THE SELECTION AND IMPLEMENTATION OF NON-LINEAR MIXED MODELS IN PHARMACOKINETIC RESEARCH: A NONMEM/PDX-POP TUTORIAL FOR STATISTICIANS

By

CAROLYN CLAYTON

B.S., Southern Adventist University, 2012

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Biostatistics Program

2017

This thesis for the Master of Science degree by

Carolyn Clayton

has been approved for the

Department of Biostatistics and Informatics

by

Samantha MaWhinney, Chair

Laura Saba

Jennifer Kiser

Date: May 26, 2017

Clayton, Carolyn (M.S. Biostatistics, Biostatistics and Informatics)

The Selection and Implementation of Non-Linear Mixed Models in Pharmacokinetic Research: A NONMEM/PDx-POP Tutorial for Statisticians

Thesis directed by Professor Samantha MaWhinney

ABSTRACT

Pharmacokinetic (PK) studies aim to develop models of the absorption, distribution, metabolism and excretion of a drug by the body. Traditional PK studies focus on minimizing inter-individual variability and obtaining initial understandings of the PK pathways, often through use of a sample of healthy subjects and strict study control protocols. In contrast, population-PK studies typically utilize a large sample from the population of interest and attempt to explain sources of variability including various demographic, genetic, or concomitant drug covariates. PK studies represent a significant step in the drug development process and inform dosing regimens, minimize toxicity, and increase drug efficacy. Drug PK may be understood as a series of compartments where organs or tissues grouped within a compartment have similar absorption and/or elimination rates. When equilibrium is quickly achieved across tissues a one-compartment model is sufficient. More often, a two- or more compartments model may be required.

Each compartmental model has associated PK parameters; for instance K_a for absorption rate, Cl for clearance, and V for volume. Because multiple measurements are taken per individual, nonlinear mixed effects models present a means to model PK parameters, account for repeated measurements, and adjust for potential covariates on each PK parameter explicitly. NONMEM, a commercially-available software, is specifically designed to analyze these types of models.

This paper aims to provide a basic introductory background in PK models, nonlinear mixed models, and NONMEM. Guidelines are given for model selection, including which parameters may require random effects or covariates, and procedures for selecting a final model. Graphical analysis guides are given to assess model fit. NONMEM options, data format requirements, and a tutorial for an associated user interface, PDx-POP, are included. Model building is demonstrated on an example dataset and output is interpreted.

The form and content of this abstract are approved. I recommend its publication.

Approved: Samantha MaWhinney

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Pharmacokinetic (PK) studies aim to model drug absorption, distribution, elimination, bioavailability, dose proportionality, or subpopulation differences in concentration levels using our understanding of biological systems (Cho, 2011) in order to increase efficacy, mitigate toxicity, and develop dosing regimens (European Medicines Agency, 1987). Traditional PK studies focus on minimizing inter-individual variability, often through use of a sample of healthy subjects and strict study protocols. In contrast, population-PK (pop-PK) studies typically utilize a larger sample that comes from the population of interest and attempt to explain the sources of variability in the sample through various demographic, genetic, or concomitant drug covariates (Food and Drug Administration, 1999, p. 2). There are multiple applications of both traditional and pop-PK. Phase I drug trials with intensive sampling may utilize traditional PK methods. These early studies aim to obtain initial information on the PK pathway, compute rough estimates for PK parameters, and determine concentrations and pharmacokinetics across various drug concentrations. Due to the small sample size typically studied, Phase I trials may include only a few covariates. Phase III and IV drug trials with sparse sampling and a large sample size from the population of interest may utilize pop-PK techniques. The inter-individual variability in the sample may aid in determining whether certain sub-populations require different dosing regimens (Food and Drug Administration, 1999, p. 5). PK parameter estimates from the Phase I trials may be used in order to improve chances of model convergence in the Phase III trial, and multiple covariates of interest are usually tested.

These PK studies represent a significant step in the drug development process because it is important to understand the concentration and flow of the drug throughout the body, both to minimize toxicity and to increase efficacy of the drug. Additional applications of these types

of studies include health risk mitigation, development of dosing frequency regimens, or prediction of dose response in a patient given age, gender, geno/phenotype, and/or another biological factor (Davidian & Giltinan, 2003; Food and Drug Administration, 2003). Usually, though not always, multiple measurements are taken per individual over time.

One method of understanding drug perfusion through the body is as a series of compartments (Fisher & Shafer, 2007, pp. 39-44). PK models can be used to estimate the involvement of groups of organ systems, called compartments. Depending on the model, the rates of absorption and elimination in various compartments (also known as perfusion rates) or the effect of certain genetic and physiological factors on the processing of the drug can also be estimated. When perfusion rates are relatively fast from the blood into other tissues, the body may be modeled as a single compartment. When the perfusion rates are slower, a more complex two-compartment model may be used instead. The change in drug concentration through these compartments is modeled with a series of differential equations.

In dose concentration studies, multiple measurements are taken per person. Mixed models allow for a correlation between measurements from the same subject, as well as estimation of covariate effects on the concentration outcome. Nonlinear mixed models utilize ordinary differential equations (ODEs) to model PK compartments so that the resulting estimated parameters are clinically interpretable.

In order to efficiently and accurately assess these models, NONMEM was developed and released as a commercial software in 1980. It includes pre-specified routines for various PK compartments and dosing situations in order to reduce the amount of coding required by the end user. PDx-POP is a graphical user interface program available from ICON PLC, the current distributers of NONMEM. PDx-POP provides certain pre-made code files to further reduce coding, and creates summary statistics and diagnostic graphs.

In order to illustrate the use of nonlinear mixed models in PK analysis in the context of NONMEM, data obtained from a drug concentration study of the antiretroviral raltegravir was analyzed. The data appears as in Figure 1 when graphed. The process of model building was demonstrated and the effect of study treatment and subject weight on PK parameters such as absorption, volume, and clearance were quantified and interpreted.

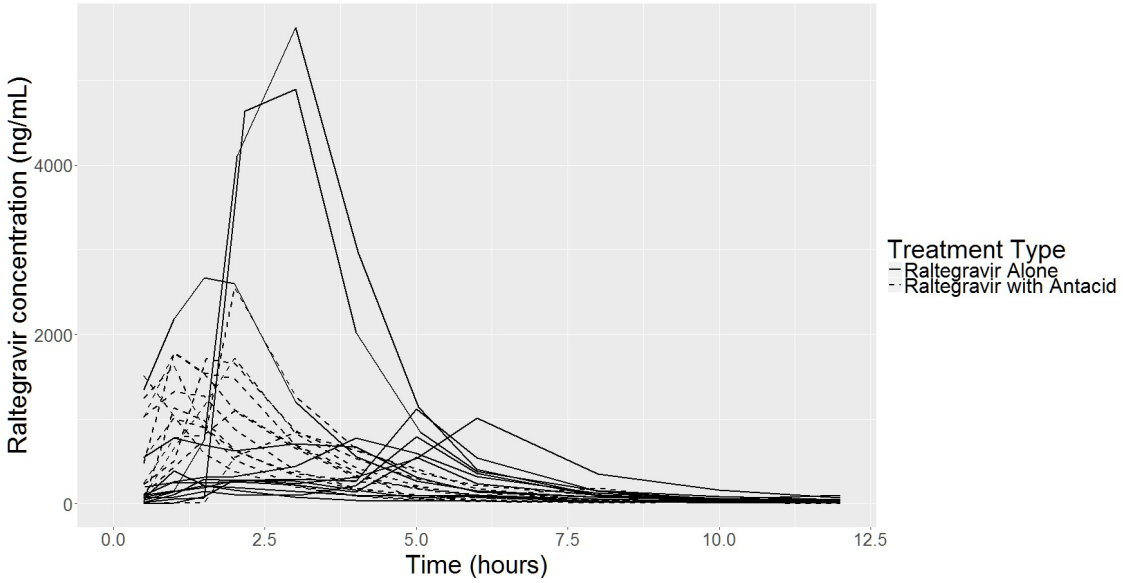

Figure 1: First 12 hours of raltegravir data. Each line represents a single treatment on a single subject. Each subject appears in the graph twice due to the crossover treatment design.

Introduction to Pharmacokinetic Models

The body can be understood as a system of multiple compartments each with varying concentrations of drug at a given time, and with different rates of absorption and elimination (referred to collectively in this paper as perfusion rates). Organs or tissues grouped within a compartment have similar perfusion rates. For example, if a drug is administered orally, the gastrointestinal tract acts as the absorption site (also called the depot). The drug then moves to the central compartment which, depending on the drug of interest, usually includes the blood, kidneys, and liver. The drug may then be absorbed from the central compartment into a peripheral compartment such as the fat, bone, and/or skin (Bourne, 2016). Drug concentration

levels may be measured most easily from the central compartment (e.g. blood plasma), but may also be measured from the peripheral compartment (e.g. cerebrospinal fluid). There are, of course, many variations of this model, including multi-compartment systems, intravenous dosing, and multiple dosing. However, this paper will focus on one- and two-compartment models with single oral dosing and measurements taken from the central compartment.

One-compartment Oral Dosing

One-compartment models are used when equilibrium is quickly achieved in the body tissues. Although all drugs initially distribute to a central compartment before moving to peripheral tissues, when the distribution is rapid enough to be considered clinically insignificant the model may be reduced to one compartment describing the entire body as a whole for simplicity (Davidian & Giltinan, 2003). For a simple graphical representation of this model, see Figure 2.

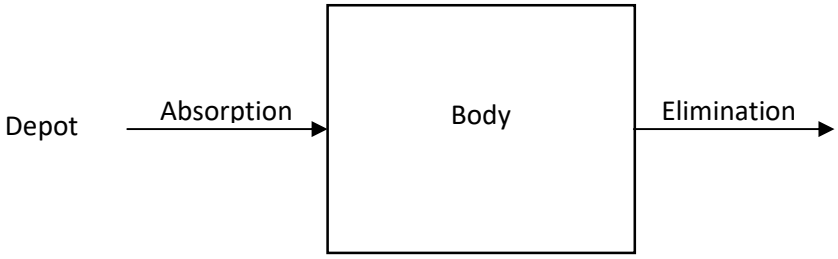

Figure 2: One-compartment oral dosing model.

The one-compartment model can be expressed via a set of ordinary differential equations (Equation 1) whose parameters are represented in Figure 3 in the one-compartment oral dosing model.

$$\frac{d\boldsymbol{a_1}}{dt} = -K_a * \boldsymbol{a_1} \tag{1}$$

$$\frac{d\boldsymbol{a_2}}{dt} = K_a * \boldsymbol{a_1} - \frac{Cl}{V} * \boldsymbol{a_2}$$

Equation 1: Ordinary differential equations for a one-compartment model with a single oral dose.

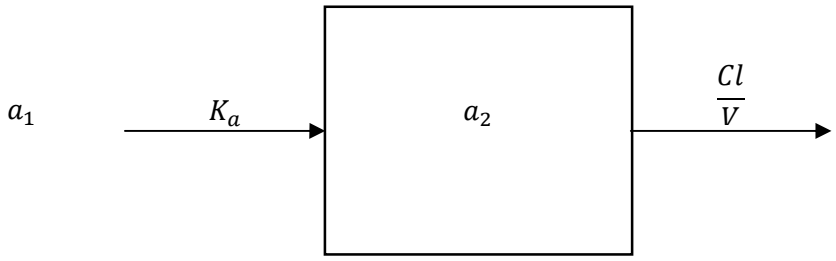

Figure 3: One-compartment oral dosing model utilizing PK parameters.

Each equation represents the change in concentration at the absorption site or in the body. The instantaneous change in drug concentration for each time point is given by the vector $\frac{d\boldsymbol{a_1}}{dt}$ for the absorption site and $\frac{d\boldsymbol{a_2}}{dt}$ for the body. The flow into and out of the absorption site and body is represented by the positive or negative signs, respectively, on each term of Equation 1. $\boldsymbol{a_1}$ is a vector of the amount at the absorption site, for example the gut, at each time. For a single dose, $\boldsymbol{a_1}$ at time 0 is the apparent amount of dose given, and the value of $\boldsymbol{a_1}$ decreases over time as the dose is absorbed into the body. $\boldsymbol{a_2}$ is a vector of the amount in the body at each time. K_a, V, and Cl are, respectively, the absorption constant, apparent volume of distribution, and apparent rate of drug clearance from the body. V and Cl are not estimable in oral dosing studies unless additional studies with IV drug administration have been performed, thus V and Cl are only apparent values, and are more accurately represented as V/F and Cl/F where F is bioavailability (Hartford, n.d.). For ease of estimation, V and Cl may be re-parameterized as *Elimination* $= K_e = \frac{Cl}{V}$.

The ordinary differential equations for a one-compartment model can be solved as in Equation 2 (Davidian & Giltinan, 2003, pp. 388, 389). Equation parameters represent the same biological processes and compartments as in Equation 1, with the addition of $C(\boldsymbol{t})$ which is the

vector of measured outcome concentration at times $\boldsymbol{t}$ for a single subject, and $Dose$ which is the amount of drug given at the administration site at time 0 assuming a single dose.

$$C(\boldsymbol{t}) = \frac{K_a * Dose}{V(K_a - \frac{Cl}{V})}\left(e^{-K_a * \boldsymbol{t}} - e^{-\frac{Cl}{V} * \boldsymbol{t}}\right) \quad (2)$$

Equation 2: Solved equation for a one-compartment model with a single oral dose.

Two-Compartment Oral Dosing

When the perfusion rates are slow, the central compartment may contain a drug concentration different enough from the peripheral compartment to be considered clinically relevant. In this case a two-compartment model may be desired. Such a two-compartment model is represented visually in Figure 4.

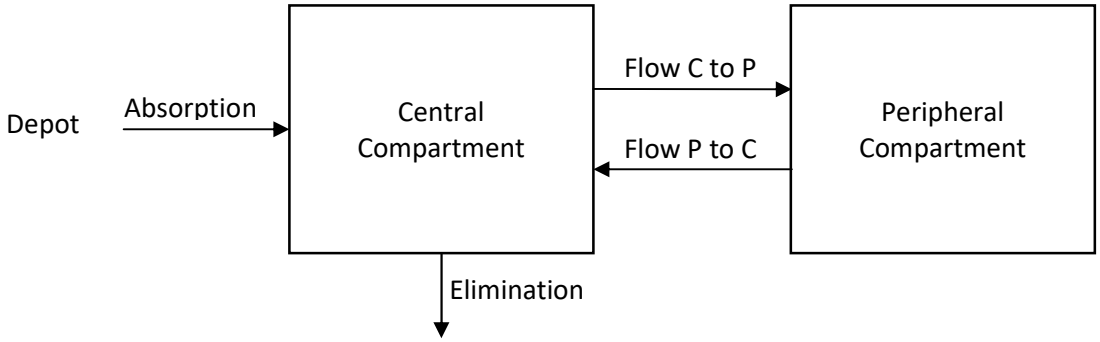

Figure 4: Two-compartment oral dosing model.

Solved equations for a two-compartment model are more complex (see Bourne, 2016); therefore we only discuss the applicable set of differential equations, as given in Equation 3. For a graphical representation of the parameters, see Figure 5.

$$\frac{d\boldsymbol{a_1}}{dt} = -K_a * \boldsymbol{a_1} \quad (3)$$

$$\frac{d\boldsymbol{a_2}}{dt} = K_a * \boldsymbol{a_1} - \frac{Cl}{V_c} * \boldsymbol{a_2} + \frac{Q}{V_p} * \boldsymbol{a_3} - \frac{Q}{V_c} * \boldsymbol{a_2}$$

$$\frac{d\boldsymbol{a_3}}{dt} = -\frac{Q}{V_p} * \boldsymbol{a_3} + \frac{Q}{V_c} * \boldsymbol{a_2}$$

Equation 3: Ordinary differential equations for a two-compartment model with a single oral dose.

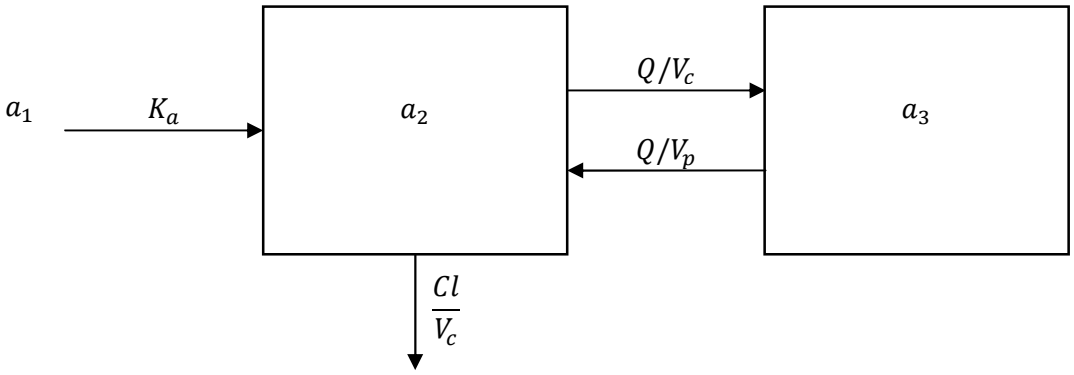

Figure 5: Two-compartment oral dosing model utilizing PK parameters.

In Equation 3, $\boldsymbol{a_1}$ is the vector of drug amount at the absorption site and $\boldsymbol{a_2}$ and $\boldsymbol{a_3}$ are the vectors of amounts in the central and peripheral compartments, respectively. K_a is the rate of absorption, Q is the rate of inter-compartmental drug transfer, and Cl is the rate of clearance from the central compartment. V_c and V_p correspond to the volume in the central and peripheral compartments, respectively. Similarly to the one-compartment model, each $\frac{d\boldsymbol{a}}{dt}$ is a vector which represents the instantaneous rate of change over time in the amount of drug in a given compartment.

Introduction to Mixed Effects Models

When to use mixed effects models

In dose concentration studies, multiple measurements per person are taken over time to determine changes in drug concentration levels. However, due to biological relevance, ease of interpretation, or model simplicity sometimes summary endpoints such as area under the concentration time curve, maximum concentration, or steady-state concentration are used instead of individual dose concentration measurements. While these summary measure analyses, which are model-independent, provide interpretable models that require relatively few assumptions (Gabrielsson & Weiner, 2012), sometimes the underlying mechanisms need to be understood in greater detail; a simple model may not adequately describe the PK

mechanism, or there may not be enough measurements per individual to calculate the summary measure (Dubois, Gsteiger, Pigeolet, & Mentré, 2010).

Linear and nonlinear mixed effects models, also known as mixed models or hierarchical models, allow the simultaneous analysis of all measurements taken instead of using summary values. When multiple measurements are taken per person or group, there is an inherent correlation which must be taken into account. If an individual has a high drug concentration at time 1, they are also likely to have a high drug concentration at time 2. A grouping unit may be an individual subject with multiple measurements, or it may be measurements taken by a certain machine, or multiple rats within a genetic strain. In other words, a group is a set of observations that share a characteristic or categorical variable level (Gelman, 2005). For purposes of this paper, we will consider a grouping unit to be an individual subject.

Some one- or two-compartment PK models require the use of nonlinear mixed models, especially when multiple intravenous or oral doses are present. The PK parameters of interest (e.g. K_a, V, and Cl) can be estimated for populations or sub-populations (categorized based on covariates such as age or concomitant medications), and can be used in multiple-dosing situations to tailor a dosing schedule which preserves consistent drug concentrations (Davidian & Giltinan, 2003, p. 389). Thus nonlinear mixed models are particularly well-suited to dose concentration, dose-response, toxicokinetic, and viral load profile studies (Davidian & Giltinan, 2003, pp. 389, 391) when the pharmacokinetic parameters are of interest instead of an individual's dose concentration curve. Although this paper will focus on dose concentration studies with a single oral dose, the implementation in NONMEM could also be adapted to the other applications and study designs mentioned.

Linear mixed effects models

Equation and Background

The notation for mixed models used in this paper is adapted from Davidian and Giltinan (2003); for simplicity and consistency, this notation is used for both linear and nonlinear mixed models, though the article referenced focuses on nonlinear models. If a translation to some terms used in NONMEM/PDx-POP documentation is desired, see Table 1 (Chapter II). For each individual, linear mixed effects models are of the form shown in Equation 4.

$$\boldsymbol{y}_i = f(\boldsymbol{t}_i, \boldsymbol{u}_i, \boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i) + \boldsymbol{\varepsilon}_i \quad (4)$$

Equation 4: General parameterization of a mixed effects model.

In this equation, $\boldsymbol{y}_i$ is the vector of the outcome (response) observations for the i^{th} individual. This vector is predicted by a function f which is constrained to be linear in the case of a linear mixed effects model and which is dependent on times $\boldsymbol{t}_i$, conditions $\boldsymbol{u}_i$ under the control of the study such as dose, $\boldsymbol{a}_i$ which includes an intercept and patient characteristics which do not change over time such as gender or weight, fixed effects coefficients $\boldsymbol{\beta}$, and random effects coefficients $\boldsymbol{b}_i$. The residual error not explained by the function is $\boldsymbol{\varepsilon}_i$. Another parameterization of Equation 4 is shown in Equation 5 where $\boldsymbol{X}_i$ is based on a combination of $\boldsymbol{u}_i$ and $\boldsymbol{a}_i$, and where $\boldsymbol{Z}_i$ is based on $\boldsymbol{a}_i$ and commonly allows for a random intercept and slope for each subject but does not necessarily have to include all the covariates from $\boldsymbol{a}_i$. In simple models, $\boldsymbol{Z}_i$ may simply include a random intercept.

