Journal of Pharmacokinetics and Biopharmaceutics, Vol. 23, No. 6, 1995

Three New Residual Error Models for Population PK/PD Analyses

Mats O. Karlsson,[1,3,4] Stuart L. Beal,[2] and Lewis B. Sheiner[1,2]

Received February 2, 1994—Final November 30, 1995

Residual error models, traditionally used in population pharmacokinetic analyses, have been developed as if all sources of error have properties similar to those of assay error. Since assay error often is only a minor part of the difference between predicted and observed concentrations, other sources, with potentially other properties, should be considered. We have simulated three complex error structures. The first model acknowledges two separate sources of residual error, replication error plus pure residual (assay) error. Simulation results for this case suggest that ignoring these separate sources of error does not adversely affect parameter estimates. The second model allows serially correlated errors, as may occur with structural model misspecification. Ignoring this error structure leads to biased random-effect parameter estimates. A simple autocorrelation model, where the correlation between two errors is assumed to decrease exponentially with the time between them, provides more accurate estimates of the variability parameters in this case. The third model allows time-dependent error magnitude. This may be caused, for example, by inaccurate sample timing. A time-constant error model fit to time-varying error data can lead to bias in all population parameter estimates. A simple two-step time-dependent error model is sufficient to improve parameter estimates, even when the true time dependence is more complex. Using a real data set, we also illustrate the use of the different error models to facilitate the model building process, to provide information about error sources, and to provide more accurate parameter estimates.

KEY WORDS: population PK/PD; residual error; intraindividual variability; autocorrelation; replicates; NONMEM.

INTRODUCTION

In pharmacokinetic data analysis using nonlinear regression, the residual error is often considered to have properties similar to those of assay error. Historically, it is easy to understand why assay error was considered

This work was supported by grants GM26676 and GM2669.

[1]Department of Pharmacy, School of Pharmacy, University of California, San Francisco, California.

[2]Department of Laboratory Medicine, School of Medicine, Room C-255, University of California, San Francisco, California 94143-0626.

[3]Division of Biopharmaceutics and Pharmacokinetics, Faculty of Pharmacy, Uppsala University, Box 580, S-751 23 Uppsala, Sweden.

[4]To whom correspondence should be addressed.

0090-466X/95/1200-0651$07.50/0

the main source, or even sole source, of data error. In the early days of pharmacokinetic (PK) studies and nonlinear regression, assays were generally less precise than they are today, and studies performed were almost exclusively well-controlled experiments using a homogeneous study group. Today many pharmacokinetic analyses are performed using data gathered from observational studies of heterogeneous patient groups, often studied at several different sites. The residual error of the fit is usually considerably greater than assay error in such population analyses (Table I). This implies that residual error may not have the simple statistical properties usually associated with assay error: independence, and error magnitude proportional only to the underlying concentration. The larger and more heteroscedastic the residual error, the more important it may be to account for it properly. Moreover, we know that several of the standard error model assumptions may not be true. Specifically:

1. Replicate determinations are treated as having the same error structure as nonreplicate observations. We expect, however, that replicates will deviate from each other to a lesser extent than will nonreplicate observations. What constitutes a replicate may differ depending on the type of experiment. For a PK study, it could be multiple blood samples at (essentially) the same time or duplicate preassay sample clean-up. For a pharmacodynamic (PD) study, it could be the multiple measurements of a PD effect such as blood pressure at (essentially) the same time.

2. Residual errors are treated as independent, although correlation often can be observed under fitted models.

3. Residual error is rarely modeled as depending directly on covariates (other than through the expected response), yet different degrees of structural model misspecification for different portions of a response–time profile may well lead to different magnitudes of residual error over time, despite similar values for the expected response.

In this study, we investigate the possibly detrimental effects on the parameter estimates of population models of considering all error to have simple properties (i.e., current practice). We also suggest extensions to the present models for residual error that can be used to modify current practice when necessary. Although the investigation focuses on population analyses, the extensions suggested are equally applicable to single-subject data. Also, although the examples are PK examples, the extensions suggested can apply as well to PD examples.

