An introduction to the full random effects model

Gunnar Yngman[1] | Henrik Bjugård Nyberg[1] | Joakim Nyberg[2] | E. Niclas Jonsson[2] | Mats O. Karlsson[1,2]

[1]Department of Pharmacy, Uppsala University, Uppsala, Sweden

[2]Pharmetheus AB, Uppsala, Sweden

Correspondence
Mats O. Karlsson, Department of Pharmacy, Uppsala University, Uppsala, Sweden.
Email: mats.karlsson@farmaci.uu.se

Abstract

The full random-effects model (FREM) is a method for determining covariate effects in mixed-effects models. Covariates are modeled as random variables, described by mean and variance. The method captures the covariate effects in estimated covariances between individual parameters and covariates. This approach is robust against issues that may cause reduced performance in methods based on estimating fixed effects (e.g., correlated covariates where the effects cannot be simultaneously identified in fixed-effects methods). FREM covariate parameterization and transformation of covariate data records can be used to alter the covariate-parameter relation. Four relations (linear, log-linear, exponential, and power) were implemented and shown to provide estimates equivalent to their fixed-effects counterparts. Comparisons between FREM and mathematically equivalent full fixed-effects models (FFEMs) were performed in original and simulated data, in the presence and absence of non-normally distributed and highly correlated covariates. These comparisons show that both FREM and FFEM perform well in the examined cases, with a slightly better estimation accuracy of parameter interindividual variability (IIV) in FREM. In addition, FREM offers the unique advantage of letting a single estimation simultaneously provide covariate effect coefficient estimates and IIV estimates for any subset of the examined covariates, including the effect of each covariate in isolation. Such subsets can be used to apply the model across data sources with different sets of available covariates, or to communicate covariate effects in a way that is not conditional on other covariates.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
Most existing methods for covariate modeling estimate fixed effects in full model approaches or stepwise model building. The performance of these methods is well-documented, but they also have well-known downsides, such as issues with correlated covariates and potential multiple-testing problems. Their

interpretation is also limited to the exact set of covariates included. A full random effects approach that addresses some of these limitations has been proposed, but a detailed presentation is lacking in literature.

WHAT QUESTION DID THIS STUDY ADDRESS?
This work provides a more detailed introduction to the full random-effects model (FREM) and demonstrates its performance compared with full fixed-effects models (FFEMs) in several real and simulated data scenarios.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
Compared with FFEMs, FREM is less restrictive in the selection of covariates to include in the analysis, and much more informative from a single estimation.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?
Covariate effects can be better described, better communicated, and more easily carried forward to new analyses. The appropriate covariate sets for inference, prediction, or comparison to other studies may be different, but they can all be extracted from a single FREM estimation.

INTRODUCTION

Covariate modeling is an integral part of pharmacometrics, where understanding the observed interindividual variability (IIV) is central. The inclusion of subject-specific predictors on model parameters can explain this variability by attributing it to individual features, and can thereby improve model predictions. This can in turn enable inference within and between populations, and enhance the ability of model-informed analyses to answer scientific and clinical questions.

Covariate models can be built by adding fixed-effects parameter-covariate relations step-by-step to gradually explain portions of the IIV with covariate effects. The parameter IIV without any covariate effects represents the total IIV parameter variability (TPV), and each included covariate-parameter relation explains a portion of this variability, leaving a gradually altered unexplained IIV parameter variability (UPV). Stepwise selection of the best covariate model from multiple competitors, according to likelihood-ratio tests, has been automated in stepwise covariate modeling.[1] Although this approach can be successful, stepwise selection processes have a multiple testing problem, and may overestimate covariate effects due to selection bias and inflated type-I error,[2,3] problems that are exacerbated if correlated covariates are considered.[4,5] Imposing stringent selection criteria ($\alpha < 0.05$) reduces the risk of type-I error, but does so at the cost of reduced power.[2,6]

The full covariate model has been suggested as an alternative to stepwise procedures (Gastonguay[5]; A full model estimation approach for covariate effects: Inference based on clinical importance and estimation precision, 2004). In this approach, which we will refer to as the full fixed-effects model (FFEM), a set of noncorrelated covariates is identified prior to model evaluation. That full set is then implemented with simultaneously estimated fixed-effect coefficients for each covariate-parameter relation, thus avoiding the multiple testing problem. If a more parsimonious model is desired from the FFEM, e.g., for prediction, then stepwise backward elimination can be performed based on statistical and clinical significance,[7] although this may curtail the benefits of the full model approach. Alternatively, the model can be reduced by simultaneously removing all relations with effect coefficients below a certain threshold,[8] which better maintains the advantages of the full model approach but does not account for correlations that may hide a larger compound effect of the removed relations.

The full random-effects model (FREM) originally proposed by Karlsson[9] also makes use of a predefined set of covariates in a full model. However, instead of capturing covariate effects in fixed-effect coefficients, FREMs takes each covariate as observations of individual deviations from a population value. It estimates a full IIV random effect covariance matrix that contains parameter IIV, covariate IIV, and the covariances between the two. The covariate effects are captured in the covariances between covariate IIV and parameter IIV. Unlike FFEM, this implementation of covariate effects does not rely on estimating multiple fixed effects on the same parameter and can therefore include correlated covariates in the analysis. There are several successful applications of FREM in literature[10–12] and a description of the handling of missing covariate data using FREM,[13] but an introduction to the method has been lacking.

An FREM model estimates the TPV and the covariate IIV, and the estimates cannot be immediately interpreted in the same way as FFEM estimates. To aid in interpreting

the results, coefficients and UPV IIV variances that are equivalent to those from a corresponding FFEM can be calculated from the FREM matrix. Although this involves an extra step compared to FFEMs, it also offers a lot of flexibility because the user can choose what covariates to include at that stage. The coefficients and parameter IIV for any covariate subset can be calculated from that same estimation of the full FREM matrix, from the full set down to the univariate coefficients, the isolated impact of each covariate. By contrast, any interpretation of the FFEM is restricted to that exact model. Each additional model must be estimated separately, which can be time-consuming, and can lead back to multiple testing issues if the models are compared statistically. The FREM is mathematically equivalent to the FFEM for a wide range of covariates and parameter-covariate relations, but challenges remain in handling time-variations, nonlinearities, and covariates with multinomial distributions.

