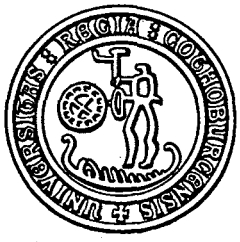

Research Report
Statistical Research Unit
Göteborg University
Sweden

Preliminary testing in a class of simple non-linear mixed models to improve estimation accuracy

Max Petzold

Research Report 2003:9
ISSN 0349-8034

Mailing address:	Fax	Phone	Home Page:
Statistical Research Unit	Nat: 031-773 12 74	Nat: 031-773 10 00	http://www.stat.gu.se/stat
P.O. Box 660 SE 405 30 Göteborg Sweden	Int: +46 31 773 12 74	Int: +46 31 773 10 00	

Preliminary testing in a class of simple non-linear mixed models to improve estimation accuracy

by Max Petzold

Department of Statistics, Göteborg University, Göteborg, Sweden.

Summary

In applied research hypothetical information about the parameters in a stochastic model sometimes can be generated from theory or previous studies. Replacing unknown parameters by constants might increase the estimation accuracy. This is especially apparent when replacing parameters in non-linear expressions. The problem is how to handle the uncertainty of the hypothetical information. Here, a pretest procedure will be examined for an unknown exponent of the explanatory variable in a simple non-linear mixed model. The optimal pretest sizes for some parameter settings are found for a minimax regret criterion based on Mean-Squared-Error. The optimal test sizes were found to be approximately valid also for the case where no subject specific components are present. The examined class of models is useful for modelling concentration-time data for drugs with rapid absorption, and a small-sample example is given to illustrate the potential gain in estimation accuracy of the pretest approach in pharmacokinetics.

Keywords: HIV; Longitudinal; Monoexponential; Pharmacokinetics; Preliminary test; Random coefficient regression; Small sample.

1. Introduction

Repeated measurements on subjects over time are collected in many research areas, and the information is then summarized for different purposes. In e.g. bioequivalence studies of drugs, measurements such as area under curve (AUC) and concentration decline rate are estimated from repeated concentration-time data on individuals.[1, 2] The proposed systematic relationship between the outcome and the explanatory variables might be non-linear due to unknown parameters.[3, 4] A disadvantage when estimating in non-linear models is the need to use either analytical approximations or numeric techniques. The knowledge about the exact properties of these techniques is limited and they are associated with practical drawbacks such as convergence problems. Enabling the use of well-known and possibly more efficient estimators by restricting the model to a given linear relationship in its unknown parameters might be valuable.

A linear model can be obtained if exact information about the parameters causing the non-linearity is incorporated. In reality only uncertain information about these parameters might be available, and the use of incorrect information can cause misleading results. However, if there is a large difference in efficiency of the estimators, the use of also slightly incorrect information can still be motivated to obtain larger estimation accuracy. The problem is then how to handle the uncertainty of the information.

In this paper a pretest procedure which incorporates hypothetical information into the model estimation will be studied. Typically, a pretest procedure consists of a preliminary test on a specific parameter. The null-hypothesis can be generated from theory and an earlier research process, while the alternative hypothesis is based on the current sample. If the null-hypothesis is not rejected, the other parameters are estimated under the null-hypothesis. The pretest size can be optimised for certain criteria.

Pretesting has a long history[5, 6] and has been extensively discussed in Judge and Bock[7]. In regression, estimation of a conditional mean after a pretest for pooling two independent samples has been studied by Bhoj et al[8] for example. If the poolability was rejected, the estimator of the mean was based on only one of the samples. In an example a large potential to improve the estimation accuracy in terms of reduced Mean-Squared-Error (*MSE*) was found for the pretest estimator compared to the one-sample estimator.

Here, the aim is to improve the estimation accuracy of the population parameters in a simple mixed model setting with an unknown parameter in a non-linear expression. Uncertain hypothetical information about the parameter will be incorporated using a pretest procedure. To the author's knowledge this has not previously been done for non-linear mixed models. In a related work[9] the significance of the quadratic term in a polynomial model of order two was tested as a preliminary test for non-linearity. However, this is a different test situation and the generalisation to longitudinal models was not treated.

In the next section, the model is described in detail and motivated by examples. The model is simple in the way that there is only one explanatory variable, and non-linear in the sense that the explanatory variable has an unknown exponent p. This model is chosen for simplicity to illustrate the potential of pretest procedures, but is useful for mono-exponential relationships between the measured response and the explanatory variable. In Section 3 the estimators of the linear model (p known) and the non-linear model (p unknown) are described. The variances of the corresponding estimators are compared over different values of p. Motivated by the differences in variance, a pretest procedure is introduced in Section 4 and the minimax regret optimal pretest sizes are found for some small-sample parameter settings. The relation to the case when no subject specific components are present is examined. A brief

example of pretesting in pharmacokinetics is given. Section 5 is a concluding discussion. The estimators are defined in Appendix I, and the asymptotic variances of the estimators are given in Appendix II.

2. The simple mixed model

In order to keep the examination of the pretest as clear as possible, the study will be limited to a simple mixed model. The model allows for random intercepts over cross-sectional units but have a fixed slope:

$$Y_{ij} = B_{0j} + \beta_1 \cdot t_i^p + e_{ij},\ i = 1 \ldots T,\ j = 1 \ldots n \qquad (1)$$

where Y_{ij} is the response at t_i for the j:th subject. The random intercept B_{0j} reflects factors which are specific for the j:th subject, and the B_{0j}'s and the e_{ij}'s are assumed to be independently and normally distributed. It then follows that the vector $\mathbf{Y}_j = (Y_{1j} \ldots Y_{Tj})'$ has a T-dimensional normal distribution. Further, let the expected value be $E(\mathbf{Y}_j \mid \tilde{\mathbf{X}})_{T\times 1} = \tilde{\mathbf{X}}(\beta_0 \mid \beta_1)'$ and let the variance be $V(\mathbf{Y}_j \mid \tilde{\mathbf{X}})_{T\times T} = \sigma_{B_0}^2 \mathbf{1}\mathbf{1}' + \sigma_e^2 \mathbf{I}$, where $x_i = t_i^p$ and $\tilde{\mathbf{X}} = \left(\mathbf{1} \mid (x_1 \ldots x_i \ldots x_T)' \right)_{T\times 2}$. The design is balanced in the sense that $\tilde{\mathbf{X}}$ is the same for all units.

Longitudinal studies are common in research areas such as pharmacokinetics and pharmacodynamics. The expression $\exp(Y_{ij})$ has a multiplicative structure according to (1) which is appropriate when the variability of the observed values increases with the size. This is a phenomenon frequently apparent in e.g.

human immunodeficiency virus (HIV) data[10] and in concentration-time data[4]. Examples of curve shapes for different values of p are given in Figure 1.

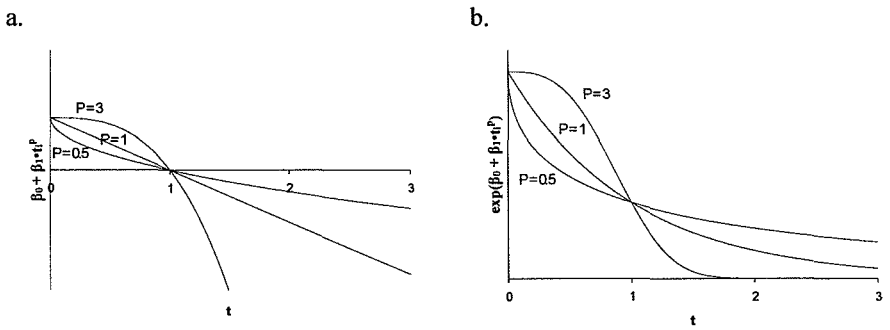

Figure 1. Examples of curve shapes for different values of p; a.) $\beta_0 + \beta_1 \cdot t_i^p$ and b.) $\exp(\beta_0 + \beta_1 \cdot t_i^p)$.