THEORY

Residual Error with Replicates

In population analyses the residual variability is a measure of the difference (or error) between the observations and their subject-specific

Table I. Survey of Population Analyses

Drug	Residual error magnitudes (% *CV*)			First author	Journal	Year	Vol.	Pages
	Estimated[a]	Assay[b]	Nonassay[c]					
Alfentanil	25	NA	24	Maitre	Anesthesiology	1987	66	3–12
Alprazolam	27–59	NA	26–57	Antal	Clin. Pharmacol. Ther.	1989	46	552–559
Bisoprolol	24–41	NA	23–40	Grevel	Clin. Pharmacokin.	1989	17	53–63
Didanosine	40–50	NA	39	Pai	J. Clin. Pharmacol.	1992	32	242–247
Digoxin	26	<10	25	Williams	Clin. Pharmacol. Ther.	1992	22	66–74
Digoxin	24	<10	23	Yukawa	J. Pharm. Pharmacol.	1992	44	761–765
Epirubicin	42	NA	41	Wade	Cancer Chem. Pharmacol.	1992	29	391–395
Flurbiprofen	40	NA	39	Aarons	J. Pharmacokin. Biopharm.	1991	19	265–269
Gentamicin	12	<10	11	Jenson	Pharmacother.	1992	12	178–182
Imipramine	25	7	25	Grasela	Clin. Pharmacol. Ther.	1987	42	433–441
Imipramine	16	<10	15	Tamayo	Eur. J. Clin. Pharmacol.	1992	43	89–92
Indomethacin	19	<5	18	Wiest	Clin. Pharmacol. Ther.	1991	49	550–557
Lisinopril	28	NA	27	Thomson	Br. J. Clin. Pharmacol.	1989	27	57–65
Lithium	16	2	15	Jermain	Clin. Pharm.	1991	10	376–381
Metoclopramide	13	9	12	Grevel	Clin. Pharmacokin.	1988	14	52–63
Mitoxantrone	21	NA	20	Launay	J. Pharm. Sci.	1989	78	877–880
Ofloxacin	7–150	4–8	7–148	Yuen	Pharmacother.	1992	12	88–92
Phenobarbitone	20	<10	19	Yukawa	J. Pharm. Pharmacol.	1992	44	755–760
Pseudoephedrine	11–30	NA	10–128	Graves	Pharmacotherapy	1989	9	372–376
Quinidine	22	4	21	Fattinger	Br. J. Clin. Pharmacol.	1991	31	279–286
Theophylline	9–60	6	9–59	Karlsson	TDM	1991	13	195–200
Theophylline	8–177	4	8–177	Moore	J. Pharmacokin. Biopharm.	1989	17	47–66
Thiopental	24	NA	23	Stanski	Anaesthesiology	1990	72	412–422
Tobramycin	21	7	21	Aarons	Br. J. Clin. Pharmacol.	1989	28	305–314
Zidovudine	20	NA	19	Collart	Dev. Pharmacol. Ther.	1992	18	71–80
Zidovudine	86	NA	84	Gitterman	Clin. Pharmacol. Ther.	1990	48	161–167

[a] If a proportional error model was used a single figure, representing the %*CV* is given. If any other error model was used, the approximate *CV*s at the lowest and highest observed concentrations are given.

[b] Assay error is given as a single figure (or NA when assay performance figures are not available in the reference) or a range, the latter when assay performance at more than one concentration is given.

[c] Nonassay error is calculated assuming no covariance between assay and nonassay error (i.e., by taking the difference between estimated and assay error on the variance scale). When no information is available, assay error is assumed to have a *CV* of 10%.

predictions. Assuming additive error, the observed concentration for the ith subject at the jth time point can be written.

$$y_{ij} = \hat{y}_{ij} + \varepsilon_{ij} \tag{1}$$

where $\hat{y}_{ij}$ is the subject-specific prediction and ε_{ij} is the error. By taking replicate measurements at a single time point it is possible to some extent to differentiate between sources of error. Variability between replicates arises from assay error and the error introduced by sample handling from the point in the handling chain where the replication is made. We will assume replication is made at the point of sampling. There are, of course, other sources of nonassay error, including imprecisely recorded sampling times and model misspecification. Assuming that at the jth time point, replicate observations are made, Eq. (1) may be extended as follows:

$$y_{ijk} = \hat{y}_{ij} + \varepsilon_{ij} + \varepsilon_{ijk} \tag{2}$$

The difference between predicted and observed concentrations at the jth time now has two components, a consistent difference (ε_{ij}) between all replicates and the prediction, and replicate-specific differences (ε_{ijk}). We will assume that ε_{ij} has 0 mean and constant variance σ_c^2 and that the ε_{ijk} have 0 mean and constant variance σ_r^2. The variance σ_c^2 incorporates all sources of variability that are constant across replicates. One may wish to consider different models for the probability distributions of the two components. Whereas we assume that the ε_{ijk} are time-invariant and uncorrelated, the possibility of time-variant and correlated errors will be considered for the ε_{ij}.