This work seeks to provide the definition and practical demonstration of the FREM method that has been missing in literature. We will describe the FREM approach in further detail and compare its performance to FFEM in real and simulated data. We will specifically examine estimation accuracy in the presence and absence of highly correlated covariates. Because of how FREM represents covariates using normal distribution means and variances, we will also investigate the impact of including non-normally distributed covariates and explore how some common parameter-covariate relation parameterizations can be implemented in FREM.

METHODS

FREM introduction

Consider an FFEM example model that expresses some quantity y for observation j in individual i as a function of two parameters (P_1, P_2), time (t), and residual error (ε),

$$y_{i,j} = P_{1,i}t_j + P_{2,i} + \varepsilon_{i,j}, \ \ \varepsilon_{i,j} \sim N(0, \sigma^2) \tag{1}$$

with IIV and linear relationships with two covariates (C_1, C_2) on both parameters,

$$P_{1,i} = \theta_{P_1} + \beta_{C_1 P_1}\left(C_{1,i} - \overline{C_1}\right) + \beta_{C_2 P_1}\left(C_{2,i} - \overline{C_2}\right) + \eta'_{P_{1,i}} \tag{2}$$

$$P_{2,i} = \theta_{P_2} + \beta_{C_1 P_2}\left(C_{1,i} - \overline{C_1}\right) + \beta_{C_2 P_2}\left(C_{2,i} - \overline{C_2}\right) + \eta'_{P_{2,i}} \tag{3}$$

$$\begin{bmatrix} \eta'_{P_{1,i}} \\ \eta'_{P_{2,i}} \end{bmatrix} \sim N\left(\vec{0}, \Omega'_{\mathrm{par}}\right) \tag{4}$$

where θ_P is the population value for parameter P, β_{CP} is the coefficient for the effect of covariate C on parameter P, $\overline{C}$ is the population mean of covariate C, $\eta'_{P,i}$ is the IIV random effect for the UPV (unexplained parameter variability after covariate inclusion) on parameter P, and Ω'_{par} is the UPV random effect covariance matrix. Note that two fixed-effect coefficients on the same parameter will not be independently identifiable if the two covariates are highly correlated.

In FREM, the data records of a covariate are considered individual observations of a dependent variable. An FREM that is equivalent to the FFEM in Equations 1–4 expresses $y_{i,j}$, $C_{1,i}$, and $C_{2,i}$,

$$\left\{ \begin{array}{c} y_{i,j} = P_{1,i}t_j + P_i + \varepsilon_{i,j}, \ \ \varepsilon_{i,j} \sim N\left(0, \sigma^2\right) \\ C_{1,i} = \theta_{C_1} + \eta_{C_1,i} \\ C_{2,i} = \theta_{C_2} + \eta_{C_2,i} \end{array} \right. \tag{5}$$

$$P_{1,i} = \theta_{P_1} + \eta_{P_1,i} \tag{6}$$

$$P_{2,i} = \theta_{P_2} + \eta_{P_2,i} \tag{7}$$

$$\begin{bmatrix} \eta_{P_1,i} \\ \eta_{P_2,i} \\ \eta_{C_1,i} \\ \eta_{C_2,i} \end{bmatrix} \sim N\left(\vec{0}, \Omega_{\mathrm{FREM}}\right) \tag{8}$$

where $\eta_{C,i}$ is the individual deviation from the covariate population value θ_C (here $\theta_C = \overline{C}$), and $\eta_{P,i}$ is the IIV random effect for the TPV (total parameter variability including IIV explained by covariates) of parameter P.

For the purposes of this work, we assume a single error-free observation of each covariate for each individual, and use normally distributed random effects to capture properties of the covariate distribution. It is expected that when covariate observations are error-free, the calculation of means, variances, and covariances, which is the aim of the estimation, is independent of the actual covariate distribution. Because of this, FREM can be used for covariates of any underlying distribution.

The FREM approach captures covariate-parameter relations in the full IIV covariance matrix Ω_{FREM}, which consists of the IIV random effect covariance matrices for parameters (Ω_{par}), covariates (Ω_{cov}), and the covariances between parameter random effects and covariate random effects $(\Omega_{\mathrm{par,cov}})$, (i.e., for this example):

$$\Omega_{\mathrm{FREM}} = \begin{pmatrix} \Omega_{\mathrm{par}} & \Omega_{\mathrm{cov,par}} \\ \Omega_{\mathrm{par,cov}} & \Omega_{\mathrm{cov}} \end{pmatrix} = \begin{pmatrix} \omega_{P_1}^2 & \omega_{P_2 P_1} & \omega_{C_1 P_1} & \omega_{C_2 P_1} \\ \omega_{P_1 P_2} & \omega_{P_2}^2 & \omega_{C_1 P_2} & \omega_{C_2 P_2} \\ \omega_{P_1 C_1} & \omega_{P_2 C_1} & \omega_{C_1}^2 & \omega_{C_2 C_1} \\ \omega_{P_1 C_2} & \omega_{P_2 C_2} & \omega_{C_1 C_2} & \omega_{C_2}^2 \end{pmatrix} \tag{9}$$

From this matrix, we can calculate the covariate-parameter relation coefficients and UPV for any combination of included covariates (i.e., the parameter distributions conditional on any subset of covariates). In the present example, the coefficients, conditional on C_1 and C_2, are available as the matrix B:

$$B = \begin{pmatrix} \beta_{C_1 P_1} & \beta_{C_2 P_1} \\ \beta_{C_1 P_2} & \beta_{C_2 P_2} \end{pmatrix} = \Omega_{\text{par,cov}} \Omega_{\text{cov}}^{-1} \qquad (10)$$

For the simplest case where the effect of one covariate is considered in isolation, the coefficients are calculated from FREM results as the ratios of the parameter-covariate covariances and the variance of the covariate, i.e., conditional only on C_1:

$$\beta_{C_1 P_1} = \omega_{C_1 P_1}/\omega_{C_1}^2, \quad \beta_{C_1 P_2} = \omega_{C_1 P_2}/\omega_{C_1}^2 \qquad (11)$$

These are the univariate coefficients that capture the effects of a single covariate without the influence of any other. The corresponding UPV, conditional only on C_1, can be similarly calculated:

$$\omega_{P_1}^{2'} = \omega_{P_1}^2 - \beta_{C_1 P_1}\omega_{C_1 P_1}, \quad \omega_{P_2}^{2'} = \omega_{C_1 P_2}^{2'} - \beta_{C_1 P_2}\omega_{C_1 P_2} \quad (12)$$

For a more in-depth description of conditional coefficients, see Supplementary Material SC.

