Automated Covariate Model Building Within NONMEM

E. Niclas Jonsson[1] and Mats O. Karlsson[1,2]

Received February 3, 1998; accepted June 13, 1998

Purpose. One important task in population pharmacokinetic/pharmacodynamic model building is to identify the relationships between the parameters and demographic factors (covariates). The purpose of this study is to present an automated procedure that accomplishes this. The benefits of the proposed procedure over other commonly used methods are (i) the covariate model is built for all parameters simultaneously, (ii) the covariate model is built within the population modeling program (NONMEM) giving familiar meaning to the significance levels used, (iii) it can appropriately handle covariates that varies over time and (iv) it is not dependent on the quality of the posterior Bayes estimates of the individual parameter values. For situations in which the computer run-times are a limiting factor, a linearization of the non-linear mixed effects model is proposed and evaluated.

Methods. The covariate model is built in a stepwise fashion in which both linear and non-linear relationships between the parameters and covariates are considered. The linearization is basically a linear mixed effects model in which the population predictions and their derivatives with respect to the parameters are fixed from a model without covariates. The stepwise procedure as well as the linearization was evaluated using simulations in which the covariates were taken from a real data set.

Results. The covariate models identified agreed well with what could be expected based on the covariates that were actually supported in each of the simulated data sets. The predictive performance of the linearized model was close to that of the non-linearized model.

Conclusions. The proposed procedure identifies covariate models that are close to the model supported by the data set as well as being useful in the prediction of new data. The linearized model performs nearly as well as the non-linearized model.

KEY WORDS: covariate model building; NONMEM; population pharmacokinetic analysis.

[1] Department of Pharmacy, Uppsala University, Box 580, S-751 23 Uppsala, Sweden.

[2] To whom correspondence should be addressed. (e-mail: mats.karlsson@biof.uu.se)

ABBREVIATIONS: y_{ij}, The ith individuals jth observation; $\widehat{y_{ij}}$, The ith individuals jth prediction based on P_i; $\bar{y}_{ij}$, The ith individuals jth prediction based on Θ; ε_{ij}, The difference between y_{ij} and $\widehat{y_{ij}}$; σ^2, The variance of ε_{ij}; X_{ij}, The ith individuals vector of independent variables (times, doses, etc); P_i, The ith individuals vector of model parameters; p_{ki}, The ith individuals kth model parameter; Θ, The vector of typical individual model parameters; θ_k, The kth typical individual parameter; η_{ki}, The difference between ith individuals kth model parameter and the typical individual parameter value given the covariates; ω_k^2, The variance of η_k; z_i, The ith individuals vector of covariate values; z_r, The vector of values for the rth covariate; m, The number of ηs in the model; z_{ir}, The ith individuals rth covariate value; c_{ir}, The ith individuals contribution of z_r on the parameter; θ_r, The coefficient for the relationship between z_r and the parameter.

INTRODUCTION

One of the important tasks in population pharmacokinetic (PK) and pharmacodynamic (PD) model building is to identify the relationship between the parameters of the model and covariates, e.g. demographic factors, clinical laboratory measurements, and indicators for pathological conditions. Originally candidate covariate relationships were identified based on residual analysis. Such graphical procedures provided limited information and the subsequent analysis often involved a large number of computer runs and were hence quite time-consuming. To obtain some guidance as to which covariate influenced which parameter, Maitre et al. (1) suggested plotting the posterior Bayes estimates of the parameters from a model without covariates versus each of the covariates. To make the covariate model building process even more efficient, Mandema et al. (2) proposed a stepwise generalized additive modeling procedure (GAM). In this procedure the covariates are regressed on the individual posterior Bayes estimates of the parameters in a stepwise fashion (one PK or PD parameter at a time) and the covariates which are significant according to the Akaike information criteria (AIC) are identified. Since it was first presented the GAM has been quite frequently used (3,4,5,6,7).