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i \quad (5)$$

Equation 5: Alternate parameterization of the linear mixed effects model.

As in Equation 4, $\boldsymbol{y}_i$ in Equation 5 is the vector of the outcome observations for the i^{th} individual. This vector has dimension $m_i \times 1$, where m_i is the number of observations on individual i. In the case of dosing regimen studies, this vector could be blood serum drug

concentrations measured in mg/L. $\boldsymbol{X_i\beta}$ is called the fixed effects term. $\boldsymbol{X}_i$ records the value of each fixed effect covariate for the individual. $\boldsymbol{X}_i$ is a matrix of dimension $m_i \times p$, where p is the number of fixed effects covariates (p may include an intercept). $\boldsymbol{\beta}$ is a vector of the sample-wide coefficients for each fixed effect covariate's effect on the outcome (i.e. these coefficients do not vary between individuals). $\boldsymbol{\beta}$ is of dimension $p \times 1$; one coefficient for each fixed covariate (including the intercept, if appropriate). These fixed effects coefficients do not vary between subjects, although they may vary with time if time is included as an interaction term. They are the best estimators of the unknown population values of each fixed covariate's effect (Bates, 2010, p. 11). $\boldsymbol{Z_i b_i}$ is called the random effects term. $\boldsymbol{Z}_i$ is a matrix of dimension $m_i \times q$ which records the value of each of the q random effects for each individual which may include a random intercept. $\boldsymbol{b}_i$ is the $q \times 1$ vector of random effects coefficients, and is distributed $N(\mathbf{0}, \boldsymbol{G}_i)$. This matrix $\boldsymbol{G}_i$ captures the between-subject variability. The error term $\boldsymbol{\varepsilon}_i$ is a vector of dimension $m_i \times 1$, and is distributed $N(\mathbf{0}, \boldsymbol{V}_i)$. This matrix $\boldsymbol{V}_i$ captures the within-subject variability. The error term can also be understood as residual variability in the outcome of each individual which is not modeled through the other coefficients. To recap mixed models are defined with dimensions as in Equation 6.

$$\underset{(m_i\times 1)}{\boldsymbol{y}_i} = \underset{(m_i\times p)}{\boldsymbol{X}_i}\ \underset{(p\times 1)}{\boldsymbol{\beta}} + \underset{(m_i\times q)}{\boldsymbol{Z}_i}\ \underset{(q\times 1)}{\boldsymbol{b}_i} + \underset{(m_i\times 1)}{\boldsymbol{\varepsilon}_i} \tag{6}$$

Equation 6: Alternate parameterization of the linear mixed effects model with dimensions.

Each covariate coefficient can be understood as the mean effect of the covariate over all individuals in the sample provided all other covariates remain constant. When the error term and random effects are included, the equation precisely describes the responses for each individual. When the error term is omitted, the predicted response of subject i is described, as seen in Equation 7 (Bates, lme4: Mixed-effects Modeling with R, 2010, p. 12).

$$E[\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{Z}_i] = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i \quad (7)$$

Equation 7: Predicted response vector of subject i, given fixed and random effects.

When the random effects are also excluded, the predicted mean response of the sample given a set of fixed effects values is represented as follows in Equation 8 (McLean, Sanders, & W., 1991, p. 55).

$$E[\boldsymbol{y}_i|\boldsymbol{X}_i] = \boldsymbol{X}_i\boldsymbol{\beta} \quad (8)$$

Equation 8: Predicted mean response vector of the sample, given fixed effects.

Thus, in a linear model the average of all modeled individual curves is the predicted population curve for a given set of fixed effects values. Note that in nonlinear models, discussed later, this is not the case (Davidian & Giltinan, 2003, p. 402).

Fixed and Random Effects

In this paper, we will consider those parameters to be fixed effects whose effect is not allowed to vary between subjects, while those whose effects do vary between subject will be considered random effects (Gelman, 2005, p. 21). Generally, fixed effects are covariates in which a single constant effect of the covariate is assumed and estimated over the whole sample. Fixed effects should at least include the primary explanatory variable, and may also include any number of covariates that we wish to adjust for such as gender, or weight at baseline. In contrast, random effect coefficient estimates are allowed to vary by individual, providing for individual dose concentration curves. Random effects may include some or all of the covariates of interest. The variances of the random effect coefficient estimates capture between-subject variability in the outcome (van de Pol & Wright, 2009, p. 754).

There are several mathematical considerations to keep in mind when deciding which covariates should have a random effect. Random effects allow covariance, and therefore correlation, between observations (SAS Institute Inc, 2009, p. 36). It is this property that renders

mixed models useful when repeated measures are present. Those covariates with a high degree of variability between subjects may be modeled as random effects. In addition, in contrast to fixed effects, random effects are estimated by partial pooling; when few data points are available for a group, the random effect estimate is partially influenced by other groups with more data (Gelman & Hill, Data Analysis Using Regression and Multilevel/Hierarchical Models, 2006). The degree of influence is determined by the number of observations, the individual mean and variance, and the sample variance. More specifically, random effects assume that each individual deviates from a common mean covariate effect which is present in the sample, and that these deviations follow a distribution. Thus, when an individual has fewer observations the missing data can be inferred based on individuals with more data. For this reason, mixed models have an assumption that any missing data in a variable is missing completely at random (i.e. the pattern of missingness in a variable is random, and does not depend on the variable with the missing data or any other variable in the data set; estimates have reduced precision but no bias) or missing at random (i.e. the pattern of missingness depends on observed values of other variables in the data set, but not on the variable with the missing data or on unobserved values of other variables; estimates have reduced precision but bias can be mitigated by adjusting for the other variables on which the pattern of missingness is dependent) (Gelman & Hill, 2006, p. 530).

Nonlinear models in PK analysis

Although nonlinear relationships may be approximated with such methods as polynomials or splines (Huang, Wu, & Zhou, 2004), the resulting estimated parameters are not clinically interpretable. Nonlinear mixed models utilize the ODEs from the PK equations in Equations 1-3 in order to estimate meaningful parameters. Models with multiple dosing,

categorical outcomes, or Michaelis-Menten elimination may especially benefit from nonlinear analysis (Beal & Sheiner, 1998, p. 10).

As with linear mixed models, nonlinear mixed models are of the form demonstrated in Equation 12 (Davidian & Giltinan, 2003).

$$\boldsymbol{y}_i = f(\boldsymbol{t}_i, \boldsymbol{u}_i, \boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{b}_i) + \boldsymbol{\varepsilon}_i \quad (12)$$

Equation 12: General form of a nonlinear mixed model.

In this equation, $\boldsymbol{y}_i$ is the vector of responses for the i^{th} subject and is of dimension $m_i \times 1$ where m_i is the number of observations for subject i, and f is a function nonlinear in the vector of parameters $\boldsymbol{\beta}$ and $\boldsymbol{b}_i$. The times for the i^{th} subject are represented by $\boldsymbol{t}_i$ and are also of dimension $m_i \times 1$. $\boldsymbol{u}_i$ is the matrix of values of dose or other study conditions for the i^{th} subject and is of dimension $m_i \times r$ where r is the number of variables under the control of the study. $\boldsymbol{a}_i$ is the matrix of values of patient characteristics for the i^{th} subject (e.g. demographics, genotype, blood pressure) and is of dimension $m_i \times \mathrm{s}$ where s is the number of patient characteristic variables. The fixed effects coefficients are designated $\boldsymbol{\beta}$ and are of dimension $p \times 1$ where p is the number of fixed effects variables and may include an intercept; generally, p is the sum of r and s. The random effects coefficients $\boldsymbol{b}_i$ are of dimension $q \times 1$ where q is the number of random effects, generally a subset of $\boldsymbol{a}_i$ and possibly an intercept, and may be distributed $N(\boldsymbol{0}, \boldsymbol{G}_i)$.The residual error in the model, $\boldsymbol{\varepsilon}_i$, is of dimension $m_i \times 1$ and distributed $N(\boldsymbol{0}, \boldsymbol{V}_i)$.

In the case of a one-compartment oral dose drug concentration study, Equation 12 is equivalent to Equation 2 where $\boldsymbol{y}_i = \mathrm{C}(\boldsymbol{t})$. Furthermore, each of the parameters of interest (in this case K_a, V, and Cl) is calculated as a combination of fixed effects (study conditions and patient covariates) and random effects. The selection process for these covariates and random effects will be discussed later. An example follows in Equation 13 of possible parameter

specifications with covariates weight and gender, as adapted from Davidian and Giltinan (2003, p. 393).

$$K_{a_i} = \beta_1 + b_{1i} \tag{13}$$

$$V_i = \beta_2 + \beta_3 * weight_i + b_{2i}$$

$$Cl_i = \beta_4 + \beta_5 * weight_i + \beta_6 * gender_i + b_{3i}$$

Equation 13: Example calculations of PK parameters with covariates and random effects.

These may be also exponentiated as in the form $K_{a_i} = exp(\beta_1 + b_{1i})$ in order to restrict the parameters to positive numbers.

Mathematically, non-linear mixed effects models simply differ from linear mixed effects models in that the function of coefficients f is nonlinear. Practically, however, because these models are more flexible, it is more difficult to estimate the coefficients. Since the integrals used often are intractable and do not have closed form solutions, numerical approximations must be used. These models tend to present convergence issues such as non-positive definite Hessian matrices (the matrices of second partial derivatives of the log-likelihood which are used for approximations or maximizations of multivariable functions) (The Hessian, 2016).

Since a large parameter space is searched and the models are complicated, starting values must also be provided for covariates. Selection of starting values can be somewhat of an art; they may be obtained from previous studies, simpler models, or graphical analysis, but often multiple combinations of values must be attempted. Convergence can be sensitive to these starting values and parameter estimates may not be robust (Davidian & Giltinan, 2003).

Fixed effects on the PK parameters may include the primary explanatory variable and any covariates of interest. Random effects may be included on some or all of the PK parameters. If the variation not explained by a fixed effect is relatively small for a certain PK parameter, i.e. if the standardized variance (percent relative standard error or %RSE) of the random effect is

small, the random effect may be assumed to be zero, and dropped for model stability (Davidian & Giltinan, 2003, p. 394)

In nonlinear mixed models the average of individual curves is not the population curve; therefore the fixed effects estimates for the parameters cannot be reported without the patient characteristic covariate values. Generally, nonlinear models are reported using typical values (estimated population medians) or the predicted response given a range of values of each covariate (Owen & Fiedler-Kelly, 2014, pp. 125, 128-130).

Model Building

While there are several recommended steps before model development, such as data quality control, and exploratory data analysis, for brevity we simply recommend referring to Owen & Fiedler-Kelly (2014), Petricoul et al. (2001), and Mandema, Verotta, and Sheiner (1992). Due to the complexity of nonlinear mixed models and corresponding difficulties with convergence, it is generally recommended to begin model selection with a simple base model, often although not always without additional covariates (Owen & Fiedler-Kelly, 2014, p. 116). The first step is to choose the compartment structure (e.g. one- versus two-compartment), using such metrics as change in objective function value (OFV; used if the models are hierarchical or nested, i.e. if the same parameters are included in both models with additional parameters in the larger model and no additional parameters in the smaller model; lower values preferred), change in Akaike Information Criterion (AIC; used if the models are not hierarchical; lower values preferred), observed versus predicted outcome measurement plots, or weighted residuals versus predicted outcome plots (Kile, et al., 2012, p. 1228). Starting value estimates for the PK parameters may be obtained from previous publications, from graphical plots (Jambhekar & Breen, 2012), or by iterating on simpler statistical models (Davidian & Giltinan, 2003, p. 408).

The difference between the OFV of two nested models approximately follows a chi-square distribution with υ degrees of freedom, where υ is the difference in number of parameters between the larger and smaller models (Owen & Fiedler-Kelly, 2014, p. 142). For example, if Model A has 9 parameters and an OFV of 3270.871, and Model B has 8 parameters and an OFV of 3275.084, at an alpha level of 0.05 a significant OFV difference would be 3.841, as given by Equation 14.

$$\chi^2_{1,\alpha=0.05} = 3.841 \tag{14}$$

Equation 14: Critical value of chi-squared distribution with one degree of freedom and alpha = 0.05. Reject if calculated chi-square statistic is greater than the critical value.

Since 4.213 is larger than 3.841, the OFV difference is considered significant; the p-value is 0.0401, and Model A should be chosen because the missing parameter in Model B significantly decreases model fit. Similarly, if two non-nested models are compared with AIC, a difference in AIC greater than 2 would be considered relevant or meaningful (Cameron & Trivedi, 2005, p. 279). Models within two AIC points of each other are considered comparable. If there is no significant difference between models, the simpler model is recommended for ease of convergence in later steps.

Random effects may be placed on each PK parameter and removed one-by-one if the model will support that degree of complexity (i.e. converge), or optionally, models with a random effect on only one or a few PK parameters may be compared and subsequently random effects may be systematically added and evaluated. The method chosen may depend on the frequency of measurements on each individual. If most measurements are obtained while drug concentration is decreasing, Cl may be the only parameter that can support a random effect. Conversely, if many measurements are obtained during absorption and near peak, K_a and V may also support a random effect. Sometimes the model may not fully converge, and may

estimate a random effect near 0. In this case, the residual error model should be examined. The appropriate residual error model (additive, proportional, or additive and proportional; these are discussed later in this paper) may facilitate model convergence even with the problematic random effect. Various starting values for the random effects and the residual error may also be considered. If the model still does not converge, it should not be inferred that the random effect does not exist; merely that the dataset is unable to support the random effect (Owen & Fiedler-Kelly, 2014, p. 130), that the starting values used were incorrect, or that the model has been misspecified in some other way.

Covariate Structures

Once the base model has been chosen, we recommend assessing potential covariates one at a time, and determining which pharmacokinetic parameter estimates, if any, should include these covariates through similar OFV comparison (Kile, et al., 2012, p. 1228). Covariates may be assessed for any number of reasons; because the covariate is a primary explanatory variable (e.g. treatment), because there is a biological rationale for inclusion (e.g. it is known that kidney function has an effect on clearance rate), or to assess a group effect (e.g. to determine whether volume varies significantly between males and females) (Mould & Upton, Basic Concepts in Population Modeling, Simulation, and Model-Based Drug Development, 2012). For more information on covariate selection see Jonsson and Karlsson (1998).

There are multiple ways covariates may be incorporated into the model. A dichotomous (binary) covariate may be added linearly, as in Equation 15, or proportionally, as in Equation 16. The interpretations for each structure vary slightly. If a dichotomous covariate is added linearly, the associated β represents the unit difference in the PK parameter (in this case Cl) associated with the non-reference group. In the case of a proportional dichotomous covariate, the associated β represents the proportional change in the PK parameter associated with the non-

reference group. Categorical covariates may be added in much the same way, with an additional term (β and indicator variable) for each non-reference level of the covariate (Owen & Fiedler-Kelly, 2014, pp. 148-150). Starting values for dichotomous or categorical covariates may be obtained from previous publication, non-compartmental analysis, or simpler models.

$$Cl = \beta_1 + \beta_2 * gender \quad (15)$$

Equation 15: Linear structure for a dichotomous covariate.

$$Cl = \beta_1 * (1 + \beta_2 * gender) \quad (16)$$

Equation 16: Proportional structure for a dichotomous covariate.

Continuous covariates may also be included in several ways. Linear, power, and exponential continuous covariate structures are shown in Equations 17, 18, and 19, respectively. Both power and exponential structures allow for greater flexibility in the covariate-parameter relationship versus a linear structure, and may be linearized through a natural log transformation (see Equations 20 and 21). βs associated with the covariate in the linear structure may be interpreted as an estimate of the change in the PK parameter (in this case Cl) for each unit change in the covariate (in this case $weight$). The β associated with the covariate in a power model represents the change in the natural log (ln) of the PK parameter for each unit change in the natural log of the covariate. The β associated with the covariate in an exponential model represents the change in the natural log of the PK parameter for each unit change of the covariate.

$$Cl = \beta_1 + \beta_2 * weight \quad (17)$$

Equation 17: Linear structure for a continuous covariate.

$$Cl = \beta_1 * weight^{\beta_2} \quad (18)$$

Equation 18: Power structure for a continuous covariate.

$$Cl = \beta_1 * e^{\beta_2 * weight} \quad (19)$$

Equation 19: Exponential structure for a continuous covariate.

$$\ln(Cl) = \ln(\beta_1) + \beta_2 * \ln(weight) \quad (20)$$

Equation 20: Linear representation of the power structure for a continuous covariate.

$$\ln(Cl) = \ln(\beta_1) + \beta_2 * weight \quad (21)$$

Equation 21: Linear representation of the exponential structure for a continuous covariate.

Starting values for continuous covariates may be obtained from previous literature, non-compartmental analysis, simpler models, or by graphical analysis. Starting estimates for a linear structure may be obtained by a scatterplot of individual estimates of the PK parameter of interest from a base model (y-axis) versus the covariate (x-axis), as in Figure 6 (Owen & Fiedler-Kelly, 2014, p. 151). By fitting a regression line, starting values for β_1 may be obtained by estimating the y-axis intercept, and starting values for β_2 may be obtained by estimating the slope. Similarly, starting values for a power model may be obtained from the exponentiated intercept and slope of a scatterplot of natural logged individual estimates of the PK parameter versus the natural logged covariate as in Figure 7 (Owen & Fiedler-Kelly, 2014, p. 153). Starting values for an exponential model may be obtained from the exponentiated intercept and non-exponentiated slope of a scatterplot of natural logged individual estimates of the PK parameter versus the covariate as in Figure 8 (Owen & Fiedler-Kelly, 2014, p. 154).

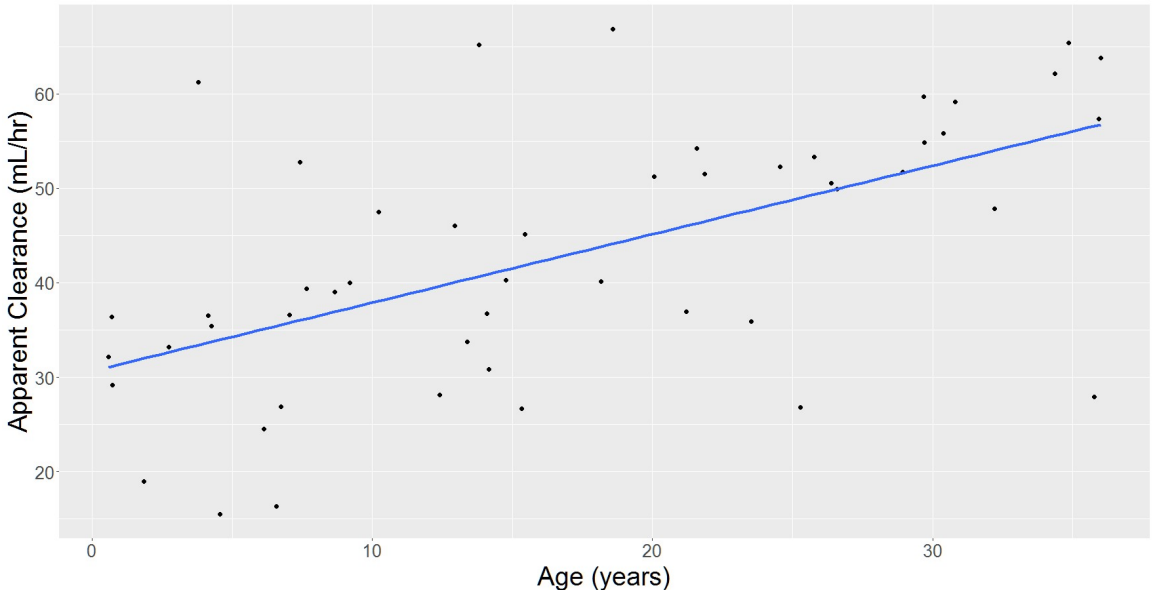

Figure 6: Estimation of fixed and random effect starting values in a linear covariate structure. The solid line represents the linear regression line. In this graph, the starting estimate for β_1 should be approximately 30 (y-intercept) and the starting estimate for β_2 (age) should be approximately 1 (slope).

Figure 7: Estimation of fixed and random effect starting values in a power covariate structure. The solid line represents the linear regression line. In this graph, the starting estimate for β_1 should be approximately $e^{3.25}$ (exponentiated y-intercept) and the starting estimate for β_2 (age) should be approximately $e^{0.2}$ (exponentiated slope).