The FREM in Equations 5–8 is equivalent to a linear covariate-parameter relation. Other FREMs that are equivalent to other fixed-effect covariate effect models can be formulated by altering the IIV parameterization in the FREM, by transforming the covariate observation data, or by a combination of the two. Four common covariate-parameter relations were implemented in this work and are shown in Table 1.

Experiments

A previously developed model of circulating neutrophil counts after docetaxel treatment with original data from 601 patients,[14] was selected as a test case for investigating FREM and comparing it to FFEM. The model incorporated five compartments (3 transit compartments), additive and proportional residual error, and four structural parameters: neutrophil baseline (base), mean transit-time (MTT), slope of drug effect (slope), and feedback mechanism power (γ). To shorten runtimes γ was fixed to 0.154. The remaining three parameters were parameterized with exponentially distributed IIV.

Five covariates were available in the original data: age (years), sex (male or female), serum level of α_1-acid-glycoprotein (AAG; g/L), performance status grade (perf; 0, 1, or 2), and prior chemotherapy (yes or no). Five patients had no performance status observations and were excluded from all experiments, leaving 596 patients with no missing covariate observations. All covariate correlations were low (<40%).

The FREM implementation considers only continuous and dichotomous covariates, and the trichotomous perf covariate was therefore dummy-coded[15] according to Equation 13 into perf01 and perf12, for a total of six covariates.

$$\text{perf01} = \begin{cases} 1, \text{if perf} = 0 \\ 1, \text{if perf} = 1 \\ 0, \text{if perf} = 2 \end{cases}, \quad \text{perf12} = \begin{cases} 0, \text{if perf} = 0 \\ 1, \text{if perf} = 1 \\ 1, \text{if perf} = 2 \end{cases} \quad (13)$$

All simulations and estimations were performed using NONMEM[16] with the stochastic simulation and estimation procedure from Perl-speaks-NONMEM.[17] Importance

TABLE 1 Equivalent FFEM and FREM parameterizations of four common continuous covariate-parameter relations

Covariate-parameter relation	FFEM IIV parameterization	FREM IIV parameterization	FREM Covariate data transformation
Linear	$P_i = \theta_P + \beta_{CP}(C_i - \overline{C}) + \eta'_{P,i}$	$P_i = \theta_P + \eta_{P,i}$ $C_i = \overline{C} + \eta_{C,i}$	–
Log-linear	$P_i = \theta_P + \beta_{CP}(\ln C_i - \overline{\ln C}) + \eta'_{P,i}$	$P_i = \theta_P + \eta_{P,i}$ $C_i = \overline{\ln C} + \eta_{C,i}$	Log-transformed
Exponential	$P_i = \theta_P e^{\beta_{CP}(C_i - \overline{C}) + \eta'_{P,i}}$	$P_i = \theta_P e^{\eta_{P,i}}$ $C_i = \overline{C} + \eta_{C,i}$	–
Power	$P_i = \theta_P \left(\dfrac{C_i}{C_{GM}}\right)^{\beta_{CP}} e^{\eta'_{P,i}}$	$P_i = \theta_P e^{\eta_{P,i}}$ $C_i = \overline{\ln C} + \eta_{C,i}$	Log-transformed

Abbreviations: β_{CP}, covariate-parameter effect coefficient; C_{GM}, covariate population geometric mean; C_i, individual covariate value; $\overline{C}$, population mean of covariate; FFEM, full fixed-effects model; FREM, full random-effects model; IIV, interindividual variability; $\overline{\ln C}$, population mean of log-transformed covariate; θ_P, population parameter value; $\eta'_{P,i}$, individual deviation from population value unexplained by covariates; η_i, total individual deviation from population value; P_i, individual parameter value.

sampling with mode a posteriori (IMPMAP) was the main estimation method. In order to estimate FREMs without residual error on the covariate model in NONMEM, an additive residual error with a very small, fixed variance was added to the covariate observation model (see Supplementary Material SA). Data analysis and transformations were performed in R[18] with figures by ggplot2[19] and corrplot.[20]

Original data comparison to FFEM

To demonstrate the FREM method, and its ability to produce coefficients and variability for any covariate subset from a single estimation, FREM coefficient values for the complete set of covariates and two covariate subsets were compared to the corresponding fixed-effects estimates. Three FFEMs were estimated with covariate effects on all estimated structural parameters, one including all six covariates, one including a three-covariate set (AAG, perf01, and perf12) and one including a single covariate (AAG). The conditional coefficients and UPV for the full set and the two subsets were calculated from a single FREM estimation including all six covariates (see Supplementary Material SA).

Simulated data comparison to FFEM

The consistency of FREM compared to FFEM was examined in a simulation experiment. An FFEM simulation model with 18 covariate-parameter relations (see Supplementary Material SB) was used to generate 150 datasets. The simulated datasets were then re-estimated using both FFEM and FREM (initialized similarly), and the covariate-parameter relation coefficients and UPV were compared with respect to accuracy and agreement between methods.

High correlation

To investigate FREM performance in the presence of highly correlated covariates, a simulation experiment was performed with AAG and a highly correlated derived binary covariate, AAG_{hi}:

$$AAG_{hi} = \begin{cases} 0, & \text{if } AAG < 1.35 \\ 1, & \text{if } AAG \geq 1.35 \end{cases} \quad (14)$$

The breakpoint (1.35) was chosen such that half of the individuals were assigned to each category, and resulted in 76% correlation between AAG and AAG_{hi}. An FFEM simulation model with six covariate-parameter relations,

AAG and AAG_{hi} on each of base, MTT, and slope (see Supplementary Material SB) was used to generate 228 datasets. Similarly initialized FREM and FFEM methods were then re-estimated on each dataset, and the covariate-parameter relation coefficients and UPV were compared.