However tractable, the covariate model building methodologies that utilize the posterior Bayes estimates has a number of drawbacks. First, they usually only consider one parameter at a time, meaning the inclusion of a covariate on a parameter will not influence the significance of the same, or other covariates on the other parameters of the model. Second, the performance of of these methods is dependent on the quality of the posterior Bayes estimates of the individual parameters (these are used as the dependent variables in the GAM analysis). The quality of the individual estimates will be influenced by (i) each individuals' design with respect to estimating the individual parameters conditional on the population model and (ii) the quality of the PK or PD observations. Third, it is hard to handle time varying covariates. In clinical studies which involve more than one study occasion, time varying covariates are common and can be important, e.g. clinical laboratory measurements or markers for disease progression. Fourth, with many of these methodologies it is hard to judge the relative importance of the covariates, especially in graphical procedures such as in Maitre's approach. Even in the GAM, which provides such a ranking, it is still necessary to check each covariate in the population model.

In the present paper we propose a stepwise procedure for covariate selection in population analysis that remedies the drawbacks of the traditional covariate model building procedures. It steps through possible covariate parameter-combinations in a forward fashion, and evaluates their importance in the *population model*. The covariate terms in the full forward model are then evaluated using backwards elimination. This approach uses less number of computer runs than an exhaustive search and is practical for models with relatively short run times. To handle situations with long run-times, we propose a linearization of the non-linear mixed effects model, i.e. a linear mixed effects model, to be used during the covariate selection. This linearization involves an approximation but is considerably faster than the full non-linear model.

0724-8741/98/0900-1463$15.00/0

METHODS

In the following we assume for simplicity of notation, additive or proportional inter- and intra-individual variability models and time constant parameters and covariates. The theory, however, makes no such assumptions.

The Population Model

In population analyses by non-linear mixed effects models, it is usually assumed that the data can be described by the model:

$$y_{ij} = f(P_i, X_{ij}) + \varepsilon_{ij} \quad (1)$$

where y_{ij} is the jth observation from the ith individual (e.g. concentrations or effect measurements), $f(\ )$ is a model (PK or PD) that relates the independent variables, X_{ij} (e.g. time and dose), to the response given the ith individuals vector of model parameters P_i. ε_{ij} accounts for the discrepancy between the model predictions, $\widehat{y_{ij}}$, and the observations. ε_{ij} is usually assumed independently symmetrically distributed with a variance σ^2 that are either constant for all observations, i.e. an additive error model, proportional to the predicted response, i.e. $\sigma^2 \widehat{y_{ij}^2}$ or a combination of the two.

The P_i's are assumed to be distributed around the their typical values in the population, Θ, according to Eq. 2:

$$p_{ki} = \theta_k + \eta_{ki} \quad (2)$$

where p_{ki} is the kth parameter in P_i and θ_k is the typical value of p_k in the population. The η_k's are the differences between the p_{ki} and the θ_k and are assumed to be independently, multivariately symmetrically distributed with mean 0 and variance ω_k^2. Often an exponential inter-individual model is used, i.e. $p_{ki} = \theta_k e^{\eta_{ki}}$ in which ω_k is approximately the coefficient of variation.

Individual parameter values can often be related to demographic factors such as age, gender, and clinical laboratory measurements (usually referred to as the covariates) in the following way:

$$p_{ki} = \theta_k g_k(z_i) + \eta_{ki} \quad (3)$$

z_i is the i*th* individuals vector of covariates and g_k is a function of z_i (and some parameters, which are notationally suppressed in Eq. 3) that describes the relationship between p_{ki} and z_i. A more detailed discussion of $g()$ is given below.

The Stepwise Procedure

For each parameter and covariate, a hierarchy of models are defined. These typically consist of: the covariate not being included in the model, it is included in a linear fashion, and it is included in a non-linear fashion (see Parameterization below).

The starting model of the stepwise procedure for each parameter and covariate is the first model in the model hierarchy. This typically is the basic model without covariates. In the first step, each model in the second level of the model hierarchy is tried in the start model, one at a time. The improvement of the fit relative to the start model, when each of the covariate models are added univariately are compared and the model with the largest improvement, given that it is significant, is kept to the next step. The start model plus the best covariate model in each step is called the current model. In subsequent steps each parameter/covariate combination is tried in the current model in the functional form given by the next level of the model hierarchy. This stepwise forward inclusion of covariates ends when there are no more significant parameter/covariate combinations (the full model). Following the forwards inclusion is a backwards elimination. During this process each covariate model in the full model is replaced by the next lower model in the model hierarchy. The model that contributes the least to the fit, given that it is not significant, is dropped and a new current model is formed of the full model with the dropped model replaced by the next lower model in the hierarchy. The backwards elimination continues until no more terms can be dropped. This procedure is similar to the stepwise procedure that the GAM uses, the difference being that the GAM includes the possibility to drop any of the covariate models in each step instead of having a full backwards elimination as a last step.