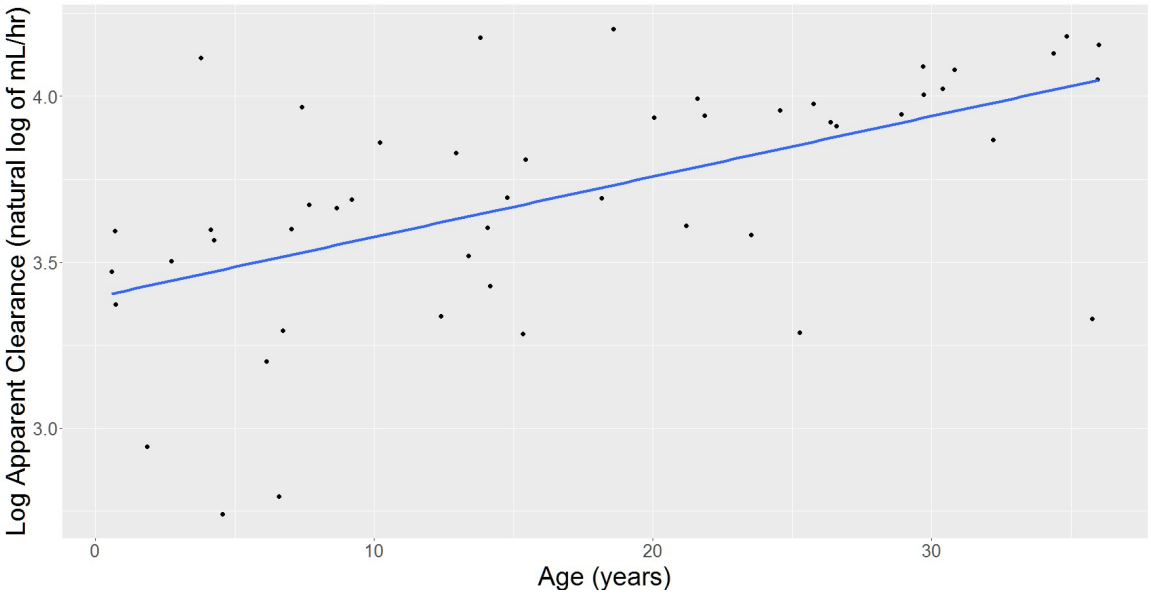

Figure 8: Estimation of fixed and random effect starting values in an exponential covariate structure. The solid line represents the linear regression line. In this graph, the starting estimate for β_1 should be approximately $e^{3.4}$ (exponentiated y-intercept) and starting estimate for β_2 (age) should be approximately 0.02 (slope).

It should be noted that neither lower nor upper bounds should be set on parameter estimates for covariates (Owen & Fiedler-Kelly, 2014, p. 149). Otherwise, for example, if a dichotomous treatment covariate is included on Cl with a lower bound of 0, the non-reference treatment group is constrained to a clearance higher than the reference treatment group. For model stability, continuous covariates may need to be median-centered, and categorical covariate groups may need to be combined if there are a large number of groups or a small number of observations per group (Kile, et al., 2012). Optionally, starting values for PK parameters in these models may be obtained from a simpler base model.

The merits of covariate inclusion or structure may be assessed via the relative reduction in unexplained between-subject variability versus a model without the covariate. This percent relative reduction may be calculated as follows in Equation 22 where Ω is the variance of the random effect of the PK parameter on which the covariate is placed. Sometimes, Ω may not be reported in model output, and so percent relative reduction can be approximated by using the

percent coefficient of variation (%CV) as given in Equation 23 (Duffull, Wright, & Winter, 2011, p. 812).

$$Percent\ relative\ reduction = \frac{\Omega\ base\ mode \quad \Omega\ full\ model}{\Omega\ base\ model} * 100 \qquad (22)$$

Equation 22: Calculation of percent relative reduction upon inclusion of a covariate. Ω represents the variance of the random effect on the PK parameter on which the covariate is placed. The model without the covariate being assessed is the base model. The model with the covariate is the full model.

$$Percent\ relative\ reduction \approx \frac{(\%CV\ base\ model/100)^2 - (\%CV\ full\ model/100)^2}{(\%CV\ base\ model/100)^2} * 100 \qquad (23)$$

Equation 23: Approximation of percent relative reduction upon inclusion of a covariate. %CV is the percent coefficient of variation. The model without the covariate being assessed is the base model. The model with the covariate is the full model.

Random Effects Structures

Random effects are commonly modeled with one of three structures: additive, proportional (also called constant coefficient of variation or CCV), or exponential. The additive structure specifies a constant variance in the random effect across values of the parameter, and is generally recommended for log-transformed PK parameters where the random effects are expected to be normally distributed (Mould & Upton, 2013, p. 5). Both proportional and exponential structures allow the variance of the random effect to increase with increasing values of the parameter. In most cases, an exponential random effects structure is preferable because it restricts parameter estimates to non-negative numbers (Owen & Fiedler-Kelly, 2014, p. 18), as is desired with PK parameters, and can accommodate right skewness (Mould & Upton, 2013, p. 5).

Residual Error Structures

Residual error is commonly modeled with one of three structures: additive, proportional (also called constant coefficient of variation or CCV), or combined additive and proportional. Combined additive and proportional error may be thought of as the combination of a fixed

measurement error (or error due to upper or lower limits of quantification in the measurement instrument) and an error proportional to the drug concentration (i.e. variance unexplained by the model may higher at higher drug concentrations) (Fisher & Shafer, 2007, pp. 11-4). The variance of either of these errors may be close enough to 0 that the model may be approximated by dropping the corresponding error, therefore using either an additive or a proportional error structure (Ette & Williams, 2007, p. 829). However, care should be taken if using only a proportional error structure, especially if the initial model has not been specified correctly (Ette & Williams, 2007, p. 829), as low drug concentrations are over-weighted because they are assigned a low residual error (Mould & Upton, 2013, pp. 8-9).

Graphical Checks

General Model Fit

One common graph used to assess model fit is a scatterplot of observed outcome measurements versus population predicted measurements. Goodness-of-fit is demonstrated by proximity and symmetry of points about the 1:1 identity line (Owen & Fiedler-Kelly, 2014, pp. 117, 118). Figure 9 gives an example of this scatterplot. In this graph the solid line represents the 1:1 identity, and the dotted line represents a lowess smooth fit[1] (ICON plc, 2016, p. 95). It should be noted that the axes may be scaled, so the 1:1 line may not appear at 45°. This graph fits reasonably well, although there are some outliers present especially in higher observed concentrations, and the lowess smooth line indicates that the model does not predict as accurately at higher concentrations.

[1] It should be noted that in the PDx-POP manual, the dotted line is defined as a "less smooth." However, upon examination of the R code used to generate this graph, a lowess smooth is utilized.

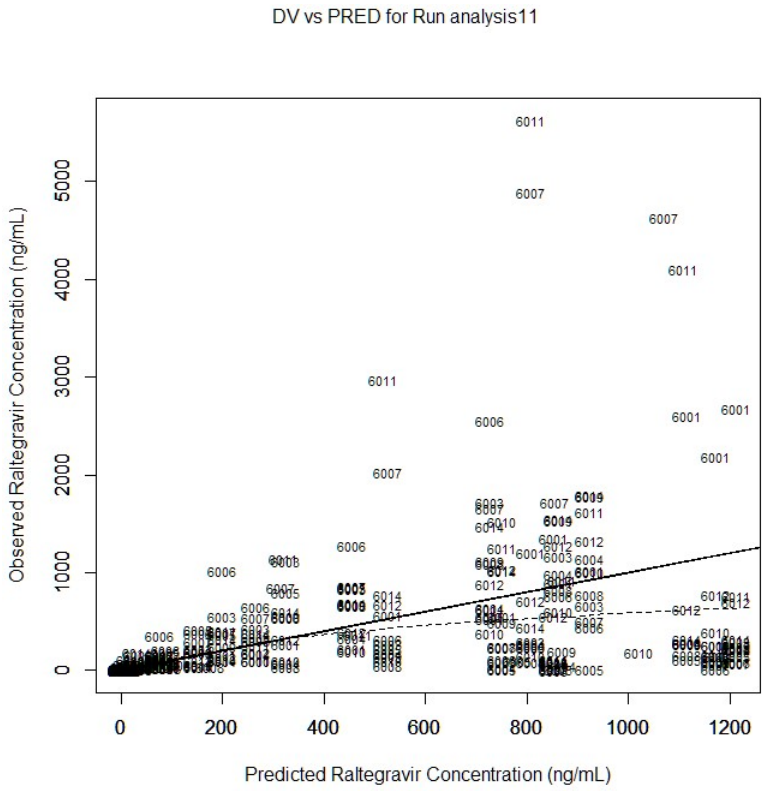

Figure 9: Scatterplot of observed outcome measurements (DV) versus predicted measurements (PRED). Solid line represents the 1:1 identity. Dotted line represents a lowess smooth fit. Each plotted number indicates a single observation for a subject, the subject ID of which is indicated by the number.

Another commonly used graph is a scatterplot of weighted residuals versus predicted measurements. Weighting standardizes the residuals so that the standard deviation among residuals is 1. For information on weighting residuals, see Owen and Fiedler-Kelly (2014, pp. 132-3, 185-6). In a well-fit model, the points should be symmetrically distributed about 0 on the y-axis. There may be a few outlier points, identified as points that lie outside of ±2 units (standard deviations) on the y-axis (Mould & Upton, 2013, pp. 12, 13); however, there should be no trend or range of predicted values in which there is a bias (Owen & Fiedler-Kelly, 2014, pp. 117, 119-121). A U-shaped plot may indicate an additional compartment is needed, and a right-skewed plot where most points fall close to or below 0 and a few points fall high above 0 may indicate a log transformation of the outcome and residual error variance is needed. Figure 10

gives an example of this scatterplot. The solid line represents a weighted residual of 0, and the dotted line represents a lowess curve. This particular plot fits relatively well, although a few positive outliers are present and there is a slight negative trend indicated by the lowess curve.

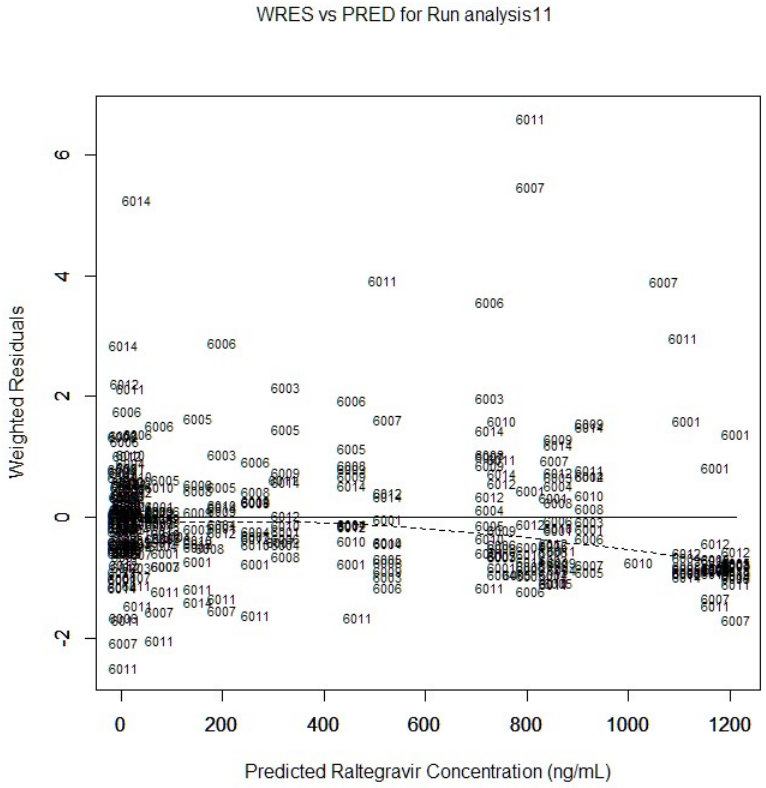

Figure 10: Scatterplot of weighted residuals (WRES) versus predicted measurements (PRED). Solid line represents weighted residual = 0. Dotted line represents a lowess smooth fit. Each plotted number indicates a single observation for a subject, the subject ID of which is indicated by the number.

Random Effects

One graphical method of assessing model fit after accounting for random effects is a plot of observed measurements versus individual predicted measurements as shown in Figure 11 (Owen & Fiedler-Kelly, 2014, p. 134). This is in contrast to the graph in Figure 9 of observed measurements versus population predicted measurements. The appropriateness and effectiveness of the random effects in the model can be determined by the proximity of points to the 1:1 identity line, the symmetry of points about it, and a lack of obvious trend away from

the identity line in the lowess smooth line. Additionally, this plot (Figure 11) should be compared to the plot of observed versus population predicted measurements (Figure 9) and should show an improvement in proximity and symmetry. It should be noted that the range of the x- and y-axes may need to be taken into account when comparing plots. In this particular example high concentrations are not modeled as well as low concentrations as evidenced by the relative distances from the 1:1 identity line, either by the fixed effects alone (Figure 9) or once random effects are accounted for (Figure 11) due to the high residual variability in larger raltegravir concentrations. However, once random effects are taken into account there is relative symmetry about the identity line, and no obvious trend in the lowess line.

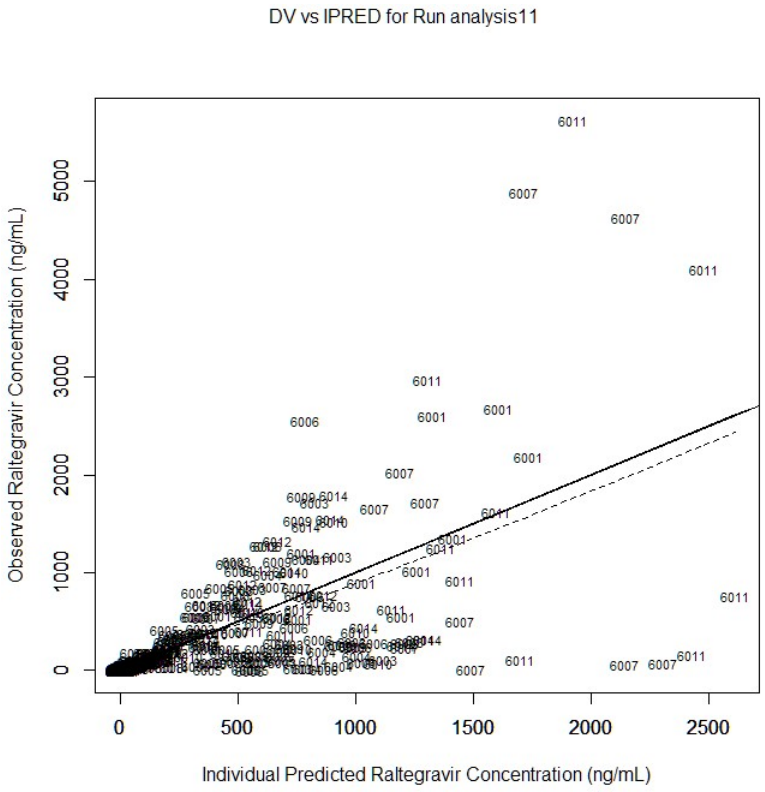

Figure 11: Scatterplot of observed outcome measurements (DV) versus individual predicted measurements (IPRED). Solid line represents the 1:1 identity. Dotted line represents a lowess smooth fit. Each plotted number indicates a single observation for a subject, the subject ID of which is indicated by the number.

Residual Error

Residual error structure fit may be evaluated using a scatterplot of residuals versus predicted measurements. In this graph, a constant spread across the predicted values may indicate an additive residual error variance model should be used. If the spread about residuals = 0 is not proportional across the predicted values (i.e. the spread is small on one end of the graph and large on the other), a proportional residual error variance may be needed. (Owen & Fiedler-Kelly, 2014, pp. 122-123). A combined additive and proportional error model may be indicated by constant error in the lower and/or upper limits of quantification (limits of measurement by the instrument or test used to measure concentration) and proportional error otherwise; however, we also recommend assessing combined error via OFV or AIC rather than solely through graphical means. Figure 12 indicates a proportional residual error is appropriate for the particular model graphed because the spread of residuals increases with predicted concentration. The solid line represents residuals of 0, and the dotted line is a lowess curve.

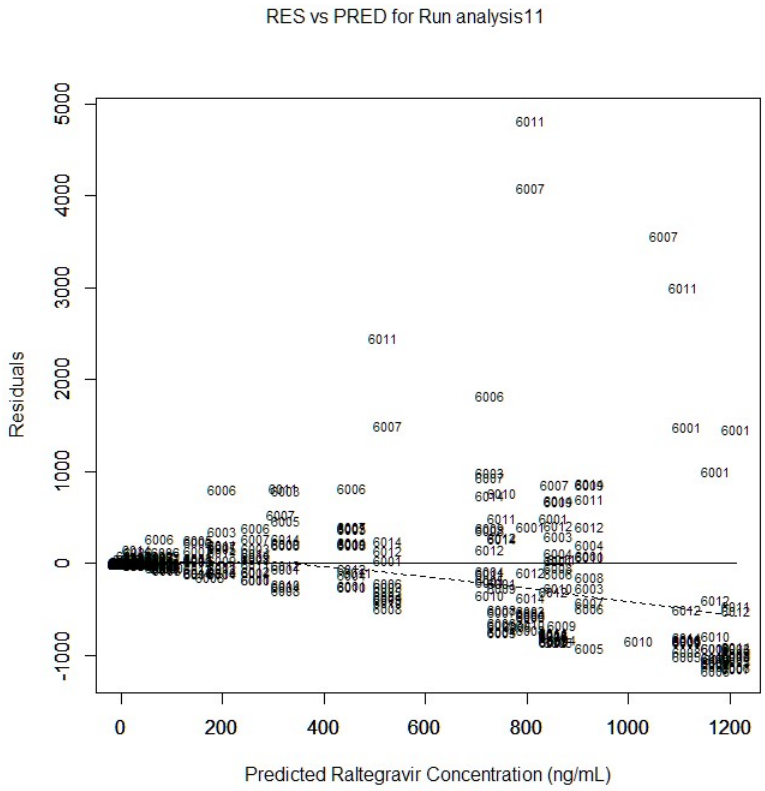

Figure 12: Scatterplot of residuals (RES) versus predicted measurements (PRED). Solid line represents residuals of zero. Dotted line represents a lowess smooth fit. Each plotted number indicates a single observation for a subject, the subject ID of which is indicated by the number.

CHAPTER II

NONMEM/PDX-POP

Background

In 1980, NONMEM was released as a commercial software for population pharmacokinetic and pharmacodynamic analysis (Goonaseelan, Mentré, & Steimer, 2005). The first iteration of the software supported first-order estimation and analysis of sparse datasets. In 1984, version II of NONMEM introduced PREDPP, a library of subroutines called ADVAN, designed to reduce the complexity of coding the models from the users' point of view (NONMEM History, n.d.). Currently, NONMEM is distributed by ICON PLC. PDx-POP is one of a number of available graphical user interfaces for NONMEM. It facilitates creation of control stream files, organizes models into projects, and creates summary statistics and plots (ICON plc, 2016). In this guide, code is indicated with `a typewriter font`, options or buttons in NONMEM/PDx-POP are indicated with an underline, and placeholder text that should be substituted with appropriate information is indicated *with italics*. Some common terms that appear in NONMEM documentation and output are given in Table 1.

Table 1: Common Notation and Terms in the NONMEM/PDx-POP Documentation or User Interface.

Term or Notation	Meaning
EV	Extravascular. Used in this context to refer to dosing methods.
IV	Intravascular. Used in this context to refer to dosing methods.
Theta θ	Fixed-effect parameter coefficient. Also referred to in this paper as β
Eta η	Random-effect parameter coefficient. Also referred to in this paper as b_i
Epsilon ε	Residual error
TV	Typical value
Control stream	A NONMEM file which contains the code necessary to execute an analysis.
Control record	A line or section of code which starts with a dollar sign ($). May be abbreviated "record" in documentation about control options.
Data record	A row in the dataset. May be abbreviated "record" in documentation about the dataset.
Event record	A row in the dataset with at least a time of measurement and the outcome of interest.
Dose record	A row in the dataset with at least a time of dose and amount of dose. This may be the same as the event record if the outcome is measured at the same time a dose is given.
Individual record	A group of rows in the dataset which share the same ID.

Data Format

For ease of import into NONMEM/PDx-POP, we recommend saving the dataset as a .csv (comma separated) file format. A new file folder in which to store the dataset may be desired, as multiple files are output from running a model. The dataset must be stored in a file folder such that the file path for the dataset does not contain any spaces (including the name of the dataset itself). Rows in the dataset are referred to in NONMEM as data records, and may be grouped and referred to as individual records if they share an ID. Data should be formatted as in Table 2.

Table 2: Data Format and Selected Variables from the First 10 Rows of the Raltegravir Dataset.

