	100	5	-0.40	2.10	-5.76	4.21	-3.37	-1.83	-0.23	0.99	2.36
15	100	6	-0.35	2.04	-5.71	4.98	-3.15	-1.73	-0.09	0.74	2.36

Figure 3-1: Results from the 1st order Taylor expansion around $\hat{b}_i$

N=25 N=50 N=100

$\square$ n_i=2, $\square$ n_i=3, $\square$ n_i=4, $\square$ n_i=5, $\square$ n_i=6
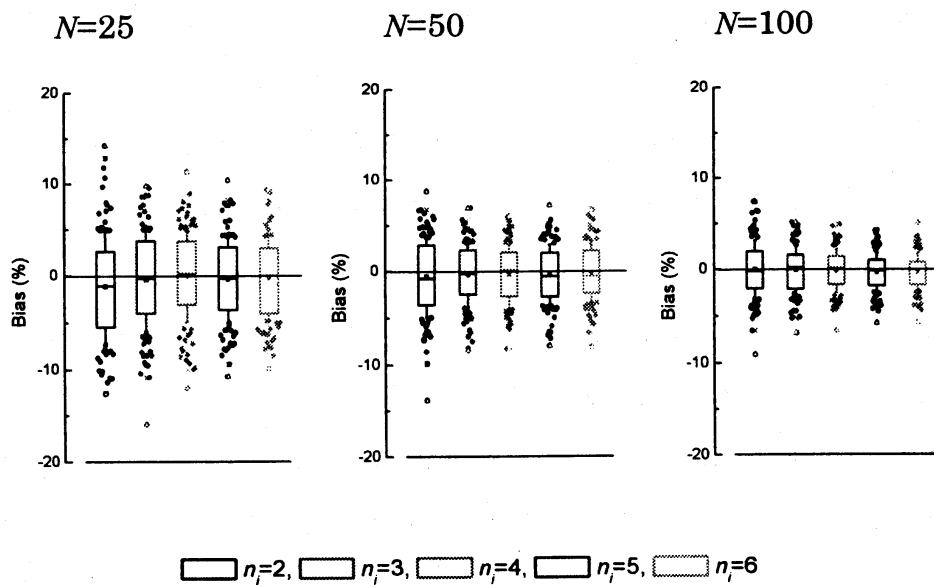
4. Conclusion

We investigate possible effects of required conditions regarding a total sample size, N,

and a number of observations per subject, n_i, on the estimated population parameters, $\hat{\beta}$, in an application of the two approximation approaches to the marginal distribution, the 1st order Taylor expansion around the expected values, 0, and around the estimated random effects, $\hat{b}_i$, in population pharmacokinetics studies.

Table 3-2 and Figure 3-1 show the effects of the approximation derived by the 1st order Taylor expansion around the estimated random effects, $\hat{b}_i$, on the estimated population parameters, $\hat{\beta}$, by changing a total sample size (N= 25, 50, and 100), and a number of observations per subject (n_i = 2, 3, 4, 5, to 6). On the condition that we observe variability in the fitted curve, $Var[b_i]$, among subjects, Figure 3-1 indicates that this approximation method produces unbiased population estimators. As the number of observations per subject increase, the confidence interval decreases accordingly. In comparison of the results from N= 25 and from N=100, the precision of the estimated population parameters on a number of observations per subject will depend on a total number of subjects. For example, in the case of N = 25, the confidence interval decreases beyond n_i = 4. On the other hand, in the case of N= 100, the size of the confidence interval does not depend on a number of observations per subject. Equation (10) suggests these relationships. That is, the precision of the estimated parameters will depend on a ratio of the number of observations per subject to the total number of subjects.

Minami, 2002 discussed the results derived from the approximation by the 1st order Taylor expansion around the expected values, 0 when we can not ignore the variations among subjects. The approximation produces biased estimators. The observed bias does not depend on the number of observations per subject. But the confidence interval of the estimators decrease as the total number of subjects increase within the same number of observations per subject. The results suggest that for a fixed total number of observations, say, 200, it would be better that 2 observations per subject are taken from 100 subjects, instead, 4 observations per subject are taken from 50 subjects to produce a smaller confidence interval.

This paper investigated the characteristics of the two types of approximation approaches to the marginal distribution of the observed data by the 1st order Taylor expansion when we observe the variability among subjects. The approximation method by the 1st order Taylor expansion around the expected values, 0, produces biased population parameter estimates, whose size of the confidence interval is subject to a number of subjects. On the other hand, the approximation method by the 1st order Taylor expansion around the estimated random effects produces unbiased estimators, whose size of the confidence interval depends on a ratio of a number of observations per subject to a total number of subjects. In actual clinical practice, it is very difficult to take many observations per subject, especially in phase III studies with possible heterogeneity of study subjects, indicating the variability of responses among the subjects. Therefore we have to deal with the problem by a study design of population pharmacokinetics studies. When we expect the variations, we should confine a homogeneous population, for which we can sample data from the population by a few number of observations per subject from many subjects. Otherwise many observations per subject have to be sampled according to an appropriate total number of subjects derived from equation (10).

We investigate a goodness of fit test to the observed data for the two types of the approximation methods because actual clinical trial settings mostly will not allow the conditions by either ignorable variation among subjects or the conditions by equation (10).

Reference:

Breslow,N.E. and Clayton,D.G.(1993). Approximate inference in generalized linear mixed models. *J. Am. Statist. Assoc.* **88**, 9-25.

International Conference on Harmonisation of (ICH) Harmonized Tripartite Guideline. Ethnic Factors in the Acceptability of Foreign Clinical Data. Recommended for Adoption at Step 4 of the ICH Progress on 5 February 1998 by the ICH Steering Committee

Laird,N.M. and Ware,J.H.(1982). Random effects models for longitudinal data. *Biometrics* **38**. 963-974.

Lindstrom, M.J. and Bates, D.M. (1990). Nonlinear mixed effects models for repeated measure data. *Biometrics* **46**, 673-87.

Minami, F.(2002). The master thesis, Kitasato University Graduate School

Sheiner, L.B. and Beal, S.L. (1980). Evaluation of methods for estimating population pharmacokinetic parameters. I. Mechaelis-Menten model: routine clinical pharmacokinetic data. *J.Pharmacokin. Biopharmac.***8**, 553-71.

Sheiner,L.B., and Beal, S.L. (1985). Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. *J. Pharmacokin. Biopharmac.* 13, 185-201.

Solomon, P.J. and Cox, D.R. (1992). Nonlinear component of variance models. *Biometrika* **79**, 1-11.

The International Conference on Harmonisation of Technical Requirements for the

Registration of Pharmaceuticals for Human Use (ICH) <http://www.ifpma.org/ich.1.html>

Vonesh, E.F. and Carter, R.L. (1992). Mixed-effects nonlinear regression for unbalanced repeated measures. *Biometrics* 48, 1-17.

Wolfinger, R. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika* 80, 791-5.