The goodness of fit is in this paper measured as the drop in the objective function value (OFV) produced by NONMEM. The difference in this value between two nested models is approximately χ^2-distributed and can be used to obtain a significance level of the improvement in fit. In this paper a p-value of 0.05 was judged significant during the forward inclusion while a stricter value of $p < 0.01$ was used during the backwards elimination. It should be noted, though, that the procedures do not rely on a specific p-value, nor is it necessary to use the difference in the OFV as the criteria to discriminate between rival covariate models. We choose to use p-values (or rather the difference in the OFV) since this is an often used criteria when assessing the value of adding more parameters to a model.

Linearization

If the time necessary to fit the model is short, the above procedure can be applied to the full population model. If, on the other hand, the run-times are long, the full model approach is not practical. In such situations we propose using a linearized version of the model.

The default algorithm in NONMEM, the first order method (FO), uses a first order Taylor series expansion around the expected value of the η's, i.e. 0 (8,9).

$$y_{ij} = \bar{y}_{ij} + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} \eta_{ki} + \varepsilon_{ij} \quad (4)$$

where $\bar{y}_{ij}$ is the model predictions based on Θ (i.e. substituting P_i with Θ in Eq. 1), m is the number of η's in the model and $\partial \bar{y}_{ij}/\partial \eta_{ki}$ is the derivative of $\bar{y}_{ij}$ with respect to η_{ki}.

The basic idea of the proposed linearization is to use the same formula as Eq. 4 but adding (linearized) terms for the covariate effects while keeping $\bar{y}_{ij}$ and $\partial \bar{y}_{ij}/\partial \eta_{ki}$ fixed to the values obtained from the fit of the basic model, i.e.

$$y_{ij} = \bar{y}_{ij} + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} g_k(z_i)^* + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} \eta_{ki}^* + \varepsilon_{ij}^* \quad (5)$$

(With the linearization, $g()$ has to be centered around zero, in contrast to the normal population model were it would be centered around one, see below.) The parameters and function with an asterisk (*) are the ones estimated when fitting Eq. 5 to the data. Eq. 5 is valid for additive models for inter- and intra-individual variability and Eq. 6 for proportional models.

$$y_{ij} = \left(\bar{y}_{ij} + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} g_p(z_i)^* + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} \times (1 + g_k(z_i)^*)\eta_{ki}^*\right)(1 + \varepsilon_{ij}^*) \quad (6)$$

In Eqs. 5 and 6, the covariate terms are added using a first order linearization. A more accurate approximation can be obtained by using a second order linearization of the covariate terms, e.g.

$$y_{ij} = \left(\bar{y}_{ij} + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} g_k(z_i)^* + \sum_{k=1}^{m} \frac{\partial^2 \bar{y}_{ij}}{2\partial \eta_{ki}^2} g_k(z_i)^{2*} + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} (1 + g_k(z_i)^*)\eta_{ki}^*\right)(1 + \varepsilon_{ij}^*) \quad (7)$$

where $\partial^2 \bar{y}_{ij}/\partial \eta_{ki}^2$ is the second derivative of $\bar{y}_{ij}$ with respect to η_{ki}.

The covariate model, in any one step, together with the second derivatives can be also used to update the first derivatives in the η-term in the above equations. This is accomplished according to Eq. 8.

$$\widehat{\frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}}} = \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} + \frac{\partial^2 \bar{y}_{ij}}{2\partial \eta_{ki}^2} g_k(z_i)^* \quad (8)$$

$\widehat{\partial \bar{y}_{ij}/\partial \eta_{ki}}$ denotes the ith individuals jth updated first derivative value with respect to η_k. Combining Eq. 7 and Eq. 8 yields:

$$y_{ij} = \left(\bar{y}_{ij} + \sum_{k=1}^{m} \frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}} g_k(z_i)^* + \sum_{k=1}^{m} \frac{\partial^2 \bar{y}_{ij}}{2\partial \eta_{ki}^2} g_k(z_i)^{2*} + \sum_{k=1}^{m} \widehat{\frac{\partial \bar{y}_{ij}}{\partial \eta_{ki}}} (1 + g_k(z_i)^*)\eta_{ki}^*\right)(1 + \varepsilon_{ij}^*) \quad (9)$$

If a covariate has a strong influence on one or more of the parameters of the model, the misspecification of the linearization can be decreased even further by obtaining $\bar{y}_{ij}$, $\partial \bar{y}_{ij}/\partial \eta_{ki}$, and $\partial^2 \bar{y}_{ij}/2\partial \eta_{ki}^2$ from a model including the covariate.

