

## TUTORIAL

# NONMEM Tutorial Part II: Estimation Methods and Advanced Examples

Robert J. Bauer<sup>1,\*</sup>

In this second tutorial on NONMEM, the examples of typical pharmacokinetic/pharmacodynamic modeling problems that occur in the pharmaceutical field will be presented, which the reader can use as a template for his or her own modeling endeavors. Each of the problems presented is challenging in some way, and the logic behind setting up each problem is discussed. Logical concepts of the problem itself as well as the technical aspect of how to set it up in NONMEM are described and demonstrated. The concepts behind the various estimation algorithms will first be described to allow the user a better understanding of how to use them.

These estimation methods include:<sup>1</sup>

The first-order (FO) method (guide II<sup>1</sup>)  
 The FO conditional estimation (FOCE)/Laplace method (guide VII<sup>1</sup>)  
 Various additional methods in NONMEM are (intro7<sup>1</sup>):  
 Monte Carlo importance sampling (IMP) expectation maximization (EM)  
 Markov Chain Monte Carlo (MCMC) stochastic approximation expectation maximization (SAEM)  
 Iterative two stage (ITS). Approximate EM method  
 MCMC Bayesian analysis

The examples consist of the following:

Example 1: Pharmacokinetic (PK) model with covariates using IMP, SAEM, and Bayesian analysis methods  
 Example 2: Modeling data that are below the limit of quantitation (LOQ)  
 Example 3: Modeling pharmacokinetic categorical response data  
 Example 4: Using ordinary differential equation (ODE) solvers to model a basic target-mediated drug disposition model  
 Example 5: Analyzing data modeled with multiple levels of mixed effects  
 Example 6: Modeling a mixture of subpopulations of parameters  
 Example 7: Modeling periodically collected urine samples

## BASIC PRINCIPLES AND HISTORY OF THE VARIOUS METHODS

### Goal of nonlinear mixed effects methods

The common goal of FO and EM methods is to determine the set of fixed effects THETAs, OMEGAs, and SIGMAs that best fit the population data, considering all possible values of individual parameters or ETAs (random effects). For

each subject, a linearized, often weighted, extended least-squares assessment of the contributions of within-subject and between-subject variability is made (FO), or an integration of the conditional density over all values of ETAs must be performed, for a given set of THETAs, OMEGAs, and SIGMAs (FO/FOCE/Laplace and EM methods). This integration is computationally difficult to do, and the various methods solve this problem in different ways.

### FIRST METHOD WAS FO (1970s<sup>2</sup>; GUIDE II<sup>1</sup>)

The FO was the first method used in population PK analysis that could simultaneously discern the variability of measured levels of drug or drug response (residual variance) within a subject and the variability of PK/pharmacodynamic (PD) parameters between subjects (intersubject variance). The FO method could determine how population PK/PD parameters related to patient characteristics, even when there were few data points per subject. FO analysis could be accomplished using the computing power and memory that was available at the time. The FO method optimizes an objective function with a linearized projection of intersubject and intrasubject effects without explicitly evaluating the integral of the joint density for each subject and hence evaluates very quickly.

### FOCE (1992) METHOD (GUIDE VII<sup>1</sup>)

Although the FO method was fast, it was very approximate. Sometimes inaccurate assessments occurred if residual error and/or intersubject variability were large. The FOCE, although also an approximate method, was more accurate for a larger variety of problems. In FOCE mixed effects modeling, an integral over all possible individual parameter values (ETAs, or random effects) is taken into consideration for each subject's joint density of observed data and ETAs when determining the best fixed effects (THETAs, OMEGAs, and SIGMAs). As mentioned previously, this integration is computationally expensive.

<sup>1</sup>Pharmacometrics R&D, ICON Clinical Research LLC, Gaithersburg, Maryland, USA. \*Correspondence: Robert J. Bauer ([robert.bauer@iconplc.com](mailto:robert.bauer@iconplc.com))

Received: January 16, 2019; accepted: April 9, 2019. doi:10.1002/psp4.12422

To alleviate the computational expense, the FOCE evaluates the mode of the joint density (most likely values of ETAs) and the first order approximation of variances of ETAs. An approximate integral using the Gaussian function centered at the mode of the joint density with the approximate variance is used as a linear approximation of the integral of the joint density with respect to ETAs and can be easily calculated. These integrations must be done for the data of each individual separately, and so FOCE takes longer to evaluate than the FO method. Thus, unlike FO, which linearizes the intersubject and within-subject variability, FOCE evaluates the intersubject effect accurately while linearizing the within-subject variability with its Gaussian function approximation.

### SECOND-ORDER CONDITIONAL ESTIMATION (LAPLACE, 1992) METHOD (GUIDE VII<sup>1</sup>)

The Laplace method is the same as FOCE except that it uses a second derivative assessment of the variance of the joint density with respect to the ETAs. The Laplace method is used when nonnormal densities are used to model some of the observed data or the prediction model imparts a high degree of nonlinearity to the joint density of the individual parameters.

### MONTE CARLO EM METHODS (2000s<sup>3-5</sup>)

These are exact methods that are able to analyze complex PK/PD problems with a greater incidence of success than FOCE. As in FOCE, the integral over all possible individual parameter values (ETAs, or random effects) is taken into consideration when determining the best fixed effects (THETAs, OMEGAs, and SIGMAs). The EM algorithms, however, perform a Monte Carlo integration to explore the entire ETA space and provide a more accurate (but more stochastic) assessment of the integral of the joint density. This is called the “expectation step.” Although the Monte Carlo expectation step can be computationally expensive, and/or highly stochastic, the update of the fixed effect parameters can be efficiently carried out with single-step algorithms if the statistical model is structured in a certain way (PHI/MU modeling). This update of the fixed effects is called the “maximization step.” The Monte Carlo EM can be more accurate than FOCE for sparse data,<sup>3,4</sup> although less precise than FOCE because the results have stochastic variability. It takes longer than FOCE for simple PK/PD problems (examples 1–3 and 6 in this tutorial) but can be more efficient than FOCE for complex PK/PD problems (such as example 4). By *simple*, this means analytical, linear PK problems and analytically evaluated PD problems, with perhaps no more than six model parameters to be estimated. Complex PK/PD problems are those that require ODEs or problems with more than six model parameters to be estimated. Efficiency reduces considerably when the model cannot be expressed in a particular fixed/random effect (PHI/MU) format (explained in the MU Referencing section). Importance sampling and SAEM are the two efficient Monte Carlo EM methods available in NONMEM.

### HOW IMP AND SAEM EXPECTATION STEPS ARE IMPLEMENTED

#### IMP: Importance Sampling Algorithm

In IMP, the expectation step is performed by creating random normal deviates of ETAs and evaluating the conditional density (joint density normalized to integrate with respect to ETAs to the value of 1) at these ETAs for each subject. The normal random sampler is typically centered at the mode or the mean of the conditional density and a variance that approximates the variance of the conditional density; hence it samples in the “important” region of the conditional density. On the first iteration, this information is not available, so the mode and its FO approximation of the variance (called *mode a posteriori* (MAP) estimation, as is done in FOCE) is obtained and used as the mean and variance for the sampler. On subsequent iterations, the MAP estimation may be repeated to obtain the normal random sampler parameters, or the conditional Monte Carlo mean and conditional Monte Carlo variance obtained from the IMP of the previous iteration may be used. With this method, weighted averages of conditional means and variances of individual parameters (or ETAs) as well as accurate (with stochastic variation) assessments of the objective functions for each individual. The weight to each ETA sample is proportional to the goodness of fit of the sample relative to the probability of the sampler to select that sample, and it reflects the properties of the exact conditional density. Usually 300–1,000 random samples are used for each subject, and about 50–200 iterations are required to approach the maximum likelihood population parameters.

#### SAEM: MCMC Sampling Algorithm

As in IMP, random samples are generated from normal proposal densities. Instead of centered at the mode (or mean) of the conditional density, the proposal density is centered at the previous sample position. New samples are accepted with a probability that is related to the conditional density (goodness of fit) at the particular ETA sample position. The variance of the proposal density is adjusted to maintain a certain average acceptance rate. This method requires more elaborate sampling strategy but is useful for highly nonnormally distributed conditional densities.

Unlike IMP that evaluates the problem in a single mode, the SAEM requires two modes of estimation. In the first mode, SAEM evaluates an unbiased but highly stochastic approximation of individual parameters (pseudo-expectation, usually two samples per individual). Population parameters are updated from individual parameters by single-iteration maximization steps that are very stable and statistically proven to improve the objective function (usually in 300–2,000 iterations). In the second mode, individual parameter samples from previous iterations are averaged together, converging toward the true conditional individual parameter means and variances. This leads to population parameters converging toward the maximum of the exact likelihood. The SAEM is not able to assess an objective function that can be used in hypothesis testing or for assessing final population goodness of fit. The objective function and standard errors are best obtained by one or a few iterations of a final IMP step

with population parameters fixed at the final SAEM values. This is expectation-only IMP because the maximization of population parameters has already been achieved by the SAEM step.

### **APPROXIMATE (LINEARIZED) LIKELIHOOD EM METHOD: ITS (1984<sup>2</sup>)**

This is an approximate conditional method that is able to analyze complex PK/PD problems with great efficiency and incidence of success. Although it is more accurate than FO, it is not as accurate as FOCE (NONMEM7 Technical Guide<sup>1,3</sup>) when data are not rich (a subject is rich in data when there are many more data points than ETAs being estimated) and/or residual variability of data is large. The efficiency reduces considerably when the model cannot be expressed in a particular fixed/random effect (PHI/MU) format (MU Referencing section). This method is considered a deterministic EM method. The ITS evaluates the conditional mode (not mean!) and first order (expected) approximation of the variance of parameters of individuals by maximizing the conditional density. This integration (expectation) step is the same as in FOCE. The parameters are updated using the mean of the conditional modes and approximate individual variances, and therefore it is less accurate than the Monte Carlo EM methods. They are updated by single-iteration maximization steps that are very stable (usually in 50–100 iterations). For rich data, ITS is almost as accurate as FOCE but much faster.

#### **Full OMEGA blocks in EM analysis**

Full OMEGA blocks are more easily updated by EM methods than by FOCE. They are preferred over diagonal OMEGAs in EM problems. Having off-diagonal elements does not necessarily make EM methods less stable, so there is no need to fix off-diagonal elements to 0 even though the standard error may be greater than the estimate. This is because every OMEGA element (including off-diagonal elements) is updated by evaluating the sample variance of the many randomly assessed individual parameters. Thus, if the covariance of clearance~volume CL~V is reported in the NONMEM output as, for example, mean +/- standard error,  $0.0018 \pm 0.0113$ , this is an unconstrained assessment that this particular off-diagonal element is not statistically different from 0. Performing a reassessment of the problem by fixing this off-diagonal element to 0 is not necessary for stabilization of the problem, as is sometimes necessary with FOCE, where the off-diagonal element is assessed by a regression process using derivatives and could challenge the stability of the estimation. If the original analysis was performed with FOCE, and if the off-diagonal element was fixed to 0 because FOCE had round-off error problems in estimation, or the \$COV step could not complete, then allow the off-diagonal to be estimated when doing EM. Then, only if the EM objective function is highly variable, or \$COV step reports non-positive-definite issues, consider fixing off-diagonal elements back to 0.

### **Usefulness and efficiency of each method**

The following lists the general properties of each method and for what types of problems can they be most useful. It should be pointed out that these are just general guidelines, so the bullet points are intended to be general rules of thumb, and each problem must be assessed individually. It is reasonable to use a trial-and-error approach, and some methods can be improved in efficiency or accuracy by modifying their options. Also, there is considerable overlap in the usefulness of each of these methods for various types of problems as we shall see in the examples.