Parameterizations

The four covariate-parameter relations defined in Table 1, linear, log-linear, exponential, and power, were applied to the two continuous covariates, age and AAG. Equivalent FREM and FFEM estimates of covariate-parameter effect coefficients and UPV were compared.

Non-normal covariate distributions

The robustness of the FREM approach when applied to covariates of non-normal distributions was tested by comparing estimates for original data and data with log-transformed observations of the two available continuous covariates, age and AAG. An FREM model with age and AAG, and log-normally distributed parameters, was estimated on the two datasets, and the estimates of covariate-parameter effect coefficients and UPV were compared.

RESULTS

Original data comparison to FFEM

The FREM and FFEM methods were successfully applied to the original data. The estimated FREM matrix is presented on correlation scale in Figure 1, together with a covariate correlation matrix from R for comparison.

Covariate-parameter relation coefficients and UPV covariance matrices for all three scenarios could be accurately calculated from this single FREM estimation, closely matching the results of the three FFEM estimations. The AAG coefficients and UPV matrices from the three FFEM models are compared to their FREM equivalents in Table 2.

Simulated data comparison to FFEM

The common model parameters, structural parameters, residual error, and IIV covariance matrix showed no relevant differences in re-estimation accuracy between FREM and FFEM, although FREM IIV covariance matrix estimates had a tendency toward higher precision, as shown in Figure 2. The mean of the root mean squared errors (RMSEs) of the 18 coefficient estimators were 0.0272
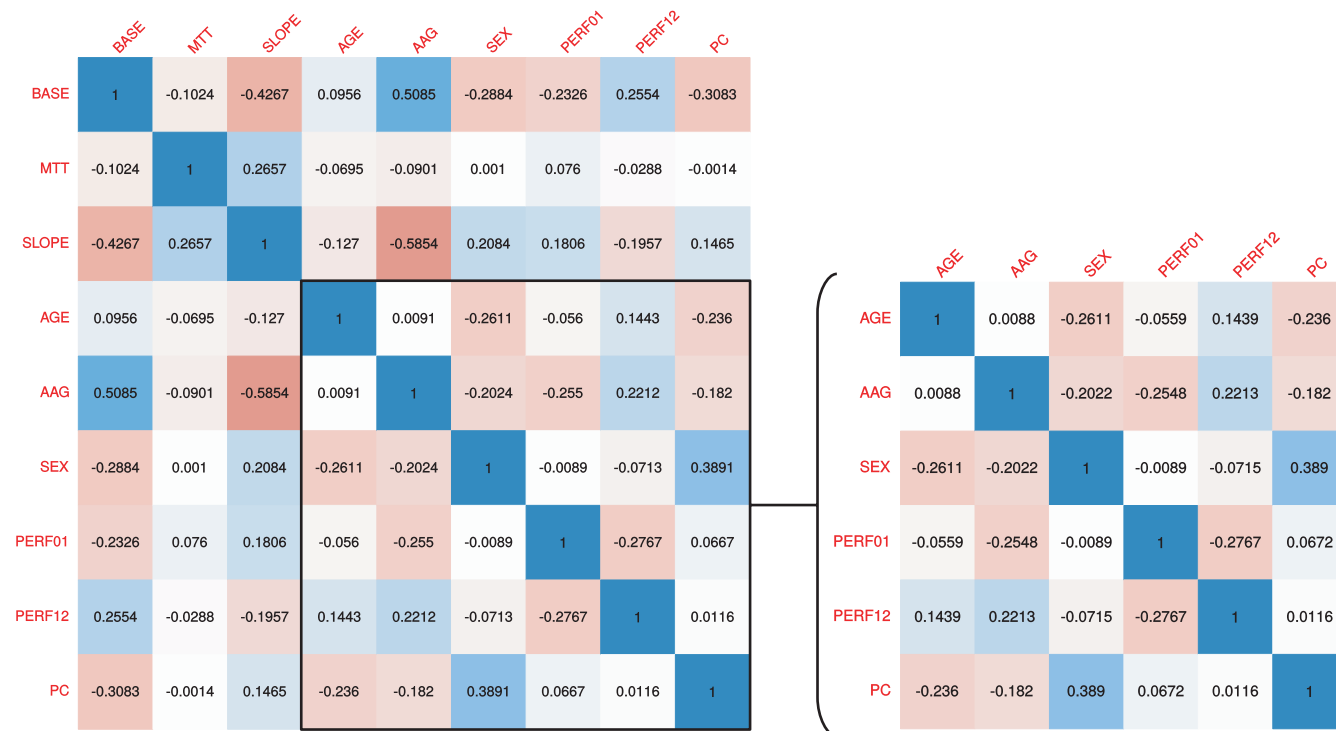
FIGURE 1 Correlation matrix of parameters (base, MTT, and slope) and covariates (age, AAG, sex, perf01, perf12, and PC; framed) from FREM estimation on original data (left), as compared to empirical correlation matrix of covariates from R (right). MTT, mean transit-time; PC, prior chemotherapy; PERF, performance status grade

TABLE 2 AAG covariate-parameter relation coefficients and UPV matrices from FREM and FFEM for three covariate sets

		AAG effect coefficients			UPV covariance matrix			
Covariate set	**Method**	$\beta_{AAG,base}$	$\beta_{AAG,MTT}$	$\beta_{AAG,slope}$		ω_{base}	ω_{MTT}	ω_{slope}
Full set (age, AAG, sex, perf01, perf12, PC)	FREM	0.290	−0.0247	−0.480	ω_{base}	0.106		
					ω_{MTT}	−0.00359	0.0224	
					ω_{slope}	−0.019	0.015	0.145
	FFEM	0.291	−0.0247	−0.481	ω_{base}	0.106		
					ω_{MTT}	−0.00356	0.0224	
					ω_{slope}	−0.019	0.015	0.144
Three covariates (AAG, perf01, perf12)	FREM	0.336	−0.0211	−0.490	ω_{base}	0.117		
					ω_{MTT}	−0.00325	0.0225	
					ω_{slope}	−0.0226	0.0154	0.148
	FFEM	0.335	−0.0215	−0.491	ω_{base}	0.117		
					ω_{MTT}	−0.00333	0.0225	
					ω_{slope}	−0.022	0.0158	0.148
One covariate (AAG)	FREM	0.372	−0.0246	−0.506	ω_{base}	0.121		
					ω_{MTT}	−0.00356	0.0226	
					ω_{slope}	−0.0247	0.0155	0.149
	FFEM	0.372	−0.0249	−0.507	ω_{base}	0.122		
					ω_{MTT}	−0.00357	0.0226	
					ω_{slope}	−0.0239	0.016	0.149