Since the coefficients of $g_k(z_i)$ will be different when estimated using the linearization compared to the corresponding non-linearized model, it is necessary to re-run the final covariate model in the regular (non-linearized) population model.

Parameterization

This section concerns the form of $g(z_i)$. In the present paper it is assumed that the relationships between the parameters and continuous covariates can be described by either a one slope model or a two slope model (i.e. two linear functions with the intersection at the median of the covariate space) and the relationship between the parameters and categorical covariates as a shift in the intercept. If a parameter is influenced by more than one covariate, the model is parameterized in a multiplicative fashion. A multi-covariate model will be assumed to be additive in the sense that interactions between covariates are not accounted for. Eq. 10 gives a general formula for the combination of covariate effects.

$$g(z_i) = \prod_{r=1}^{q} (1 + c_{ir}) \quad (10)$$

c_{ir} is the ith individual's contribution to the effect of the covariate z_r on the parameter in question (further details are given below).

It is usually a good idea to center the relationships between continuous covariates and parameters around the value of the covariate for the typical individual in the population. For a continuous covariate this is done according to Eq. 11 (assuming a linear relationship).

$$c_{ir} = \theta_r(z_{ir} - median(z_r)) \quad (11)$$

where z_{ir} is the ith individuals covariate value and θ_r is the fractional change in the parameter in question per unit change in the covariate. Categorical covariates are usually not centered in the normal population model but is handled by one or more *if* statements. For example, c_{ir} for a bivariate covariate with levels a and b can be parameterized according to Eq. 12.

$$c_{ir} = \begin{cases} 0 & \text{if } z_{ir} = a \\ \theta_r & \text{if } z_{ir} = b \end{cases} \quad (12)$$

The meaning of θ_r in Eq. 12 is the fractional change in the parameter for individuals of category b.

This parameterization will make the model numerically more stable as well as give the parameters of the model a more relevant meaning compared to when the covariates are not centered (e.g. since the product of all covariate effects for the typical individual will be 1, θ_k in Eq. 3 will be the parameter estimate for the typical individual with respect to the covariates).

With the linearization it is *necessary* to center the covariates, because the predictions from the basic model, the $\bar{y}_{ij}$, are fixed in the linearization, and are the predictions from the typical individual. This also means the categorical covariates must be centered. For the same bivariate covariate used as an example in Eq. 12, centering can be accomplished in the following way

$$c_{ir} = \begin{cases} \theta_r f_a & \text{if } z_{ir} = b \\ -\theta_r(1 - f_a) & \text{if } z_{ir} = a \end{cases} \quad (13)$$

where f_a is the fraction of the individuals with a as the value of the covariate. Categorical covariates with more than two levels can not be centered in this way.

Since the linearization requires $g()$ to be 0 for the typical individual, it will also be necessary to modify Eq. 10 (which will be 1 for the typical individual). This is done simply by subtracting 1 from the product of all covariate contributions (Eq. 14).

$$g(z_i) = -1 + \prod_{r=1}^{q} (1 + c_{ir}) \quad (14)$$

Given the covariates are centered as described above, the linear approximation will be close to the non-linear model for weak covariate relationships and for individuals that have almost the same covariate values as the typical individual. The approximation will be less exact for strong covariate relationships and for individuals with extreme covariate values.

Evaluation of the Stepwise Procedure

The stepwise procedure and linearization were evaluated using simulations. The benefit of simulations is the true model is known. The drawback, on the other hand, is it is impossible to

obtain the same complexity (e.g. with respect to the covariance structure between parameters, covariates and observations) as in a real data set. As a compromise, the data were simulated using covariates from a real data set.