The $\exp(Y_{ij})$-model has been used in HIV-studies.[11, 12] In such studies it is normally assumed that two infected cell compartments (productively infected cells and long-lived infected cells) can be identified based on plasma viral load data. Each of these two compartments is believed to produce a viral decay during treatment with potent antiviral therapies which can be modeled similar to $\exp(Y_{ij})$. β_1 is then the decay rate for the compartment, and it has been suggested that the decay rate reflects the potency (efficacy) of the antiviral therapies.[13] Accurate estimation of the decay rates is thus important in bioequivalence studies of different regimens.[14]

Further, in pharmacokinetics models similar to $\exp(Y_{ij})$ have been suggested for concentration-time studies. In e.g. Davidian and Giltinan[15] Chap 5.2.4 the model was used over each of the two apparent monoexponential phases of plasma concentrations following intravenous injection of indomethacin. The model has also been suggested for modelling the total concentration-time

relationships for drugs where instantaneous distribution between plasma and tissues takes place. In Gibaldi and Perrier[16] Chap. 1 examples of modelling concentration in plasma after an intravenous dose of prednisolone, and in serum and heart tissue after an oral dose of dipyridamole are given. The population AUC can be found by integrating the expected value of $\exp(Y_{ij})$ over time (cf. Gibaldi and Perrier[16] p. 14). Since $\exp(Y_{ij})$ has a lognormal distribution longitudinal models which separate the between-subject variance from the within-subject variance are required when estimating the AUC.

In the referred studies above a strong assumption of linearity was imposed when the exponent p was put to unity, i.e. t^1. Even if this value is well-founded, it may be incorrect in a new study where the circumstances might have changed slightly. The common more flexible non-linear alternative, i.e. the use of a general real-valued exponent p estimated from the current data, was discussed early by Box and Tidwell[17] for a monotonic model like (1). The latter model was recently motivated and further developed for applications in medicine and epidemiology[18, 19]. However, introducing an unknown p means that we have a model which is non-linear in its parameters, and that we have an extra parameter to estimate. The estimator of the latter model can thus be suspected to be less efficient. By introducing a pretest procedure for p we can combine the use of uncertain hypothetical information with a more data-dependent flexible modeling when H_0 is rejected. An example where the pretest procedure is used for accurate and reliable estimation in bioequivalence studies is given in Section 4.3.

For a more detailed characterization of viral decay rates or drug concentrations in body more detailed models are needed. In such case the parameters in (1) are regarded as functions of underlying differential equations.[3, 10] For simplicity this kind of models is not treated here.

3. Relative efficiency of the estimators

Different estimators of the unknown parameters in (1) have to be used for the linear and non-linear models, respectively. For the linear model (p known), the Maximum Likelihood (ML) estimators from Appendix IA can be used. These estimators are denoted with a single hat. For the non-linear model (p unknown) simultaneous ML estimators can be found by maximizing the likelihood directly.[20] Having a balanced design it is easy to find the ML solution from the derivative of the likelihood function with respect to p, cf. Appendix IB. The estimators in the non-linear case are denoted with a double hat.

It is the purpose of this section to examine the differences in variance of the corresponding estimators in the linear and non-linear cases, respectively. The size of the differences indicates the potential gain of using a pretest approach. The relative precision of the estimators of a parameter θ is examined using the ratio $R = V(\hat{\theta})/V(\hat{\hat{\theta}})$.

In the examples, the parameter settings were chosen by practical experience. The study was concentrated to the interval $0.5 \le p \le 3$ which covers a wide range of curve shapes adequate for the applications in this paper, see Figure 1. The t_i's were chosen as equally spaced on the interval [1,10]. It should be noted that the sample design is important for the efficiency of the estimators. This has been studied specially for pretesting in linear regression[21] and more generally also in non-linear regression[22], but is not treated here.

3.1 Asymptotic variances of the ML estimators

The asymptotic variances and covariances (n large) of the ML estimators can be obtained from the expectations of the second order derivatives of the Likelihood in (2), cf. Appendix II. It can easily be seen that the asymptotic

variances for the linear and non-linear model differ, except for the estimators of $\sigma^2_{B_0}$ and σ^2_e which have the same variance in both cases. It can further be seen that the *R*-ratio for the estimators of β_1 does not depend on the variances, while the *R*-ratios for β_0 and the mean function $E[Y_{ij}] = \beta_0 + \beta_1 x_i$ do depend on the quotient $\sigma^2_{B_0} / \sigma^2_e$ but not on the absolute magnitude of the variances. Finally, from Appendix II it can be seen that none of the three ratios depend on n.

However, to get an idea about the losses in efficiency, the actual *R*-ratios over p have to be calculated in an example. Here, three quotients of the variances were chosen, $\sigma^2_{B_0} / \sigma^2_e = 0.5,\ 1.0$ and 2.0, which are of the same order as the estimated quotient $\hat{\sigma}^2_{B_0} / \hat{\sigma}^2_e = 1.98$ from the example in Section 4.3.

When calculating the *R*-ratios they were found to be markedly below unity for the estimators of β_0 and β_1. Considering the former, it can be seen from Figure 2a that R is an increasing function of p which starts at a small value but ends at a fairly large value. For the estimators of β_1, the value of R was found to be small with a local maximum of $1.1 \cdot 10^{-2}$ at $p = 0.90$ (cf. Figure 2b). The large difference in variances of the β_1-estimators is quite remarkable. However, it is important to remember that many common functions such as the mean curve and the tolerance limits do not involve β_0 and β_1 separately. From Figure 2c it can be seen that the *R*-ratio for $E[Y_{1j}]$ is fairly large for all values of p, and always larger than the corresponding *R*-ratios for β_0 and β_1, respectively. This means that the relative increase in variance in the non-linear case is smaller for the combined estimator than for the single estimators. This result is due to the fact that the two estimators compensate each other which results in a relatively large *R*-ratio for the combined estimator. For simplicity

only results for the first time point x_1 are given here. From Figure 2a and 2c it can also be seen that R is positively related with the quotient of the variances, i.e. a relatively smaller σ_e^2 tends to decrease the difference in variance of the estimators.