C		Data Desc: Raltegravir Dataset									
C	ID	TREAT	TIME	DV	AMT	sex	race	weight	height	age	EVID
	6001	1	0	.	4.00E+05	0	0	63.25	65.5	33	1
	6001	1	0.5	1349	0	0	0	63.25	65.5	33	0
	6001	1	1	2179	0	0	0	63.25	65.5	33	0
	6001	1	1.5	2670	0	0	0	63.25	65.5	33	0
	6001	1	2	2603	0	0	0	63.25	65.5	33	0
	6001	1	3	1206	0	0	0	63.25	65.5	33	0
	6001	1	4	560	0	0	0	63.25	65.5	33	0
	6001	1	5	268	0	0	0	63.25	65.5	33	0
	6001	1	6	153	0	0	0	63.25	65.5	33	0
	6001	1	8	77.3	0	0	0	63.25	65.5	33	0

The first column must be blank, except for a `C` in any row containing metadata, variable names, or data which is to be excluded from analysis. The first row must contain `Data Desc:` `x` (where `x` is a short text description of the data file). For an example, see the first row in Table 2. The second row must contain variable names. Rows should be sorted first by `ID`, and then by `TIME`. Certain variable (column) names must follow a strict format. These include `ID` for subject ID, `TIME` for dose or observation time (which, for NONMEM, should be formatted as relative time – i.e. hours or days or weeks since initial observation), `DV` for outcome (dependent variable) measurement, and `AMT` for dose amount. If any data record is not an observation time (i.e. an outcome measurement was not taken at that time), an `MDV` (the abbreviation for missing dependent variable) column should be included, where 1 indicates an observation was made at that time and 0 indicates an observation was not made. If the study is a crossover design (i.e. subjects receive more than one treatment), an `EVID` variable should be used instead of `MDV`, where 0 is an observation record, 1 is a dosing record, and 4 resets all compartments (Owen & Fiedler-Kelly, 2014, pp. 75-6). Thus, `EVID` should be 1 on all dosing times, except on the first dosing time of any new treatment in which case it should be 4 (See Appendix B). A maximum of

20 columns may be included in the dataset; however, by including a CONT (continuation) variable in the data set it is possible to include multiple records of the same event time. For more information regarding CONT, see Beal, Boeckmann, and Sheiner (1992, p. 5).

There should be no missing values; even if a missing value is indicated in the dataset (e.g. by a period), it will be replaced by a zero in the analysis and graphics output (Beal, Boeckmann, & Sheiner, 1992, p. 5).

In order to access the dataset and any associated control streams (discussed next) or output, a Project must be created in PDx-POP. This is accomplished by clicking Add New under the Projects / Data Tab. If a folder has been previously created for the dataset, Use Existing Directory should be selected. Navigate to the desired folder and click Accept Selected Directory. The name of the file folder will now be listed in the Project Name dropdown list on the upper-left side of the PDx-POP Projects / Data Tab. Each time PDx-POP is closed and re-opened, the project will need to be re-selected from the dropdown list.

Creating and Editing Control Streams

A control stream is a file which contains the lines of code which determine which model will be run and what options are selected for the model. It can be created either via a wizard in PDx-POP, or manually by typing out the code. To edit a control stream from within PDx-POP, it must be moved to the Control Streams Selected box (see Figure 13) in the Model / Run Tab, highlighted, then the Edit button pressed. It can also be edited outside PDx-POP via Windows Notepad or another text editing program. If a duplicate of the control stream is desired, the control stream should be highlighted in the Control Streams Selected box, then the Copy button pressed. The project may need to be re-selected under the Projects / Data tab in order to refresh the Control Streams Available list. Note that whichever way a duplicate is created, any

instances of the old file name in the control stream should be manually changed to the new name.

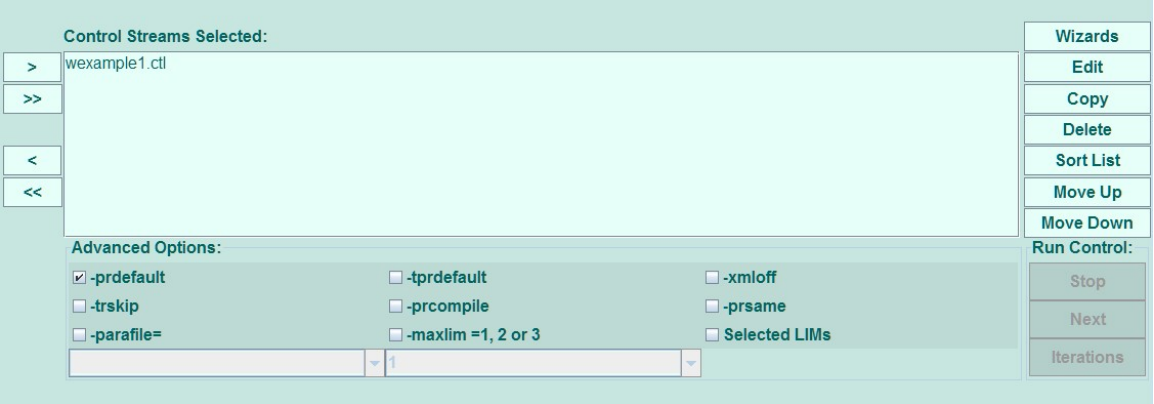

Figure 13: Control Streams Selected Box and associated control buttons in the Model/Run Tab.

The control stream has several coding conventions and formatting requirements. Each control stream section (called a control record) begins with a `$`. The name of this control record and any options within it must be in all upper-case letters. Each line of code must be less than or equal to 80 characters, although multiple lines may be utilized. A semicolon (`;`) indicates the subsequent text in the line is a comment. In general, the ordering of control records and the options within them is not restricted. An example control stream is given in Figure 14. For more information on control stream formatting and various control record options, see Boeckmann, Beal, and Sheiner (1992, pp. 11-44).

```
;Model Desc: basic two cmt additive shift trt cl random vc cl q
;Project Name: raltest
;Project ID: NO PROJECT DESCRIPTION

;Project ID: NO PROJECT DESCRIPTION

$PROB RUN# analysis14
$INPUT C ID TREAT TIME DV AMT EVID sex race weight height age HC
$DATA RAL_FOR_NONMEM.CSV IGNORE=C
$SUBROUTINES ADVAN4 TRANS4
$PK
   TVCL=THETA(1)
   CL=TVCL*EXP(ETA(1))+THETA(6)*TREAT
   TVV2=THETA(2)
   V2=TVV2*EXP(ETA(2))
   TVQ=THETA(3)
   Q=TVQ*EXP(ETA(3))
   TVV3=THETA(4)
   V3=TVV3
   TVKA=THETA(5)
   KA=TVKA
   S2=V2

$ERROR
   Y = F + F*ERR(1)
 IPRE=F
 W1= 1
 W2= F
 IRES= DV-IPRE
 IWRE=IRES/(W1+W2)

$EST METHOD=1 INTERACTION PRINT=5 MAX=9999 SIG=3    MSFO=analysis14.msf
$THETA
  (0, 104);[CL]
  (0, 102);[V2]
  (0, 14.5);[Q]
  (0, 668);[V3]
  (0, 0.481);[KA]
  5;[trt]
$OMEGA
  0.04 ;[P] omega(1,1)
  0.04 ;[P] omega(2,2)
  0.04 ;[P] omega(3,3)
$SIGMA
  0.04 ;[P] sigma(1,1)
$COV
$TABLE ID TIME IPRE ONEHEADER NOPRINT FILE=analysis14.tab
```

Figure 14: Example of a control stream.

In order to create a control stream via a wizard, click on the Wizards box under the Model / Run Tab in PDx-POP, then click on either PK Model Wizard or New Methods Wizard, depending on the version of PDx-POP used. This wizard provides the general structure for a control stream, although usually the control stream will need to be fine-tuned after creation. Most of the relevant options are located under the Model Setup Tab in the upper-left corner (displayed by default when the wizard is opened). A control file name must be entered into the

Enter Control File Name field (and optionally a model description may also be included in the Enter Model Description box), and a data file must be chosen by clicking on the Choose Data File box and navigating to the file (see Figure 15).

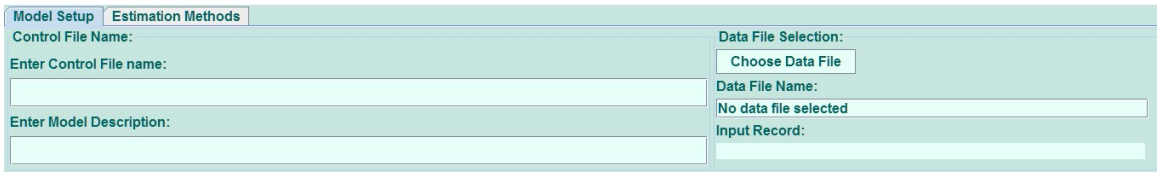

Figure 15: Control File name and Data File Selection sections of the Model Setup tab in the wizard.

The compartment structure is chosen under the Cmt Selection and Parameter Selection sections (see Figure 16). The Cmt (compartment) section influences the `$SUBROUTINES ADVAN` record listed in the control stream. The Parameter section influences both the `TRANS` option in the subroutine, which specifies the particular parameterization of the model, and the `$THETA` record, discussed later. For example, `$SUBROUTINES ADVAN1 TRANS1` contains K as the only parameter to be estimated. This is similar to `$SUBROUTINES ADVAN1 TRANS2`, where the parameters estimated are Cl and V, where Cl/V is equivalent to K. For descriptions of commonly utilized ADVAN subroutines and associated TRANS options, see Table 3. For more information on ADVAN subroutines, refer to Beal, Boeckmann, and Sheiner (1992, pp. 77-92).

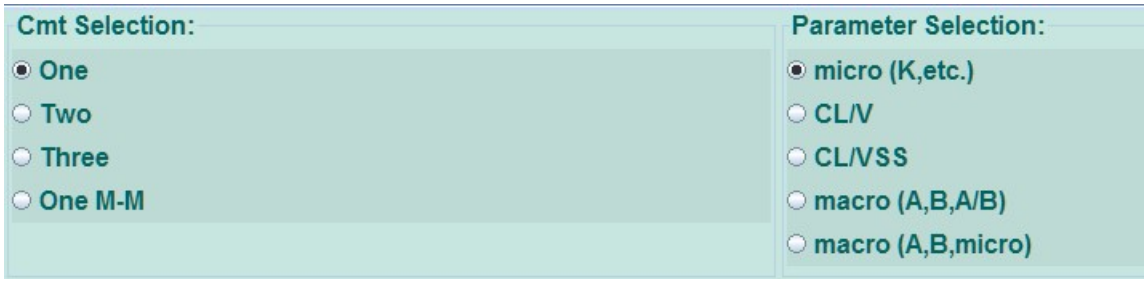

Figure 16: Cmt Selection and Parameter Selection sections of the Model Setup tab in the wizard.

Table 3: Frequently used ADVAN and TRANS subroutines.

ADVAN	Description	Main PK Parameters and Associated TRANS Options
ADVAN1	One-compartment with linear elimination. Used for intravascular dosing.	Cl and V are available via TRANS2. Alternatively, K, an elimination rate constant, may be the main parameter in TRANS1, where $K = Cl/V$.
ADVAN2	One-compartment with linear elimination and first-order absorption. Used for oral, intramuscular, or subcutaneous dosing.	K_a, Cl, and V are available via TRANS2. Alternatively, the model may be parameterized with K_a and K via TRANS1, where $K = Cl/V$.
ADVAN3	Two-compartment with linear elimination. Used for intravascular dosing.	Cl, $V1$, Q, and $V2$ are available viaTRANS4, where $V1$ and $V2$ are equivalent to V_c and V_p, respectively. Multiple parameterizations/TRANS subroutines are available.
ADVAN4	Two-compartment with linear elimination and first-order absorption. Used for oral, intramuscular, or subcutaneous dosing.	K_a, Cl, $V2$, Q, and $V3$ are available via TRANS4, where $V2$ and $V3$ are equivalent to V_c and V_p, respectively. Multiple parameterizations/TRANS subroutines are available.

The PK parameterization is chosen under the Parameter Selection section; choose Cl/V in order to replicate Equations 1, 2, or 3 (Chapter I). This section influences the Theta Estimates section in the wizard, and the `$THETA` section in the control stream. The dosing type is chosen in the Dosing Selection section (see Figure 17). The option chosen will depend on the form of drug administration utilized in the study. Intravenous (IV) should be chosen if the drug was injected into a vein. Extravascular (EV) should be chosen if the drug was administered non-intravenously (e.g. orally, intradermally, or topically) or a combination of intravenously and non-intravenously. The extravascular option appears as EV or EV & IV in the wizard and replicates the parameters in Equations 1, 2, or 3 (Chapter I). The Dosing Selection section also influences Theta Estimates and the `$THETA` record.

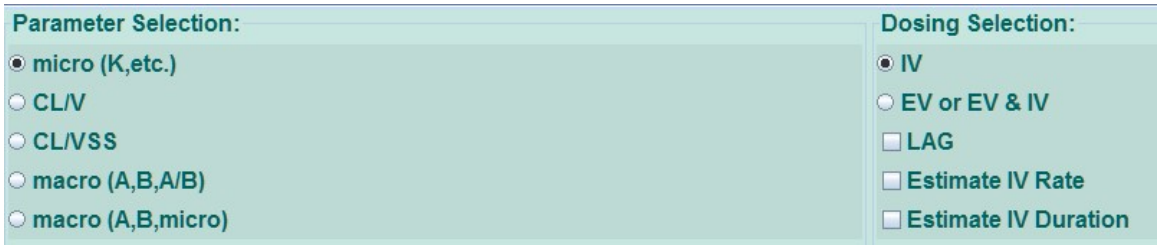

Figure 17: Parameter Selection and Dosing Selection sections of the Model Setup tab in the wizard.

The residual error structure may be chosen in the Residual Error Model section (see Figure 18); this influences the `$SIGMA` record. The estimation type is chosen in the Estimation Options section (see Figure 19) and influences the $EST record. Both the Residual Error Model and Estimation Options sections are discussed more in-depth later. The Population button should be chosen under the Modeling Options section (see Figure 20), unless the dataset contains information on only one subject, in which the Individual button should be chosen.

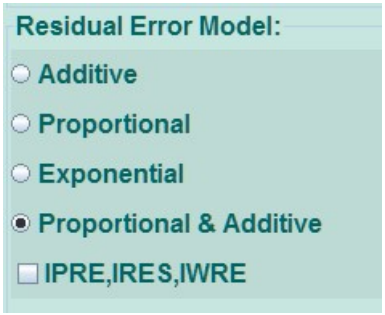

Figure 18: Residual Error Model section of the Model Setup tab in the wizard.

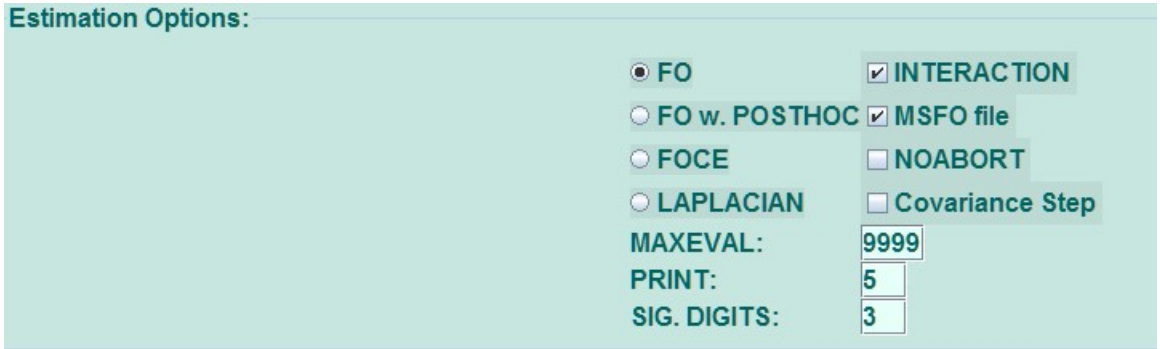

Figure 19: Estimation Options section of the Model Setup tab in the wizard.

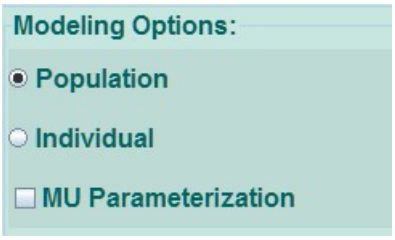

Figure 20: Modeling Options section of the Model Setup tab in the wizard.

Once the above options are selected, the starting estimates for the PK parameters should be entered under the Theta Estimates section (see Figure 21) in the Initial boxes. Lower and Upper bounds for the PK parameter search space may also be entered in this section. Double-click on parameter limits or initial estimates to change them, then click on another cell of the same table to register the change. These changes will be reflected in the `$THETA` record. If only an initial estimate is listed, the `$THETA` record will be of the form `64;[CL]` where the initial estimate is `64` and the name of the theta is `CL`. If a lower bound and initial estimate are specified, `$THETA` will be of the form `(0, 64);[CL]` where the lower bound is `0`. An upper bound may only be specified if a lower bound has also been specified. In this case, `$THETA` will be of the form `(0, 64, 100);[CL]` where the upper bound is `100`.

Theta Estimates:

Parameter	Lower	Initial	Upper
K	NB	1	NB
V	NB	1	NB

Figure 21: Theta Estimates section of the Model Setup tab in the wizard.

The Omega Estimates section (see Figure 22) gives starting values for random effects and influences the `$OMEGA` record. Similarly, the Residual Variance Estimates section (see Figure 23) gives starting values for residual error and influences the `$SIGMA` record. Both sections require double-clicking on a cell to edit, then clicking on another cell of the same table to register the change.

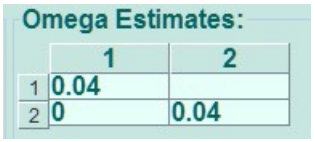

Figure 22: Omega Estimates section of the Model Setup tab in the wizard.

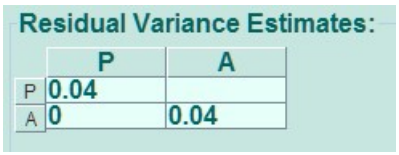

Figure 23: Residual Variance Estimates section of the Model Setup tab in the wizard.

If any parameter or random effects tables are desired in the output, they must be requested by checking the appropriate box under the Table Records section (see Figure 24), or added manually in the code. By default, the wizard does not produce %RSE and 95% CI estimates. If these are desired, the Covariance box must be checked in the General section of the Estimation Methods Tab; however, in order to register this change, if PK Model Wizard was chosen the Create Control Stream button must be clicked with the Estimation Methods tab selected and displayed. It should also be noted that the appropriate FO/FOCE/LAPLACIAN section option should be chosen or re-chosen if the Create Control Stream button is to be clicked with the Estimation Methods Tab displayed, as the currently-displaying tab over-writes options chosen in the de-selected tab. If the New Methods Wizard was chosen, the FO/FOCE/LAPLACIAN section and Create Control Stream button only exist on the Estimation Methods Tab, so this is less of a consideration. Alternatively a `$COV` record may be manually added to the control stream after creation.

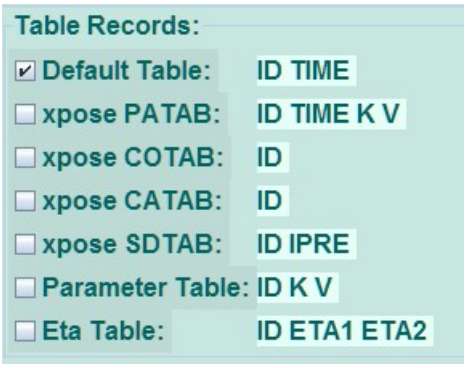

Figure 24: Table Records section of the Model Setup tab in the wizard.

PK and Theta Records

Once the control stream has been created, the structure of the fixed effect (`THETA`) estimates may be further refined. By default the control stream output by the wizard contains a `$PK` record of the form

```
TVCL=THETA(1)
CL=TVCL*EXP(ETA(1))
TVV=THETA(2)
V=TVV*EXP(ETA(2))
TVKA=THETA(3)
KA=TVKA*EXP(ETA(3))
```

where each of the PK parameters has an associated log-normally distributed random effect (`ETA`). For information on forms which this `ETA` may take, see Owen and Fiedler-Kelly (2014, pp. 41-43). Any of these random effects may be removed from the parameter; however, subsequent `ETA`s should be re-numbered so that all present `ETA`s are consecutively numbered from 1.

If covariates are desired on the PK parameters, they may be manually added to the control stream in any number of structural forms, some of which were discussed earlier in this paper. For more information on adding covariates, see Owen and Fiedler-Kelly (2014, pp. 148-159). For example, if a linear dichotomous treatment variable is desired on `CL` and a power

continuous weight variable is desired on KA, the $PK record may appear as follows. Note that the double asterisks (**) indicate an exponentiation.

```
TVCL=THETA(1)
CL=TVCL*EXP(ETA(1))+THETA(4)*TREAT
TVV=THETA(2)
V=TVV*EXP(ETA(2))
TVKA=THETA(3)
KA=TVKA*EXP(ETA(3))*(WEIGHT**THETA(5))
```

Parameter starting estimates and lower and upper bounds may be set in the $THETA record. By default, the code may appear as below, where 104 is the starting estimate and *pn* is placeholder text for the parameter name (e.g. CL).

```
$THETA
  104;[pn]
```

Lower and upper bounds may be set with code of the form (*lb*, *sv*, *ub*);[*pn*] where *lb* is the lower bound, *sv* is the starting value, and *ub* is the upper bound. A parameter estimate may be fixed (i.e. not estimated) by including code of the form *sv* FIX;[*pn*].