Ten data sets of 64 individuals each were simulated. The sampling design was 0.5, 2, 4 and 6 hours post dose in half of the individuals and 0.5, 2, 8 and 12 hours post dose in the other half. The covariates were taken from the data set of DRUG C in the paper by Mandema et al. (2) and included AGE (age, median = 56 years, range = 24–69), HT (height, median = 173 cm, range = 140–188), WT (body weight, median = 85 kg, range = 51–137), SEX (gender, 42 males and 22 females), SMOK (smoking, yes = 16, no = 48), RACE (race, caucasian = 44, black = 20), PROP (co-treatment with propranolol, no = 54, yes = 10), HCTZ (co- treatment with hydrochlorotiazide, yes = 35, no = 29) and CON (other co-medication, yes = 51, no = 13). The same covariate values were used in all data sets.

The drug-levels were generated from a one compartment model with first order absorption under steady-state conditions. Clearance (CL) was related to AGE (in a non-linear fashion) and SEX, while the volume of distribution (V) was related to HCTZ and WT. The relationship between CL and AGE was formulated as a "hockey-stick" model, i.e. a constant CL for AGE-values below the median AGE and a linear decrease in CL for AGE-values higher than the median AGE. The covariate relationships were constructed so that the ratios between the parameter values for the lower and higher extremes (of the covariates) were approximately 0.5, 0.8, 0.55 and 0.6 for AGE on CL, SEX on CL, HCTZ on V and WT on V respectively. The covariate models were parametrizised according to Eqs. 11 and 13. The data were generated using exponential inter-individual variability models (the η-models) and a proportional residual error model (the ε-model). The parameter values used to simulate the data are summarized in Table I.

The stepwise procedure was applied to three different ways of formulating the mixed effects model. The first was the regular population model (hereafter referred to as the non-linear model), the second was the linearized model (Eq. 9, hereafter referred to as the linear model) and the third was the linearized model using the $\bar{y}_{ij}$ and $\partial\bar{y}_{ij}/\partial\eta_{ki}$ obtained from a model including the most important covariate (as indicated by the corresponding linear model), hereafter referred to as the linear+ model.

Table I. Parameter Values Used in the Simulations

Parameter	Value
CL	20
V	100
Ka	2
ω_{CL}	25%
ω_V	25%
ω_{Ka}	20%
σ	15%
AGE on CL[a]	−0.0385
SEX on CL	0.195
HCTZ on V	−0.560
WT on V	0.006

[a] The value refers to the slope for AGE values > the median value of age (56 years).

To determine which of the true covariate relationships were actually supported by each of the ten simulated data sets, the true covariate model was evaluated using a backwards elimination similar to the one used in the stepwise procedure (see above). This model will be referred to as the "supported" model, while the covariate model that was used to generate the data will be called the nominal model.

One way to assess the similarity between covariate models is to compare the actual covariate terms included. This might, however, not be a relevant criteria with respect to the usefulness of the covariate models, partly since two models with different covariate terms might still be equivalent in terms of prediction due to correlations between the covariates and partly because identical models might produce different results simply because the covariate term coefficients are different. The predictive performance of the methods were therefore evaluated in a new data set. The new data set consisted of 12800 individuals generated by replicating the original data set 200 times (with new η's and ε's). The predictive performance was measured as the root mean squared error (RMSE) between the log of the "observed" concentration and the log of the predictions for the typical individual.

The mixed effects models, linear as well as non-linear, were fitted using NONMEM. The stepwise procedure was written in the S language as implemented in S-PLUS version 3.4 (MathSoft Inc. 1997).

RESULTS

Table II compares the covariate selection to the nominal covariate model. There were no difference in the selection of HCTZ on V and only minor differences in the selection of AGE on CL although the linear method did not find the covariate in the correct functional form as often as the other methods. The linear model also performed worse in the selection of the other covariates, compared to the non-linear and the linear+ models. The non-linear model included a total of 16 covariates not in the nominal model compared to 8 and 10 for the linear and linear+ model respectively.

Figure 1 is a box and whiskers plot of the RMSE for the different methods. Included is also the RMSE of the basic model (i.e. no covariate relationships) for each data set. The median RMSE for the supported, non-linear, linear+, linear and basic models were 0.47, 0.48, 0.48, 0.50 and 0.59 respectively. The RMSE of the nominal model with the true parameter values (Table I) was 0.44.

Table II. Summary of the Covariate Selection

Method	AGE	SEX	HCTZ	WT	False CL	False V
Non-linear	9(7)	6	10	7(6)	7	9
Linear	8(3)	2	10	2(2)	4	4
Linear+	8(6)	4	10	5(5)	5	5
Supported	9(6)	5	10	5	—	—

Note: The figures are the number of times each of the methods identified any of the covariates in the nominal covariate model. The figures in parentheses are the number of times the covariate entered the model in the correct functional form (e.g. non-linear for AGE on CL). The two rightmost columns gives the number of times covariates not in the nominal covariate model was found (CL and V respectively).