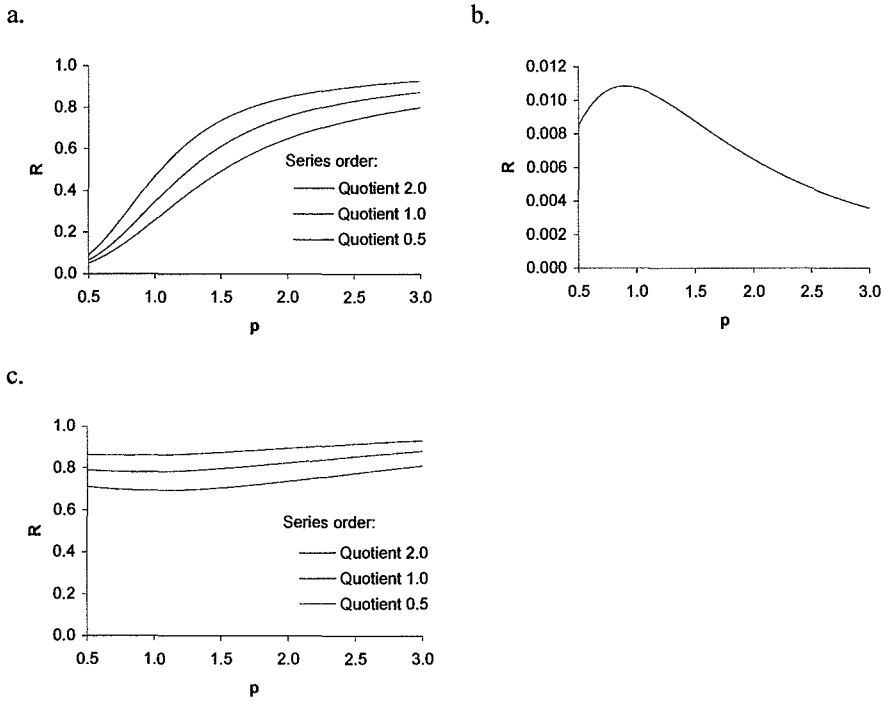

Figure 2. The R-ratio in (2) when estimating a.) β_0, b.) β_1 and c.) $E[Y_{1j}]$ for the three quotients of $\sigma_{B_0}^2 / \sigma_e^2$ where $n = T = 10$.

3.2 Small sample results

The small-sample situation is important in many research areas such as pharmacokinetics where the sample sizes and the number of observations per subject often are limited in early phases due to practical and ethical considerations To verify that the results for the asymptotic variances in Section 3.1 are relevant also in the given small sample situation, a simulation study

(100.000 replicates) was performed for $p = 0.5, 0.6, \ldots, 3.0$. The numeric estimation procedure introduced in Appendix IB was confirmed to give unbiased estimators, with estimated variances close to the asymptotic variances from Appendix II. In the non-linear case, the maximal relative differences in variance were found to be about 2% for the estimators of β_0, β_1 and $E[Y_{1j}]$, respectively. The corresponding differences in the linear case were found to be less than 0.1%. Thus, the results in Section 3.1 are assumed to be applicable in the given small sample situation.

4. Incorporating uncertain hypothetical information in the estimation procedure

In Section 3 it was found that the estimators of the non-linear model can have a considerably larger variance than the corresponding estimators of the linear model. The loss in efficiency was also found to depend on the value of p. In many situations no information about p is available, and the less efficient estimators of the non-linear model have to be used. Methods for improving the estimation accuracy in such situations include shrinkage and penalized likelihoods which were covered in a recent review.[23] However, in this paper situations where hypothetical information about p is available are treated. The results in Section 3 then point out that it can be valuable to incorporate such information facilitating the use of the efficient estimators of the linear model.

A common way to incorporate hypothetical information in general is to use Bayesian methods, and different aspects of this have been studied extensively in pharmacokinetics[24, 25]. In the Bayesian approach we have prior beliefs about the distributions of the population parameters, and we incorporate the new sample still using estimators of the non-linear model to obtain our posterior distributions. An example of the latter can be seen when modelling the viral

decay in a HIV study[12, 26], but this is not the situation studied in this paper. Here, our aim is to find a more accurate estimation approach by using the estimators of the linear model when appropriate. When the beliefs about p is based on a previous sample the empirical Bayes approach can be used as an alternative where the estimates of the linear and the non-linear models are weighted.[27] However, in this paper our hypothetical information is considered as generated from theory and a previous research process, and not from a single sample.

Here, a frequentistic pretest approach will be used to incorporate the hypothetical information about p. First, a pretest estimator is introduced which is based on both the estimators of the linear model calculated given a hypothetical p_0, and the estimators of the non-linear model. Second, since p_0 is not necessary equal to p an optimal test size is discussed in terms of the *MSE* of the estimator. This is a combined measure of variance and bias which captures the accuracy, or closeness, of the estimator. Last, the pretest estimator is examined using the suggested optimal test size in a concentration-time data example.

4.1 The pretest approach

Having a hypothetically value p_0 the null-hypothesis $H_0 : p = p_0$ can be tested. A pretest estimator of a parameter θ in model (1) can be defined as:

$$\hat{\theta}^{\alpha} = \begin{cases} \hat{\theta} \text{ if } H_0 \text{ is not rejected at test size } \alpha \in [0,1] \\ \hat{\hat{\theta}} \text{ otherwise} \end{cases}$$

where the linear model under H_0 is estimated given p_0. Estimating e.g. β_0 the pretest estimator will be $\hat{\beta}_0^{\alpha} = \hat{\beta}_0$ when H_0 is not rejected, and $\hat{\beta}_0^{\alpha} = \hat{\hat{\beta}}_0$ otherwise.

Let the mean squared error of $\hat{\theta}^{\alpha}$ be $MSE(\alpha) = E\left[(\hat{\theta}^{\alpha} - \theta)^2\right]$. Further, let the hypothesis error be $d = p_0 / p$ which is equal to unity if H_0 is true. Some general properties of $MSE(\alpha)$ then follows from Judge et al[28] Chap. 3.3.1a-c which to some extent explain the results in the sequel:

1. $\alpha = 0$ and $\alpha = 1$ correspond to estimating under H_0 and H_1, respectively.
2. $MSE(0) < MSE(\alpha)$ when $\alpha > 0$ and $d = 1$.
3. $MSE(\alpha) \to MSE(1)$ when $\alpha > 0$ and $d \to \pm\infty$, i.e. for large deviations from H_0.
4. For $0 < \alpha < 1$ the $MSE(\alpha)$ first increases and obtains a maximum larger than $MSE(1)$ and then monotonically decreases to $MSE(1)$ as d diverges from unity.
5. $MSE(\alpha) \geq \min\{MSE(0), MSE(1)\}$, i.e. the lower limit of $MSE(\alpha)$ will always be the MSE of the estimators of either the linear or the non-linear models.

The distribution of $\hat{\theta}^{\alpha}$ is complicated, and the $MSE(\alpha)$-curves were here examined using simulations over a d-interval sufficiently large to catch the relevant information. H_0 was rejected for large values of the pretest statistic $-2\log\lambda$ where λ is the ratio of the ML functions under H_0 and H_1, respectively. Using critical values from the chi-squared distribution with one degree of freedom the test was found to approximately hold the nominal test

size. Over all examined parameter settings in the sequel, the maximal difference was 8% of the nominal test size α.

As an illustrative example, some $MSE(\alpha)$'s for $E[Y_{1j}]$ when $p=1$ and $\sigma^2_{\beta_0}/\sigma^2_e = 0.05/0.10 = 0.5$ are given in Figure 3. For $d=1$, i.e. $p_0 = p = 1$, $MSE(0)$ is smallest and $MSE(1)$ is largest as expected. This was found to be true for all d on the interval [0.961;1.038], with the equality $MSE(0) = MSE(1)$ at the end points. It can also be seen that $MSE(\alpha)$ for all $\alpha > 0$ tend to be equal to $MSE(1)$ already for relatively small deviations of d from unity. Note that the curves in Figure 3 are not fully symmetric. The general findings in Figure 3 are valid also for the other parameter settings in this paper.