Omega Record

The $OMEGA record appears by default from the wizard as below, with an entry for each random effect present in the model. Letters that appear in brackets after the semicolon (;[*n*]) indicate a name for the variance, P for proportional random effect variance, A for additive random effect variance, and F for off-diagonal covariance (if it appears. This may be set in the wizard as the off-diagonal cells in the Omega Estimates table; see Figure 22) (ICON plc, 2016, p. 270).

```
$OMEGA
  0.04 ;[P] omega(1,1)
  0.04 ;[P] omega(2,2)
  0.04 ;[P] omega(3,3)
```

Error and Sigma Records

Additive residual error may be accessed by clicking on Additive in the Residual Error Model section of the wizard (see Figure 18), or via the following code.

```
$ERROR
Y = F + ERR(1)
```

Proportional residual error may be accessed by clicking on Proportional in the Residual Error Model section, or via the following code.

```
$ERROR
Y = F + F*ERR(1)
```

Combined additive and proportional residual error may be accessed by clicking on Proportional & Additive in the Residual Error Model section, or via the following code.

```
$ERROR
Y = F + F*ERR(1) + ERR(2)
```

If certain diagnostic plots are desired, additional code should be included in the `$ERROR` record to weight and calculate individual residuals. The weighting scheme can be changed by changing the definition of `W`. For information on weighting schemes, see Owen and Fiedler-Kelly (2014, pp. 132-3). An example of a `$ERROR` record with a weighting structure for combined additive and proportional error appears below.

```
$ERROR
IPRED = F
W = SQRT(THETA(X)**2 + (THETA(Y)*IPRED)**2)
IRES = DV-IPRED
IWRES = IRES/W
Y = IPRED + W*ERR(1) + ERR(2)
```

The `$SIGMA` record gives estimates for the variance of the residual errors. It is accessed via the table in the Residual Variance Estimates section of the wizard, and by default the initial variance estimates are set to 0.04 and covariance estimates set to 0. `$SIGMA` may also be modified in the code such that one entry in `$SIGMA` appears for each *n* in `ERR(`*n*`)` in the `$ERROR` record. The order of entries in the `$SIGMA` record are important, and correspond to

the ERR(*n*) entries in the $ERROR record (i.e. the first entry gives the variance estimate for ERR(1)). The code that follows illustrates the default code generated by the wizard for variance estimates in a combined additive and proportional residual error structure. In this code, letters that appear in brackets after the semicolon (;[*n*]) indicate a name for the variance, P for proportional error variance, A for additive error variance, and F for off-diagonal covariance (if it appears. This may be set in the wizard as the off-diagonal cells in the Residual Variance Estimates table; see Figure 23) (ICON plc, 2016, p. 270).

```
$SIGMA
  0.04 ;[P] sigma(1,1)
  0.04 ;[A] sigma(2,2)
```

Estimation Methods

Due to the nonlinear nature of the mixed model, and because each pharmacokinetic parameter can have associated coefficients and random effects, the calculation of a maximum likelihood is often intractable. Thus, approximations to the likelihood, also called estimation methods, must be used. There are three main estimation options available in NONMEM: first-order conditional estimation (FOCE), first-order (FO), and the Laplacian method. Each of these estimation methods minimizes the approximation to the -2 log likelihood of the model; this minimum value is called the objective function value (OFV) and the approximation is called the objective function. The difference between these methods arises from the different types of objective functions used (Boeckmann, Beal, & Sheiner, 1992, p. 36). The OFV is used to judge model fit, with smaller numbers indicating a superior fit. However, care should be taken when comparing models across estimation methods; because the OFV of each estimation method is calculated using a different objective function, OFVs are only equivalent when comparing models with the same estimation method (Beal & Sheiner, 1998, p. 11).

First-order Conditional Estimation

First-order conditional estimation (FOCE) and FOCE with interaction (FOCEI) perform well (Bonate, 2011, p. 258) because they are able to estimate both population parameters and individual random effects. Both FOCE and FOCEI are recommended versus other estimation methods when the model used is very nonlinear (as in PK modeling), when there is multiple dosing per individual, or when there is large between subject variability (Beal & Sheiner, 1998, p. 3). The advantages presented by the use of this method are less evident when few measurements are obtained per individual (i.e. when there is sparse data) (Beal & Sheiner, 1998, p. 4). FOCEI allows for heteroscedastic within subject variance and is generally recommended, especially if proportional or combined additive and proportional residual error structures are used. This is in contrast with FOCE alone which assumes homoscedastic within-subject variance. The necessity for FOCEI versus FOCE may be relaxed if a transformation such as a natural log transformation sufficiently corrects heteroscedasticity in the residuals; in this case, FOCE may be recommended due to relative ease of model convergence (Lukas & Piotrovskij, 2001).

The FOCE method is obtained via the `$EST METHOD=1` record or in the wizard via the Estimation Options section, then clicking on FOCE and un-checking INTERACTION. Similarly, FOCEI may be obtained via the `$EST METHOD=1 INTERACTION` record or by following the above instructions for FOCE, checking the INTERACTION box instead. FOCEI is the default FOCE setting in the wizard.

First-order Estimation

First-order (FO) estimation was the first estimation method developed for NONMEM (NONMEM History, n.d.). This method produces estimates of fixed effects parameters but unlike FOCE does not estimate random effects for each individual (Boeckmann, Beal, & Sheiner, 1992, p. 36), such that the random effects are assumed to be zero (Beal & Sheiner, 1998, p. 6).

Instead, individual random effect estimates can be obtained posthoc through the empirical Bayes posterior density, although these random effect estimates may be biased (Mould & Upton, 2013, p. 2). It was first developed and recommended because of the relatively small computation time required. Usage guidelines for this method indicate that it can be useful for initial model building, especially if a large number of models are tried or if the model is particularly new or complex (Beal & Sheiner, 1998, pp. 12, 13). However, even if FO estimation has been used in the model-building process, we recommend re-running the final model with FOCE or FOCEI to check for improvement in the bias of estimates or variance-covariance.

FO estimation is accessed via the `$EST` record (Boeckmann, Beal, & Sheiner, 1992, p. 36) or in the wizard under the Estimation Options section, then clicking on FO. Posthoc random effect estimates can be obtained by adding `$EST POSTHOC` to the control stream or in the wizard by clicking on FO w. POSTHOC under the Estimation Options section.

The Laplacian Method

The most general estimation method, the Laplacian method, uses unconstrained conditional estimates of the random effects (Vonesh, 1996; Wolfinger, 1993). The Laplacian method (Beal & Sheiner, 1998, p. 5) is recommended for use when the outcome is categorical or discrete ordinal (Beal & Sheiner, 1998, p. 12), which is generally not the case in PK models. The Laplacian method should not perform worse than the FOCE method; although it can be more computationally intensive. It may also be preferable if the model is very nonlinear in the random effects (Beal & Sheiner, 1998, pp. 10, 12).

This method can be accessed via the `$ESTIMATION LAPLACIAN` record (Boeckmann, Beal, & Sheiner, 1992, p. 36) or by clicking on the LAPLACIAN button under the Estimation Options section in the wizard. By default in the wizard, it is assumed there is an interaction between the individual subject random effect parameters and the error terms, i.e.

that there is heteroscedastic intra-individual variance. However, this may be disabled by unchecking the INTERACTION box in the Estimation Options section.

Output

Once the analysis has been run, the run result information will appear in the run log of the Output Tab (see Figure 25) with the file name, date and time of the analysis, name of the data file used, OFV, minimization status (yes or no), whether estimates hit a boundary (yes or no), covariance step status (yes, no, not implemented, or aborted), and model description (ICON plc, 2016, p. 156). If the model has converged successfully, an OFV estimate will appear and the MIN BND COV results will be `Y N Y`. If the model has not converged successfully, there may be no OFV estimate, and `Y N N`, `N N N`, or `NR N N` may appear. `Y N N` or `Y N NI` may also appear if the model has converged successfully but no `$COV` record was specified in the control stream.

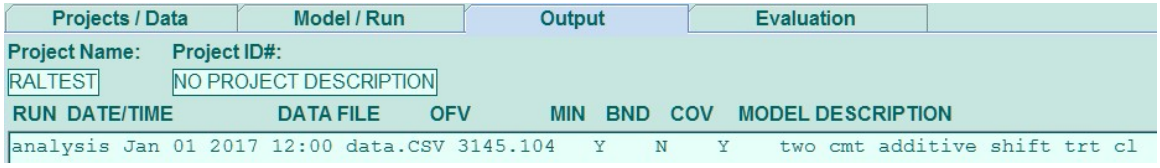

Figure 25: Run log of the Output Tab.

Two common errors may be generated if starting values are mis-specified or if the model cannot support a particular specification. In the Output Window of the Model Tab, an error may appear as follows:

```
ERROR IN TRANS ROUTINE: CL IS NEGATIVE
```

In the summary (.sum) file an error may appear as follows:

```
PARAMETER ESTIMATE IS NEAR ITS BOUNDARY
THIS MUST BE ADDRESSED BEFORE THE COVARIANCE STEP CAN BE
IMPLEMENTED
```

In these cases, at least one of the `$THETA`, `$OMEGA`, or `$SIGMA` estimates is near a boundary (either 0 or as-specified, as applicable) or is two orders of magnitude different than the starting

value. In the case of the second error, the cause is often that a random effect variance estimate is near 0 and may need to be dropped in order for the model to converge (Owen & Fiedler-Kelly, 2014, pp. 190-1).

Additional possible solutions if the model fails to converge are to try FO instead of FOCE estimation, change starting values (especially of `$OMEGA` or `$SIGMA`), change upper and/or lower bounds, reduce model complexity, reduce complexity of the residual error structure (e.g. additive instead of combined additive and proportional), or fix a parameter estimate in `$THETA`, `$OMEGA`, or `$SIGMA` (e.g. `0.5 FIXED;[KA]`) (Fisher & Shafer, 2007, pp. 20-1).

To view results from a model, select the model from the run log (see Figure 25), then select the desired results from the Output Selection section (see Figure 26). The NONMEM Results File option opens a .res file which contains detailed information about model options, specifications, iteration results, and parameter estimate results. The Output Summary option opens a relatively succinct .sum file which contains a model description and parameter estimates. Report Graphs opens S-Plus or R (The R Foundation, n.d.) and can produce graphs such as Figures 9-12 (Chapter I) for model diagnostics.

Figure 26: Output selection section of the Output Tab.

CHAPTER III

EXAMPLE OF A DATASET ANALYZED IN NONMEM

Introduction to the Dataset

The absorption of raltegravir (RAL), an antiretroviral drug used to treat human immunodeficiency virus (HIV), varies based on gastric pH (Arab-Alameddine, et al., 2012, p. 2959). Additionally RAL may compete with divalent metals present in the body due to its mechanism of divalent metal ion chelation in the cell (Kiser, et al., 2010, p. 4999), thereby decreasing bioavailability of the drug. Some antacids contain divalent metals such as magnesium or aluminum, and these may be consumed by HIV-positive patients. In order to better characterize the effect of metals delivered via antacid on PK parameters for RAL, 12 healthy (HIV sero-negative) subjects were measured 13 times over 48 hours (at 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10, 12, 24, and 48 hours) on two treatments: a single RAL dose of 400mg, and a single RAL dose of 400 mg plus a single dose of 30 ml antacid which contained 2.7 g magnesium hydroxide and 3 g aluminum hydroxide dried gel. The order of treatments was randomized. Treatments are coded in the dataset as raltegravir alone = 0 and raltegravir with antacid = 1. Subjects fasted 8 hours before and 4 hours after each dose, and blood samples were drawn at regular intervals following drug administration. Since each subject received both treatments, a washout period (5-12 days) preceded the second treatment (Kiser, et al., 2010, p. 5000). Additional covariates such as sex (female = 0), race (White = 0 or Hispanic = 1), weight (kg), height (in), and age (yrs) were also collected. Selected variables from the first 10 rows of the data can be seen in Table 1 (Chapter II), a graph of the first 12 hours of data is shown in Figure 1 (Chapter I), and the dataset used is reproduced in its entirety in Appendix B. Output generated in the analysis of this dataset is reproduced in Appendix C. It should be noted that analysis of this dataset is not intended to

corroborate, challenge, or duplicate any past or future publications utilizing this dataset or to present any clinically actionable outcomes, and is for instructional purposes only.

Model Building

The general outline of the analysis plan for this dataset is as follows. First, one- and two-compartment models were built and it was determined which PK parameters should include a random effect. Then for illustrative purposes a dichotomous variable, treatment (TREAT), and a continuous variable, weight (WEIGHT) were analyzed. Once one- vs two-compartment models were compared and selected, linear and exponential TREAT structures were tested on various parameters. Once a model with TREAT was selected, linear and power WEIGHT was included on various parameters. Then the selected TREAT model was compared with the TREAT and WEIGHT models to obtain a final model selection. For simplicity, no interactions between TREAT and WEIGHT were considered. For an overview of models analyzed see Table 4, and for an in-depth table of results from each model see Appendix A.

Table 4: Summary of models run.

Model #	Description	Converged	OFV (AIC)
Choosing Random Effects for a One-compartment Model			
1	Additive and proportional error. $Cl = \theta_1 * exp(\eta_1)$ $V = \theta_2 * exp(\eta_2)$ $K_a = \theta_3 * exp(\eta_3)$	Yes	3380.097 (3396.1)
2	Additive and proportional error. $Cl = \theta_1$ $V = \theta_2 * exp(\eta_1)$ $K_a = \theta_3 * exp(\eta_2)$	Yes	3510.184 (3524.18)
Choosing Random Effects for a Two-compartment Model			
3	Additive and proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4 * exp(\eta_4)$ $Q = \theta_5 * exp(\eta_5)$	No	3270.871 (3294.87)
4	Additive and proportional error. $Cl = \theta_1$ $K_a = \theta_2 * exp(\eta_1)$ $V_c = \theta_3 * exp(\eta_2)$ $V_p = \theta_4 * exp(\eta_3)$ $Q = \theta_5 * exp(\eta_4)$	No	3402.856 (3424.86)

Table 4: Summary of models run cont'd.

Model #	Description	Converged	OFV (AIC)
Choosing Random Effects for a Two-compartment Model cont'd			
5	Additive and proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2$ $V_c = \theta_3 * exp(\eta_2)$ $V_p = \theta_4 * exp(\eta_3)$ $Q = \theta_5 * exp(\eta_4)$	No	3275.084 (3297.08)
6	Additive and proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3$ $V_p = \theta_4 * exp(\eta_3)$ $Q = \theta_5 * exp(\eta_4)$	No	3280.613 (3302.61)
7	Additive and proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	3270.871 (3292.87)
8	Additive and proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4 * exp(\eta_4)$ $Q = \theta_5$	No	3290.593 (3312.59)
9	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	Yes	3270.871 (3290.87)
10	Proportional error. $Cl = \theta_1$ $K_a = \theta_2 * exp(\eta_1)$ $V_c = \theta_3 * exp(\eta_2)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_3)$	Yes	3424.624 (3442.62)
Choosing Treatment Structures for a Two-Compartment Model			
11	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	Yes	3145.104 (3167.10)
12	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2) + \theta_6 * TREAT$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	Yes	3171.470 (3193.47)

Table 4: Summary of models run cont'd.

Model #	Description	Converged	OFV (AIC)
Choosing Treatment Structures for a Two-Compartment Model cont'd			
13	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3) + \theta_6 * TREAT$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	.
14	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4 + \theta_6 * TREAT$ $Q = \theta_5 * exp(\eta_4)$	Yes	3214.600 (3236.60)
15	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4) + \theta_6 * TREAT$	Yes	3260.697 (3282.70)
16	Proportional error. $Cl = \theta_1 * exp(\eta_1) * \exp(\theta_6 * TREAT)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	3172.528 (3194.53)
17	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2) * \exp(\theta_6 * TREAT)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	3197.899 (3219.90)
18	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3) * \exp(\theta_6 * TREAT)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	Yes	3154.911 (3176.91)
19	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4 * \exp(\theta_6 * TREAT)$ $Q = \theta_5 * exp(\eta_4)$	Yes	3214.600 (3236.60)
20	Proportional error. $Cl = \theta_1 * exp(\eta_1)$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4) * \exp(\theta_6 * TREAT)$	Yes	3234.808 (3256.81)

Table 4: Summary of models run cont'd.

Model #	Description	Converged	OFV (AIC)
Choosing Treatment Structures for a Two-Compartment Model cont'd			
21	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3) * \exp(\theta_7 * TREAT)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_3)$	No	.
Choosing Weight Structures for a Two-Compartment Model with Treatment			
22	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT + \theta_7 * WEIGHT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	3145.348 (3169.35)
23	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2) + \theta_7 * WEIGHT$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	.
24	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3) + \theta_7 * WEIGHT$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	Yes	3144.649 (3168.65)
25	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4 + \theta_7 * WEIGHT$ $Q = \theta_5 * exp(\eta_4)$	No	3141.590 (3165.59)
26	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4) + \theta_7 * WEIGHT$	No	.
27	Proportional error. $Cl = \theta_1 * exp(\eta_1) * WEIGHT^{\theta_7} + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	3144.795 (3168.80)
28	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2) * WEIGHT^{\theta_7}$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	Yes	3144.739 (3168.74)

Table 4: Summary of models run cont'd.

Model #	Description	Converged	OFV (AIC)
Choosing Weight Structures for a Two-Compartment Model with Treatment cont'd			
29	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3) * WEI^{\theta_7}$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4)$	No	3140.615 (3164.61)
30	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4 * WEIGHT^{\theta_7}$ $Q = \theta_5 * exp(\eta_4)$	No	3225.959 (3249.96)
31	Proportional error. $Cl = \theta_1 * exp(\eta_1) + \theta_6 * TREAT$ $K_a = \theta_2 * exp(\eta_2)$ $V_c = \theta_3 * exp(\eta_3)$ $V_p = \theta_4$ $Q = \theta_5 * exp(\eta_4) * WEIGHT^{\theta_7}$	No	3156.603 (3180.60)

Table 4 footnotes: Model is listed as converged if both parameter estimates and variance estimates are obtained. Rows with models which did not converge are greyed out. Bolded model numbers indicate the chosen model in each section. The section on Choosing Weight Structures does not have a bolded model because neither of the models that converged had an AIC meaningfully better than the model chosen in the section on Choosing Treatment Structures.

To begin analysis of this dataset, simple one- and two-compartment models were compared using change in OFV if the models were nested and using difference in AIC if the models were not nested. Starting values for PK parameters were derived from Arab-Alameddine, et al (2012), and a lower limit of 0 was placed on all PK parameters. For the one-compartment model, starting values were set to $K_a = 0.65$, $V = 281$ (the sum of the central and peripheral compartment volume estimates from Arab-Alameddine, et al (2012)), and $Cl = 64$. For the two-compartment model, starting values were set to $K_a = 0.65$, $V_C = 138$, $V_p = 143$, $Cl = 64$, and $Q = 9.3$. Initial estimates for the OMEGA (random effect) and SIGMA (residual error) matrices were set to 0.04 in each diagonal cell and 0 in each off-diagonal cell,

unless otherwise noted. Each model was estimated using FOCE with interaction, and the $COV record was manually included after the control stream was created.

The one- and two-compartment models were initially analyzed with random effects on all parameters (or as many parameters as possible, if a full model would not converge), then the random effect with the smallest percent relative standard error (%RSE) was removed, and OFV or AIC was compared. This process was repeated until the OFV or AIC in the simpler model was significantly larger than the more complex model. A one-compartment model with random effects on Cl, V, and K_a (Model 1) was compared with a one-compartment model with random effects on V and K_a (Model 2; the random effect on Cl was dropped because it had the lowest %RSE). Model 1 was chosen due to a significantly lower OFV ($\chi^2_{1,\alpha=0.05} = 3.84$; the OFV difference between the two models is 130.087, p < 0.0001). Since the larger model was chosen, no additional smaller models were analyzed. Eight two-compartment models (Models 3-10) with random effects on various combinations of parameters were then compared via change in AIC. Model 3 with random effects on all parameters did not converge, so Models 4-8 with random effects on all parameters but one were compared. Model 7 with no random effect on V_p had the lowest AIC of the group, but did not converge due to a small estimate for additive residual error. Consequently Model 9 was based on Model 7 with an adjustment to only include proportional residual error. This model converged. To test whether additional random effects should be removed, Model 10 with no random effects on V_p or Cl was analyzed (because the random effect for Cl in Model 9 had the smallest %RSE). Model 10 had a larger AIC than Model 9. Accordingly Model 9 with random effects on Cl, K_a, V_c, and Q and with proportional residual error was chosen. Then Model 1 (one-compartment) and Model 9 (two-compartment) were compared using difference in AIC; Model 9 was chosen.

Next, two-compartment models with linear or exponential TREAT on Cl, K_a, V_c, V_p, or Q were compared. The chosen model in the previous step (Model 9) with the previously selected random effects and residual error was used as the base model for this step. Starting values for this step were obtained from parameter estimates in Model 9. Models 11-21, two-compartment models with linear and exponential TREAT on various parameters, were compared. Not all models converged successfully. Of the models that converged, Model 11 with a linear TREAT on Cl was chosen via AIC comparison.