The FREM results were calculated from a single estimation using the full set of covariates, while the FFEM results are estimated separately for each covariate set.
Abbreviations: AAG, serum level of α1-acid-glycoprotein; FFEM, full fixed-effects model; FREM, full random-effects model; MTT, mean transit-time; perf, performance status grade; UPV, unexplained parameter variability.
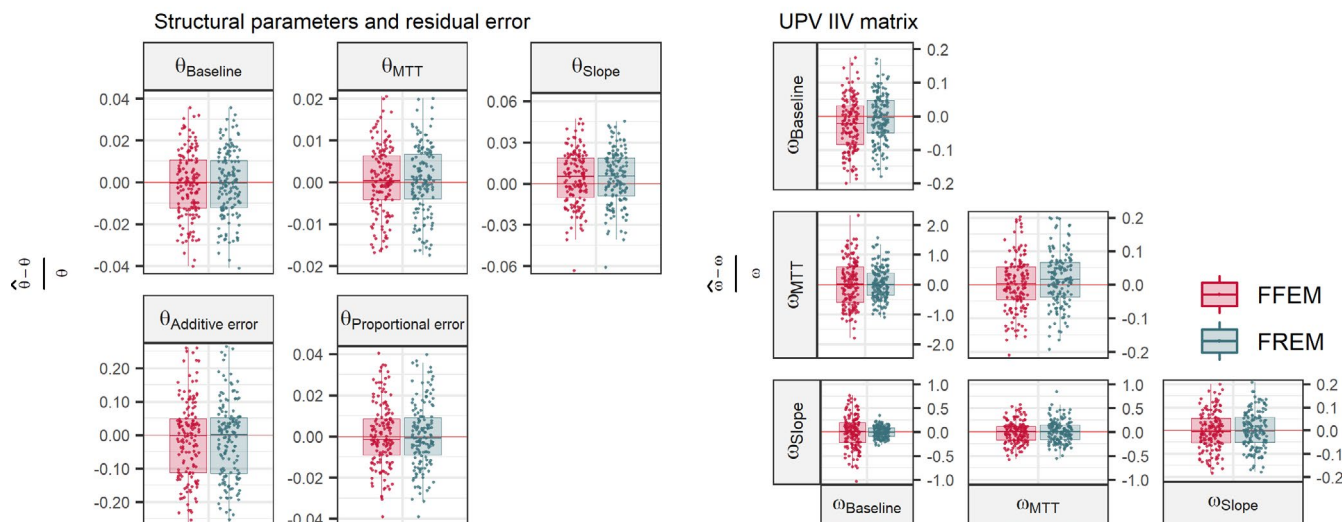
FIGURE 2 Re-estimation accuracy for structural parameter population values and residual error (left), and the unexplained parameter variability (UPV) covariance matrix (right). Accuracy is here calculated as estimated mean value minus the true value, divided by the true value. FFEM, full fixed-effects model; FREM, full random-effects model; IIV, interindividual variability

and 0.0271 for FFEM and FREM, respectively. The ratio of the mean of the standard deviation (of the coefficient estimates) was 1.0015 (FFEM/FREM). See Figure 3 for a graphical overview of the FFEM, multivariate FREM (conditional on all 6 covariates), and univariate FREM coefficient estimation accuracy. The higher precision of the univariate FREM coefficients is due to correlated covariates being disregarded, and the univariate coefficients describing the effect of that single covariate in isolation.

High correlation

Multivariate coefficients and UPV estimates from FREM and FFEM were very similar. The mean of the RMSEs of the six coefficient estimators were 0.0412 for both FFEM and FREM, and the ratio (FFEM/FREM) of the mean of the standard deviation of the coefficient estimators was 1.0005. Univariate covariate-parameter coefficients calculated from the FREM estimation, demonstrates higher estimation precision than multivariate estimates (see Figure 4), showing how univariate coefficients can be utilized in the presence of highly correlated covariates. A similar trend was seen for the previous experiment, see Figure 3, although the effect is not as pronounced there, due to the covariates being less correlated.

Parameterizations

The four covariate-parameter relation implementations all produced coefficients and UPV estimates to the

equivalent FFEM models, as shown in Table 3. Parameter uncertainty was small for both methods, at most 4.77% coefficient of variation in FREM ($\omega_{AAG,MTT}$) and 4.27% in FFEM ($\beta_{AAG,MTT}$) for log-normally distributed parameters.

Non-normal covariate distributions

The expected empirical age–AAG covariance matrix, given the current FREM parameterization, was accurately estimated using both original and log-transformed data. See the linear and log-linear FREM results in Table 3, where accurate estimates of coefficients and UPV are dependent on accurate covariate covariance matrices.

DISCUSSION

Experiments in both original and simulated data show that both FREM and FFEM can accurately estimate the base model parameters, the IIV, and the covariate effects. The results of the simulation data comparison suggest that the FREM estimate of the parameter IIV covariance matrix is slightly more precise, particularly the off-diagonal elements of the UPV IIV covariance matrix (see Figure 2). This is likely due to FREM estimating the TPV, which is more precise than the UVP IIV obtained after including structural covariate relationships. The same behavior was observed in the high correlation experiment (results not shown).