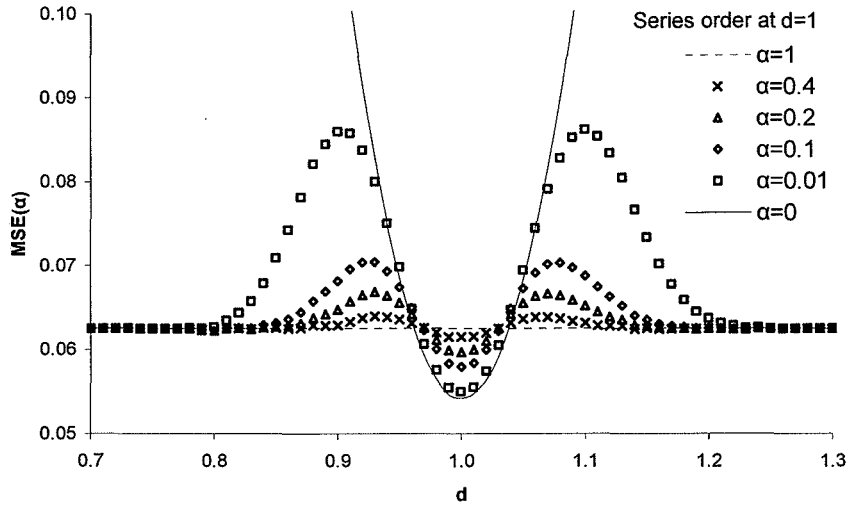

Figure 3. $MSE(\alpha)$ for $E[Y_{1j}]$ when $p=1$.

In Figure 3, p was equal to unity. However, from Figure 4a it can be seen that the value of d where $MSE(0) = MSE(1)$ for $E[Y_{1j}]$ is a decreasing function of p for $d > 1$. Approximately this gives the length of the d-interval where $MSE(0) < MSE(\alpha)$ for $\alpha > 0$ which thus can be seen to be decreasing with larger values of p and to be very short for $p \geq 2$. This length has in a further simulation study been found not to be dependent on $\sigma^2_{\beta_0}$, but to be a decreasing function of n and T and to be an increasing function of σ^2_e. This result is to some extent surprising since we know from Section 3 that the R-values are dependent on the quotient $\sigma^2_{\beta_0} / \sigma^2_e$, i.e. both variances.

4.2 Optimal pretest size

There are several common optimality criteria for pretesting, see e.g. Judge and Bock[7] Chap. 3.3.3. Here, a minimax regret solution will be considered where the regret, as a function of d, is defined as $REG(\alpha) = MSE(\alpha) - \min\{MSE(0), MSE(1)\}$.[29] The minimax optimal value of α, say α', is the value which minimises the maximum possible regret; $\sup_d REG(\alpha') \leq \sup_d REG(\alpha)$ for $d \in (-\infty, +\infty)$ and all $\alpha \geq 0$. This criterion is reasonable for many applications since it avoids large losses in estimation accuracy both when d is close to unity and when it diverges more. Considering a situation where p_0 might be well-founded from an extensive previous research it is obvious that e.g. a very risk adverse criterion that just minimises max(MSE) would be inappropriate. The solution for the latter criterion would be to use $\alpha' = 1$, since $\max(MSE(1)) \leq \max(MSE(\alpha))$ over $d \in (-\infty, +\infty)$. This would be a relatively inefficient solution if we can assume that p_0 is close to p.

The maximum regret of $E[Y_{1j}]$ can be seen in Figure 4b as a function of α. In this example the minimum of the maximum regret-function can be found approximately for $\alpha' = 0.17$. Using the minimax regret optimality criterion the value of α' is determined by the shape of the *MSE*-curves. For $d = 1$ the *MSE*-values are solely determined by the differences in variance of the estimators, and estimators with small *R*-values (cf. Figure 2) will then have relatively large regrets. When $d \neq 1$ the shape of the *MSE*-curves is more complicated since it also depends on the bias, and α' thus have to be examined in a simulation study.

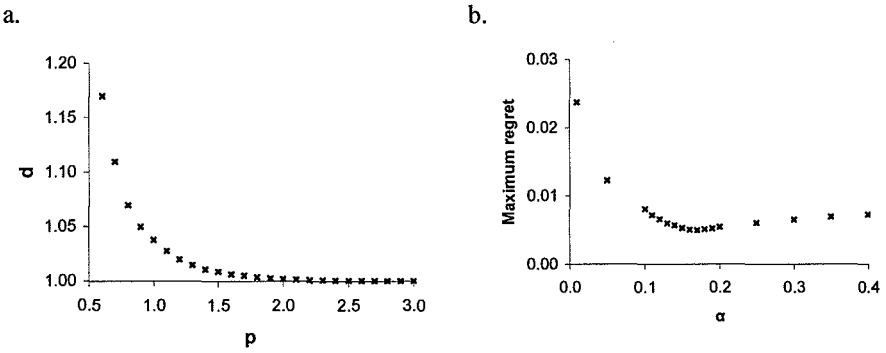

Figure 4. In a.) the values of d where $MSE(0) = MSE(1)$ over p for $E[Y_{1j}]$, and in b.) the maximum *MSE* regret over d as a function of α when $p = 1$.

The minimax regret solution α' was examined in a simulation study for a variety of practically relevant parameter settings. The results do not indicate that α' depends on $\sigma^2_{\beta_0}$ and σ^2_e, but on n, T and p for which the values are given in Table 1. The dependences are expressed differently for the different parameters. From Table 1 the value of α' can be seen to be relatively large and stable for the estimator of $E[Y_{1j}]$ over p. This can probably partly be explained by the corresponding relatively large and stable value of R in Figure 2c. Compared to $E[Y_{1j}]$ we can see that α' for β_0 and β_1 is smaller for small

values of p, and approximately the same for larger values of p. From Figure 2a-b we can then see that the explanation of the results in Table 1 is more complicated than only the differences in variance. Extreme parameter values were not treated here, but $n=5, T=100$ and $n=100, T=5$ were included for comparing α' in pre-clinical and clinical study settings, respectively. However, for the cases studied here no differences regarding α' were found. This is interesting since the values of n and T affect both the MSE's and variance and bias of the estimators to different extent. The main conclusion from Table 1 is that the level of α' is rather stable for the different settings and that it is relatively large compared to the nominal test sizes normally chosen in ordinary test situations.

Table 1. Minimax regret optimal test sizes α' found in a simulation study.

p	n,T	α' for β_0	α' for β_1	α' for $E[Y_{1j}]$
0.5	$n=T=5$	0.10	0.09	0.18
	$n=T=10$	0.14	0.13	0.18
	$n=100,\ T=5$	0.16	0.16	0.19
	$n=5,\ T=100$	0.16	0.16	0.19
1.0	$n=T=5$	0.16	0.15	0.17
	$n=T=10$	0.15	0.15	0.17
	$n=100,\ T=5$	0.16	0.16	0.19
	$n=5,\ T=100$	0.16	0.16	0.19
≥2.0	$n=T=5$	0.17	0.17	0.17
	$n=T=10$	0.17	0.17	0.17
	$n=100,\ T=5$	0.17	0.17	0.19
	$n=5,\ T=100$	0.17	0.17	0.19

Having a balanced design, the estimation of the population parameters in mixed models is related to the case when no subject specific components are present, especially for $\sigma^2_{B_0} = 0$. This relation was first indicated in Section 3 where it was found that the *R*-ratio for β_1 does not depend on the variances, while the *R*-ratios for β_0 and $E[Y_{ij}]$ depend on only the quotient $\sigma^2_{B_0}/\sigma^2_e$. In Table 1 a stronger indication of the relation was found since the results were found not to be dependent on the variances at all. The generalizability of the results in Table 1 was further revealed in a study where $\sigma^2_{B_0}$ was put to zero. Approximately the same results as in Table 1 were obtained also in this particular situation.