#### **Monte Carlo IMP EM**

- Complex PK/PD problems with ODEs and/or many parameters
- Sparse (fewer data points per subject than ETAs to be estimated) or rich data
- Can be less accurate than SAEM with highly categorical data or very sparse data
- Can track progress of improvement in true objective function with each iteration
- Results can vary stochastically, typically by about 25% of SE
- Can handle full OMEGA blocks well
- Less efficient when some or many THETAs may not be MU referenced (described later)

#### **SAEM**

- Categorical data
- Very sparse, sparse, or rich data
- Complex PK/PD problems with many parameters (may sometimes reach true objective function only within  $\pm 10$  units of optimum and can take longer than IMP)
- Cannot assess true objective function during its progress, must finish analysis with IMP assessment of objective function
- Results can vary stochastically, typically by about 25% of SE
- Can handle full OMEGA blocks well
- Less efficient when some or many THETAs may not be MU referenced

#### **ITS**

- Rich data
- Rapid, exploratory method
- Can be used as preanalysis to facilitate IMP or SAEM
- Requires less fuss with adjusting options than SAEM or IMP
- Results are highly reproducible to  $\pm 4$  digits
- Can have large bias or instability for some problems
- Can handle full OMEGA blocks well
- Less efficient when some or many THETAs may not be MU referenced

## FOCE

- Rich and semirich data
- Does not require MU referencing
- Good for many THETAS with no ETAs associated with them
- More accurate than ITS
- Requires less fuss with adjusting options than SAEM or IMP
- Results are highly reproducible to  $\pm 4$  digits
- Does not handle full OMEGA blocks as easily as EM
- If convergence criterion reduced (along with reducing precision criteria for ODE problems), then time of analysis can be reduced by twofold to threefold in some cases. Using MU referencing and the FAST option, additional increases in speed occur.

In theory, all problems can be solved by EM methods. For Monte Carlo EM methods to be efficient, however, they rely on a particular statistical structure of the population analysis problem (gray box process) in which fixed effects parameters (THETAs) are involved only in defining the mean ( $\mu$ ) of the normal population distribution of individual parameters (PHI). That is, most THETAs need to be referenced in a PHI/MU format (MU referencing). If the mu-referencing process is too complicated for a problem and/or there are many parameters to be estimated that cannot be or are not easily MU referenced, then FOCE/Laplace would be the preferred method of analysis of such a problem.

### MU referencing

THETAs that define typical values of individual parameters are “mu related” (Greek letter  $\mu$  to indicate mean of individual parameters). So, THETAs that are associated with ETAs can be MU referenced. Fixed effect THETAs that have no intersubject variability associated with them but serve as a shared value among subjects are not mu related. To best utilize EM methods, it is best to create MU reference equations in NMTRAN to describe these mu relations.

The user supplies information on how THETA parameters are associated arithmetically with the ETAs and individual parameters:

$$\text{Phi}_j = \mu_j(\text{theta}) + \text{eta}(j)$$

For each parameter  $j$  that has an ETA associated with it, and  $\mu_j$  is a function of THETA.

The association of one or more THETAs with ETA(1) must be identified by a variable called MU\_1.

For ETA(2), the variable is MU\_2, etc.

For example, suppose the original code is

```
TVCL=(THETA(1)*(AGE/50)**THETA(2))
CL=TVCL*EXP(ETA(5))
TVV=THETA(3)
V=TVV*EXP(ETA(3))
```

This code can be modified to include intermediate variables MU\_x, which inform NONMEM of the relationship of certain THETAs with certain ETAs:

```
LTVCL= LOG(THETA(1))+THETA(2)*LOG(AGE/50)
MU_5= LTVCL
LTVV=LOG(THETA(3))
MU_3=LTVV
CL=EXP(MU_5+ETA(5))
V=EXP(MU_3+ETA(3))
```

In the above example, LTVCL stands for natural log of typical value clearance. Because the individual parameters CL and V are log-normally distributed ( $\exp(\text{ETA})$ ), the MUs are the log of the typical values so that they are arithmetically associated with the ETAs. Even better, one may impose a linear relationship between MU and THETA:

```
LTVCL= THETA(1)+THETA(2)*LOG(AGE/50)
MU_5= LTVCL
LTVV=THETA(3)
MU_3=LTVV
CL=EXP(MU_5+ETA(5))
V=EXP(MU_3+ETA(3))
```

In such cases, THETA(1) is no longer the typical value of CL, but the logarithm of the typical value (LTVCL) of CL, just as MU\_5 is. The initial THETA values on the \$THETA record would need to be log transformed. Any former lower bound of 0 should not be included, as the  $\log()$  imposes a natural 0 boundary on the original value. One can avoid transposing the inputted THETAs while still having linear MU referencing by using the \$THETA1 and \$THETA2 record (intro7<sup>1</sup>).

Linear MU referencing increases the robustness of the EM algorithms to converge toward a minimum, as the THETAs will then be updated by one-step linear regression rather than by one-step nonlinear regression. Furthermore, lower ISAMPLE ( $\leq 3$ ) values are needed for stable updates with SAEM with linear MU referencing. IMP is not as sensitive in this regard because of its typically higher ISAMPLE value ( $\geq 300$ ).

Additional information about MU referencing and their rules are given in **Supplementary Materials S1—Part A**.

### MCMC BAYESIAN ANALYSIS

The goal of MCMC full Bayesian analysis is to obtain a sample distribution of probable population parameters (posterior distribution), which is unlike the goal of maximization methods such as FO, FOCE/Laplace, and EM, which seek to find point (single-valued) estimates of population parameters that best fit the data. Usually 10,000–30,000 samples are required (intro7.pdf<sup>1,6</sup>). The samples are not statistically independent, but when the analysis is properly performed, they are uncorrelated overall.

Unlike maximization methods (FO, FOCE/Laplace, EM), maximum likelihood parameters are not obtained, but with problems of sufficient data, the sample mean parameters are similar to maximum likelihood values. Also, a maximum likelihood objective function is not obtained, but a distribution of joint probability densities is obtained, from which

95% confidence bounds can be constructed and tested for overlap with those of alternative models. Criteria such as Bayes information criterion can also be used for hypothesis testing.

Although the FOCE and EM methods obtain only approximate SEs to the parameters at evaluated point estimates using Fischer information matrix process, the MCMC Bayesian methods average the many samples of population parameters with frequencies in proportion to their ability to explain the data, and also, empirical SEs are evaluated from the variance of these parameters. With these properties, MCMC Bayesian analysis can be used as a substitute for bootstrap. MCMC Bayesian analysis takes longer than maximization analysis but provides more information and is faster than bootstrap. Use MCMC Bayesian analysis to embellish information on your final model.

All problems can be solved by MCMC Bayesian analysis. In practice, the more the problem is linearly MU modeled, the more the Bayesian analysis can rely on the efficient Gibbs sampling method (gray box process). The fewer parameters that are linearly MU modeled, the more Bayesian analysis relies on the less efficient but general Metropolis-Hastings sampling method (black box process) and potentially can be less stable. However, the no-u-turn Hamiltonian (NUTS) Metropolis-Hastings sampling method references 7,8 can be very efficient if analytical derivatives are supplied (**Supplementary Materials S1—Part B**).

### Prior information

Prior information of population parameters is important for MCMC Bayesian analysis in order for them to be stable. Priors to OMEGAs are the most important, and therefore it is best to supply at least weakly informative priors to them. Although prior information may also be added to maximum likelihood (FO, FOCE, EM) estimations in NONMEM and could be useful for supplying informative priors based on previous studies, the maximization methods have sufficient statistical power for obtaining their goals of point estimates to not require uninformative or weakly informative priors for most problems. The MCMC Bayesian analysis, however, has less statistical power to obtain its goal because it does not seek just point estimates of the population parameters but their entire posterior distribution. A general rule in statistics is that the more you ask of a fixed set of data, the less power there is for each piece of information requested. See **Supplementary Materials S1—Part C** on how to provide prior information in NONMEM.

### Termination testing

For EM analyses (IMP, SAEM, BAYES), because of the stochastic nature of the results, a deterministic method of stopping the analysis when the values change by no more than the third digit cannot be used because the values will always be changing stochastically at that precision. Instead, a statistical test is performed on whether the general change in the parameter varies by more than the stochastic noise. See **Supplementary Materials S1—Part D** on how to control the statistical test for convergence on Monte Carlo estimation methods.

The following examples make use of some of the concepts described previously.

### EXAMPLE 1: PK MODEL WITH COVARIATES USING IMP, SAEM, AND BAYESIAN ANALYSIS METHODS

In the previous tutorial, example 504.ctl was used to demonstrate how the effect of covariates could be assessed on various PK parameters using the FOCE with interaction (FOCEI) method. This same model is easily rendered for EM analysis by identifying for NONMEM the “MU” parameters, the arithmetic means associated with the random effects. This is done conveniently for this example by log transforming the typical values. Although not essential, THETA(1) and THETA(2) are reparameterized as the log of the original parameters, so as to express the MUs completely linearly in all THETAs, as recommended earlier, and removing the need for placing lower bounds of 0 on THETA(1) and THETA(2) (**Supplementary Materials S2**, 504\_its.ctl, 504\_saem.ctl, 504\_foce.ctl). Please see **Table 1**, code item 1, showing how the logarithm of the typical values are defined and associated with the MU\_ variables.

For example 504\_its.ctl, the ITS is used. The \$EST record for this is listed in **Table 1**, code item 2. Although more approximate than FOCE, ITS is very fast and can be used as a preliminary analysis. For example 504\_saem.ctl, the SAEM method is used, followed by IMP to obtain the final objective function (**Table 1**, code item 3). A convenient tool in NONMEM is available in which a subsequent \$EST step uses as initial values the final estimates of the immediately previous \$EST step. Here, the SAEM estimation is followed by the evaluation of the marginal likelihood objective function using IMP's expectation step only (EONLY = 1), but not allowing the IMP step to perform maximization updates on the fixed effects, retaining the fixed effects values at the final answer of the SAEM method. Furthermore, the conditional mean and variance of the individual parameters from the last step of the SAEM are used as the parameters to the importance sampler's proposal density (MAPITER = 0); that is, MAP estimation is not necessary for the first iteration because we have information from a previous estimation step. Subsequent IMP iterations borrow the conditional mean and variance of the previous iteration for the proposal density parameters. Five iterations of IMP are requested to obtain a stochastic sampling of the objective function value so one can get a sense of the extent of the Monte Carlo noise in the result.

An FOCE assessment was included to compare the results (504\_foce.ctl). Note that the final parameters from the ITS, SAEM/IMP, and FOCE are very similar, with the ITS objective function nearly identical to the FOCE, although the IMP objective function is slightly different from the FOCE result by about eight units (**Table 2**). When the FOCE objective function is evaluated at the population results of SAEM/IMP, the FOCE objective function differs by about two units. This suggests a small amount of nonlinearity of the conditional density, that is, the conditional density is not exactly normally distributed, and this results in slightly different assessments of the objective function as well as slightly different positions, although the population

parameters are not statistically different between SAEM and FOCEI. Thus, for this example, the FOCEI and SAEM results are statistically similar, but one should always use the objective function from the same analysis method when making statistical comparisons between models.