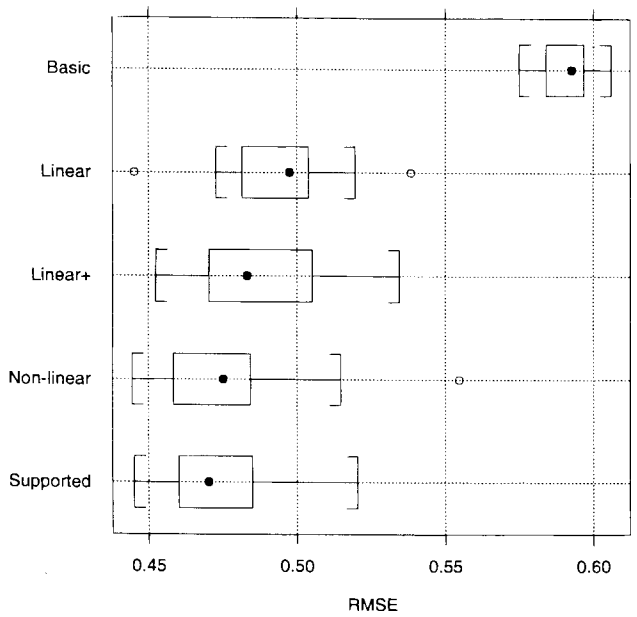

Fig. 1. A box and whiskers plot of the root mean squared error (RMSE) between the true concentrations and the predictions obtained with the basic, linear, linear+, non-linear and supported covariate models respectively. The dot in each box is the median. The length of each box is the inter-quartile range. The whiskers extend to the point that is less/greater than or equal to upper/lower quartile times 1.5. Any value beyond the whiskers is indicated by a circle.

The supported and the non-linear model had the lowest RMSE in three data sets each while the linear and the linear+ had the lowest RMSE in two data sets each. The basic model had the highest RMSE in all of the data sets.

The difference in run time between the non-linear linear and the two linear methods was approximately a factor of four, i.e. the non-linear model took about four times as long to run.

DISCUSSION

This paper presents a stepwise covariate model building algorithm that avoids drawbacks present with many other covariate model building procedures, for example the GAM (2). The present procedure builds the covariate model within the population model which means that covariate effects can be included and evaluated simultaneously on all parameters; the significance levels used for judging the importance of the covariates will have a familiar meaning and there is no dependence on the posterior Bayes estimates of the individual parameter values. The latter can be a major problem with sparse, poorly designed individual sampling schemes. The present procedure also has the benefit of being able to handle covariates that vary between study occasions, although that aspect is not addressed in this study.

The main drawback of building covariate models in the proposed way is the potentially long computer run-times. As a remedy to this problem we also evaluated an approximation (i.e. linearization) of the non-linear mixed effects model which decreased the run-times by a factor of four. This factor can be expected to increase with more complicated structural models. This is because the run-times of the linearized model will only increase with the amount of data, number of covariates, and the number of parameters for which a covariate model is built, i.e. *not* the size or complexity of the structural model. The run-times for the non-linear model, on the other hand, will also increase with the complexity of the structural model, a relationship that tends to be non-constant, i.e. a doubling of the number of parameters will more than double the run-times.

Even with the linearization, the run-times of the stepwise procedure will be much longer than running, for example, the GAM on each of the parameters of the model or performing a graphical analysis of the posterior Bayes estimates and covariates. With these methods, it is then necessary to critically evaluate the covariate relationships found. With the stepwise procedure, on the other hand, the evaluation is an integrated part of the methodology. Regardless of the approach taken when evaluating the covariate model found by the posterior Bayes estimates methods, it will involve a fair number of runs and it is not unlikely the evaluation procedure will take as long as the (linearized) stepwise procedure. Apart from the relative run-times of the existing methodologies and the proposed procedure, there are other aspects of the covariate model building strategies that need to be investigated further, for example the performance of the found covariate models in a data set not used to derive the covariate model and the sensitivity to structural and statistical model misspecification.