4.3 An example of pretesting in pharmacokinetics

To illustrate the potential gain of using a pretest procedure we will consider an early phase study of the pharmacokinetics of indomethacin following bolus intravenous injection of the same dose in 6 individuals.[30] For each subject, plasma concentrations of indomethacin were measured at 11 time points ranging from 15 minutes to 8 hours post-injection. This data set was used by Davidian and Giltinan[15] in Chap. 5.2.4 where they represented each of the two apparent exponential phases of drug disposition by a mono-exponential function like $\exp(Y_{ij})$. Here, the 5 observations equally spaced between 15 minutes and 75 minutes are chosen as the first exponential phase for illustrative purposes. In Figure 5, the logarithm of the measured concentrations and the estimated mean functions $E[Y_{ij}]$ for the linear model (where $p_0 = 1$ as in Davidian and Giltinan) and the non-linear model ($\hat{\hat{p}} = 0.461$) are shown. The sums of squares of errors for the two mean functions are relatively similar, 2.01 and 1.90 respectively, which verifies that the chosen time interval is reasonable for representing the first phase.

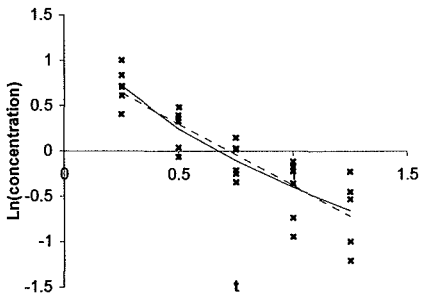

Figure 5. $E[Y_{ij}]$ for $\hat{\hat{p}} = 0.461$ (solid line) and $p_0 = 1$ (dashed line) and the logarithm of the concentrations in the example.

The gain of using the suggested pretest procedure can be examined by comparing the *MSE*'s for the linear and non-linear models, respectively, with the *MSE* of the pretest estimators using the minimax regret optimal test size. In a simulation study the parameter settings $\sigma^2_{B_0} / \sigma^2_e = 0.049/0.025$ and $p = 0.461$ were used as estimated from the data in Figure 5. Three different null-hypotheses were tested, $p_0 = 0.461,\ 0.5,\ 1.0$, where $p_0 = 0.5$ was included since it would be a natural ad-hoc value from looking at the data. The optimal test sizes were found in Table 1 for $p = 0.5$.

The results can be found in Table 2. For $p_0 = p = 0.461$ the results were expected from the results in Figure 2. The $MSE(0)$'s from estimating β_0 and β_1 should then be much smaller than the corresponding $MSE(\alpha')$'s and $MSE(1)$'s, while a smaller difference was expected when estimating $E[Y_{1j}]$. From the results in Figure 4a we know that the value of d can diverge relatively much from unity when p is small and still having the inequality $MSE(0) < MSE(\alpha)$ for $\alpha > 0$. In Table 2 this is clear regarding the results when $p_0 = 0.5$. $MSE(0)$ is then still smallest, but the differences have decreased. For $p_0 = 1$ we can see that the power is almost 100% which gives

$MSE(\alpha') \approx MSE(1)$. However, all three $MSE(0)$'s are now very large, even when estimating $E[Y_{1j}]$. The decrease in *MSE* from the stabilizing effect of using a fixed p_0 is now erased by the bias of the estimators.

Table 2. The ratio $MSE(\alpha)/MSE(\alpha')$ where $p = 0.461$.

Estimated parameter:	α' from Table 1:	$p_0 = 0.461$		$p_0 = 0.5$		$p_0 = 1.0$	
		$\alpha = 1$	$\alpha = 0$	$\alpha = 1$	$\alpha = 0$	$\alpha = 1$	$\alpha = 0$
β_0	0.10	1.63	0.062	1.19	0.18	1.00	5.88
β_1	0.09	1.66	0.024	1.20	0.16	1.00	6.60
$E[Y_{1j}]$	0.18	1.04	0.91	1.02	0.91	1.00	5.92

The main conclusions from Table 2 are 1.) also incorporating relatively poor hypothetical information can decrease the *MSE* substantially compared to always estimating the non-linear model, and 2.) the risk in terms of *MSE* of using the linear model can be large also for reasonable choices of p_0. However, it can be argued that the main interest is in the mean function for which the *MSE* only to a limited degree depends on the choice of model, and that the *MSE's* when estimating β_0 and β_1 separately are of less interest. This is partly true, since $E[Y_{ij}]$ is important for calculating e.g. AUC in both pharmacokinetics and in pharmacodynamics.[31] Nevertheless, also the single estimator of β_0 is important when e.g. adjusting the AUC for baseline.[32], and β_1 is e.g. used for estimating viral decay rates[32] and for shelf-life studies of drugs[33].

5. Discussion

It was the purpose of this paper to discuss accurate estimation strategies in non-linear mixed models. To emphasise on the inferential issues of the pretest, the study was limited to a simple non-linear mixed model setting. First it was found that the relative difference in variance of the estimators of the linear model (p known) and the non-linear model (p unknown) can be considerably large. The difference is especially large for some estimators and some values of p. The problem was then how to utilize the efficient estimators of the linear model when p is unknown. Since uncertain hypothetical information about p might be available in e.g. bioequivalence studies, a pretest procedure based on both models was introduced. Optimal pretest sizes were found for a minimax regret criterion, and the benefits of the pretest approach were illustrated in a study of the pharmacokinetics of indomethacin. For $d=1$ it was found that the *MSE* was decreased by up to 40% when using the pretest estimator instead of applying the non-linear model. The decrease was even larger, 85%, when comparing the pretest estimator to a reasonable linear model suggested in the literature. The suggested optimal pretest sizes were also found to be approximately valid for the case when no subject specific components are present. Having a balanced design this result was to some extent expected since the estimators of the longitudinal model then coincide, cf. Longford[34] Chap. 2.7. However, in Section 3 it was found that the difference in efficiency of some of the estimators depends on both $\sigma^2_{\beta_0}$ and σ^2_e, which could have influenced the generalizability.

In Section 4 the linear model was tested versus the less restricted non-linear model. The assumptions about the restrictions are crucial for the subsequent inference. Here, p_0 was assumed to be generated from theory and an earlier research procedure and to be exogenous, i.e. independent of the stochastic variation in the present data sample. It should not be mixed up with

endogenous values from formal and informal modelling procedures. An example of the latter is when the value of p_0 is found by plotting and comparing the data with known functions (often from a limited ad hoc set as the square root, the linear, the quadratic and the cubic). Since this is a kind of pretesting in itself, the inference of the following formal pretest procedure will then be unclear. It should also be clarified that no absolute knowledge about p is assumed here, since it would then be inefficient to use pretesting.