The number of samples collected for each subject ISAMPLE is by default 2 for SAEM, but should be increased up to 10 if data are sparse or data are nonnormally distributed, such as categorical data or below-detection data. For IMP, ISAMPLE should be increased from its default of 300 when there are many ETAs or the objective function has large stochastic fluctuations. When there are fewer data points than there are ETAs to be estimated (sparse data) or data are categorical, then DF (degrees of freedom) should be set to a nonzero number to use the t-distribution sampler and/or decrease the acceptance rate IACCEPT from the default value of 0.4 to about 0.2. Alternatively, option AUTO = 1 may be selected as is done for some of these examples. This is an option that requests NONMEM to make decisions about ISAMPLE for SAEM and IACCEPT, DF, and ISAMPLE for IMP for each subject. Furthermore, because SAEM uses an MCMC algorithm so that samples of the ETAs generated are correlated, the correlation interval CINTERVAL is also evaluated by AUTO = 1 to assure that when NONMEM tests for convergence, the appropriate spread of iterations is considered when assessing whether all parameters no longer change systematically and only

random (stochastic) variations remain. **Supplementary Materials S1—Part E** shows how this example may be rendered for Bayesian analysis.

For Monte Carlo methods, stochastic reproducibility of results can be improved by setting RANMETHOD = P to allow reproducibility during parallelization of a problem. The AUTO = 1 feature performs additional complex analysis and decisions, and using this feature may result in lack of stochastic reproducibility.

Notice that the MATRIX = R option is generally requested in these examples. The R matrix is the Fisher information matrix constructed from the second derivative of the objective function with respect to the various parameters estimated. The more approximate S matrix (score matrix) is constructed from the first derivatives of each individual's objective function contribution with respect to the various parameters estimated and can be requested with MATRIX = S. The variance-covariance of the estimates will be evaluated as  $R^{-1}$  or  $S^{-1}$ , respectively. If no MATRIX option is provided, the variance-covariance is evaluated as  $R^{-1}S R^{-1}$ . The S matrix is more quickly evaluated, which may be significant especially for ODE-type problems (such as example 4 below), and is sufficiently accurate if there are more subjects (preferably about two times more) than the dimension of the variance-covariance. If computation time is not a concern, I prefer MATRIX = R as it is more consistent with other statistical software.

**Table 1** Control stream code segments showing relevant code

Item	Code
Example 1	
1	<pre>\$PK ... LTVCL=THETA (1) +LOG (WT/70) *THETA (3)       +LOG (AGE/50) *THETA (5) +SEX*THETA (7) LTVV=THETA (2) +LOG (WT/70) *THETA (4)       +LOG (AGE/50) *THETA (6) +SEX*THETA (8) MU_1=LTVCL MU_2=LTVV CL=EXP (MU_1+ETA (1)) V=EXP (MU_2+ETA (2))</pre>
2	\$EST METHOD=ITS INTERACTION AUTO=1 PRINT=20 NOABORT
3	\$EST METHOD=SAEM AUTO=1 NITER=500 PRINT=20
	\$EST METHOD=IMP EONLY=1 PRINT=1 NITER=5 ISAMPLE=1000 MAPITER=0
Example 2	
4	<pre>\$PK MU_1 = THETA (1) MU_2 = THETA (2) MU_3 = THETA (3) MU_4 = THETA (4) CL = EXP (MU_1 + ETA (1)) V1 = EXP (MU_2 + ETA (2)) Q = EXP (MU_3 + ETA (3)) V2 = EXP (MU_4 + ETA (4))</pre>

(Continues)

Table 1 (Continued)

Item	Code
5	<pre> \$ERROR IEPRED=A(1)/S1 LLOQ=-3.5 SD = THETA(5) ; DEL can be 10E-10 or smaller DEL=1.0E-30 ; The following coding for LOG() prevents numerical errors ; from occurring when F&lt;=0 IPRED = LOG(ABS(IEPRED)+DEL) IF(COMACT==1) PREDV=IPRED DUM = (LLOQ - IPRED)/SD ; Adding DEL to CUMD prevents it from becoming 0, which is not good when NONMEM evaluates -2*LOG(CUMD) CUMD = PHI(DUM)+DEL TYPE=1 IF(DV&lt;LLOQ) TYPE=2 IF(MDV==1) TYPE=0 IF(TYPE.EQ.2) DV_LOQ=LLOQ IF (TYPE.NE. 2.OR.NPDE_MODE==1) THEN     F_FLAG = 0     Y = IPRED + SD * ERR(1) ENDIF IF (TYPE.EQ. 2.AND.NPDE_MODE==0) THEN     F_FLAG = 1     Y = CUMD     MDVRES=1 ENDIF </pre>
6	<pre> \$EST METHOD=IMP LAPLACE INTERACTION AUTO=1 PRINT=5 RANMETHOD=S2 </pre>
Example 3	
7	<pre> \$ERROR     IPRED=A(2)/V     EXPP=THETA(4)+IPRED*THETA(5) ; Put a limit on this, as it will be exponentiated, ; to avoid floating overflow     IF(EXPP.GT.40.0) EXPP=40.0 IF (TYPE.EQ.0) THEN ; PK Data     F_FLAG=0     Y=F+F*ERR(1); a prediction ELSE ; Categorical data     F_FLAG=1 ; IF EXPP&gt;40, then A&gt;1.0d+17, A/B=1, and Y=DV     AA=EXP(EXPP)     B=1+AA     Y=DV*AA/B+(1-DV)/B ; a likelihood ENDIF </pre>

(Continues)

Table 1 (Continued)

Item	Code
Example 4 8	<pre> \$DES DADT(1) = -(K10+K12)*A(1) + K21*A(2) - VM*A(1)*A(3)/(A(1)+KM) DADT(2) = K12*A(1) - K21*A(2) DADT(3) = -(VM-K30)*A(1)*A(3)/(A(1)+KM) - K30*A(3) + K03 </pre>
9	<pre> \$PK ... A_0(3)=K03/K30 </pre>
10	<pre> \$ERROR ETYP=1 IF(CMT.NE.1) ETYP=0 CP=A(1)/S1 CR=A(3)/S3 IPRE=CP IF(CMT.NE.1) IPRE=CR Y = IPRE + IPRE*ETYP*EPS(1) + IPRE*(1.0-ETYP)*EPS(2) </pre>
Example 5 11	<pre> \$ABBR REPLACE ETA(OCC_CL)=ETA(5,7,9) \$ABBR REPLACE ETA(OCC_V)=ETA(6,8,10) </pre>
12	<pre> \$LEVEL SID=(3[1],4[2]) </pre>
13	<pre> \$PK ... CL=EXP(MU_1+ETA(1)+ETA(3)+ETA(OCC_CL)) V=EXP(MU_2+ETA(2)+ETA(4)+ETA(OCC_V)) ;Individual omegas (1,2) \$OMEGA BLOCK(2) .1 -.0001 .1 ;SID OMEGAS (3,4) \$OMEGA BLOCK(2) .3 -.0001 .3 ; inter-occasion omegas for occasion 1 (5,6) \$OMEGA BLOCK(2) .03 -.0001 .03 ; inter-occasion omegas for occasion 2 and 3 ; SAME(n) means repeat block structure n times; ; and omega parameters used for occasions 2 and 3 are shared ; with those of occasion 1. \$OMEGA BLOCK(2) SAME(2) </pre>
Example 6 15	<pre> \$PK ... VCM=DEXP(MU_1+ETA(1)) K10M=DEXP(MU_2+ETA(2)) VCF=DEXP(MU_3+ETA(3)) K10F=DEXP(MU_4+ETA(4)) Q=1 IF(MIXNUM.EQ.2) Q=0 V=Q*VCM+(1.0-Q)*VCF K=Q*K10M+(1.0-Q)*K10F </pre>



Table 1 (Continued)

Item	Code
16	<pre>;Initial OMEGA block 1, for sub-population 1 \$OMEGA BLOCK(2)     .04; [p]     .01; [f]     .027; [p] ;Initial OMEGA block 2, for sub-population 2 \$OMEGA BLOCK(2)     .05; [p]     .01; [f]     .06; [p]</pre>
Example 7	
17	<pre>\$PK ... CL=EXP(MU_1+ETA(1)) CLR=EXP(MU_3+ETA(3)) F2=CLR/CL</pre>
18	<pre>\$ERROR CP=A(1)/V UVL=UVOL IF(UVL&lt;=0.0) UVL=1.0 CU=A(2)/UVL IF(CMT==1) THEN   IPRE=CP   W=IPRE+1.0E-10   Y=IPRE + W*EPS(1) ELSE   IPRE=CU   W=IPRE+1.0E-10   Y=IPRE + W*EPS(2) ENDIF</pre>

**EXAMPLE 2: MODELING DATA THAT ARE BELOW THE LOQ**

The following example models data that are below the LOQ (see ref. 9). Normally distributed, below quantifiable limits are given the following likelihood evaluation:

$$p(x < \text{LOQ}) = \int_{-\infty}^{(\text{LOQ}-f)/\text{sd}} \exp(-y^2/2) dy = \Phi((\text{LOQ}-f)/\text{sd})$$

For data below the LOQ, their data likelihood contribution is expressed as the probability that the data value is normally distributed and below the LOQ. Because this partial integral of the normal density is not itself a normal density, NONMEM must be informed by setting F\_FLAG = 1 that the returned Y value is not the predicted function but is a user-defined data likelihood for that datum. As we shall see, the F\_FLAG = 1 option may be generally used to allow the user to provide any data likelihood that is not normally distributed. In such cases, the Laplace option must be set.

See **Supplementary Materials S3** for the complete control stream of this example. The first part of the control stream ad3tr4\_loqb\_lap.ctl (Laplace method) and ad3tr4\_loqb\_imp.ctl (IMP method) consists of a typical setup with MU referencing (**Table 1**, code item 4). Consideration to modeling the LOQ data is confined to the \$ERROR block (**Table 1**, code item 5). The individual predicted value (IPRED) and standard deviation (SD) of the residual error are defined, followed by evaluation of the Gaussian function PHI(), the result stored in the variable CUMD (cumulative distribution). In addition to IPRED, it is desired to store the population predicted value, which is essentially the value of IPRED evaluated when ETA = 0, and this is signaled by NONMEM on the occasion that COMACT = 1, which is saved in the variable PREDV (population prediction variable). Normally, the NONMEM reserved variable PRED (population prediction) will be automatically calculated for the user when the data are the typical normally distributed type and calculating PREDV is unnecessary. However, when data are modeled with a user-specified likelihood, the PRED value will be set to an uninformative value of 1. One may wish to use the more informative PREDV.

Table 2 Final parameter estimates and standard errors for Example 1: Statistical assessment of covariates added to a base pharmacokinetic model using FOCE, importance sampling, SAEM  
Statistic/Estimation method, and Bayesian analysis methods

	THETA1	THETA2	THETA3	THETA4	THETA5	THETA6	THETA7	THETA8	SIGMA(1:1)	OMEGA(1:1)	OMEGA(2:1)	OMEGA(2:2)	OBJ
Mean													
ITS	1.112	3.464	0.671	1.32	-0.523	0.048	-0.1	-0.054	0.050	0.032	0.001	0.047	1,058.542
FOCEI	1.109	3.478	0.66	1.321	-0.534	0.052	-0.101	-0.055	0.050	0.031	0.001	0.047	1,058.304
SAEM (IMP)	1.07	3.467	0.644	1.311	-0.544	0.060	-0.101	-0.057	0.050	0.03	0.0003	0.049	1,051.790
BAYES	1.07	3.467	0.641	1.315	-0.557	0.055	-0.103	-0.052	0.055	0.029	-0.002	0.046	FOCE OBJ = 1,060.360 661.634 FOCE OBJ = 1,061.227
SE													
ITS	0.05	0.044	0.2	0.25	0.135	0.149	0.063	0.081	0.007	0.008	0.01	0.015	
FOCEI	0.038	0.049	0.16	0.202	0.103	0.135	0.057	0.071	0.007	0.009	0.008	0.014	
SAEM (IMP)	0.039	0.050	0.163	0.208	0.105	0.133	0.058	0.073	0.007	0.009	0.008	0.014	
BAYES	0.039	0.050	0.165	0.204	0.106	0.132	0.057	0.073	0.008	0.009	0.008	0.014	19.475

BAYES, Bayesian analysis; FOCEI, first-order conditional estimation; FOCEI, FOCE with interaction; IMP, importance sampling; ITS, iterative two stage; OBJ, objective function; SAEM, stochastic approximation expectation maximization; SE, standard error.