A unique feature of FREM is its ability to provide coefficient estimates for any subset of covariates from a single estimation. The first conditional coefficients to consider
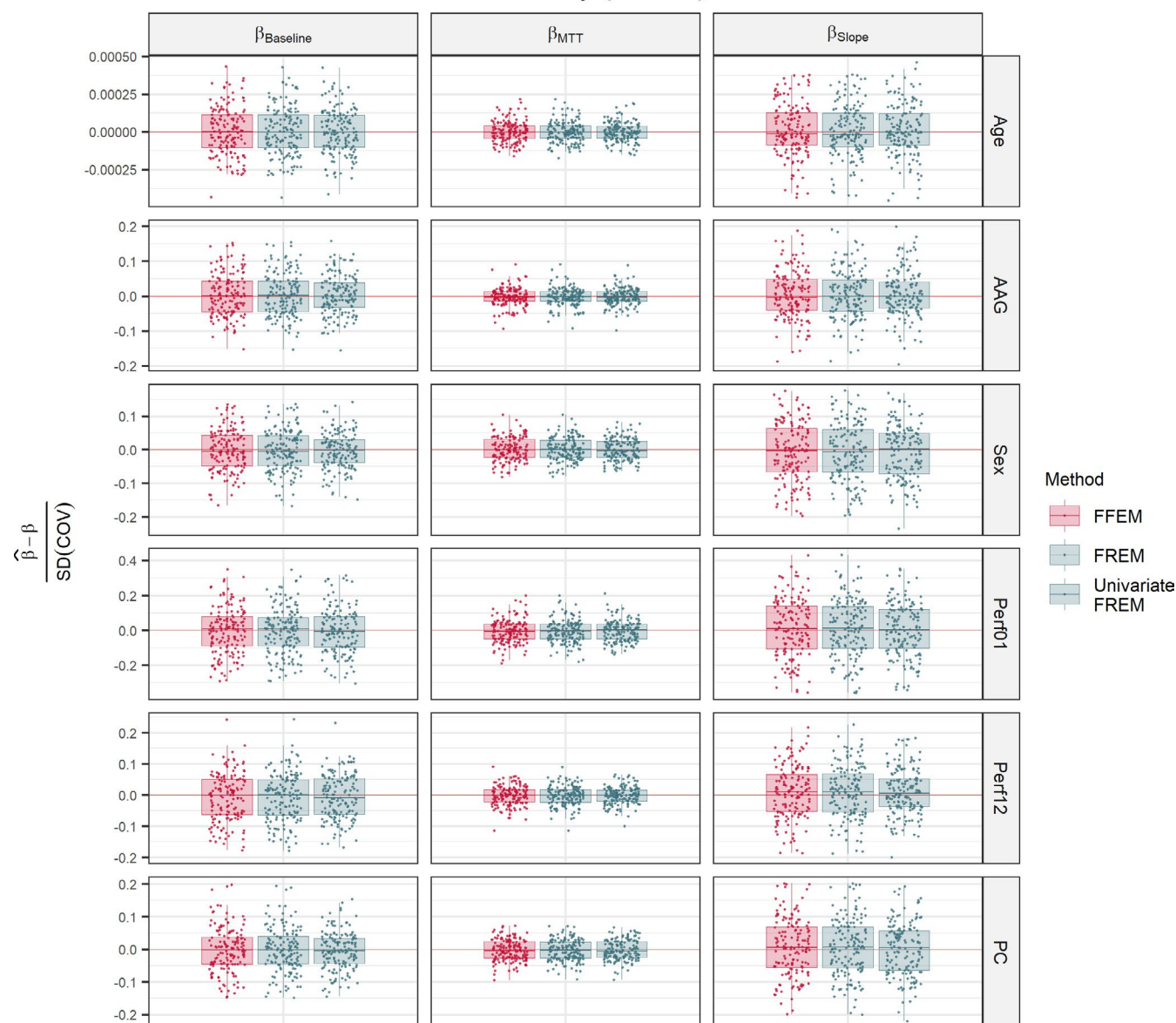
FIGURE 3 Re-estimation accuracy for parameter-covariate relation coefficients estimated by FFEM and FREM, as well as univariate FREM coefficients calculated from the FREM estimation. Each column of panels presents one parameter, and each row one covariate. Accuracy is here calculated as estimated coefficient value minus the true coefficient value, divided by the standard deviation of the covariate. The reason for normalizing to the standard deviation is to be able to compare univariate and multivariate coefficients. FFEM, full fixed-effects model; FREM, full random-effects model; PC, prior chemotherapy; Perf, performance status grade

are the univariate FREM coefficients. These provide precise measures of the explanatory effect of each covariate in isolation, making them very useful tools in covariate analyses. They are especially helpful in communicating covariate effects, because it is not necessary to include any caveats or conditional assumptions regarding other covariates. The accuracy of these univariate coefficients is seen in Figures 3 and 4, where the multivariate coefficients are less accurate due to the inclusion of correlated covariates. One aim of removing correlated covariates

in FFEM is to obtain coefficient estimates that are essentially univariate, but their highly correlated nature means that either many covariates must be excluded, or some correlations must be accepted. To acquire similar metrics in fixed-effects methods, the effects of each covariate must instead be estimated separately, resulting in a large number of models, and potentially introducing bias and uncertainty.

Beyond the univariate coefficients, we have also shown how the results of any covariate subset can be retrieved