Finally, Models 22-31 with linear and power WEIGHT structures on various parameters were compared. In this step, Model 11 with linear TREAT on Cl was used as a base model. Of the models that converged, no models with WEIGHT had an AIC meaningfully better than Model 11 without WEIGHT. Since Model 11 had a lower AIC and was also less complex, it was chosen as the final model in this analysis.

Results

Graphical Checks

The DV vs PRED plot seen in Figure 9 (Chapter I) demonstrates a reasonably good fit of Model 11. Some outliers are present, but the majority of points is in close proximity to the identity line and symmetrical about it. Similarly, the WRES vs PRED plot given in Figure 10 (Chapter I) indicates there are some outliers, although the majority of weighted residual points are symmetrical about the prediction line with no trend or strong skew that would indicate a bias or necessitate a log transform. However, because of the outliers, further analysis may compare a log-transformed model against the final model to confirm. The DV vs IPRED plot given in Figure 11 (Chapter I) indicates that the random effects appear to improve model fit, as evidenced by the improved individual predicted concentration (x-axis) range as compared to Figure 10; however, even once the random effects are accounted for, there are still multiple

data points that are not well-predicted with this model, and predicted concentrations are systematically lower than observed concentrations as shown by the lowess line. Figure 12 (Chapter I) of RES vs PRED indicates that proportional residual error is indeed appropriate for this model due to the wider spread of points at higher predicted values.

Interpretations

The raltegravir dataset was best described by a two-compartment model with first-order absorption and a treatment effect on clearance. Clearance (Cl), central compartment volume (V_c), inter-compartmental clearance (Q), and absorption (K_a) estimates were allowed to vary by individual. A proportional residual error structure was deemed appropriate. The addition of treatment as a linear covariate on clearance offered a significant improvement in fit versus a model without treatment (ΔOFV 125.767. $\chi^2_{1,\alpha=0.05}$, $p < 0.001$). Apparent clearance varied significantly by treatment; subjects on raltegravir with antacid had a clearance estimate 50.8 ml/hr higher than subjects on raltegravir alone (95% CI 38.3, 63.3). Unexplained inter-individual variance in clearance was reduced by 7.7% versus a model without treatment (Percent relative reduction $\approx \frac{(\%CV\ base/100)^2 - (\%CV full/100)^2}{(\%CV\ base/100)^2} * 100 = \frac{0.297 - .274}{0.297} * 100 = 7.7$). The addition of linear or power weight on any of the PK parameters did not significantly improve fit as assessed by change in AIC. Parameter estimates and confidence intervals are given in Table 5. It should be noted that the increase in apparent clearance in subjects on raltegravir with antacid may be due to decreased bioavailability of the drug due to the competition with antacid rather than an increase in clearance itself (Kiser, et al., 2010, p. 5002).

Table 5: Final Model Parameter Estimates and 95% Confidence Intervals (Given in Parentheses).

Clearance (ml/hr) (95% CI)	**Absorption (hr^{-1}) (95% CI)**	**Central Compartment Volume (ml) (95% CI)**	**Peripheral Compartment Volume (ml) (95% CI)**	**Inter-Compartmental Clearance (ml/hr) (95% CI)**	**Linear Treatment on Clearance (ml/hr) (95% CI)**
76.6 (49.2, 104)	0.597 (0.443, 0.751)	95.9 (47.9, 144)	374 (83.9, 664)	11.5 (5.37, 17.6)	50.8 (38.3, 63.3)

CHAPTER IV

CONCLUSION

PK studies are an important step in the drug development process, and nonlinear mixed effects models aid in the estimation of PK parameters of interest and the identification of covariates which affect them. NONMEM specializes in the analysis of these types of models and PDx-POP provides a user interface for the creation of control streams which specify model structure.

Care should be taken when designing studies for PK analysis, especially if multiple covariates are of interest or the models are expected to be complex. Nonlinear mixed models are particularly prone to convergence issues and may present difficulties with interpretations. Additionally, NONMEM contains many more procedures and options than were discussed. This paper is meant to be a basic introduction to PK models, mixed models, and NONMEM. For a next step in understanding, we recommend referring to Owen and Fiedler-Kelly (2014), Fisher and Shafer (2007), Davidian and Giltinan (2003), Aarons and Ogungbenro (2010), Pinheiro and Bates (2000), a list of NONMEM resources (GM, 2008), and the NONMEM user guides (n.d.).

REFERENCES

Aarons, L., & Ogungbenro, K. (2010). Optimal Design of Pharmacokinetic Studies. *Basic & Clinicial Pharmacology & Toxicology, 106*, 250-255. doi:10.1111/j.1742-7843.2009.00533.x

Arab-Alameddine, M., Fayet-Mello, A., Lubomirov, R., Neely, M., di Iulio, J., & al., e. (2012). Population Pharmacokinetic Analysis and Pharmacogenetics of Raltegravir in HIV-Positive and Healthy Individuals. *Antimicrobial Agents and Chemotherapy, 56*(6), 2959-2966. doi:10.1128/AAC.05424-11

Bates, D. M. (2010, June 25). *lme4: Mixed-effects Modeling with R.* Retrieved from R-Forge: http://lme4.r-forge.r-project.org/lMMwR/lrgprt.pdf

Bates, D. M. (2011, January 11). Mixed Models in R Using the lme4 Package Part 5: Generalized Linear Mixed Models. Retrieved from http://lme4.r-forge.r-project.org/slides/2011-01-11-Madison/5GLMM.pdf

Beal, S. L., & Sheiner, L. B. (1998). *NONMEM Users Guide - Part VII Conditional Estimation Methods.* Retrieved from ICON: https://nonmem.iconplc.com/nonmem720/guides

Beal, S. L., Boeckmann, A. J., & Sheiner, L. B. (1992). *NONMEM Users Guide - Part VI PREDPP Guide.* Retrieved August 8, 2016, from ICON: https://nonmem.iconplc.com/nonmem720/guides/vi.pdf

Boeckmann, A. J., Beal, S. L., & Sheiner, L. B. (1992). *NONMEM Users Guide - Part IV NM-TRAN Guide.* Retrieved from Icon: https://nonmem.iconplc.com/nonmem720/guides

Boeckmann, A. J., Sheiner, L. B., & Beal, S. L. (2011). NONMEM Users Guide - Part V Introductory Guide. Ellicott City, MD, USA. Retrieved from https://nonmem.iconplc.com/nonmem720/guides/v.pdf

Bonate, P. L. (2011). *Pharmacokinetic-Pharmacodynamic Modeling and Simulation* (2 ed.). New York, New York, USA: Springer Science+Business Media, LLC. doi:10.1007/978-1-4419-9485-1

Bourne, D. W. (2016). *Chapter 19 Multi-Compartment Pharmacokinetic Models*. Retrieved from Boomer: https://www.boomer.org/c/p4/c19/c1902.html

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications.* New York, New York, USA: Cambridge University Press. Retrieved from www.cambridge.org/9780521848053

Categorical Mixed Models. (2008, July 26). Retrieved from York University: http://wiki.math.yorku.ca/index.php/Categorical_Mixed_Models

Cho, S. J. (2011, September). Principles and Applications of Pharmacokinetics in Drug Development. Retrieved from http://www.fda.gov/downloads/Drugs/NewsEvents/UCM275443.pdf

Davidian, M., & Giltinan, D. M. (2003). Nonlinear Models for Repeated Measurement Data: An Overview and Update. *Journal of Agricultural, Biological, and Environmental Statistics, 8*(4), 387-419. doi:10.1198/1085711032697

Dubois, A., Gsteiger, S., Pigeolet, E., & Mentré, F. (2010). Bioequivalence Tests Based on Individual Estimates Using Non-compartmental or Model-Based Analyses: Evaluation of Estimates of Sample Means and Type I Error for Different Designs. *Pharmaceutical Research, 27*(1), 92-104. doi:10.1007/s11095-009-9980-5

Duffull, S. B., Wright, D. F., & Winter, H. R. (2011). Interpreting Populations Pharmacokinetic-pharmacodynamic Analyses - A Clinical Viewpoint. *British Journal of Clinical Pharmacology, 71*(6), 807-814. doi:10.1111/j.1365-2125.2010.03891.x

Dunne, A. (2005). *Introduction to Categorical Data Analysis.* Retrieved from Population Approach Group Europe: http://www.page-meeting.org/page/page2005/PAGE2005O16.pdf

Ette, E. I., & Williams, P. J. (2007). *Pharmacometrics: The Science of Quantitative Pharmacology* (1 ed.). Hoboken, New Jersey, USA: John Wiley & Sons, Inc. Retrieved from http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471677833.html

European Medicines Agency. (1987, February). Pharmacokinetic Studies in Man. Retrieved from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003127.pdf

Fielding, S., Fayers, P. M., & Ramsay, C. R. (2009). Investigating the Missing Data Mechanism in Quality of Life Outcomes: A Comparison of Approaches. *Health and Quality of Life Outcomes, 7*(57). doi:10.1186/1477-7525-7-57

Fisher, D., & Shafer, S. (2007, March 7-11). Fisher/Shafer NONMEM Workshop Pharmacokinetic and Pharmacodynamic Analysis with NONMEM: Basic Concepts. Ghent, Belgium. Retrieved December 6, 2016, from https://www.scribd.com/document/319808093/Shafer-Nonmem-2

Food and Drug Administration. (1999, 02). *Guidance for Industry Population Pharmacokinetics.* Retrieved from FDA.gov: http://www.fda.gov/downloads/Drugs/.../Guidances/UCM072137.pdf

Food and Drug Administration. (2003, 04). Guidance for Industry Exposure-Response Relationships - Study Design, Data Analysis, and Regulatory Applications. Rockville, MD, USA. Retrieved from http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm072109.pdf

Gabrielsson, J., & Weiner, D. (2012). Non-compartmental Analysis. *Methods in Molecular Biology, 929*, 377-389. doi:10.1007/978-1-62703-050-2_16

Gelman, A. (2005). Analysis of Variance – Why it is More Important Than Ever. *The Annals of Statistics, 33*(1), 1-53. doi:10.1214/009053604000001048

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York, NY, USA: Cambridge University Press. Retrieved from www.cambridge.org/9780521867061

Generalized Linear Model. (n.d.). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Generalized_linear_model#Link_function

GM. (2008, November 3). *NONMEM Resources*. Retrieved from PK/PD: http://pk-pd.blogspot.com/2008/11/nonmem-resources.html

Goonaseelan, C. P., Mentré, F., & Steimer, J.-L. (2005). Non-Linear Mixed Effects Modeling – From Methodology and Software Development to Driving Implementation in Drug Development Science. *Journal of Pharmacokinetics and Pharmacodynamics, 32*(2), 161-183. doi:10.1007/s10928-005-0062-y

Hartford, A. (n.d.). Introduction to Pharmacokinetic/Pharmacodynamic Modeling: Concepts and Methods. Retrieved from https://www.scribd.com/document/245682325/Pharmacokinetic-Pharmacodynamic-Modeling-Simulation-pdf

Hedeker, D. (2005). Generalized Linear Mixed Models. In B. S. Everitt, & D. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science.* John Wiley & Sons, Inc. doi:10.1002/0470013192.bsa251

Hedeker, D. (n.d.). Mixed-Effects Models for Count Data. Chicago, IL, USA. Retrieved from http://www.ssicentral.com/supermix/Documentation/PoisSDLS.pdf

Huang, J. Z., Wu, C. O., & Zhou, L. (2004). Polynomial Spline Estimation and Inference for Varying Coefficient Models with Longitudinal Data. *Statistica Sinica, 14*(3), 763-788. Retrieved from http://www.jstor.org/stable/24307415

ICON plc. (2016, July 06). PDx-POP Manual Version 5.2. Hanover, MD, USA. Retrieved from https://nonmem.iconplc.com/pdxpop/PDx-Pop5.2/manual.pdf

Jacqmin-Gadda, H. (n.d.). Introduction to Mixed Model and Missing Data Issues in Longitudinal Studies. Bordeaux, France. Retrieved from http://christophe.genolini.free.fr/recherche/atelierATelecharger/atelier205_1_HJG.pdf

Jambhekar, S. S., & Breen, P. J. (2012). *Basic Pharmacokinetics.* London, UK: Pharmaceutical Press. Retrieved from http://www.pharmpress.com/product/9780853699804/basic-pharmacokinetics

Jonsson, E. N., & Karlsson, M. O. (1998). Automated Covariate Model Building Within NONMEM. *Pharmaceutical Research, 15*(9), 1463-1468. doi:10.1023/A:1011970125687

Kile, D. A., MaWhinney, S., Aquilante, C. L., Rower, J. E., Castillo-Mancilla, J. R., & Anderson, P. L. (2012). A Population Pharmacokinetic-Pharmacogenetic Analysis of Atazanavir. *AIDS Research adn Human Retroviruses, 28*(10), 1227-1234. doi:10.1089/aid.2011.0378

Kiser, J. J., Bumpass, J. B., Meditz, A. L., Anderson, P. L., Bushman, L., Ray, M., . . . Brundage, R. (2010). Effect of Antacids on the Pharmacokinetics of Raltegravir in Human Immunodeficiency Virus-Seronegative Volunteers. *Antimicrobial Agents and Chemotherapy, 54*(12), 4999-5003. doi:10.1128/AAC.00636-10

Lukas, J. C., & Piotrovskij, V. (2001, 02 26-7). FO vs FOCE vs FOCE With Interaction. Retrieved 01 16, 2017, from http://www.cognigencorp.com/nonmem/nm/98feb262001.html

Mandema, J. W., Verotta, D., & Sheiner, L. B. (1992). Building Population Pharmacokinetic-Pharmacodynamic Models. I. Models for Covariate Effects. *Journal of Pharmacokinetics and Biopharmaceutics, 20*(5), 511-528. doi:10.1007/BF01061469

Maxwell, S. E., Delaney, H. D., & Kelley, K. (2003). *Designing Experiments and Analyzing Data.* New York, New York, USA: Psychology Press Taylor & Francis Group. Retrieved from https://www.routledge.com/Designing-Experiments-and-Analyzing-Data-A-Model-Comparison-Perspective/Maxwell-Delaney-Kelley/p/book/9780805837186

McLean, R. A., Sanders, W. L., & W., S. W. (1991). A Unified Approach to Mixed Linear Models. *The American Statistician, 45*(1), 54-64. doi:10.2307/2685241

Mould, D. R., & Upton, R. N. (2012). Basic Concepts in Population Modeling, Simulation, and Model-Based Drug Development. *Pharmacometrics and Systems Pharmacology, 1*(9), e6. doi:10.1038/psp.2012.4

Mould, D. R., & Upton, R. N. (2013). Basic Concepts in Population Modeling, Simulation, and Model-Based Drug Development Part 2: Introduction to Pharmacokinetic Modeling Methods. *CPT: Pharmacometrics and Systems Pharmacology, 2*(4), e38. doi:10.1038/psp.2013.14

NONMEM History. (n.d.). Retrieved August 30, 2016, from ICON PLC: http://www.iconplc.com/innovation/solutions/nonmem/history/

NONMEM User Guides. (n.d.). Retrieved December 6, 2016, from ICON PLC: https://nonmem.iconplc.com/nonmem720/guides

Owen, J. S., & Fiedler-Kelly, J. (2014). *Introduction fo Population Pharmacokinetic / Pharmacodynamic Analysis with Nonlinear Mixed Effects Models.* Hoboken, NJ: John Wiley & Sons, Inc.

Petricoul, O., Claret, L., Barbolosi, D., Iliadis, A., & Pouzzo, C. (2001). Information Tools for Exploratory Data Analysis in Population Pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics, 28*(6), 577-599. doi:10.1023/A:1014464505261

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS.* New York, New York, USA: Springer-Verlag New York, Inc. Retrieved from http://link.springer.com/book/10.1007/b98882

SAS Institute Inc. (2009). SAS/STAT 9.2 User's Guide, Second Edition. Cary, NC, USA. Retrieved August 18, 2016, from SAS Institute Inc: https://support.sas.com/documentation/cdl/en/statug/63033/PDF/default/statug.pdf

Schabenberger, O. (2005). *Introducing the GLIMMIX Procedure for Generalized Linear Mixed Models.* Carey, NC: SAS Institute. Retrieved from http://www2.sas.com/proceedings/sugi30/toc.html

see Bourne. (2016).

Seltman, H. J. (2015, 09 08). Expermental Design and Analysis. Retrieved from http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Sheu, C.-F. (2002). Fitting mixed-effects models for repeated ordinal outcomes with the NLMIXED procedure. *Behavior Research Methods, Instruments, & Computers, 34*(2), 151-157. doi:10.3758/bf03195436

The Hessian. (2016). Retrieved from KhanAcademy: https://www.khanacademy.org/math/multivariable-calculus/applications-of-multivariable-derivatives/quadratic-approximations/a/the-hessian

The R Foundation. (n.d.). *The R Project for Statistical Computing*. Retrieved from R Project: https://www.r-project.org/

van de Pol, M., & Wright, J. (2009). A Simple Method for Distinguishing Within- Versus Between-Subject Effects Using Mixed Models. *Animal Behaviour, 77*(3), 753-758. doi:10.1016/j.anbehav.2008.11.006

Vonesh, E. F. (1996). A Note on the Use of Laplace's Approximation for Nonlinear Mixed-Effects Models. *Biometrika, 83*(2), 447-452. doi:10.1093/biomet/83.2.447

Wolfinger, R. D. (1993). Laplace's Approximation for Nonlinear Mixed Models. *Biometrika, 80*(4), 791-795. doi:10.1093/biomet/80.4.791

Ying, G.-s., & Liu, C. (2006). Statistical Analysis of Clustered Data using SAS System. Retrieved from http://lexjansen.com/nesug/nesug06/an/da01.pdf