There are several optimality criteria[7] of pretests and generally a relatively large test size is suggested. Three common criteria are: 1.) The minimax regret criterion which was used by Ohtani[35] when pretesting a linear hypothesis with the aim of estimating the residual variance. An optimal test size of about 30-70% was then suggested. The minimax regret criterion was used in this paper due to its appealing minimisation of the maximal loss compared to using the optimal solution for given d. Compared to the two following criteria, this gives a solution which can be considered as more useful for situations where relatively accurate information about p might be available. 2.) Defining the relative accuracy as $MSE(1)/MSE(\alpha)$ another criterion is to choose the test size that gives the maximum relative accuracy among the test sizes which guarantee at least a certain minimum relative accuracy over d. In Bhoj et al[8] and in Khan et al[36] examples were given where the optimal test sizes were found to be 20% and 35% respectively. However, when using the single hypothesis pretest procedure in this paper it can easily be seen that the value of α' would then solely be determined by the subjectively chosen minimum relative accuracy. Both the subjectiveness and the fact that the optimisation is not at all based on the regret when d is close to unity can be questioned. 3.) A third criterion would be to minimize the average $MSE(\alpha)$, or equivalently the average regret, over d.[37] However, optimising over $d \in (-\infty,\infty)$ we will always chose $\alpha' = 1$ since $MSE(\alpha) \to MSE(1)$ when $\alpha > 0$ and $d \to \pm\infty$. This

is not a satisfactory solution in situations where p_0 is likely to be close to p, which can be expected in areas of extensive research. A pretest size of 100% is then inefficient since it does not enable the use of the efficient estimators of the linear model although the risk of a large *MSE* is small.

In this paper only the properties of the pretest estimator has been examined. Another important issue is the inference of a main test following after a pretest. For example, the inference when testing the decay rates of two drugs after pretesting the exponent p will not be straight forward. As noted by Greenland[38] in a discussion of reanalysis of epidemiologic databases using pretesting[39], one has to construct confidence intervals and interpret tests results obtained from a likelihood function chosen by preliminary testing carefully. It has e.g. been shown that pretest estimators potentially have asymptotic non-normality.[40]

Further, it is well-known that the size of the main test may be inflated by the pretest procedure. This has been studied, e.g. for pretesting for non-linearity in regression followed by a main test for association between two variables, in Grambsch et al[9]. They found that the size of the main test increased by roughly 50% and suggested simple modifications of standard practice to protect the size with only minimal loss of power. Pretest procedures are sometimes motivated by giving a larger power of the following main test. The difference in size and power between a test following after a pretest procedure and the test in the complete family has been studied by Albers et al[41]. They used a preliminary test of equality of variances in two samples followed by a main test of the means. They found that the power change is often mainly nothing else but a factor times the size change. This implies that a larger power is only obtained if the size exceeds the nominal size. However, it has been shown that an increased sample size can reduce the consequences of pretesting with respect to size and power.[42]

Here, a class of simple non-linear regression models has been examined for a balanced design. A study of more general models including also other aspects of modelling such as prediction outside the study interval and tolerance intervals would be valuable. The study could also be extended to other parameters than the exponent and compared to the use of linearization procedures[15].

Acknowledgement

I would like to thank Robert Jonsson PhD, Prof. Marianne Frisén, Thomas Holgersson PhD and Magnus Pettersson PhLic for valuable comments during the progress of the paper.

Appendix I – The estimators

Some denotations to be used below:

$$\bar{Y}_j = T^{-1}\sum_{i=1}^{T} Y_{ij}\,, \quad \bar{Y} = n^{-1}\sum_{j=1}^{n} \bar{Y}_j\,, \quad S_{\bar{Y}\bar{Y}} = \sum_{j=1}^{n} (\bar{Y}_j - \bar{Y})^2\,, \quad S_{YY} = \sum_{j=1}^{n}\sum_{i=1}^{T} (Y_{ij} - \bar{Y}_j)^2\,,$$

$$\bar{x} = \sum_{i=1}^{T} x_i / T\,,\ S_{xx} = \sum_{i=1}^{T} (x_i - \bar{x})^2 \text{ and } S_{xY} = \sum_{j=1}^{n}\sum_{i=1}^{T} (x_i - \bar{x})(Y_{ij} - \bar{Y}_j)\,.$$

A. Estimators of the linear model. The estimators in the linear case, denoted by a single hat, are well-known and are given here for completeness. [43, 44] Having a balanced design and a known p, the uniformly minimum variance unbiased estimators of the regression parameters in (1) are $\hat{\beta}_0 = \frac{1}{n}\sum_{j=1}^{n} \hat{\beta}_{0j}$ and $\hat{\beta}_1 = \frac{1}{n}\sum_{j=1}^{n} \hat{\beta}_{1j}$. Here $\hat{\beta}_{0j}$ and $\hat{\beta}_{1j}$ are the ordinary least squares estimators of β_{0j} and β_{1j} which are obtained by only using the data from the j:th subject. The variances are estimated as $\hat{\sigma}^2_{\beta_0} = \frac{S_{\bar{Y}\bar{Y}}}{n-1} - \frac{\hat{\sigma}^2_e}{T}$ and $\hat{\sigma}^2_e = \frac{S_{YY} - \hat{\beta}_1 S_{xY}}{n(T-1)-1}$.

It should be noted that the probability of a negative estimate of $\sigma^2_{\beta_0}$ is $P\left(\hat{\sigma}^2_{\beta_0} < 0\right) = P\left(F_{n-1,n(T-1)-1} < (1 + T\cdot\sigma^2_{\beta_0}/\sigma^2_e)^{-1}\right)$. However, the properties of $\hat{\sigma}^2_{\beta_0}$ are not examined in the paper.

B. Estimators of the non-linear model. The ML estimators in the non-linear case are denoted by a double hat and are defined as the estimators in Appendix IA but given the estimator $\hat{\hat{p}}$ of p. To obtain $\hat{\hat{p}}$, start with the Likelihood of the observations in (1) [45]:

$$L=(\sigma_{B_0}^2 T+\sigma_e^2)^{-n/2}\sigma_e^{-n(T-1)/2}\cdot$$
$$\cdot\exp\left\{\frac{-T\sum_{j=1}^{n}(\overline{Y}_j-\beta_0-\beta_1\overline{x})^2}{2(\sigma_{B_0}^2 T+\sigma_e^2)}-\frac{(S_{YY}-2\beta_1 S_{xY}+\beta_1^2 S_{xx})}{2\sigma_e^2}\right\}. \qquad (2)$$

Recall that $x_i=t_i^p$. The estimator $\hat{\hat{p}}$ is then obtained by solving the equation $\frac{\delta \log L}{\delta p}=\frac{\delta}{\delta p}\left(2\beta_1 S_{xY}-\beta_1^2 S_{xx}\right)=0$ where the unknown β_1 parameter is simultaneously replaced by its estimator. The resulting expression can be written as

$$\frac{\delta L}{\delta p}=\frac{\sum_{j=1}^{n}\sum_{i=1}^{T}Y_{ij}x_i-nT\cdot\overline{Y}\cdot\overline{x}}{\sum_{i=1}^{T}x_i^2-T\cdot\overline{x}^2}-\frac{\sum_{j=1}^{n}\sum_{i=1}^{T}Y_{ij}x_i\ln t_i-n\overline{Y}\sum_{i=1}^{T}x_i\ln t_i}{\sum_{i=1}^{T}x_i^2\ln t_i-\overline{x}\sum_{i=!}^{T}x_i\ln t_i}=0. \qquad (3)$$

In the simulation studies, Section 3.2 and 4, this approach was found to be very time-efficient when using a numeric grid search procedure over p in (3).