FIGURE 4 Re-estimation accuracy for multivariate parameter-covariate relation coefficients estimated by FFEM and FREM, as well as univariate FREM coefficients calculated from the FREM estimation. Accuracy is here normalized to the standard deviation of the covariate in order to compare the multivariate and univariate coefficients. FFEM, full fixed-effects model; FREM, full random-effects model; MTT, mean transit-time
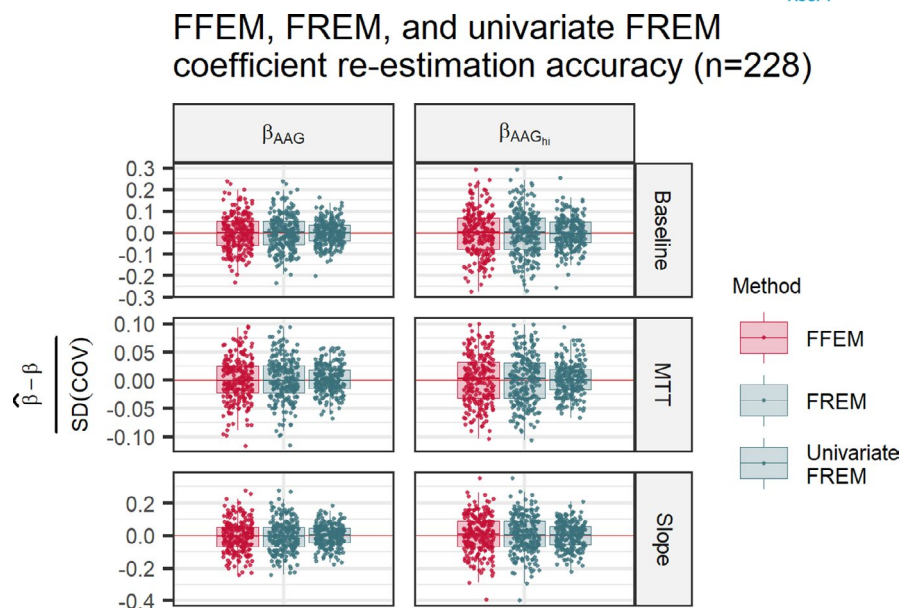
from a single FREM estimation. The ability to include a large set of covariates in the FREM estimation, and then condition the analysis on different subsets, has several applications. The results of the full set may, for example, be more relevant for inference, whereas a limited set may be more useful for prediction. High quality, covariate-rich data may provide the opportunity to estimate many predictive relationships in a particular model, but only some of these may be available in a specific population where the model is to be applied. Such adjustments are possible from the single FREM estimation. The impact of knowing different sets of covariates can also be elucidated without making changes to the model, which in turn can help consolidate data from different trials or guide the design of future trials.

The final subset to consider is the empty set (i.e., the FREM IIV estimates conditioned on no covariates). It may be argued that this is obtainable from estimating the base model without covariate effects, but two subtleties need to be considered: (i) the estimation uncertainty may differ, especially because covariate observations add information; and (ii) TPV is not necessarily equal to explained parameter variability + unexplained parameter variability.[21] The differences of uncertainty in FFEM and FREM, as well as the different approaches to retrieve uncertainty of the calculated measures of FREM (conditional coefficients and variability), have not been explored and warrants further study. Either propagation of uncertainty or empirical simulation retrieval may be attempted. For the latter, the differences in explained and unexplained parameter variability have not been explored in depth either. However, the perspectives that the natural separation of TPV and UPV provide is an advantage of FREM.

Although the effects of any covariate subset on all parameters can be extracted from the FREM estimates,

it is not trivial to exclude specific relations. This means that physiologically improbable effects, such as creatinine clearance on absorption rate, will be present in the FREM method as long as there is a non-zero correlation between individual values of the parameter and the covariate. Keeping all relations maintains the benefits of the full model approach, acknowledges imperfections in the data, and avoids bias. Specific covariate effects with strong support, such as maturation of clearance or allometric scaling, can be implemented before FREM is applied.

We have shown how to implement four common parameter-covariate relations in FREM and obtain coefficient estimates equivalent to fixed-effect implementations. Although additional relations can undoubtedly be implemented, there are relations that cannot be formulated for FREM. This does not exclude the use of FREM, but it means that calculated coefficients have a different scale than those from an FFEM method would. This restriction does, however, come with some subtle advantages. For example, a parameter restricted to [0,1] with logit-normally distributed IIV requires no extra consideration for restricting the effects to the same domain.

As observed in the non-normal covariate distribution experiment, covariates are not required to be normally distributed for FREM to produce the expected results. This is also supported by the accurate estimation of coefficients for categorical covariates in other experiments. All that is required is that the estimation method can produce the arithmetic mean and variance of the covariates (i.e., the normal distribution parameterization), even if the true distribution is not a multivariate normal.

The presented examples do not address data with multiple covariate observations per subject or missing covariate observations. Multiple observations can simply

TABLE 3 Age and AAG covariate-parameter relation coefficients and UPV matrices from FREM and FFEM implementations of four common parameterizations: linear, log-linear, exponential, and power.

Relation	Method	Age effect coefficients			AAG effect coefficients			UPV covariance matrix			
		Base	MTT	Slope	Base	MTT	Slope		Base	MTT	Slope
Linear	FREM	0.0188	−0.0837	−0.104	2.25	−2.12	−7.99	Base	3.66		
								MTT	−1.69	154	
								Slope	−1.76	25.1	47.4
	FFEM	0.0188	−0.0831	−0.104	2.25	−2.03	−7.98	Base	3.65		
								MTT	−1.63	155	
								Slope	−1.77	24.6	47.4
Log-linear	FREM	0.801	−3.78	−4.64	3.03	−2.18	−12.2	Base	3.88		
								MTT	−2.03	156	
								Slope	−2.08	26.7	46
	FFEM	0.811	−3.76	−4.62	3.03	−2.17	−12.2	Base	3.86		
								MTT	−2.13	156	
								Slope	−2.07	26.8	46.8
Exponential	FREM	0.00378	−0.000960	−0.00587	0.373	−0.0252	−0.507	Base	0.120		
								MTT	−0.00314	0.0224	
								Slope	−0.0214	0.0154	0.145
	FFEM	0.00364	−0.000962	−0.00564	0.372	−0.0245	−0.505	Base	0.120		
								MTT	−0.00315	0.0225	
								Slope	−0.0213	0.0154	0.145
Power	FREM	0.158	−0.0460	−0.242	0.532	−0.0251	−0.713	Base	0.123		
								MTT	−0.00381	0.0228	
								Slope	−0.0255	0.0169	0.152
	FFEM	0.159	−0.0419	−0.239	0.533	−0.0253	−0.715	Base	0.123		
								MTT	−0.00384	0.0228	
								Slope	−0.0256	0.017	0.153

Abbreviations: AAG, serum level of α1-acid-glycoprotein; FFEM, full fixed-effects model; FREM, full random-effects model; MTT, mean transit-time; UPV, unexplained parameter variability.

be ignored by using a baseline or average value, but there may be valuable information in these variations that can be extracted in several ways. Random variations can be captured by introducing a residual error term in the FREM covariate model, and any time-dependence of time-varying covariates can be modeled there as well. It is also possible to consider observations at different occasions as separate covariates and estimating their covariance with interoccasion random effects.[11] For missing covariate observations, there is an inherent support for the handling of these in the covariate covariance matrix estimated by FREM. This topic is further explored by Nyberg et al.[13]