APPENDIX A

SUMMARY AND RESULTS OF MODELS IN THE ANALYSIS OF THE RALTEGRAVIR DATASET

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
1	1 cmt, random effects on all parameters; additive and proportional error	CL = THETA(1) *EXP(ETA(1)) V = THETA(2) *EXP(ETA(2)) KA = THETA(3) *EXP(ETA(3))	Yes	3380.097 (3396.1)	250 (192, 308)	120 (85.7, 154)	1.09 (0.667, 1.51)	.	.	.	.	.	Starting values (SV): V = 281, Cl = 64, Ka = 0.65, Omega (all random effects) = [0.04, 0.04, 0.04], Sigma (additive and proportional error) = [0.04, 0.04].
2	1 cmt, random effects on V, Ka; additive and proportional error	CL = THETA(1) V = THETA(2) *EXP(ETA(1)) KA = THETA(3) *EXP(ETA(2))	Yes	3510.184 (3524.18)	173 (105, 241)	88.7 (63.6, 114)	3.63 (-1.86, 9.12)	.	.	.	.	.	SV same as Model 1. %RSE for Ka was 77.1%, and the 95% CI for Ka includes 0.
3	2 cmt, random effects on all parameters; additive and proportional error	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) *EXP(ETA(4)) KA = THETA(5) *EXP(ETA(5))	No	3270.871 (3294.87)	.	104 (., .)	0.511 (., .)	107 (., .)	662 (., .)	15.0 (., .)	.	.	SV: Cl = 64, Ka = 0.65, Vc = 138, Vp = 143, Q = 9.3, Omega (all random effects) = [0.04, 0.04, 0.04, 0.04, 0.04], Sigma (additive and proportional error) = [0.04, 0.04]. Model partially converged.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
4	2 cmt, random effects on Vc, Vp, Q, Ka; additive and proportional error	CL = THETA(1) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) *EXP(ETA(4)) KA = THETA(5) *EXP(ETA(5))	No	3402.856 (3424.86)	.	74.3 (., .)	0.752 (., .)	21.8 (., .)	141 (., .)	7.51 (., .)	.	.	SV same as Model 3. Model partially converged.
5	2 cmt, random effects on Cl, Vc, Vp, Q; additive and proportional error	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) *EXP(ETA(4)) KA = THETA(5)	No	3275.084 (3297.08)	.	104 (., .)	0.481 (., .)	102 (., .)	668 (., .)	14.5 (., .)	.	.	SV same as Model 3. Model partially converged.
6	2 cmt, random effects on Cl, Ka , Vp, Q; additive and proportional error	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) *EXP(ETA(4)) KA = THETA(5) *EXP(ETA(5))	No	3280.613 (3302.61)	.	101 (., .)	0.511 (., .)	95.3 (., .)	650 (., .)	14.3 (., .)	.	.	SV same as Model 3. Model partially converged.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
7	2 cmt, random effects on Cl, Ka, Vc, Q; additive and proportional error	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(5))	No	3270.871 (3292.87)	.	104 (., .)	0.511 (., .)	107 (., .)	662 (., .)	15.0 (., .)	.	.	SV same as Model 3. Model partially converged.
8	2 cmt, random effects on Cl, Ka, Vc, Vp; additive and proportional error	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) V3 = THETA(4) *EXP(ETA(4)) KA = THETA(5) *EXP(ETA(5))	No	3290.593 (3312.59)	.	94.1 (., .)	0.552 (., .)	114 (., .)	984 (., .)	20.8 (., .)	.	.	SV same as Model 3. Model partially converged.
9	2 cmt, random effects on Cl, Ka, Vc, Q; proportional error. Based on Model 7, which had the lowest AIC of the models with random effects on all parameters but one	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(5))	Yes	3270.871 (3290.87)	.	104 (73.8, 134)	0.511 (0.434, 0.588)	107 (67.4, 147)	662 (426, 895)	15.0 (7.59, 22.4)	.	.	SV same as Model 3.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
10	2 cmt, random effects on Ka, Vc, Q; proportional error	CL = THETA(1) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(5))	Yes	3424.624 (3442.62)	.	71.2 (56.9, 85.5)	0.601 (0.447, 0.755)	3.83 (-6.85, 14.5)	310 (16.0, 604)	7.41 (2.90, 11.9)	.	.	SV same as Model 3. %RSE for Vc was 142% and 95%CI for Vc included 0.
11	Model 9 with linear TREAT on Cl	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	Yes	3145.104 (3167.10)	.	76.6 (49.2, 104)	0.597 (0.443, 0.751)	95.9 (47.9, 144)	374 (83.9, 664)	11.5 (5.37, 17.6)	50.8 (38.3, 63.3)	.	SV: Cl = 104, Ka = 0.481, Vc = 102, Vp = 668, Q = 14.5. SV: trt=5. SV based on parameter estimates from Model 9.
12	Model 9 with linear TREAT on Ka	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4)) +THETA(6)*TREAT	Yes	3171.470 (3193.47)	.	104 (68.1, 140)	0.319 (0.218, 0.420)	140 (63.6, 216)	667 (316, 1020)	15.2 (7.38, 23.0)	0.794 (0.263, 1.33)	.	SV same as Model 11. SV: trt=0.1.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
13	Model 9 with linear TREAT on Vc	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) +THETA(6)*TREAT Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	.	.	.	.	.	.	.	.	.	SV same as Model 11. SV: trt = -50. Model did not converge.
14	Model 9 with linear TREAT on Vp	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) +THETA(6)*TREAT KA = THETA(5) *EXP(ETA(4))	Yes	3214.600 (3236.60)	.	95.2 (66.6, 124)	0.549 (0.490, 0.608)	103 (64.2, 142)	518 (267, 769)	20.7 (12.1, 29.3)	4160 (1060, 7260)	.	SV same as Model 11. SV: trt = 500.
15	Model 9 with linear TREAT on Q	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) +THETA(6)*TREAT *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	Yes	3260.697 (3282.70)	.	101 (71.6, 130)	0.543 (0.436, 0.650)	101 (.49.5, 153)	428 (-119, 975)	16.2 (7.24, 25.2)	-7.86 (-18.8, 3.04)	.	SV same as Model 11. SV: trt = 0.1. %RSE for Vp and TREAT were 65.2% and 70.7%, respectively. 95%CIs for both parameters included 0.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
16	Model 9 with exponential TREAT on Cl	CL = THETA(1) *EXP(ETA(1)) *EXP(THETA(6)*TREAT) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	3172.528 (3194.53)	.	72.2 (., .)	0.629 (., .)	96.8 (., .)	366 (., .)	10.8 (., .)	0.540 (., .)	.	SV same as Model 11. SV: trt = 0.5. Omega = [0.04, 0.004, 0.04, 0.04]. Model partially converged.
17	Model 9 with exponential TREAT on Ka	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4)) *EXP(THETA(6)*TREAT)	No	3197.899 (3219.90)	.	92.9 (., .)	0.499 (., .)	179 (., .)	526 (., .)	12.1 (., .)	30.0 (., .)	.	SV same as Model 11. SV: trt = 30.
18	Model 9 with exponential TREAT on Vc	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) *EXP(THETA(6)*TREAT) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	Yes	3154.911 (3176.91)	.	99 (64.7, 133)	0.699 (0.586, 0.812)	411 (180, 642)	524 (63.4, 985)	10.6 (3.15, 18.0)	-1.97 (-2.24, -1.70)	.	SV same as Model 11. SV: trt = 0.5.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
19	Model 9 with exponential TREAT on Vp	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) *EXP(THETA(6)*TREAT) KA = THETA(5) *EXP(ETA(4))	Yes	3214.600 (3236.60)	.	95.2 (66.0, 124)	0.5549 (0.490, 0.608)	103 (65.8, 140)	518 (269, 767)	20.7 (12.0, 29.4)	2.20 (1.77, 2.63)	.	SV same as Model 11. SV: trt = 0.5.
20	Model 9 with exponential TREAT on Q	CL = THETA(1) *EXP(ETA(1)) V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) *EXP(THETA(6)*TREAT) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	Yes	3234.808 (3256.81)	.	101 (67.5, 135)	0.539 (0.471, 0.607)	105 (58.4, 152)	561 (175, 947)	19.3 (6.31, 32.3)	-0.877 (-1.26, -0.491)	.	SV same as Model 11. SV: trt = 0.5.
21	Model 9 with linear TREAT on Cl and exponential TREAT on Q (combination of the two TREAT models with the lowest AIC)	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) *EXP(THETA(7)*TREAT) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	.	.	.	.	.	.	.	.	.	SV same as Model 11. SV: trt on Cl = 50, trt on Vc = -2. Model did not converge.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
22	Model 11 with linear weight on Cl	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT +THETA(7)*weight V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	3145.348 (3169.35)	.	57.3 (., .)	0.610 (., .)	85.5 (., .)	361 (., .)	11.4 (., .)	50.0 (., .)	0.256 (., .)	SV same as Model 11. SV: trt = 5, wt = -0.5. Omega = [0.04, 0.04, 0.04, 0.004]. Model partially converged.
23	Model 11 with linear weight on Ka	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4)) +THETA(7)*weight	No	.	.	.	.	.	.	.	.	.	SV same as Model 11. SV: trt = 5, wt = 0.1.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
24	Model 11 with linear weight on Vc	CL = THETA(1) *EXP(ETA(1))+THETA(6) *TREAT V2 = THETA(2) *EXP(ETA(2)) +THETA(7) *weight Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	Yes	3144.649 (3168.65)	.	76.4 (48.4, 104)	0.604 (0.471, 0.737)	50.8 (-95.8, 197)	368 (93.6, 642)	11.4 (5.62, 17.2)	50.8 (38.1, 63.5)	0.641 (-0.841, 2.12)	SV same as Model 11. SV: trt = 5, wt = 0.1. Omega = [0.004, 0.004, 0.004, 0.004]. %RSE for Vc and WEIGHT were 147% and 117%, respectively. 95%CIs for both parameters included 0.
25	Model 11 with linear weight on Vp	CL = THETA(1) *EXP(ETA(1))+THETA(6) *TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) +THETA(7) *weight KA = THETA(5) *EXP(ETA(4))	No	3141.590 (3165.59)	.	76.76 (., .)	0.658 (., .)	107 (., .)	73.9 (., .)	12.1 (., .)	49.7 (., .)	4.07 (., .)	SV same as Model 11. SV: trt = 5, wt = -0.5. Omega = [0.4, 0.4, 0.4, 0.004]. Model partially converged.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
26	Model 11 with linear weight on Q	CL = THETA(1) *EXP(ETA(1))+THETA(6) *TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) +THETA(7) *weight V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	3141.912 (3165.91)	.	76.76 (., .)	0.653 (., .)	107 (., .)	332 (., .)	9.46 (., .)	50.0 (., .)	0.0347 (., .)	SV same as Model 11. SV: trt = 5, wt = -0.1.Model partially converged.
27	Model 11 with power weight on Cl	CL = THETA(1) *EXP(ETA(1)) *(weight**THETA(7)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	3144.795 (3168.80)	.	47.0 (., .)	0.595 (., .)	95.4 (., .)	375 (., .)	11.5 (., .)	50.7 (., .)	0.112 (., .)	SV same as Model 11. SV: trt = 5, wt = 0.1. Omega = [0.4, 0.4, 0.4, 0.4]. Sigma = [0.4]. Model partially converged.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
28	Model 11 with power weight on Ka	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4)) *(weight**THETA(7))	Yes	3144.739 (3168.74)	.	76.5 (50.6, 102)	0.116 (-0.737, 0.969)	95.9 (36.9, 155)	380 (78.2, 682)	11.5 (5.58, 17.4)	50.8 (38.2, 63.4)	0.395 (-1.35, 2.14)	SV same as Model 11. SV: trt = 5, wt = 0.1. Omega = [0.4, 0.4, 0.4, 0.4]. Sigma = [0.4]. %RSE for Ka and WEIGHT was 375% and 225%, respectively. 95%CIs for both parameters included 0.
29	Model 11 with power weight on Vc	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) *(weight**THETA(7)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	3140.615 (3162.61)	.	76.7 (., .)	0.653 (., .)	1.02 (., .)	329 (., .)	12.0 (., .)	50.1 (., .)	1.12 (., .)	SV same as Model 11. SV: trt = 5, wt = 0.1. Model partially converged.

#	Model Description	Equation	Con-verged?	OFV (AIC)	V Est (95% CI)	Cl Est (95% CI)	Ka Est (95% CI)	Vc Est (95% CI)	Vp Est (95% CI)	Q Est (95% CI)	TRT Est (95% CI)	WT Est (95% CI)	Notes
30	Model 11 with power weight on Vp	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) V3 = THETA(4) *(weight**THETA(7)) KA = THETA(5) *EXP(ETA(4))	No	3225.959 (3249.96)	.	65.6 (., .)	4190000 (., .)	188 (., .)	674 (., .)	6.95 (., .)	30.2 (., .)	-0.226 (., .)	SV same as Model 11. SV: trt = 5, wt = 1. Model partially converged.
31	Model 11 with power weight on Q	CL = THETA(1) *EXP(ETA(1)) +THETA(6)*TREAT V2 = THETA(2) *EXP(ETA(2)) Q = THETA(3) *EXP(ETA(3)) *(weight**THETA(7)) V3 = THETA(4) KA = THETA(5) *EXP(ETA(4))	No	3156.603 (3180.6)	.	73.4 (., .)	0.571 (., .)	82.2 (., .)	376 (., .)	26.6 (., .)	49.1 (., .)	-0.222 (., .)	SV same as Model 11. SV: trt = 5, wt = 1. Model partially converged.

Table 2 Notes: Cmt stands for compartment. In the two-compartment model, V2 is equivalent to V_c, and V3 is equivalent to V_p. THETAs indicate fixed effects parameters. ETAs indicate random effects parameters. Periods (".") indicate a parameter estimate that is not applicable or not calculated. Convergence is listed as "No" when there is no convergence or partial convergence. Models which did not converge are greyed out. Starting values are abbreviated SV. Starting values for ETA variances are set via the `$OMEGA` record, and are set to 0.04 unless otherwise stated. Starting values for residual error variances are set via the `$SIGMA` record, and are set to 0.04 unless otherwise stated.

APPENDIX B

RALTEGRAVIR DATASET

C	Data Desc: Raltegravir Dataset											
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6001	0	0	.	4.00E+05	1	0	0	63.25	65.5	33	1
	6001	0	0.5	1349	0	0	0	0	63.25	65.5	33	1
	6001	0	1	2179	0	0	0	0	63.25	65.5	33	1
	6001	0	1.5	2670	0	0	0	0	63.25	65.5	33	1
	6001	0	2	2603	0	0	0	0	63.25	65.5	33	1
	6001	0	3	1206	0	0	0	0	63.25	65.5	33	1
	6001	0	4	560	0	0	0	0	63.25	65.5	33	1
	6001	0	5	268	0	0	0	0	63.25	65.5	33	1
	6001	0	6	153	0	0	0	0	63.25	65.5	33	1
	6001	0	8	77.3	0	0	0	0	63.25	65.5	33	1
	6001	0	10	49.1	0	0	0	0	63.25	65.5	33	1
	6001	0	12	41.9	0	0	0	0	63.25	65.5	33	1
	6001	0	24.26667	21.9	0	0	0	0	63.25	65.5	33	1
	6001	0	48.33333	10.7	0	0	0	0	63.25	65.5	33	1
	6001	1	0	.	4.00E+05	4	0	0	63.25	65.5	33	1
	6001	1	0.5	545	0	0	0	0	63.25	65.5	33	1
	6001	1	1	1014	0	0	0	0	63.25	65.5	33	1
	6001	1	1.5	895	0	0	0	0	63.25	65.5	33	1
	6001	1	2	522	0	0	0	0	63.25	65.5	33	1
	6001	1	3	215	0	0	0	0	63.25	65.5	33	1
	6001	1	4	92.7	0	0	0	0	63.25	65.5	33	1
	6001	1	5	43.2	0	0	0	0	63.25	65.5	33	1

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6001	1	6	26.2	0	0	0	0	63.25	65.5	33	1
	6001	1	8	13.6	0	0	0	0	63.25	65.5	33	1
	6001	1	10	14.9	0	0	0	0	63.25	65.5	33	1
	6001	1	12	12.8	0	0	0	0	63.25	65.5	33	1
	6001	1	24	2.5	0	0	0	0	63.25	65.5	33	1
	6001	1	48	1.25	0	0	0	0	63.25	65.5	33	1
	6003	0	0	.	4.00E+05	1	0	0	58.9	68.5	19	1
	6003	0	0.5	17.2	0	0	0	0	58.9	68.5	19	1
	6003	0	1	112	0	0	0	0	58.9	68.5	19	1
	6003	0	1.5	289	0	0	0	0	58.9	68.5	19	1
	6003	0	2	276	0	0	0	0	58.9	68.5	19	1
	6003	0	3	288	0	0	0	0	58.9	68.5	19	1
	6003	0	4	269	0	0	0	0	58.9	68.5	19	1
	6003	0	5	1117	0	0	0	0	58.9	68.5	19	1
	6003	0	6	544	0	0	0	0	58.9	68.5	19	1
	6003	0	8	147	0	0	0	0	58.9	68.5	19	1
	6003	0	10.18333	47.8	0	0	0	0	58.9	68.5	19	1
	6003	0	12.13333	45.1	0	0	0	0	58.9	68.5	19	1
	6003	0	25.26667	17.9	0	0	0	0	58.9	68.5	19	1
	6003	0	47.61667	15.3	0	0	0	0	58.9	68.5	19	1
	6003	1	0	.	4.00E+05	4	0	0	58.9	68.5	19	1
	6003	1	0.5	84.3	0	0	0	0	58.9	68.5	19	1