In this particular setup, any observed value DV (dependent variable) that is less than the LOQ (in this case the  $\log(\text{LOQ}) = \text{LLOQ}$ , because the  $\log(\text{DV})$  is fitted in the model) will be designated TYPE = 2, and those data above LOQ are designated TYPE = 1. During estimation (when NPDE\_MODE = 0, that is, normalized prediction distribution error (NPDE) is not being calculated) when TYPE = 1, NONMEM is informed by setting F\_FLAG = 0 that the returned Y value is to be treated as the predicted value to a normally distributed dependent variable, which is the usual interpretation of data. When TYPE = 2, NONMEM is informed by setting F\_FLAG = 1 that the returned Y value is to represent the likelihood of the datum itself, which in this case is CUMD.

As of NONMEM 7.4, setting MDVRES = 1 (MDVRES stands for MDV (missing dependent variable indicator) during residual error analysis) for the nonnormal data (F\_FLAG = 1) assures that weighted residual diagnostics (CWRES<sup>10</sup>) are calculated for the collection of normal data (F\_FLAG = 0) in that subject and excludes from the weighted residual evaluation any data for which MDVRES = 1 (or MDV = 1). In previous versions, if there was at least one nonnormal datum in the collection of data considered for weighted residual assessment, no weighted residuals for any of the data for that subject were calculated.

Normalized prediction distribution error (NPDE) values for LOQ data may also be calculated (based on refs. 11–13). To implement this, use the NPDE\_MODE switch and set reserved variable DV\_LOQ as shown in the previous code. When NPDE\_MODE = 1, NONMEM is executing the \$ERROR block during the evaluation of NPDEs for the items in the \$TABLE records (so, not during estimation), and if the datum is <LOQ (TYPE = 2), then user code should return the predicted value IPRED just as it does with normal data. In addition, with DV\_LOQ is set, NONMEM knows to use that as the “observed value” limit.

Notice that the \$TABLE record requests that NPDE values be generated from ESAMPLE = 1,000 samples. In general, the most extreme value generated in NPDE is limited by ESAMPLE as follows:

$$1/\text{ESAMPLE} = \Phi(\text{NPDE}_{\text{low}})$$

so selecting a sufficiently high ESAMPLE is necessary so that extreme NPDE values are sufficiently detected. Usually it is sufficient that the NPDE extreme values be -3.09 to 3.09 (default ESAMPLE = 300) to accurately reflect those NPDE beyond the 95% range (NPDE = -1.96, 1.96) without significant truncation. An ESAMPLE of 10,000 will give you a range of NPDE = -3.72, 3.72.

A portion of table ad3tr4\_loqb\_imp.tab is listed in **Table 3**. Note that CWRES values for all normal data exists, but not for records whose DV < LOQ (bolded in **Table 3**), whereas the NPDE values are calculated for all data records. In **Figure 1** one can see how providing additional NPDEs from the LOQ data can fill out the lower x-regions of a scatter plot.

The conditional Laplace method is susceptible to ending the analysis with round-off errors, as in this example. The Monte Carlo EM analysis method does not have this

**Table 3** A portion of the *post hoc* results from Example 2 (ad3tr4\_loqb\_imp.tab): Modeling data that are below the limit of quantitation

ID	TIME	DV	IPRED	PRED	PREDV	CWRES	NPDE
1.0000E+00	0.0000E+00	-2.9934E+01	3.0116E+00	2.9972E+00	2.9972E+00	0.0000E+00	0.0000E+00
1.0000E+00	2.0000E-01	2.7061E+00	2.6978E+00	2.7013E+00	2.7013E+00	2.9914E-02	1.2819E-01
1.0000E+00	1.0000E+00	1.4796E+00	1.5042E+00	1.5732E+00	1.5732E+00	-1.5716E-01	1.9934E-01
1.0000E+00	4.0000E+00	-6.6033E-01	-7.0388E-01	-6.2029E-01	-6.2029E-01	1.1124E-01	-6.7731E-02
1.0000E+00	7.0000E+00	-1.2313E+00	-1.1889E+00	-1.0987E+00	-1.0987E+00	-2.1749E-01	8.5329E-02
<b>1.0000E+00</b>	<b>4.0000E+01</b>	<b>-5.1486E+00</b>	<b>-5.8988E+00</b>	<b>1.0000E+00</b>	<b>-5.5438E+00</b>	<b>0.0000E+00</b>	<b>-2.8193E-01</b>

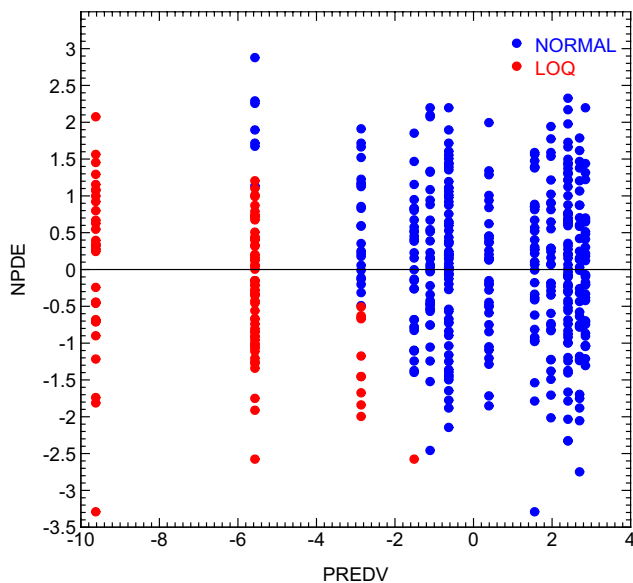
The record in bold contains a DV that is less than LOQ, for which CWRES is not calculated, but NPDE is calculated.

CWRES, conditional weighted residuals; DV, dependent variable; ID, subject identification number; IPRED, individual predicted value; NPDE, normalized prediction distribution error; PRED, predicted value; PREDV, Population Prediction Variable.

difficulty and can be particularly effective with problems with LOQ data. Control stream ad3tr4\_loqb\_imp.ctl uses the IMP estimation method (Table 1, code item 6). Again, the AUTO feature is used, which adjusts IACCEPT and other option values to allow handling of highly nonnormal conditional densities that are likely to occur with nonnormal data likelihoods. Also, RANMETHOD = S2 turns on quasi-random (Sobol) sampling to reduce stochastic noise in the objective function evaluation (intro7<sup>1</sup>).

### EXAMPLE 3: MODELING PK CATEGORICAL RESPONSE DATA

The following is another example in which the user must define the likelihood for the data. In this example, PK samples are collected and interpreted in the usual manner with a normal data likelihood. Response data are also available, 0 indicating treatment failure, and 1 indicating treatment success. The first portion models the standard one-compartment model with absorption using the convenient library module ADVAN2 in NONMEM and also providing



**Figure 1** Normalized prediction distribution error (NPDE) plotted against population prediction variable (PREDV) for Example 2: Modeling data that are below the limit of quantitation (LOQ). The NPDE values for data < LOQ are shown in red and, as expected, occupy the lower PREDV region of the plot.

MU-reference equations for EM analysis (wexample10\_saem.ctl, wexample10\_bayes.ctl, wexample10\_lap.ctl, wexample10\_imp.ctl; **Supplementary Materials S4**).

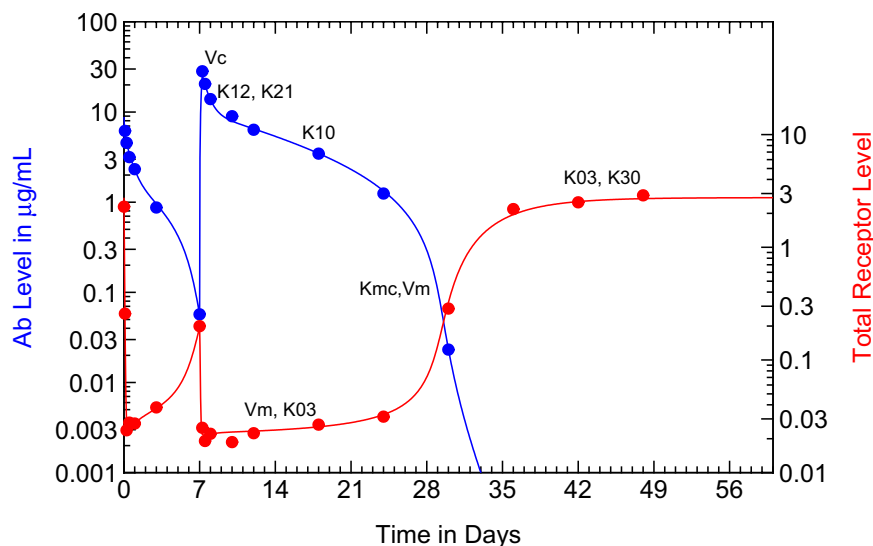
In the \$ERROR block, the PK data (TYPE = 0) is modeled as usual normally distributed data (F\_FLAG = 0 is set), and its prediction function (mean to the normal density) is returned as the Y value, whereas when the data are clinical response (TYPE = 1), a logistic regression function is used to match up the concentration of drug with the increased probability of positive response, and a likelihood is returned as the Y value (Table 1, code item 7). A linear function (EXPP = THETA(4)+F\*THETA(5)) relates the drug concentration A(2)/V linearly to drug effectiveness, which in turn is used in a logistic regression function ( $\exp(x)/(1 + \exp(x))$ ) to convert this into a probability of effectiveness. When DV = 1, Y is returned with the probability that the drug was effective, and when DV = 0, Y is returned with the probability that the drug was not effective. Notice that extremes of likelihood values (0 or 1) are returned depending on the following scenarios:

	Very low drug levels	Very high drug levels
No response (DV = 0)	Y = 1	Y = 0
Response (DV = 1)	Y = 0	Y = 1

Thus the model tests that with drug present, a high incidence of treatment success occurs, and with little or no drug present, a low incidence of treatment success occurs.

The remainder of the control stream file contains typical records for estimation and output. There may be more than two categories of responses for clinical data, for which the user's model must return the appropriate likelihood for each of those occurrences, and the likelihoods should total to 1 to statistically consider all possibilities.

Notice in these examples that an ITS step is first performed, followed by the estimation with the method of interest. In wexample10\_lap.ctl, an ITS estimation is performed first and then the classical Laplace estimation. As mentioned previously, Laplace can be sometimes unstable in its search for the minimal objective function, so having the Laplace estimation begin at the population parameter values where the ITS ended can stabilize the analysis. Similarly, in wexample10\_bayes.ctl, an ITS estimation is performed first and then BAYES estimation. This can be beneficial to improving stability and reducing the burn-in time required for BAYES, as the BAYES method will use as initial values the final population and individual parameters that were assessed by



**Figure 2** Antibody levels (pharmacokinetic) and total receptor levels (pharmacodynamic) are plotted against time for Example 4: Using ordinary differential equation solvers to model a basic target-mediated drug disposition model. Parameter symbols are located where they impact the curve shape the most to give some sense of their contributions. Ab, antibody; Kmc, concentration of half-maximal rate of internalization; K12, rate constant of transfer from compartment 1 to compartment 2; K21, rate constant of transfer from compartment 2 to compartment 1; K10, rate constant of elimination from compartment 1; K03, rate of production of receptor (compartment 3); K30, rate constant of elimination of receptor (compartment 3); Vc, volume of distribution of central compartment; Vm, maximal rate of internalization.