Appendix II - The asymptotic variances

The asymptotic variances and covariances (n large) of the estimators in Appendix I can be derived from the expectations of the second order derivatives of the Likelihood in (2).[45] For a given p, the following expressions for the estimators of the linear model were obtained ($T\geq 2$):

$$\left.\begin{aligned} &V(\hat{\beta}_0)=\frac{1}{n}\left[\sigma_{B_0}^2+\sigma_e^2\left(\frac{1}{T}+\frac{\bar{x}^2}{S_{xx}}\right)\right],\quad V(\hat{\beta}_1)=\frac{\sigma_e^2}{nS_{xx}},\quad Cov(\hat{\beta}_0,\hat{\beta}_1)=-\frac{\sigma_e^2\bar{x}}{nS_{xx}},\\ &V(\hat{\sigma}_{B_0}^2)=\frac{2}{n(T-1)T^2}\left[(T-1)\left(\sigma_{B_0}^2T+\sigma_e^2\right)^2+\sigma_e^4\right],\quad V(\hat{\sigma}_e^2)=\frac{2\sigma_e^4}{n(T-1)},\\ &V(\hat{\beta}_0+\hat{\beta}_1x_i)=\frac{1}{n}\left[\sigma_{B_0}^2+\frac{\sigma_e^2}{T}\right]+\frac{\sigma_e^2}{nS_{xx}}(x_i-\bar{x})^2 \end{aligned}\right\} \quad (4)$$

while all covariances between $(\hat{\beta}_0,\hat{\beta}_1)$ and $(\hat{\sigma}_{B_0}^2,\hat{\sigma}_e^2)$ are zero.

Further, for the estimators of the non-linear model the following expressions in (5) and (6) were obtained. First put: $M=\frac{1}{T}\sum_{i=1}^{T}x_i\ln t_i$, $S_1=\sum_{i=1}^{T}x_i^2\ln t_i-T\cdot M\cdot\bar{x}$, $S_2=\sum_{i=1}^{T}(x_i\ln t_i)^2-T\cdot M^2$ and let $Q_{\beta_1}=1-\frac{S_1^2}{S_{xx}S_2}$. Then

$$\left.\begin{aligned} &V(\hat{\hat{\beta}}_0)=\frac{1}{n}\left[\sigma_{B_0}^2+\sigma_e^2\left\{\frac{1}{T}+\left(\frac{\bar{x}^2}{S_{xx}}+\frac{M}{S_2}\left(M-\frac{2\bar{x}S_1}{S_{xx}}\right)\right)\frac{1}{Q_{\beta_1}}\right\}\right],\\ &V(\hat{\hat{\beta}}_1)=\frac{\sigma_e^2}{nS_{xx}}\cdot\frac{1}{Q_{\beta_1}}=V(\hat{\beta}_1)\cdot\frac{1}{Q_{\beta_1}},\\ &V(\hat{\hat{p}})=\frac{\sigma_e^2}{n\beta_1^2S_2}\cdot\frac{1}{Q_{\beta_1}} \end{aligned}\right\} \quad (5)$$

and

$$\left.\begin{aligned} Cov(\hat{\hat{\beta}}_0,\hat{\hat{\beta}}_1) &= \frac{\sigma_e^2}{nS_{xx}S_2}\cdot\frac{(M\cdot S_1-\overline{x}\cdot S_2)}{Q_{\beta_1}}, \\ Cov(\hat{\hat{\beta}}_0,\hat{\hat{p}}) &= \frac{\sigma_e^2}{nS_{xx}S_2}\cdot\frac{(\overline{x}\cdot S_1-M\cdot S_{xx})}{\beta_1 Q_{\beta_1}}, \\ Cov(\hat{\hat{\beta}}_1,\hat{\hat{p}}) &= -\frac{\sigma_e^2}{nS_{xx}S_2}\cdot\frac{S_1}{\beta_1 Q_{\beta_1}} \end{aligned}\right\} \quad (6)$$

while the asymptotic variances of $\hat{\hat{\sigma}}^2_{B_0}$ and $\hat{\hat{\sigma}}^2_e$ are the same as the corresponding variances in (4), and all covariances between $(\hat{\hat{\beta}}_0,\hat{\hat{\beta}}_1,\hat{\hat{p}})$ and $(\hat{\hat{\sigma}}^2_{B_0},\hat{\hat{\sigma}}^2_e)$ are zero. Finally, using the previous results the variance of the estimated mean function was approximated (Taylor series) by

$$V(\hat{\hat{\beta}}_0+\hat{\hat{\beta}}_1 x_i)\cong V(\hat{\hat{\beta}}_0)+x_i^2V(\hat{\hat{\beta}}_1)+(\beta_1 x_i\cdot\ln t)^2V(\hat{\hat{p}})+$$
$$+2x_i[\mathrm{cov}(\hat{\hat{\beta}}_0,\hat{\hat{\beta}}_1)+\beta_1\cdot\ln t\cdot\mathrm{cov}(\hat{\hat{\beta}}_0,\hat{\hat{p}})+\beta_1 x_i\cdot\ln t\cdot\mathrm{cov}(\hat{\hat{\beta}}_1,\hat{\hat{p}})].$$

References

1. Zariffa, N. M. D., Patterson, S. D., Boyle, D. and Hyneck, M. 'Case studies, practical issues and observations on population and individual bioequivalence', *Statistics in Medicine*, **19**, 2811-2820 (2000).
2. Mauger, D. T. and Chinchilli, V. M. 'An alternative index for assessing profile similarity in bioequivalence trials', *Statistics in Medicine*, **19**, 2855-2866 (2000).
3. Gabrielsson, J. and Weiner, D. *Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications*, 3 ed., Swedish Pharmaceutical Society, Stockholm, 2000:924.
4. Racine-Poon, A. and Wakefield, J. 'Statistical methods for population pharmacokinetic modelling', *Statistical Methods in Medical Research*, **7**, 63-84 (1998).
5. Bancroft, T. A. 'Analysis and Inference for Incompletely Specified Models Involving the Use of Preliminary Test(s) of Significance', *Biometrics*, **20**, 427-442 (1964).
6. Bancroft, T. A. and Han, C.-P. 'Inference based on conditional specification: A note and a bibliography', *International Statistical Review*, **45**, 117-128 (1977).
7. Judge, G. G. and Bock, M. E. *The statistical implications of pre-test and Stein-rule estimators in econometrics*. Studies in mathematical and managerial economics North-Holland, Amsterdam, 1978:340.
8. Bhoj, D. S. and Ahsanullah, M. 'Estimation of a Conditional Mean in a Linear-Regression Model after a Preliminary Test on Regression Coefficient', *Biometrical Journal*, **36**, 153-163 (1994).
9. Grambsch, P. M. and Obrien, P. C. 'The Effects of Transformations and Preliminary Tests for Nonlinearity in Regression', *Statistics in Medicine*, **10**, 697-709 (1991).
10. Ding, A. A. and Wu, H. L. 'A comparison study of models and fitting procedures for biphasic viral dynamics in HIV-1 infected patients treated with antiviral therapies', *Biometrics*, **56**, 293-300 (2000).
11. Perelson, A. S., Essunger, P., Cao, Y. Z., et al. 'Decay characteristics of HIV-1-infected compartments during combination therapy', *Nature*, **387**, 188-191 (1997).
12. Wu, H. L. and Wu, L. 'Identification of significant host factors for HIV dynamics modelled by non-linear mixed-effects models', *Statistics in Medicine*, **21**, 753-771 (2002).
13. Wein, L. M., D'Amato, R. M. and Perelson, A. S. 'Mathematical analysis of antiretroviral therapy aimed at HIV-1 eradication or maintenance of low viral loads', *Journal of Theoretical Biology*, **192**, 81-98 (1998).