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6003	1	1	655	0	0	0	0	58.9	68.5	19	1
	6003	1	1.5	1159	0	0	0	0	58.9	68.5	19	1
	6003	1	2	1719	0	0	0	0	58.9	68.5	19	1
	6003	1	3	828	0	0	0	0	58.9	68.5	19	1
	6003	1	4	429	0	0	0	0	58.9	68.5	19	1
	6003	1	5	217	0	0	0	0	58.9	68.5	19	1
	6003	1	6	87.3	0	0	0	0	58.9	68.5	19	1
	6003	1	8	28.8	0	0	0	0	58.9	68.5	19	1
	6003	1	10	13.8	0	0	0	0	58.9	68.5	19	1
	6003	1	12	13.3	0	0	0	0	58.9	68.5	19	1
	6003	1	23.83	5.29	0	0	0	0	58.9	68.5	19	1
	6003	1	47.83	2.5	0	0	0	0	58.9	68.5	19	1
	6004	0	0	.	4.00E+05	1	0	0	55.5	65.5	21	0
	6004	0	0.516667	44.3	0	0	0	0	55.5	65.5	21	0
	6004	0	1.016667	145	0	0	0	0	55.5	65.5	21	0
	6004	0	1.5	191	0	0	0	0	55.5	65.5	21	0
	6004	0	2	276	0	0	0	0	55.5	65.5	21	0
	6004	0	3	265	0	0	0	0	55.5	65.5	21	0
	6004	0	4	165	0	0	0	0	55.5	65.5	21	0
	6004	0	5	73.2	0	0	0	0	55.5	65.5	21	0
	6004	0	6	96.8	0	0	0	0	55.5	65.5	21	0
	6004	0	8	27.6	0	0	0	0	55.5	65.5	21	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6004	0	10	21.8	0	0	0	0	55.5	65.5	21	0
	6004	0	12	27.4	0	0	0	0	55.5	65.5	21	0
	6004	0	24	2.5	0	0	0	0	55.5	65.5	21	0
	6004	0	48.16667	1.25	0	0	0	0	55.5	65.5	21	0
	6004	1	0	.	4.00E+05	4	0	0	55.5	65.5	21	0
	6004	1	0.5	23.2	0	0	0	0	55.5	65.5	21	0
	6004	1	1	1135	0	0	0	0	55.5	65.5	21	0
	6004	1	1.5	980	0	0	0	0	55.5	65.5	21	0
	6004	1	2	641	0	0	0	0	55.5	65.5	21	0
	6004	1	3	332	0	0	0	0	55.5	65.5	21	0
	6004	1	4	174	0	0	0	0	55.5	65.5	21	0
	6004	1	5.583333	62.9	0	0	0	0	55.5	65.5	21	0
	6004	1	6.083333	37.5	0	0	0	0	55.5	65.5	21	0
	6004	1	8.166667	18.6	0	0	0	0	55.5	65.5	21	0
	6004	1	10.08333	14.3	0	0	0	0	55.5	65.5	21	0
	6004	1	12	7.86	0	0	0	0	55.5	65.5	21	0
	6004	1	23.83333	6.03	0	0	0	0	55.5	65.5	21	0
	6004	1	47.83333	2.5	0	0	0	0	55.5	65.5	21	0
	6005	0	0	.	4.00E+05	1	1	0	82.6	68.25	29	0
	6005	0	0.5	039	0	0	1	0	82.6	68.25	29	0
	6005	0	1	1801	0	0	1	0	82.6	68.25	29	0
	6005	0	1.5	1058	0	0	1	0	82.6	68.25	29	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6005	0	2	8859	0	0	1	0	82.6	68.25	29	0
	6005	0	3.083333	264	0	0	1	0	82.6	68.25	29	0
	6005	0	4	9767	0	0	1	0	82.6	68.25	29	0
	6005	0	5	3297	0	0	1	0	82.6	68.25	29	0
	6005	0	6	5457	0	0	1	0	82.6	68.25	29	0
	6005	0	8	7889	0	0	1	0	82.6	68.25	29	0
	6005	0	10	6792	0	0	1	0	82.6	68.25	29	0
	6005	0	12	8734	0	0	1	0	82.6	68.25	29	0
	6005	0	24.41667	5755	0	0	1	0	82.6	68.25	29	0
	6005	0	48.16667	2.5	0	0	1	0	82.6	68.25	29	0
	6005	1	0	.	4.00E+05	4	1	0	82.6	68.25	29	0
	6005	1	0.5	4759	0	0	1	0	82.6	68.25	29	0
	6005	1	1	5793	0	0	1	0	82.6	68.25	29	0
	6005	1	1.5	8036	0	0	1	0	82.6	68.25	29	0
	6005	1	2	0111	0	0	1	0	82.6	68.25	29	0
	6005	1	3	7802	0	0	1	0	82.6	68.25	29	0
	6005	1	5	7784	0	0	1	0	82.6	68.25	29	0
	6005	1	6	9415	0	0	1	0	82.6	68.25	29	0
	6005	1	8	1214	0	0	1	0	82.6	68.25	29	0
	6005	1	10	5134	0	0	1	0	82.6	68.25	29	0
	6005	1	11.91667	1465	0	0	1	0	82.6	68.25	29	0
	6005	1	23.75	2.5	0	0	1	0	82.6	68.25	29	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6005	1	48.75	1.25	0	0	1	0	82.6	68.25	29	0
	6006	0	0	.	4.00E+05	1	1	0	64.3	72	19	0
	6006	0	0.5	0.1	0	0	1	0	64.3	72	19	0
	6006	0	1	9	0	0	1	0	64.3	72	19	0
	6006	0	1.5	87	0	0	1	0	64.3	72	19	0
	6006	0	2	285	0	0	1	0	64.3	72	19	0
	6006	0	3	239	0	0	1	0	64.3	72	19	0
	6006	0	4	317	0	0	1	0	64.3	72	19	0
	6006	0	5	536	0	0	1	0	64.3	72	19	0
	6006	0	6	1011	0	0	1	0	64.3	72	19	0
	6006	0	8	353	0	0	1	0	64.3	72	19	0
	6006	0	10	4	0	0	1	0	64.3	72	19	0
	6006	0	12	72.4	0	0	1	0	64.3	72	19	0
	6006	0	24.38333	4	0	0	1	0	64.3	72	19	0
	6006	0	48.38333	3301	0	0	1	0	64.3	72	19	0
	6006	1	0	.	4.00E+05	4	1	0	64.3	72	19	0
	6006	1	0.5	111	0	0	1	0	64.3	72	19	0
	6006	1	1	439	0	0	1	0	64.3	72	19	0
	6006	1	1.5	754	0	0	1	0	64.3	72	19	0
	6006	1	2	2552	0	0	1	0	64.3	72	19	0
	6006	1	3	1272	0	0	1	0	64.3	72	19	0
	6006	1	4	650	0	0	1	0	64.3	72	19	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6006	1	5	378	0	0	1	0	64.3	72	19	0
	6006	1	6	212	0	0	1	0	64.3	72	19	0
	6006	1	8	101	0	0	1	0	64.3	72	19	0
	6006	1	10	58.8	0	0	1	0	64.3	72	19	0
	6006	1	12	30	0	0	1	0	64.3	72	19	0
	6006	1	24.03333	11.5	0	0	1	0	64.3	72	19	0
	6006	1	48.28333	6.19	0	0	1	0	64.3	72	19	0
	6007	0	0	.	4.00E+05	1	0	0	65	67.3	29	0
	6007	0	0.5	14.4	0	0	0	0	65	67.3	29	0
	6007	0	1	55.5	0	0	0	0	65	67.3	29	0
	6007	0	1.5	71.6	0	0	0	0	65	67.3	29	0
	6007	0	2.166667	4634	0	0	0	0	65	67.3	29	0
	6007	0	3	4891	0	0	0	0	65	67.3	29	0
	6007	0	4	2022	0	0	0	0	65	67.3	29	0
	6007	0	5.066667	848	0	0	0	0	65	67.3	29	0
	6007	0	6	385	0	0	0	0	65	67.3	29	0
	6007	0	8	152	0	0	0	0	65	67.3	29	0
	6007	0	10	89.8	0	0	0	0	65	67.3	29	0
	6007	0	12	45	0	0	0	0	65	67.3	29	0
	6007	0	24.03333	16.4	0	0	0	0	65	67.3	29	0
	6007	0	48.11667	2.5	0	0	0	0	65	67.3	29	0
	6007	1	0	.	4.00E+05	4	0	0	65	67.3	29	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6007	1	0.5	229	0	0	0	0	65	67.3	29	0
	6007	1	1	504	0	0	0	0	65	67.3	29	0
	6007	1	1.533333	1714	0	0	0	0	65	67.3	29	0
	6007	1	2	1661	0	0	0	0	65	67.3	29	0
	6007	1	3	861	0	0	0	0	65	67.3	29	0
	6007	1	4	532	0	0	0	0	65	67.3	29	0
	6007	1	5	298	0	0	0	0	65	67.3	29	0
	6007	1	6	144	0	0	0	0	65	67.3	29	0
	6007	1	8	36.9	0	0	0	0	65	67.3	29	0
	6007	1	10	20.3	0	0	0	0	65	67.3	29	0
	6007	1	12	10.9	0	0	0	0	65	67.3	29	0
	6007	1	24	5.07	0	0	0	0	65	67.3	29	0
	6007	1	48.03333	2.5	0	0	0	0	65	67.3	29	0
	6008	0	0	.	4.00E+05	1	1	0	82.95	73.25	26	0
	6008	0	0.5	9697	0	0	1	0	82.95	73.25	26	0
	6008	0	1	8246	0	0	1	0	82.95	73.25	26	0
	6008	0	1.5	7258	0	0	1	0	82.95	73.25	26	0
	6008	0	2	6601	0	0	1	0	82.95	73.25	26	0
	6008	0	3	9466	0	0	1	0	82.95	73.25	26	0
	6008	0	4	5846	0	0	1	0	82.95	73.25	26	0
	6008	0	5	5378	0	0	1	0	82.95	73.25	26	0
	6008	0	6.25	6185	0	0	1	0	82.95	73.25	26	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6008	0	7.583333	0331	0	0	1	0	82.95	73.25	26	0
	6008	0	10	2928	0	0	1	0	82.95	73.25	26	0
	6008	0	11.83333	2259	0	0	1	0	82.95	73.25	26	0
	6008	0	24.66667	2.5	0	0	1	0	82.95	73.25	26	0
	6008	0	48.08333	1.25	0	0	1	0	82.95	73.25	26	0
	6008	1	0	.	4.00E+05	4	1	0	82.95	73.25	26	0
	6008	1	0.5	1123	0	0	1	0	82.95	73.25	26	0
	6008	1	1	0151	0	0	1	0	82.95	73.25	26	0
	6008	1	1.5	9562	0	0	1	0	82.95	73.25	26	0
	6008	1	2	.145	0	0	1	0	82.95	73.25	26	0
	6008	1	3	0613	0	0	1	0	82.95	73.25	26	0
	6008	1	4	3784	0	0	1	0	82.95	73.25	26	0
	6008	1	5	3081	0	0	1	0	82.95	73.25	26	0
	6008	1	6	9708	0	0	1	0	82.95	73.25	26	0
	6008	1	8	5161	0	0	1	0	82.95	73.25	26	0
	6008	1	10	048	0	0	1	0	82.95	73.25	26	0
	6008	1	12	2.5	0	0	1	0	82.95	73.25	26	0
	6008	1	24.58333	1.25	0	0	1	0	82.95	73.25	26	0
	6008	1	48.41667	0.1	0	0	1	0	82.95	73.25	26	0
	6009	0	0	.	4.00E+05	1	0	0	56.8	64.75	59	0
	6009	0	0.516667	168	0	0	0	0	56.8	64.75	59	0
	6009	0	1	4165	0	0	0	0	56.8	64.75	59	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6009	0	1.5	5926	0	0	0	0	56.8	64.75	59	0
	6009	0	2	5809	0	0	0	0	56.8	64.75	59	0
	6009	0	3	2685	0	0	0	0	56.8	64.75	59	0
	6009	0	4	6858	0	0	0	0	56.8	64.75	59	0
	6009	0	5	9193	0	0	0	0	56.8	64.75	59	0
	6009	0	6	715	0	0	0	0	56.8	64.75	59	0
	6009	0	8	6	0	0	0	0	56.8	64.75	59	0
	6009	0	10	3724	0	0	0	0	56.8	64.75	59	0
	6009	0	12	5791	0	0	0	0	56.8	64.75	59	0
	6009	0	24	8466	0	0	0	0	56.8	64.75	59	0
	6009	0	47.91667	2.5	0	0	0	0	56.8	64.75	59	0
	6009	1	0	.	4.00E+05	4	0	0	56.8	64.75	59	0
	6009	1	0.5	7743	0	0	0	0	56.8	64.75	59	0
	6009	1	1	.144	0	0	0	0	56.8	64.75	59	0
	6009	1	1.5	.92	0	0	0	0	56.8	64.75	59	0
	6009	1	2	.444	0	0	0	0	56.8	64.75	59	0
	6009	1	3	1611	0	0	0	0	56.8	64.75	59	0
	6009	1	4	2683	0	0	0	0	56.8	64.75	59	0
	6009	1	5	5507	0	0	0	0	56.8	64.75	59	0
	6009	1	6	7431	0	0	0	0	56.8	64.75	59	0
	6009	1	8	3034	0	0	0	0	56.8	64.75	59	0
	6009	1	10	9422	0	0	0	0	56.8	64.75	59	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6009	1	11.66667	2017	0	0	0	0	56.8	64.75	59	0
	6009	1	24.66667	2.5	0	0	0	0	56.8	64.75	59	0
	6009	1	47.46667	1.25	0	0	0	0	56.8	64.75	59	0
	6010	0	0	.	4.00E+05	1	1	1	77.8	67	37	0
	6010	0	0.5	3475	0	0	1	1	77.8	67	37	0
	6010	0	1	5558	0	0	1	1	77.8	67	37	0
	6010	0	1.5	6139	0	0	1	1	77.8	67	37	0
	6010	0	2.333333	9858	0	0	1	1	77.8	67	37	0
	6010	0	3	1608	0	0	1	1	77.8	67	37	0
	6010	0	4	5628	0	0	1	1	77.8	67	37	0
	6010	0	5	1797	0	0	1	1	77.8	67	37	0
	6010	0	6	9291	0	0	1	1	77.8	67	37	0
	6010	0	8	2102	0	0	1	1	77.8	67	37	0
	6010	0	10	6551	0	0	1	1	77.8	67	37	0
	6010	0	12	1037	0	0	1	1	77.8	67	37	0
	6010	0	24.28333	2946	0	0	1	1	77.8	67	37	0
	6010	0	48.58333	2.5	0	0	1	1	77.8	67	37	0
	6010	1	0	.	4.00E+05	4	1	1	77.8	67	37	0
	6010	1	0.5	.965	0	0	1	1	77.8	67	37	0
	6010	1	1.033333	.131	0	0	1	1	77.8	67	37	0
	6010	1	1.5	7648	0	0	1	1	77.8	67	37	0
	6010	1	2	141	0	0	1	1	77.8	67	37	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6010	1	3	6034	0	0	1	1	77.8	67	37	0
	6010	1	4	5186	0	0	1	1	77.8	67	37	0
	6010	1	5	1444	0	0	1	1	77.8	67	37	0
	6010	1	6	4257	0	0	1	1	77.8	67	37	0
	6010	1	8	7076	0	0	1	1	77.8	67	37	0
	6010	1	10	0531	0	0	1	1	77.8	67	37	0
	6010	1	12	1345	0	0	1	1	77.8	67	37	0
	6010	1	24.16667	2.5	0	0	1	1	77.8	67	37	0
	6010	1	48.8	1.25	0	0	1	1	77.8	67	37	0
	6011	0	0	.	4.00E+05	1	0	0	55	66.5	29	1
	6011	0	0.5	875	0	0	0	0	55	66.5	29	1
	6011	0	1	9097	0	0	0	0	55	66.5	29	1
	6011	0	1.5	4271	0	0	0	0	55	66.5	29	1
	6011	0	2.033333	.037	0	0	0	0	55	66.5	29	1
	6011	0	3	.495	0	0	0	0	55	66.5	29	1
	6011	0	4.033333	.643	0	0	0	0	55	66.5	29	1
	6011	0	5.033333	.542	0	0	0	0	55	66.5	29	1
	6011	0	6	7534	0	0	0	0	55	66.5	29	1
	6011	0	8	9327	0	0	0	0	55	66.5	29	1
	6011	0	10	7934	0	0	0	0	55	66.5	29	1
	6011	0	12	0209	0	0	0	0	55	66.5	29	1
	6011	0	23.9	2506	0	0	0	0	55	66.5	29	1

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6011	0	47.9	2.5	0	0	0	0	55	66.5	29	1
	6011	1	0	.	4.00E+05	4	0	0	55	66.5	29	1
	6011	1	0.5	.377	0	0	0	0	55	66.5	29	1
	6011	1	1	.358	0	0	0	0	55	66.5	29	1
	6011	1	1.466667	0142	0	0	0	0	55	66.5	29	1
	6011	1	2	9627	0	0	0	0	55	66.5	29	1
	6011	1	2.95	4335	0	0	0	0	55	66.5	29	1
	6011	1	4	7807	0	0	0	0	55	66.5	29	1
	6011	1	5	5197	0	0	0	0	55	66.5	29	1
	6011	1	6	5168	0	0	0	0	55	66.5	29	1
	6011	1	8.016667	2142	0	0	0	0	55	66.5	29	1
	6011	1	10	6874	0	0	0	0	55	66.5	29	1
	6011	1	12	9791	0	0	0	0	55	66.5	29	1
	6011	1	23.96667	0229	0	0	0	0	55	66.5	29	1
	6011	1	47.41667	2.5	0	0	0	0	55	66.5	29	1
	6012	0	0	.	4.00E+05	1	0	1	56.6	65	27	1
	6012	0	0.5	8729	0	0	0	1	56.6	65	27	1
	6012	0	1	0439	0	0	0	1	56.6	65	27	1
	6012	0	1.5	0715	0	0	0	1	56.6	65	27	1
	6012	0	2	3973	0	0	0	1	56.6	65	27	1
	6012	0	3	6392	0	0	0	1	56.6	65	27	1
	6012	0	4	1092	0	0	0	1	56.6	65	27	1

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6012	0	5	6405	0	0	0	1	56.6	65	27	1
	6012	0	6	6884	0	0	0	1	56.6	65	27	1
	6012	0	8	8277	0	0	0	1	56.6	65	27	1
	6012	0	10	3254	0	0	0	1	56.6	65	27	1
	6012	0	12	9655	0	0	0	1	56.6	65	27	1
	6012	0	24	2.5	0	0	0	1	56.6	65	27	1
	6012	0	48.11667	1.25	0	0	0	1	56.6	65	27	1
	6012	1	0	.	4.00E+05	4	0	1	56.6	65	27	1
	6012	1	0.5	.167	0	0	0	1	56.6	65	27	1
	6012	1	1	.031	0	0	0	1	56.6	65	27	1
	6012	1	1.5	.904	0	0	0	1	56.6	65	27	1
	6012	1	2	4712	0	0	0	1	56.6	65	27	1
	6012	1	3	9971	0	0	0	1	56.6	65	27	1
	6012	1	4	9868	0	0	0	1	56.6	65	27	1
	6012	1	5	3413	0	0	0	1	56.6	65	27	1
	6012	1	6	3032	0	0	0	1	56.6	65	27	1
	6012	1	8	323	0	0	0	1	56.6	65	27	1
	6012	1	10	8139	0	0	0	1	56.6	65	27	1
	6012	1	12	4856	0	0	0	1	56.6	65	27	1
	6012	1	23.83333	2.5	0	0	0	1	56.6	65	27	1
	6012	1	47.75	1.25	0	0	0	1	56.6	65	27	1
	6014	0	0	.	4.00E+05	1	0	0	60.35	63.3	51	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6014	0	0.5	4752	0	0	0	0	60.35	63.3	51	0
	6014	0	1	631	0	0	0	0	60.35	63.3	51	0
	6014	0	1.5	961	0	0	0	0	60.35	63.3	51	0
	6014	0	2	4844	0	0	0	0	60.35	63.3	51	0
	6014	0	3	1976	0	0	0	0	60.35	63.3	51	0
	6014	0	4	0975	0	0	0	0	60.35	63.3	51	0
	6014	0	5	8339	0	0	0	0	60.35	63.3	51	0
	6014	0	6	2146	0	0	0	0	60.35	63.3	51	0
	6014	0	8	6833	0	0	0	0	60.35	63.3	51	0
	6014	0	10	4663	0	0	0	0	60.35	63.3	51	0
	6014	0	12	408	0	0	0	0	60.35	63.3	51	0
	6014	0	24	2295	0	0	0	0	60.35	63.3	51	0
	6014	0	47.93333	6276	0	0	0	0	60.35	63.3	51	0
	6014	1	0	.	4.00E+05	4	0	0	60.35	63.3	51	0
	6014	1	0.5	.644	0	0	0	0	60.35	63.3	51	0
	6014	1	1	.681	0	0	0	0	60.35	63.3	51	0
	6014	1	1.5	.992	0	0	0	0	60.35	63.3	51	0
	6014	1	2	.337	0	0	0	0	60.35	63.3	51	0
	6014	1	3	2675	0	0	0	0	60.35	63.3	51	0
	6014	1	4	1364	0	0	0	0	60.35	63.3	51	0
	6014	1	5	6125	0	0	0	0	60.35	63.3	51	0
	6014	1	6	3115	0	0	0	0	60.35	63.3	51	0

C		Data Desc: Raltegravir Dataset										
C	ID	TREAT	TIME	DV	AMT	EVID	sex	race	weight	height	age	HC
	6014	1	8	6622	0	0	0	0	60.35	63.3	51	0
	6014	1	10	5884	0	0	0	0	60.35	63.3	51	0
	6014	1	12	1933	0	0	0	0	60.35	63.3	51	0
	6014	1	24	2.5	0	0	0	0	60.35	63.3	51	0
	6014	1	47.95	1.25	0	0	0	0	60.35	63.3	51	0

Appendix B Notes: Header rows in this reproduction of the dataset are repeated on each page for clarity, although they are not repeated in the dataset itself.

APPENDIX C

NONMEM SUMMARY FROM FINAL MODEL

```
PDx-Pop 5.2 Multiple Estimation Method Run Summary File Run No:
analysis14
[15NOV2015 Revision]
DataFile:  RAL_FOR_NONMEM.CSV
MODEL DEFINITION:
 ADVAN4 TRANS4

TVCL=THETA(1)
CL=TVCL*EXP(ETA(1))+THETA(6)*TREAT
TVV2=THETA(2)
V2=TVV2*EXP(ETA(2))
TVQ=THETA(3)
Q=TVQ*EXP(ETA(3))
TVV3=THETA(4)
V3=TVV3
TVKA=THETA(5)
KA=TVKA*EXP(ETA(4))
S2=V2

Y = F + F*ERR(1)
IPRE=F
W1= 1
W2= F
IRES= DV-IPRE
IWRE=IRES/(W1+W2)

TABLES CREATED:
 analysis11.tab
 PATABanalysis11
 COTABanalysis11
 CATABanalysis11
 SDTABanalysis11
 analysis11.par
 analysis11.eta

Estimation Methods:

1. Results for First Order Conditional Estimation with
Interaction

TERMINATION STATUS:
 MINIMIZATION SUCCESSFUL
 NO. OF FUNCTION EVALUATIONS USED:      323
 NO. OF SIG. DIGITS IN FINAL EST.:  3.5

 ETABAR IS THE ARITHMETIC MEAN OF THE ETA-ESTIMATES,
```

AND THE P-VALUE IS GIVEN FOR THE NULL HYPOTHESIS THAT THE TRUE MEAN IS 0.

ETABAR:	-7.0919E-03	2.6709E-02	-2.3738E-02	1.3404E-02
SE:	1.4680E-01	1.4845E-01	1.1929E-01	4.8315E-02
N:	12	12	12	12
P VAL.:	9.6147E-01	8.5722E-01	8.4227E-01	7.8145E-01
ETAshrink(%):	1.0000E-10	1.3736E+01	3.5790E+00	1.7391E+01
EBVshrink(%):	2.4986E+00	1.4505E+01	1.0372E+01	2.3914E+01
EPSshrink(%):	3.3437E+00			

MINIMUM VALUE OF OBJECTIVE FUNCTION: 3145.104
COVARIANCE STEP SUCCESSFUL

	FINAL ESTIMATE	%RSE	95% CONFIDENCE INTERVAL LBOUND	UBOUND		DESCRIPTOR/ VARIABILITY
THETA						
1	76.6	18.3%	49.2	104		CL
2	95.9	25.5%	47.9	144		V2
3	11.5	27.2%	5.37	17.6		Q
4	374	39.6%	83.9	664		V3
5	0.597	13.1%	0.443	0.751		KA
6	50.8	12.6%	38.3	63.3		TRT
OMEGA					INTERINDIVIDUAL VARIABILITY	
1,1	0.274	33.4%	0.0947	0.453	CV =	52.3%
2,2	0.388	55.2%	-0.0314	0.807 *	CV =	62.3%
3,3	0.200	76.5%	-0.0999	0.500 *	CV =	44.7%
4,4	0.0448	82.8%	-0.0279	0.118 *	CV =	21.2%
SIGMA					RESIDUAL VARIABILITY	
1,1	0.422	12.7%	0.317	0.527	CV =	65.0%

*Indicates 95% confidence interval that includes zero
%RSE is percent relative standard error (100% x SE/EST)

Akaike Information Criterion: 3167.1
Schwarz Bayesian Criterion: 3208.24

APPENDIX D

NONMEM CODE FOR FINAL MODEL

```
;Model Desc: two cmt additive shift trt cl
;Project Name: raltest
;Project ID: NO PROJECT DESCRIPTION

;Project ID: NO PROJECT DESCRIPTION

$PROB RUN# analysis11
$INPUT C ID TREAT TIME DV AMT EVID sex race weight height age HC
$DATA RAL_FOR_NONMEM.CSV IGNORE=C
$SUBROUTINES ADVAN4 TRANS4
$PK
   TVCL=THETA(1)
   CL=TVCL*EXP(ETA(1))+THETA(6)*TREAT
   TVV2=THETA(2)
   V2=TVV2*EXP(ETA(2))
   TVQ=THETA(3)
   Q=TVQ*EXP(ETA(3))
   TVV3=THETA(4)
   V3=TVV3
   TVKA=THETA(5)
   KA=TVKA*EXP(ETA(4))
   S2=V2

$ERROR
   Y = F + F*ERR(1)
 IPRE=F
 W1= 1
 W2= F
 IRES= DV-IPRE
 IWRE=IRES/(W1+W2)

$EST METHOD=1 INTERACTION PRINT=5 MAX=9999 SIG=3
MSFO=analysis11.msf
$THETA
  (0, 104);[CL]
  (0, 102);[V2]
  (0, 14.5);[Q]
  (0, 668);[V3]
  (0, 0.481);[KA]
  5;[trt]
$OMEGA
  0.04 ;[P] omega(1,1)
  0.04 ;[P] omega(2,2)
  0.04 ;[P] omega(3,3)
  0.04 ;[P] omega(4,4)
$SIGMA
  0.04 ;[P] sigma(1,1)
```

```
$COV
$TABLE ID TIME IPRE ONEHEADER NOPRINT FILE=analysis11.tab
$TABLE ID TIME CL V2 Q V3 KA ONEHEADER NOPRINT
FILE=PATABanalysis11
$TABLE ID ONEHEADER NOPRINT FILE=COTABanalysis11
$TABLE ID ONEHEADER NOPRINT FILE=CATABanalysis11
$TABLE ID IPRE ONEHEADER NOPRINT FILE=SDTABanalysis11
$TABLE ID CL V2 Q V3 KA FIRSTONLY NOAPPEND NOPRINT
FILE=analysis11.par
$TABLE ID ETA1 ETA2 ETA3 FIRSTONLY NOAPPEND NOPRINT
FILE=analysis11.eta
$SCAT DV VS PRED UNIT
$SCAT WRES VS PRED
```