ITS. For `wexample10_saem.ctl`, one or a few iterations of ITS is first performed to obtain individual parameters evaluated at the mode of their posteriors, which can facilitate the SAEM analysis.

#### EXAMPLE 4: USING ODE SOLVERS TO MODEL A BASIC TARGET-MEDIATED DRUG DISPOSITION MODEL

This example shows one of the most versatile ways to model a PK/PD problem, with ODE, using mass transfer equations as a means to describe the interactions of molecules and rates of movements from one physiological or biochemical component to the other. The following is a basic target-mediated drug disposition model. Consider an antibody with central volume of distribution VC (compartment 1) that distributes to the peripheral compartment (compartment 2) with rate constants K12 and K21, a linear rate constant of elimination K10 associated with the high-capacity receptor via the Fc component of the antibody, and an interaction with a target cell membrane receptor (compartment 3, total receptor) for which the antibody was designed to bind with high affinity via its (Fab')<sub>2</sub> component, which results in the internalization and elimination of the antibody-receptor complex with maximal internalization constant Vm and half-maximal internalization rate occurring at plasma antibody levels of Kmc. The receptor is continually replenished by the cell with first order production rate K03 and first order elimination rate K30 in the absence of antibody. This is readily expressed as a series of mass transfer equations that take into account each of these actions by simple equation components in record \$DES, shown in **Supplementary Materials S5** (`r2complb_imp.`

`ctl`, `r2complb_foce.ctl`) and **Table 1**, code item 8. Notice that the initial value of the receptor level is set using the reserved variable `A_0(3)`, to `K03/K30` (**Table 1**, code item 9). Also, the PK data are modeled with residual variance `SIGMA(1,1)`, and PD data are modeled with another residual variance `SIGMA(2,2)` (**Table 1**, code item 10). **Figure 2** shows a typical PK profile and total receptor profile for a subject receiving an infusion of loading-dose antibody during the first week followed by a higher dose the second week. Parameter symbols are located where they impact the curve shape the most to give some sense of their contributions. Although there are eight parameters to the model, there is sufficient detail in the PK and PD curves to allow for their evaluation in a suitably rich data set.

For target-mediated drug disposition models, the IMP method is particularly efficient.<sup>14</sup> When the analysis starts far from reasonable initial values, the mode of the conditional (MAP) estimation during the first iteration in IMP is facilitated by allowing a Monte Carlo sampling of MCETA-1 ETAs to be tested, and the ETA vector providing the lowest initial objective function is used as the initial position for the MAP estimation. Also, the TOL was set to 6, and SIGL (significant digit precision for lower level calculations) should also be set to 6, as described in the manual.

For a data set with 50 subjects, this problem is solved several fold faster by IMP than by FOCEI, but improvements in the efficiency of FOCE evaluation have been made over the years, such as the introduction of the SIGL option, which allows one to reduce the convergence criterion NSIG (significant digit precision for population parameter estimates) and the ODE evaluation tolerance TOL, which can make the FOCE method also more efficient.<sup>14</sup> Also, the use of MU referencing with the FAST option allows the evaluation

Table 4 A portion of the final output results for Example 5 (superid30\_1\_foc6.tab): Analyzing data modeled with multiple levels of mixed effects

ID	SID	OCC	TIME	IPRED	ETA1	ETA2	ETA3	ETA4	ETA5	ETA6	ETA7	ETA8	ETA9	ETA10
1.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	5.7690E+01	-1.2951E-01	-2.7152E-01	-5.3382E-01	-7.1304E-01	9.0991E-02	-1.6888E-01	-8.3666E-02	8.6754E-02	-4.1063E-02	-1.1915E-02
1.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	5.7053E+01	-1.2951E-01	-2.7152E-01	-5.3382E-01	-7.1304E-01	9.0991E-02	-1.6888E-01	-8.3666E-02	8.6754E-02	-4.1063E-02	-1.1915E-02
1.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	5.6423E+01	-1.2951E-01	-2.7152E-01	-5.3382E-01	-7.1304E-01	9.0991E-02	-1.6888E-01	-8.3666E-02	8.6754E-02	-4.1063E-02	-1.1915E-02
...														
2.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	4.4818E+01	-7.0114E-02	-1.4319E-01	-5.3382E-01	-7.1304E-01	-1.1614E-01	-4.4747E-02	1.0101E-01	4.5861E-02	-3.3181E-03	-5.0617E-02
2.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	4.4486E+01	-7.0114E-02	-1.4319E-01	-5.3382E-01	-7.1304E-01	-1.1614E-01	-4.4747E-02	1.0101E-01	4.5861E-02	-3.3181E-03	-5.0617E-02
2.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	4.4156E+01	-7.0114E-02	-1.4319E-01	-5.3382E-01	-7.1304E-01	-1.1614E-01	-4.4747E-02	1.0101E-01	4.5861E-02	-3.3181E-03	-5.0617E-02
...														
3.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	2.8086E+01	-1.1209E-01	9.8214E-02	-5.3382E-01	-7.1304E-01	-1.9225E-02	1.8118E-01	7.7868E-02	-7.4813E-02	-1.0431E-01	-6.4336E-02
3.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	2.7948E+01	-1.1209E-01	9.8214E-02	-5.3382E-01	-7.1304E-01	-1.9225E-02	1.8118E-01	7.7868E-02	-7.4813E-02	-1.0431E-01	-6.4336E-02
3.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	2.7811E+01	-1.1209E-01	9.8214E-02	-5.3382E-01	-7.1304E-01	-1.9225E-02	1.8118E-01	7.7868E-02	-7.4813E-02	-1.0431E-01	-6.4336E-02
...														
4.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	3.2419E+01	-9.2289E-02	1.6778E-01	-5.3382E-01	-7.1304E-01	6.7656E-03	-3.1853E-02	-5.2640E-02	-8.5933E-02	3.9829E-03	1.8459E-01
4.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	3.2227E+01	-9.2289E-02	1.6778E-01	-5.3382E-01	-7.1304E-01	6.7656E-03	-3.1853E-02	-5.2640E-02	-8.5933E-02	3.9829E-03	1.8459E-01
4.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	3.2036E+01	-9.2289E-02	1.6778E-01	-5.3382E-01	-7.1304E-01	6.7656E-03	-3.1853E-02	-5.2640E-02	-8.5933E-02	3.9829E-03	1.8459E-01
...														
5.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	3.2916E+01	2.0926E-01	1.5725E-01	-5.3382E-01	-7.1304E-01	1.2382E-03	-3.6522E-02	8.7817E-03	-2.9378E-02	5.8401E-02	1.1359E-01
5.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	3.2649E+01	2.0926E-01	1.5725E-01	-5.3382E-01	-7.1304E-01	1.2382E-03	-3.6522E-02	8.7817E-03	-2.9378E-02	5.8401E-02	1.1359E-01
5.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	3.2385E+01	2.0926E-01	1.5725E-01	-5.3382E-01	-7.1304E-01	1.2382E-03	-3.6522E-02	8.7817E-03	-2.9378E-02	5.8401E-02	1.1359E-01
...														
6.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	4.7660E+01	-1.4427E-01	-1.1702E-01	-5.3382E-01	-7.1304E-01	7.9282E-02	-1.3239E-01	-1.3833E-01	-9.0787E-02	3.6565E-03	1.8711E-01
6.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	4.7232E+01	-1.4427E-01	-1.1702E-01	-5.3382E-01	-7.1304E-01	7.9282E-02	-1.3239E-01	-1.3833E-01	-9.0787E-02	3.6565E-03	1.8711E-01
6.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	4.6808E+01	-1.4427E-01	-1.1702E-01	-5.3382E-01	-7.1304E-01	7.9282E-02	-1.3239E-01	-1.3833E-01	-9.0787E-02	3.6565E-03	1.8711E-01
...														
7.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	4.6352E+01	1.1683E-02	-1.2005E-01	-5.3382E-01	-7.1304E-01	-2.7490E-02	-1.0154E-01	-7.6464E-02	5.8465E-02	1.1414E-01	-1.9799E-03
7.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	4.5931E+01	1.1683E-02	-1.2005E-01	-5.3382E-01	-7.1304E-01	-2.7490E-02	-1.0154E-01	-7.6464E-02	5.8465E-02	1.1414E-01	-1.9799E-03
7.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	4.5514E+01	1.1683E-02	-1.2005E-01	-5.3382E-01	-7.1304E-01	-2.7490E-02	-1.0154E-01	-7.6464E-02	5.8465E-02	1.1414E-01	-1.9799E-03
...														
8.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	4.1899E+01	1.7997E-01	-5.5437E-02	-5.3382E-01	-7.1304E-01	5.2661E-02	-6.5148E-02	-8.6443E-02	7.8233E-02	1.0205E-01	-4.2697E-02
8.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	4.1458E+01	1.7997E-01	-5.5437E-02	-5.3382E-01	-7.1304E-01	5.2661E-02	-6.5148E-02	-8.6443E-02	7.8233E-02	1.0205E-01	-4.2697E-02
8.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	4.1022E+01	1.7997E-01	-5.5437E-02	-5.3382E-01	-7.1304E-01	5.2661E-02	-6.5148E-02	-8.6443E-02	7.8233E-02	1.0205E-01	-4.2697E-02
...														
9.0000E+00	1.0000E+00	1.0000E+00	0.0000E+00	4.0748E+01	-3.4621E-02	2.5466E-02	-5.3382E-01	-7.1304E-01	1.7205E-02	-1.1819E-01	4.9575E-02	5.0070E-02	-8.0644E-02	7.9295E-02
9.0000E+00	1.0000E+00	1.0000E+00	1.0000E-01	4.0423E+01	-3.4621E-02	2.5466E-02	-5.3382E-01	-7.1304E-01	1.7205E-02	-1.1819E-01	4.9575E-02	5.0070E-02	-8.0644E-02	7.9295E-02
9.0000E+00	1.0000E+00	1.0000E+00	2.0000E-01	4.0100E+01	-3.4621E-02	2.5466E-02	-5.3382E-01	-7.1304E-01	1.7205E-02	-1.1819E-01	4.9575E-02	5.0070E-02	-8.0644E-02	7.9295E-02
...														
1.0000E+01	1.0000E+00	1.0000E+00	0.0000E+00	3.3914E+01	-7.9089E-03	1.3188E-01	-5.3382E-01	-7.1304E-01	-7.2529E-02	-4.1041E-02	7.4692E-02	-6.3879E-02	-1.1558E-02	1.5417E-01
1.0000E+01	1.0000E+00	1.0000E+00	1.0000E-01	3.3703E+01	-7.9089E-03	1.3188E-01	-5.3382E-01	-7.1304E-01	-7.2529E-02	-4.1041E-02	7.4692E-02	-6.3879E-02	-1.1558E-02	1.5417E-01
1.0000E+01	1.0000E+00	1.0000E+00	2.0000E-01	3.3492E+01	-7.9089E-03	1.3188E-01	-5.3382E-01	-7.1304E-01	-7.2529E-02	-4.1041E-02	7.4692E-02	-6.3879E-02	-1.1558E-02	1.5417E-01
...														
1.1000E+01	2.0000E+00	1.0000E+00	0.0000E+00	9.0996E+00	-1.0569E-01	-7.6361E-02	4.6715E-04	7.1010E-01	-1.2187E-02	5.9678E-02	9.9922E-02	-7.5580E-03	-1.2245E-01	-7.5073E-02
1.1000E+01	2.0000E+00	1.0000E+00	1.0000E-01	9.0745E+00	-1.0569E-01	-7.6361E-02	4.6715E-04	7.1010E-01	-1.2187E-02	5.9678E-02	9.9922E-02	-7.5580E-03	-1.2245E-01	-7.5073E-02
1.1000E+01	2.0000E+00	1.0000E+00	2.0000E-01	9.0495E+00	-1.0569E-01	-7.6361E-02	4.6715E-04	7.1010E-01	-1.2187E-02	5.9678E-02	9.9922E-02	-7.5580E-03	-1.2245E-01	-7.5073E-02
...														
1.2000E+01	2.0000E+00	1.0000E+00	0.0000E+00	7.1985E+00	-1.9556E-02	1.0727E-01	4.6715E-04	7.1010E-01	-9.3251E-02	1.1041E-01	6.0611E-02	-1.3951E-01	2.0220E-02	6.9816E-02
1.2000E+01	2.0000E+00	1.0000E+00	1.0000E-01	7.1827E+00	-1.9556E-02	1.0727E-01	4.6715E-04	7.1010E-01	-9.3251E-02	1.1041E-01	6.0611E-02	-1.3951E-01	2.0220E-02	6.9816E-02
1.2000E+01	2.0000E+00	1.0000E+00	2.0000E-01	7.1669E+00	-1.9556E-02	1.0727E-01	4.6715E-04	7.1010E-01	-9.3251E-02	1.1041E-01	6.0611E-02	-1.3951E-01	2.0220E-02	6.9816E-02