14. Ding, A. A. and Wu, H. 'Assessing antiviral potency of anti-HIV therapies in vivo by comparing viral decay rates in viral dynamic models', *Biostatistics*, **2**, 13-29 (2001).
15. Davidian, M. and Giltinan, D. M. *Nonlinear Models for Repeated Measurement Data*. Monographs on Statistics and Applied Probability Chapman & Hall, London, 1995:359.
16. Gibaldi, M. and Perrier, D. *Pharmacokinetics*, 2 ed., Marcel Dekker, New York, 1982:494.
17. Box, G. E. P. and Tidwell, P. W. 'Transformation of the independent variables', *Technometrics*, **4**, 531-550 (1962).
18. Royston, P. 'A useful monotonic non-linear model with applications in medicine and epidemiology', *Statistics in Medicine*, **19**, 2053-2066 (2000).
19. Royston, P. 'A strategy for modelling the effect of a continuous covariate in medicine and epidemiology', *Statistics in Medicine*, **19**, 1831-1847 (2000).
20. Vonesh, E. F. and Chinchilli, V. M. *Linear and nonlinear models for the analysis of repeated measurements*, Dekker, New York, 1997:560.
21. Benda, N. 'Pre-test estimation and design in the linear model', *Journal of Statistical Planning and Inference*, **52**, 225-240 (1996).
22. Mentre, F., Mallet, A. and Baccar, D. 'Optimal design in random-effects regression models', *Biometrika*, **84**, 429-442 (1997).
23. van Houwelingen, J. C. 'Shrinkage and penalized likelihood as methods to improve predictive accuracy', *Statistica Neerlandica*, **55**, 17-34 (2001).
24. Wakefield, J. 'The Bayesian analysis of population pharmacokinetic models', *Journal of the American Statistical Association*, **91**, 62-75 (1996).
25. Sheiner, L. and Wakefield, J. 'Population modelling in drug development', *Statistical Methods in Medical Research*, **8**, 183-193 (1999).
26. Putter, H., Heisterkamp, S. H., Lange, J. M. A. and de Wolf, F. 'A Bayesian approach to parameter estimation in HIV dynamical models', *Statistics in Medicine*, **21**, 2199-2214 (2002).
27. Susarla, V. *Empirical Bayes Theory*,Kotz S, Johnson NL, eds. Encyclopedia of Statistical Sciences. Vol. 2 Wiley, New York, 1982:613.
28. Judge, G. G., Griffiths, W. E., Hill, R. C., Lütkepohl, H. and Lee, T.-C. *The theory and practise of econometrics*, 2 ed. Wiley series in probability and mathematical statistics Wiley, New York, 1985:1019.
29. Sawa, T. and Hiromatsu, T. 'Minimax regret significance points for a preliminary test in regression analysis', *Econometrica*, **41**, 1093-1101 (1973).
30. Kwan, K. C., Breault, G. O., Umbenhauer, E. R., McMahon, F. G. and Duggan, D. E. 'Kinetics of Indomethacin Absorption, Elimination, and

Enterohepatic Circulation in Man', *Journal of Pharmacokinetics and Biopharmaceutics*, **4**, 255-280 (1976).

31. Weinberg, J. M. and Lagakos, S. W. 'Linear rank tests under general alternatives, with application to summary statistics computed from repeated measures data', *Journal of Statistical Planning and Inference*, **96**, 109-127 (2001).
32. Wu, H. L. and Ding, A. A. 'Design of viral dynamic studies for efficiently assessing potency of Anti-HIV therapies in AIDS clinical trials', *Biometrical Journal*, **44**, 175-196 (2002).
33. Murphy, J. R. and Hofer, J. D. 'Establishing shelf life, epiry limits, and release limits', *Drug Information Journal*, **36**, 769-781 (2002).
34. Longford, N. T. *Random coefficient models*. Oxford statistical science series Oxford University Press, Oxford, 1993:270.
35. Ohtani, K. 'Optimal Levels of Significance of a Pre-Test in Estimating the Disturbance Variance after the Pre-Test for a Linear-Hypothesis on Coefficients in a Linear-Regression', *Economics Letters*, **28**, 151-156 (1988).
36. Khan, S. and Saleh, A. 'On the comparison of the pre-test and shrinkage estimators for the univariate normal mean', *Statistical Papers*, **42**, 451-473 (2001).
37. Toyoda, T. and Wallace, T. D. 'Optimal critical values for pre-testing in regression', *Econometrica*, **44**, 365-375 (1976).
38. Greenland, S. 'Cautions in the Use of Preliminary-Test Estimators - Comment', *Statistics in Medicine*, **8**, 669-673 (1989).
39. Michalek, J. E., Mihalko, D. and Tripathi, R. C. 'Cautions on the Reanalysis of Epidemiologic Databases', *Statistics in Medicine*, **8**, 653-664 (1989).
40. Sen, P. K. 'Asymptotic properties of maximum likelihood estimators based on conditional specification', *The Annals of Statistics*, **7**, 1019-1033 (1979).
41. Albers, W., Boon, P. C. and Kallenberg, W. C. M. 'Size and power of pretest procedures', *Annals of Statistics*, **28**, 195-214 (2000).
42. Albers, W., Boon, P. C. and Kallenberg, W. C. M. 'The asymptotic behavior of tests for normal means based on a variance pre-test', *Journal of Statistical Planning and Inference*, **88**, 47-57 (2000).
43. Swamy, P. A. V. B. *Statistical inference in random coefficient regression models*. Lecture notes in operation research and mathematical systems Springer-Verlag, Berlin, 1971:209.
44. Rao, C. R. 'The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves', *Biometrika*, **52**, 447-468 (1965).
45. Hsiao, C. *Analysis of panel data*, 2 ed., Cambridge University Press, Cambridge, 2003:366.

Research Report

2002:4	Persson, A.:	Prediction of work resumption among men and women with lower back- and neck pain in a Swedish population.
2002:5	Persson, A.:	Prediction of work resumption in theory and practice.
2002:6	Andersson, E., Bock, D. & Frisén, M.:	Detection of turning points in business cycles.
2002:7	Andersson, E., Bock, D. & Frisén, M.:	Some statistical aspects on methods for detection of turning points in business cycles.
2002:8	Andersson, E., Bock, D. & Frisén, M.:	Statistical surveillance of cyclical processes with application to turns in business cycles.
2002:9	Holgersson, T.:	Testing for non-normality in multivariate regression with nonspherical disturbances.
2003:1	Holgersson, T. & Shukur, G.:	Testing for multivariate heteroscedasticity.
2003:2	Holgersson, T.:	Testing for multivariate autocorrelation.
2003:3	Petzold, M. & Sonesson, C.:	Detection of intrauterine growth restriction.
2003:4	Bock, D.:	Similarities and differences between statistical surveillance and certain decision rules in finance.
2003:5	Holgersson, T. & Lindström, F.:	A comparison of conditioned versus unconditioned forecasts of the VAR(1) process.
2003:6	Bock, D.:	Early warnings for turns in business cycles and finance.
2003:7	Lindström, F.:	On prediction accuracy of the first order vector auto regressive process.
2003:8	Petzold, M. & Jonsson, R.:	Maximum Likelihood Ratio based small-sample tests for random coefficients in linear regression.