Estimation efficiency was not an aim of this work, and all estimations used identical IMPMAP settings for parsimonious reasons. Both gradient-based estimation methods, such as first-order conditional estimation (FOCE), and expectation maximization (EM) methods, such as IMPMAP, can be suitable for FREM estimation. The authors have observed minimization issues with FOCE in some models. This may be caused by the FREM matrix having a small positive-definite space, and can sometimes be alleviated by using an EM method. Another reason to consider EM methods is that the FREM method adds additive random effects with known means and sampling variances, which enables efficient Markov chain Monte Carlo sampling. The FREM may also be suitable for linearization around the population estimates,[22] a principle that has previously been successfully applied to step-wise covariate modeling.[23]

CONCLUSION

We have proposed FREM, a full random effects approach to covariate modeling. We have demonstrated its unique advantages over full models of fixed effects, and have discussed its current limitations regarding the data and covariate-parameter relations that it supports. The advantages include covariate correlation management and conditional interpretations of covariate effects. Our results support FREM as a more informative alternative to FFEM. We have also highlighted areas where further study is warranted, such as time-variant covariates, missing data, and covariates measured with error.

CONFLICT OF INTEREST
The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS
G.Y. and H.B.N. wrote the manuscript. G.Y., H.B.N., J.N., E.N.J., and M.O.K. designed the research. G.Y., H.B.N., and J.N. performed the research and analyzed the data.

REFERENCES

1. Jonsson EN, Karlsson MO. Automated covariate model building within NONMEM. *Pharm Res*. 1998;15(9):1463-1468. http://doi.org/10.1023/a:1011970125687

2. Ribbing J, Jonsson EN. Power, selection bias and predictive performance of the population pharmacokinetic covariate model. *J Pharmacokinet Pharmacodyn*. 2004;31(2):109-134. http://doi.org/10.1023/b:jopa.0000034404.86036.72

3. Ivaturi V, Hooker AC, Karlsson MO. Selection bias in pre-specified covariate models. In PAGE Abstract #2228 (Athens, Greece, 2011).

4. Bonate PL. The effect of collinearity on parameter estimates in nonlinear mixed effect models. *Pharm Res*. 1999;16(5):709-717. http://doi.org/10.1023/a:1018828709196

5. Gastonguay MR. A full model estimation approach for covariate effects: Inference based on clinical importance and estimation precision. *AAPS J*. 2004;6:W4354.

6. Wählby U, Jonsson EN, Karlsson MO. Assessment of Actual Significance Levels for Covariate Effects in NONMEM. *J Pharmacokin Pharmacodyn*. 2001;28(3):231-252. http://doi.org/10.1023/a:1011527125570

7. Gastonguay MR. Full covariate models as an alternative to methods relying on statistical significance for inferences about covariate effects: a review of methodology and 42 case studies. In PAGE Abstract #2229 (Athens, Greece, 2011).

8. Hu C, Adedokun O, Ito K, Raje S, Lu M. Confirmatory population pharmacokinetic analysis for bapineuzumab phase 3 studies in patients with mild to moderate Alzheimer's disease. *J Clin Pharmacol*. 2015;55(2):221-229. http://doi.org/10.1002/jcph.393

9. Karlsson MO. A full model approach based on the covariance matrix of parameters and covariates. In PAGE Abstract #2455 (Venice, Italy, 2012).

10. Novakovic AM, Krekels EHJ, Munafo A, Ueckert S, Karlsson MO. Application of item response theory to modeling of expanded disability status scale in multiple sclerosis. *AAPS J*. 2017;19(1):172-179. http://doi.org/10.1208/s12248-016-9977-z

11. Brekkan A, Lopez-Lazaro L, Yngman G, et al. A population pharmacokinetic-pharmacodynamic model of Pegfilgrastim. *AAPS J*. 2018;20(5). http://doi.org/10.1208/s12248-018-0249-y

12. Abrantes JA, Solms A, Garmann D, Nielsen EI, Jönsson S, Karlsson MO. Integrated modelling of factor VIII activity kinetics, occurrence of bleeds and individual characteristics in haemophilia A patients using a full random effects modelling (FREM) approach. In PAGE Abstract #8646 (2018).

13. Nyberg J, Karlsson MO, Jonsson EN. Implicit and efficient handling of missing covariate information using full random effects modelling. In PAGE Abstract #9181 (2019).

14. Friberg LE, Henningsson A, Maas H, Nguyen L, Karlsson MO. Model of chemotherapy-induced Myelosuppression with parameter consistency across drugs. *J Clin Oncol*. 2002;20(24):4713-4721. http://doi.org/10.1200/jco.2002.02.140

15. Draper NR, Smith H. "Dummy" variables. *Applied Regression Analysis, Third Edition*. Hoboken, NJ: John Wiley & Sons, Inc.; 1998; 299-325.

16. Beal S, Sheiner LB, Boeckmann A, Bauer RJ. *NONMEM 7.4 User's Guides (1989–2018)*. Icon Development Solutions, 2017.

17. Keizer RJ, Karlsson MO, Hooker A. Modeling and simulation workbench for NONMEM: Tutorial on Pirana, PsN, and

Xpose. *CPT: Pharmacomet Syst Phar.* 2013;2(6):50. http://doi.org/10.1038/psp.2013.24

18. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing; 2017.

19. Wickham H. *Ggplot2: Elegant graphics for data analysis*; 2009.

20. Wei T, Simko V. Corrplot: Visualization of a correlation matrix (Version 0.77); 2016. https://cran.r-project.org/package=corrplot.

21. Hennig S, Karlsson MO. Concordance between criteria for covariate model building. *J Pharmacokin Pharmacodyn.* 2014;41(2):109-125. http://doi.org/10.1007/s10928-014-9350-8

22. Yngman G, Nordgren R, Freiberga S, Karlsson MO. Linearization of full random effects modeling (FREM) for time-efficient automatic covariate assessment. In PAGE Abstract #8750 (Montreux, Switzerland, 2018).

23. Khandelwal A, Harling K, Jonsson EN, Hooker AC, Karlsson MO. A fast method for testing covariates in population PK/ PD models. *AAPS J.* 2011;13(3). http://doi.org/10.1208/s12248-011-9289-2

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.