ETA1-10=ETA(f) to ETA(10), empirical bayes estimates; IPRED, individual predicted value; OCC=occasion number; SID=site identification number.

**Table 5** Outputted table results for Example 6: Modeling a mixture of subpopulations of parameters (pmixture\_foce.par)

ID	V	K	BESTSUB
1.0000E+00	4.2434E+01	9.3660E-02	1.0000E+00
2.0000E+00	6.6534E+01	1.1262E-01	1.0000E+00
3.0000E+00	6.1668E+01	5.0276E-01	2.0000E+00
4.0000E+00	7.4587E+01	9.5481E-02	1.0000E+00
5.0000E+00	5.0625E+01	1.1113E-01	1.0000E+00
6.0000E+00	5.9876E+01	4.2720E-01	2.0000E+00
7.0000E+00	5.8508E+01	1.3610E-01	1.0000E+00
8.0000E+00	6.9308E+01	1.2559E-01	1.0000E+00
9.0000E+00	7.0448E+01	4.0392E-01	2.0000E+00
1.0000E+01	6.2197E+01	1.2000E-01	1.0000E+00
1.1000E+01	7.4701E+01	8.9449E-02	1.0000E+00
1.2000E+01	7.6044E+01	7.6652E-01	2.0000E+00
1.3000E+01	6.9676E+01	9.4314E-02	1.0000E+00
1.4000E+01	6.4067E+01	1.4138E-01	1.0000E+00
1.5000E+01	5.3923E+01	8.0854E-01	2.0000E+00
1.6000E+01	6.7635E+01	1.2107E-01	1.0000E+00
1.7000E+01	1.0518E+02	1.0739E-01	1.0000E+00
1.8000E+01	9.1576E+01	5.1767E-01	2.0000E+00
1.9000E+01	9.4717E+01	7.6706E-02	1.0000E+00
2.0000E+01	8.7513E+01	7.7070E-02	1.0000E+00
2.1000E+01	5.7781E+01	4.5237E-01	2.0000E+00
2.2000E+01	6.9630E+01	1.0151E-01	1.0000E+00
2.3000E+01	5.8963E+01	1.5230E-01	1.0000E+00
2.4000E+01	5.9655E+01	8.3048E-01	2.0000E+00
2.5000E+01	6.0324E+01	1.0135E-01	1.0000E+00
2.6000E+01	7.3017E+01	1.1475E-01	1.0000E+00
2.7000E+01	6.6592E+01	4.7002E-01	2.0000E+00
2.8000E+01	7.1897E+01	1.0585E-01	1.0000E+00
2.9000E+01	7.1601E+01	1.1429E-01	1.0000E+00
3.0000E+01	5.1493E+01	3.3186E-01	2.0000E+00
3.1000E+01	4.8920E+01	1.1613E-01	1.0000E+00
3.2000E+01	9.2087E+01	1.1539E-01	1.0000E+00
3.3000E+01	5.8832E+01	5.1238E-01	2.0000E+00
3.4000E+01	7.7011E+01	1.0821E-01	1.0000E+00
3.5000E+01	6.5992E+01	9.3086E-02	1.0000E+00
3.6000E+01	7.3862E+01	2.8765E-01	2.0000E+00
3.7000E+01	7.7355E+01	1.1107E-01	1.0000E+00
3.8000E+01	6.3286E+01	9.9980E-02	1.0000E+00
3.9000E+01	7.8039E+01	6.8947E-01	2.0000E+00

BESTSUB=index to best fitting sub-population; ID=subject identification number; K=rate constant of elimination; V=volume of distribution.

of analytical derivatives with respect to the population parameters, which also increases accuracy of evaluating the variance-covariance of estimates during the \$COV step. The example r2complb\_foce.cti was evaluated with low SIGL(=6) low NSIG(=1), with low tolerance setting for ODE integration (TOL = 6) and using the FAST option, resulting in the estimation completing in 475 seconds. For a more conservative precision assessment, with SIGL = 9, NSIG = 3, TOL = 9, the FAST option completes the estimation in 1,022 seconds (r2complb\_foce3.cti), which is still much faster, and with a cleaner termination status, than when not using the FAST

option (23,777 seconds, r2complb\_foce4.cti). Of course, individual rates of improvement will vary depending on the model and data.

### EXAMPLE 5: ANALYZING DATA MODELED WITH MULTIPLE LEVELS OF MIXED EFFECTS

Additional nested levels of random effects may be added in NONMEM. In the following example the PK parameters may vary within an individual from one dosing event to the next, called interoccasion variability. Furthermore, there are the usual random effects between individuals (interindividual variability) as well as random effects explaining variability between one clinical site and the next (intersite variability). With the residual variance, this consists of four hierarchical random effects and errors. The code of superid30\_1\* gives an example of how this may be conveniently coded using some of the symbolic declarations available since NONMEM 7.3 (superid30\_1\_bayes.cti, superid30\_1\_saem.cti, superid\_1\_foce.cti; **Supplementary Materials S6**). The dosing events within a given subject are demarcated by the OCC (occasion) data item, which has values of 1, 2, 3, etc., to indicate occasion number. The \$ABBR RERPLACE records map symbolic references OCC\_CL and OCC\_V to the ETAs that are appropriate for each occasion. For example, for PK parameter CL, OCC\_CL maps to ETA(5), ETA(7), or ETA(9), depending on whether the OCC data item is 1, 2, or 3, respectively. This allows three separate CLs to be independently created, one for each occasion, within a particular individual. A similar mapping is performed for parameter V using OCC\_V to map ETA(6), ETA(8), or ETA(10) for the three occasions using the \$ABBR REPLACE record (**Table 1**, code item 11). Random effects that scope across individuals such as intersite variability in this example are mapped using the \$LEVEL record (**Table 1**, code item 12). The SID (site identification number) data item has separate values in the data set for each site, and those subjects sharing the same site value will share the same random effect ETA value. NONMEM is informed that ETA(3) is a CL ETA that changes only with every site and is associated by nesting with CL ETA(1), which varies with each subject. They associate together because they provide random effects to the same individual parameter. Similarly, volume ETA(4) is another eta that changes only with every site and is associated by nesting with intersubject volume random effect ETA(2). This relationship is reflected in the way the individual parameters are modeled using these ETAs in the \$PK record (**Table 1**, code item 13). The hierarchical random effects levels are each given OMEGA block descriptions (**Table 1**, code item 14). To permit a BAYES analysis, uninformative prior OMEGAs are supplied with the same pattern as the to-be-estimated OMEGAs (superid30\_1\_bayes).

The estimation records shown in **Supplementary Materials S6** are examples of how one may perform ITS, SAEM, IMP, BAYES, and FOCE on this type of problem. These are followed by the \$COV and \$TABLE records. Of particular importance is that the FNLETA must be set to 0 to have final SID ETAs retain their shared values among subjects of a common SID value when outputted in a table (in this example, superid30\_1\*.tab). Also, for FOCE, options

Table 6 NONMEM, partial phm file results for Example 6: Modeling a mixture of subpopulations of parameters (pmixture\_force.phm)

TABLE NO. 1: First-order conditional estimation with interaction: Problem = 1 Subproblem = 0 Superproblem1 = 0 Iteration1 = 0 Superproblem2 = 0 Iteration2 = 0													
ID	SUBPOP	PMIX	ETA(1)	ETA(2)	ETA(3)	ETA(4)	ETC (1,1)	ETC (2,1)	ETC (2,2)	ETC (3,3)	ETC (4,3)	ETC (4,4)	OBJ
1	1	1.00000E+00	-4.93563E-01	-7.16567E-02	0.00000E+00	0.00000E+00	1.82140E-03	-1.33878E-03	2.97276E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-14.7965423227514
1	2	4.67808E-13	0.00000E+00	0.00000E+00	-5.35659E-01	-1.63579E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.85406E-03	-1.33784E-03	2.82839E-03	40.5984872879239
2	1	1.00000E+00	-4.37870E-02	1.12666E-01	0.00000E+00	0.00000E+00	1.83753E-03	-1.14689E-03	2.12519E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-27.45668591873074
2	2	1.42478E-09	0.00000E+00	0.00000E+00	-8.52734E-02	-1.45695E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.86034E-03	-1.13173E-03	2.00410E-03	11.8952271143952
3	1	1.88263E-21	-1.25051E-01	1.60474E+00	0.00000E+00	0.00000E+00	1.86450E-03	-2.72658E-04	1.15078E-04	4.93103E-02	-1.10284E-02	5.99658E-02	41.2497512311461
3	2	1.00000E+00	0.00000E+00	0.00000E+00	-1.40094E-01	-1.18895E-02	4.61817E-02	-8.57633E-03	2.72419E-02	1.86913E-03	-2.72204E-04	1.14364E-04	-55.5798855094779
4	1	1.00000E+00	7.04636E-02	-5.24048E-02	0.00000E+00	0.00000E+00	1.82337E-03	-1.31803E-03	2.87198E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-30.5564906816249
4	2	2.61403E-11	0.00000E+00	0.00000E+00	2.91913E-02	-1.61671E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.85486E-03	-1.31447E-03	2.72696E-03	16.79219076339984
5	1	1.00000E+00	-3.17050E-01	9.93389E-02	0.00000E+00	0.00000E+00	1.83656E-03	-1.16029E-03	2.17821E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-12.5394921766005
5	2	1.54579E-10	0.00000E+00	0.00000E+00	-3.61141E-01	-1.46635E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.86007E-03	-1.14184E-03	2.04093E-03	31.2547378153604
6	1	5.53963E-17	-1.49876E-01	1.43913E+00	0.00000E+00	0.00000E+00	1.86469E-03	-3.21643E-04	1.60099E-04	4.93103E-02	-1.10284E-02	5.99658E-02	32.0133384541306
6	2	1.00000E+00	0.00000E+00	0.00000E+00	-1.69586E-01	-1.74743E-01	4.61817E-02	-8.57633E-03	2.72419E-02	1.86934E-03	-3.20334E-04	1.58334E-04	-44.2371064095188
7	1	1.00000E+00	-1.72338E-01	3.02066E-01	0.00000E+00	0.00000E+00	1.84871E-03	-9.68943E-04	1.48974E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-26.9334957081190
7	2	2.47355E-07	0.00000E+00	0.00000E+00	-2.12555E-01	-1.27566E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.86444E-03	-9.51460E-04	1.40747E-03	2.1049824200648
8	1	1.00000E+00	-2.94763E-03	2.21691E-01	0.00000E+00	0.00000E+00	1.84453E-03	-1.04190E-03	1.73401E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-21.2148466908645
8	2	7.50375E-08	0.00000E+00	0.00000E+00	-5.24174E-02	-1.33823E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.86324E-03	-1.01054E-03	1.59067E-03	10.2093031386916
9	1	2.32945E-16	1.11717E-02	1.38345E+00	0.00000E+00	0.00000E+00	1.86471E-03	-3.39888E-04	1.78874E-04	4.93103E-02	-1.10284E-02	5.99658E-02	26.5010173168367
9	2	1.00000E+00	0.00000E+00	0.00000E+00	-6.98870E-03	-2.30785E-01	4.61817E-02	-8.57633E-03	2.72419E-02	1.86939E-03	-3.38783E-04	1.77081E-04	-46.8768468381362
10	1	1.00000E+00	-1.11194E-01	1.76185E-01	0.00000E+00	0.00000E+00	1.84181E-03	-1.08493E-03	1.88839E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-26.2675357733099
10	2	6.10413E-09	0.00000E+00	0.00000E+00	-1.51094E-01	-1.39829E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.86188E-03	-1.07033E-03	1.78824E-03	10.1746567796667
11	1	1.00000E+00	7.19914E-02	-1.17659E-01	0.00000E+00	0.00000E+00	1.81638E-03	-1.38883E-03	3.22615E-03	4.93103E-02	-1.10284E-02	5.99658E-02	-28.9248184755368
11	2	5.02444E-12	0.00000E+00	0.00000E+00	2.67665E-02	-1.67209E+00	4.61817E-02	-8.57633E-03	2.72419E-02	1.85243E-03	-1.38321E-03	3.03126E-03	21.7221908489041
12	1	4.92953E-85	3.25971E-01	5.45635E-01	0.00000E+00	0.00000E+00	1.85739E-03	-7.72211E-04	9.33345E-04	4.93103E-02	-1.10284E-02	5.99658E-02	313.955053188483
12	2	1.00000E+00	0.00000E+00	0.00000E+00	6.94565E-02	4.09859E-01	4.61817E-02	-8.57633E-03	2.72419E-02	1.86850E-03	-1.78488E-04	4.92198E-05	-75.6803309494078