The stepwise procedure involves a large number of NONMEM runs and requires an automated routine to be practical and efficient. Such a routine will also have the benefit of being consistent in the way covariate models are built, something that simplifies the writing of reports and similar documents and will reduce the frequency of user-induced errors. Also the information gained when trying intermediate models, not included in the full or final models, may be useful, for example, to characterize the absence of a relationship. The code for the routine used for this work is available upon request to the authors. In addition, a more general and exportable version of the routine is under development.

In terms of selecting covariate relationships, the linearization seems to be a reasonable approximation to the non-linear mixed effects model. Although their performance did not quite match the non-linear model in terms of selecting true covariates, the linear models also selected fewer false relationships. Based on the more relevant criteria of predicting concentrations for new individuals of the same population, the linear models performed similarly to the non-linear model. The small trend observed indicates a better performance for the non-linear model is what can be expected due to the approximations of the linear models.

An alternative to using the predictions and derivatives from the basic model is to obtain them from a model already including the most important covariate(s). The way it was done in the present study was to first run the linear model to identify the most important covariate and then use that covariate to obtain the predictions and derivatives. An alternative is to automatically recalculate the derivatives and predictions based on the current model. This would involve the estimation of all parameters of a non-linear mixed effects model each time the current model changes, i.e. as many times as there are covariate relationships in the full model. Yet another alternative is to perform such recalculation only for the first covariate relationship(s) included.

One drawback of the linearization is the necessity of having the first and second derivatives available. Fortunately, they can be obtained from NONMEM without having to do the algebra manually (9).

In conclusion, we have presented a stepwise procedure for covariate model building in population analysis of PK and PD data. It evaluates all the covariate in a stepwise fashion *within* the population model and selects the covariates which are significant according to the same criteria that is used during other population model building phases. The advantages over other covariate model building procedures is it can handle time-varying covariates, is not dependent on the posterior Bayes estimates, and it considers covariate effects on all parameters at the same time. As a complement we also propose a linearization of the non-linear mixed effects model (the example used in this paper) which performed nearly as well as the non-linear mixed effects model.

ACKNOWLEDGMENTS

During this work, Niclas Jonsson was supported by a grant from Pfizer Central Research, Sandwich, Kent, UK.

REFERENCES

1. P. O. Maitre, M. Bührer, D. Thomson, and D. R. Stanski. A three-step approach combining bayesian regression and NONMEM population analysis: application to midazolam. *J. Pharmacokin. Biopharm.* **19**:377–384, 1991.
2. J. W. Mandema, D. Verotta, and L. B. Sheiner. Building population pharmacokinetic-pharmacodynamic models. I. Models for covariate effects. *J. Pharmacokin. Biopharm.* **20**:511–528, 1992.
3. C. F. Minto, T. W. Schneider, T. D. Egan, E. Youngs, *et al.* Influence of age and gender on the pharmacokinetics and pharmacodynamics of remifentanil, I. Model development. *Anasthesiology* **86**:10–23, 1997.
4. T. W. Schneider, C. F. Minto, H. Bruckert, and J. W. Mandema. Population pharmacodynamic modeling and covariate detection for central neural blockade. *Anasthesiology* **85**:502–512, 1996.
5. M. O. Karlsson, R. E. Port, M. J. Ratain, and L. B. Sheiner. A population-model for the leukopenic effect of etoposide. *Clin. Pharmacol. Ther.* **57**:325–334, 1995.
6. P. Burtin, E. V. Jacz-Aigrain, P. Girard, R. Lenclen, J-F. Magny, P. Betremieux, C. Tehiry, L. Desplanques, and P. Mussat. Population pharmacokinetics of midazolam in neonates. *Clin. Pharmacol. Ther.* **56**:615–625, 1994.
7. J-M. Gries, I. F. Troconiz, D. Verotta, M. Jacobsen, and L. B. Sheiner. A pooled analysis of CD4 response to zidivudine and zalcitabine treatment in patients with AIDS and ADS-related complex. *Clin. Pharmacol. Ther.* **61**:70–82, 1997.
8. S. L. Beal. Population pharmacokinetic data and parameter estimation based on their first two statistical moments. *Drug Metab. Rev.* **15**(1&2):173–193, 1984.
9. S. L. Beal and L. B Sheiner (Eds.). *NONMEM users guides.* NONMEM project group, University of California, San Fransisco, CA, 1992.