ETA(x)=empirical bayes estimate to xth eta; ETC(x,y)=conditional variance between ETA(x) and ETA(y); ID=subject identification number; OBJ=objective function value for a given subject and sub-population; PMIX=probability of subject belonging to this sub-population; SUBPOP=sub-population index.

**Table 7** Subset of records to data file (urine.dat) for example 7, with urine voiding records

ID	TIME	AMT	RATE	DV	MDV	EVID	CMT	UVOL
1.0000E + 00	0.0000E + 00	1.0000E + 03	0.0000E + 00	0.0000E + 00	1.0000E + 00	1.0000E + 00	1.0000E + 00	0.0000E + 00
1.0000E + 00	0.0000E + 00	0.0000E + 00	0.0000E + 00	0.0000E + 00	1.0000E + 00	2.0000E + 00	2.0000E + 00	0.0000E + 00
...								
1.0000E + 00	5.0000E + 00	0.0000E + 00	0.0000E + 00	9.5704E + 00	0.0000E + 00	0.0000E + 00	-2.0000E + 00	1.0000E + 01
1.0000E + 00	5.0000E + 00	0.0000E + 00	0.0000E + 00	0.0000E + 00	1.0000E + 00	2.0000E + 00	2.0000E + 00	0.0000E + 00
...								
1.0000E + 00	1.0000E + 01	0.0000E + 00	0.0000E + 00	5.8533E + 00	0.0000E + 00	0.0000E + 00	-2.0000E + 00	1.2000E + 01
1.0000E + 00	1.0000E + 01	0.0000E + 00	0.0000E + 00	0.0000E + 00	1.0000E + 00	2.0000E + 00	2.0000E + 00	0.0000E + 00
...								

AMT, amount of drug; CMT, compartment number; DV, dependent variable; EVID=event ID; ID, subject identification number; MDV, missing dependent variable indicator; UVOL, urine volume.

SLOW and NONINFETA = 1 must be set so that proper gradients are evaluated for updating the SID ETAs. **Table 4** shows a part of superid30\_1\_foce.tab that outputs the ID, SID, OCC, TIME, IPRED, and ETAs. Notice that the SID-level ETAs 3 and 4 do not change until SID changes.

Only one of the ETAs belonging to an individual parameter requires a nonzero MU associated with the typical value for that parameter. Here, the MU\_ variables are associated with the ETAs to intersubject variability (such as MU\_1 as typical value for CL and indexed to ETA(1), the intersubject variability). The other ETAs, for intersite variability and interoccasion variability, implicitly have 0 valued MU\_'s, hence MU\_3, MU\_5, etc., are not modeled. NMTRAN issues MU referencing warnings, which can be ignored. Notice that once again, ITS estimation can be used as a partial burning-in/estimation step to facilitate SAEM, BAYES, and FOCE.

#### EXAMPLE 6: MODELING A MIXTURE OF SUBPOPULATIONS OF PARAMETERS

The distribution of model parameters sometimes may be best described as coming from two or more subpopulations among the subjects. When this occurs, it is desirable to determine if a particular categorical covariate, such as sex, can explain the reason for these subpopulations. Sometimes none of the covariates available can explain the subpopulation distributions, and it is then desired that individuals be tested for their relative goodness of fit between these subpopulations. The example in **Supplementary Materials S7** (pmixture\_foce.ctl, pmixture\_saem.ctl, pmixture\_bayes.ctl) with a one-compartment model and two subpopulations of V and K demonstrates how this is set up in NONMEM. In the \$PK record, two sets of V and K1 are set up for the analysis, one pair for each subpopulation, each pair distributed according to that subpopulation's mean MU and ETAs (**Table 1**, code item 15). In this example, subpopulation 1 is modeled using MU\_1 for mean to LOG(V), and MU\_2 for mean to log(K), with their respective ETAs for their random effects, whereas subpopulation 2 parameters are modeled with means MU\_3 and MU\_4, and their respective ETAs. NONMEM requests individual parameters to be modeled for subpopulation 1 or 2 for each call to the user model via the parameter MIXNUM. It then determines the goodness of fit of the subject's data considering each subpopulation, and the MIXEST value is set to the subpopulation number that best fits that subject's data.

The subpopulations are given separate block OMEGAs, as there should be no correlation between parameters across subpopulations (**Table 1**, code item 16). As usual, priors to OMEGAs are set up to match the block pattern of the estimated OMEGAs. The *post hoc* best-fitting V, K for each subject and the subpopulation to which they belong (BESTSUB = MAXEST) may be outputted, as indicated in the \$TABLE record, outputting to file pmixture\*.par. **Table 5** lists a part of pmixture\_foce.par. In addition, the conditional mean (for Monte Carlo EM) or mode (for FOCE/Laplace/ITS) individual PHIs and their variances for each subpopulation are listed in pmixture\*.phm (see **Table 6**), along with the probability of belonging to a subpopulation PMIX. For each subject, MIXEST is related to PMIX by:

MIXEST=sub-population with largest PMIX value.

The phi(x) values in the.phm file are related to the V and K in the pmixture.par file as follows:

```
If BESTSUB=1:
V=exp(phi(1))
K=exp(phi(2))
If BESTSUB=2:
V=exp(phi(3))
K=exp(phi(4))
```

The conditional mean and variances averaged among the subpopulations using the PMIX value as weight is listed in pmixture\*.phi (not shown). For this example, the subjects are extremely likely to belong to one subpopulation or another, so these values do not differ from the records in pmixture.phm whose PMIX value is close to 1.

#### EXAMPLE 7: MODELING PERIODICALLY COLLECTED URINE SAMPLES

When urine data are collected, the model must account for the urine compartment to be emptied and then filled. Complete urine data consists of time of start of an empty bladder, time of emptying of bladder measured, concentration of drug in urine, and total volume of urine collected at that time. In NONMEM, a particular compartment X may be emptied (turned off), that is, its amount A(x) set to 0, when a data record is encountered in which its CMT (compartment



number) item is minus  $x$ . A subsequent data record may be used to indicate the start of its refilling or turned on with a positive  $x$ . Usually, this would be immediately after urine collection (and bladder voiding) of the previous  $-x$  record. For example, a series of data records may be seen in **Table 7**.

For the ADVAN series of models, the  $N+1$ th compartment is called the output compartment. For ADVAN1, compartment 2 is the output (urine) compartment. The output compartment by default is turned off and is turned on when a CMT = 2 record appears. As this record is used just to turn on “bladder filling” and there are no data associated with it, its EVID is set to 2, a nondose, nonobservation event. The next record of interest is the urine collection record at time = 5, where the urine concentration is recorded as DV, urine volume (amount of urine voided) UVOL is recorded, and the CMT =  $-2$ , to indicate that, in addition to this being a urine collection record, the compartment is to be “turned-off,” that is, amount of compartment 2 set to 0 (bladder emptied) immediately after. Then, the next record should be EVID = 2, CMT = 2, at the same time, to indicate that the output compartment is to be turned on again, that is, the bladder is to be refilled in anticipation of the next urine collection at time = 10.

The model listed in **Supplementary Materials S8** uses the urine data to determine the portion of drug that is renally cleared (total clearance is CL, renal clearance is CLR), and IMP is used as the estimation method. The CLR is related to the total CL by  $F2 = CL/CLR$ , where F2 is a reserved variable that determines what fraction of eliminated drug from compartment 1 goes to the urine compartment 2 (urine.cti; **Table 1**, code item 17). The \$ERROR block assigns the predicted values, CP (concentration of plasma), or CU (concentration of urine), and residual error (EPS(1) or EPS(2)) to the appropriate data (**Table 1**, code item 18).

This problem may also be estimated using one of the differential equation solver ADVANs (ADVAN6, ADVAN8, ADVAN9, ADVAN13, ADVAN14, and ADVAN15). The control stream model needs to be modified at the \$SUBROUTINES record and \$DES record (urine2.cti; **Supplementary Materials S8**). The SIGL should match the TOL, and NSIG should be  $3\times$  less than SIGL for ODE problems. Note that output compartment is NCOMPARTMENTS+1, so the output compartment is still 2.

## CONCLUSION

This tutorial listed examples of typical types of PK/PD modeling problems that occur in the pharmaceutical field, which the reader can use as a template for his or her own modeling endeavors in NONMEM. It is hoped that with these examples and the description of the general concepts about the properties and behaviors of the various statistical algorithms, the user is able to construct models of varying complexity and wisely choose the appropriate estimation method.

**Supporting Information.** Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website ([www.psp-journal.com](http://www.psp-journal.com)).

**Supplementary Materials S1.** (file NONMEM\_Tutorial\_II\_S1.pdf)

A PDF file containing:

Part A: Additional Information on MU-Referencing.

Part B: No-U-Turn Hamiltonian Bayes Method.

Part C: Providing Prior Information for Bayesian Analysis.

Part D: Termination Testing.

Part E: Setting up Example 1 for Bayesian Analysis (504b.cti).

**Supplementary Materials S2.** (Example 1, file NONMEM\_Tutorial\_I\_S2.zip):

The zip file contains the following files for example 1:

501.csv: data file.

504\_foce.cti: control stream file.

504\_foce.res: NONMEM report file.

504\_foce.tab: Parameters outputted by user-specified \$TABLE record.

504\_foce.ext: Raw output file.

504\_foce.cov: Variance-Covariance matrix of estimates to THETAs, OMEGAs, and SIGMAs.

504\_foce.cor: Fully informative correlation matrix of estimates, with standard errors as diagonal elements, and correlation values on the off-diagonal elements.

504\_foce.coi: Inverse covariance matrix (Fisher information matrix).

504\_foce.phi: Individual ETAs (eta()), their variances (etc()), and individual objective function values.

504\_its.cti: control stream file.

504\_its.res: NONMEM report file.

504\_its.tab: Parameters outputted by user-specified \$TABLE record.

504\_its.ext: Raw output file.

504\_its.cov: Variance-Covariance matrix of estimates to THETAs, OMEGAs, and SIGMAs.

504\_its.cor: Fully informative correlation matrix of estimates, with standard errors as diagonal elements, and correlation values on the off-diagonal elements.

504\_its.coi: Inverse covariance matrix (Fisher information matrix).

504\_its.phi: Individual ETAs (eta()), their variances (etc()), and individual objective function values.

504\_saem.cti: control stream file.

504\_saem.res: NONMEM report file.

504\_saem.tab: Parameters outputted by user-specified \$TABLE record.

504\_saem.ext: Raw output file.

504\_saem.cov: Variance-Covariance matrix of estimates to THETAs, OMEGAs, and SIGMAs.

504\_saem.cor: Fully informative correlation matrix of estimates, with standard errors as diagonal elements, and correlation values on the off-diagonal elements.

504\_saem.coi: Inverse covariance matrix (Fisher information matrix).

504\_saem.phi: Individual ETAs (eta()), their variances (etc()), and individual objective function values.

504\_bayes.cti: control stream file.

504\_bayes.res: NONMEM report file.

504\_bayes.tab: Parameters outputted by user-specified \$TABLE record.

504\_bayes.ext: Raw output file.

504\_bayes.cov: Variance-Covariance matrix of estimates to THETAs, OMEGAs, and SIGMAs.

504\_bayes.cor: Fully informative correlation matrix of estimates, with standard errors as diagonal elements, and correlation values on the off-diagonal elements.

504\_bayes.coi: Inverse covariance matrix (Fisher information matrix).

504\_bayes.phi: Individual ETAs (eta()), their variances (etc()), and individual objective function values.

504\_nuts.cti: control stream file.

504\_nuts.res: NONMEM report file.

504\_nuts.tab: Parameters outputted by user-specified \$TABLE record.

504\_nuts.ext: Raw output file.

504\_nuts.cov: Variance-Covariance matrix of estimates to THETAs, OMEGAs, and SIGMAs.

504\_nuts.cor: Fully informative correlation matrix of estimates, with standard errors as diagonal elements, and correlation values on the off-diagonal elements.

504\_nuts.coi: Inverse covariance matrix (Fisher information matrix).

504\_nuts.phi: Individual ETAs (eta()), their variances (etc()), and individual objective function values.

**Supplementary Materials S3. (Example 2, file NONMEM\_Tutorial\_II\_S3.zip):**

The zip file contains the following files for example 2:

ad3tr4.csv.

ad3tr4\_loqb\_lap.ctl.

ad3tr4\_loqb\_lap.res.

ad3tr4\_loqb\_lap.ext.

ad3tr4\_loqb\_lap.tab.

ad3tr4\_loqb\_lap.phi.

ad3tr4\_loqb\_lap.coi.

ad3tr4\_loqb\_lap.cov.

ad3tr4\_loqb\_imp.ctl.

ad3tr4\_loqb\_imp.res.

ad3tr4\_loqb\_imp.ext.

ad3tr4\_loqb\_imp.tab.

ad3tr4\_loqb\_imp.phi.

ad3tr4\_loqb\_imp.coi.

ad3tr4\_loqb\_imp.cov.

**Supplementary Materials S4. (Example 3, file NONMEM\_Tutorial\_II\_S4.zip):**

The zip file contains the following files for example 3:

wexample10.csv.

wexample10\_lap.ctl.

wexample10\_lap.res.

wexample10\_lap.ext.

wexample10\_lap.tab.

wexample10\_lap.phi.

wexample10\_lap.coi.

wexample10\_lap.cov.

wexample10\_saem.ctl.

wexample10\_saem.res.

wexample10\_saem.ext.

wexample10\_saem.tab.

wexample10\_saem.phi.

wexample10\_saem.coi.

wexample10\_saem.cov.

wexample10\_bayes.ctl.

wexample10\_bayes.res.

wexample10\_bayes.ext.

wexample10\_bayes.tab.

wexample10\_bayes.phi.

wexample10\_bayes.coi.

wexample10\_bayes.cov.

**Supplementary Materials S5. (Example 4, file NONMEM\_Tutorial\_II\_S5.zip):**

The zip file contains the following files for example 4:

r2comp.csv.

r2complb\_imp.ctl.

r2complb\_imp.res.

r2complb\_imp.ext.

r2complb\_imp.phi.

r2complb\_imp.coi.

r2complb\_imp.cov.

r2complb\_foce.ctl.

r2complb\_foce.res.

r2complb\_foce.ext.

r2complb\_foce.phi.

r2complb\_foce.coi.

r2complb\_foce.cov.

r2complb\_foce3.ctl.

r2complb\_foce3.res.

r2complb\_foce3.ext.

r2complb\_foce3.phi.

r2complb\_foce3.coi.

r2complb\_foce3.cov.

r2complb\_foce4.ctl.

r2complb\_foce4.res.

r2complb\_foce4.ext.

r2complb\_foce4.phi.

r2complb\_foce4.coi.

r2complb\_foce4.cov.

**Supplementary Materials S6. (Example 5, file NONMEM\_Tutorial\_II\_S6.zip):**

The zip file contains the following files for example 5:

Superid30.csv.

superid30\_1\_foce.ctl.

superid30\_1\_foce.res.

superid30\_1\_foce.ext.

superid30\_1\_foce.tab.

superid30\_1\_foce.phi.

superid30\_1\_foce.coi.

superid30\_1\_foce.cov.

superid30\_1\_imp.ctl.

superid30\_1\_imp.res.

superid30\_1\_imp.ext.

superid30\_1\_imp.tab.

superid30\_1\_imp.phi.

superid30\_1\_imp.coi.

superid30\_1\_imp.cov.

superid30\_1\_bayes.ctl.

superid30\_1\_bayes.res.

superid30\_1\_bayes.ext.

superid30\_1\_bayes.tab.

superid30\_1\_bayes.phi.

superid30\_1\_bayes.coi.

superid30\_1\_bayes.cov.

**Supplementary Materials S7. (Example 6, file NONMEM\_Tutorial\_II\_S7.zip):**

The zip file contains the following files for example 6:

Pmixture\_foce.csv.

pmixture\_foce.ctl.

pmixture\_foce.res.

pmixture\_foce.ext.

pmixture\_foce.phi.

pmixture\_foce.phm.

pmixture\_foce.cov.

pmixture\_foce.coi.

pmixture\_foce.cor.

pmixture\_saem.ctl.  
pmixture\_saem.res.  
pmixture\_saem.ext.  
pmixture\_saem.phi.  
pmixture\_saem.phm.  
pmixture\_saem.cov.  
pmixture\_saem.coi.  
pmixture\_saem.cor.  
pmixture\_bayes.ctl.  
pmixture\_bayes.res.  
pmixture\_bayes.ext.  
pmixture\_bayes.phi.  
pmixture\_bayes.phm.  
pmixture\_bayes.cov.  
pmixture\_bayes.coi.  
pmixture\_bayes.cor.

**Supplementary Materials S8. (Example 7, file NONMEM\_Tutorial\_II\_S8.zip):**

The zip file contains the following files for example 7:

urine.dat.  
urine.ctl  
urine.res.  
urine.tab.  
urine.ext.  
urine.phi.  
urine.coi.  
urine.cov.  
urine.cor.  
urine2.ctl.  
urine2.res.  
urine2.tab.  
urine2.ext.  
urine2.phi.  
urine2.coi.  
urine2.cov.  
urine2.cor.

**Funding.** ICON Clinical Research LLC. licenses the NONMEM and PDx-Pop software.

**Conflict of Interest.** The author is a paid employee of ICON Clinical Research LLC. and is the present developer of NONMEM.

1. Beal, S.L., Sheiner, L.B., Boeckmann, A.J. & Bauer, R.J. (Eds). *NONMEM 7.4 users guides* (ICON plc, Gaithersburg, MD, 1989–2018). <<https://nonmem.iconplc.com/nonmem743/guides>>.

2. Steimer, J.L., Mallet, A., Golmard, J.L. & Boisvieux, J.F. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed-effect model. *Drug Metab. Rev.* **15**, 265–292 (1984).
3. Bauer, R.J., Guzy, S. & Ng, C.M. A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. *AAPS J* **9**, E60–E83 (2007).
4. Bauer, R.J. & Guzy, S. Monte Carlo parametric expectation maximization (MC-PEM) method for analyzing population pharmacokinetic/pharmacodynamic data. In *Advanced Methods of Pharmacokinetic and Pharmacodynamic Systems Analysis*, Vol. 3 (ed. D'Argenio, D.Z.) 135–163 (Kluwer Academic Publishers, Boston, MA, 2004).
5. Lavielle, M. SAEM in MATLAB: an alternative to linearization. Population Approach Group Europe (PAGE) Meeting, Uppsala, Sweden, June 17–18, 2004.
6. PKBUGS software, version 2.0 [computer program] (MRC Biostatistics Unit, Cambridge, UK).
7. Hoffman, M.D. & Gelman, A. The No-U-Turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**, 1593–1623 (2014).
8. Stan Modeling Language. *User's guide and reference manual*. Stan Reference Manual. Stan Development Team, V.2.18 (July 2018) <https://mc-stan.org/>.
9. Ahn, J.E., Karlsson, M.O., Dunne, A. & Ludden, T.M. Likelihood based approaches to handling data below the quantification limit using NONMEM VI. *J. Pharmacokinet. Phar.* **35**, 401–402 (2008).
10. Hooker, A.C., Staatz, C.E. & Karlsson, M.O. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. *Pharm. Res.* **24**, 2187–2197 (2007).
11. Comets, E., Brendel, K. & Mentre, F. Computing normalized prediction distribution errors to evaluate nonlinear mixed effects models: the npde add-on package for R. *Comput. Methods Programs Biomet.* **90**, 154–166 (2008).
12. Brendel, K., Comets, E., Laffont, C., Laveille, C. & Mentré, F. Metrics for external model evaluation with an application to the population pharmacokinetics of glimepiride. *Pharm. Res.* **23**, 2036–2049 (2006).
13. Nguyen, T.H.T., Comets, E. & Mentré, F. Extension of NPDE for evaluation of nonlinear mixed effect models in presence of data below the quantification limit with applications to HIV dynamic model. *J. Pharmacokinet Pharmacodyn.* **39**, 499–518 (2012).
14. Gibansky, L., Gibansky, E. & Bauer, R. Comparison of NONMEM 7.2 estimation methods and parallel processing efficiency on a target-mediated drug disposition model. *J. Pharmacokinet Pharmacodyn.* **39**, 17–36 (2012).

© 2019. ICON Clinical Research LLC. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals, Inc. on behalf of the American Society for Clinical Pharmacology & Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.