Auto-correlated Errors

In the case that a pharmacokinetic model is misspecified, and frequent sampling is done, the data often show regions of consistent deviations (of the same sign) between the predicted and observed concentrations. One of the many models that have been suggested for describing such time-correlated errors is the AR(1) model (1), the inclusion of which into linear (2) and nonlinear (3) mixed-effects models has previously been suggested. In the AR(1) model the positive correlation between two errors, ε_{t_1} and ε_{t_2}, decreases exponentially with the time separating the two observations according to

$$\text{corr}(\varepsilon_{t_1}, \varepsilon_{t_2}) = \exp(-|t_1 - t_2| / t_{\text{corr}}) \tag{3}$$

where t_{corr} is a constant determining how fast the correlation decreases with time. In what follows, this model is used, but other models, involving, for

example, a hyperbolic or biexponential decrease with time, may be equally appropriate.

In a final model (obtained after some model building) some serial correlation of errors may be considered unavoidable. In the context of population PK data analysis, little is known about the impact of the degree of correlation on parameter estimates when the correlation is ignored. With single-subject data it is known that by ignoring correlation in the errors, the precision of parameter estimates can be overestimated (4) and that standard errors can be underestimated (5).

To diagnose model misspecification, examination of residuals is the standard approach. Serially correlated residuals may indicate misspecification of at least some part of the model. However, it is not easy to make plots for population data sets, with many subjects and complex dosing and sampling patterns, that can reveal serial correlation. The use of a model that specifically allows the correlation to be estimated can therefore be of diagnostic value.

Time-Dependent Error Variance

Sources of variability such as erroneous dosing and sampling history and model misspecification may in PK data introduce errors whose variances are time-dependent (in addition, perhaps to depending on the expected concentration). Inappropriately recorded sampling times introduce larger errors when the concentrations are rapidly changing than when they are not. Pharmacokinetic absorption models are generally less well specified than disposition models, resulting in potentially larger error magnitudes in the ascending limb of a concentration–time curve after an oral dose, than in the descending limb (even at points where the expected concentrations are similar). The error variance of PD data can also be time-dependent, an example may be PD data from *in situ* animal models, which tend to be less reliable as time progresses. In neither of these situations is time the error source per se, but time can be used as a surrogate variable. When only single dose or steady-state PK data are considered, as in the present investigation, the modeling of error magnitude as a function of time, or time after dose, is straightforward. If more complex dosing patterns are present, it may be difficult to find a relationship with time that can adequately describe time-dependent variances, although it has been suggested that $d\hat{y}/dt$ be used as a surrogate variable (where $\hat{y}$ is already a function of t) (6).

Of the many possible models for describing the time dependency of error variance that one might propose, we choose to use a simple step function where the time point for the change in variance is either fixed or estimated.

METHODS

Simulations

To investigate previously used error models, and the new ones proposed here, simulations of nine different conditions were performed. These conditions can be grouped into three groups of three. Simulations CONT1, CONT2, and CONT3 represent control simulations, made to study the estimates of σ_r^2 and σ_c^2 under basic conditions. Simulations DUPL, CORR, and TIME represent positive controls of the more elaborate error models and are made using extensive duplication of samples (DUPL), autocorrelated errors (CORR), and time-varying error magnitude (TIME). Simulations SAMPL, BIEXP, and ERRAT represent attempts to mimic the type of errors that may occur in real data sets due to error in sampling times (SAMPL), error in the specification of the input model (BIEXP), and time-varying pharmacokinetic parameters (ERRAT). The basic characteristics of the simulations are as follow. Differences from these basic conditions are discussed later.

A one-compartment disposition model with first-order absorption and with the following parameters for the typical individual is used: clearance ($\tilde{CL}$) = 10 L/hr, volume of distribution ($\tilde{V}$) = 100 L, first-order absorption rate constant ($\tilde{k}_a$) = $1 \cdot hr^{-1}$. The individual parameter values (P_i, for example) are obtained using a log-normal distribution around the population (geometric) average value ($\tilde{P}$), according to

$$P_i = \tilde{P} \exp(\eta_i^P) \tag{4}$$

where η^P is a normally distributed, independent random variable with zero mean and variance ω_P^2. The values used for ω_{CL}, ω_V, and ω_{ka} are 0.3. The residual error is of the form

$$y_{ijk} = \hat{y}_{ij} \exp(\varepsilon_{ij} + \varepsilon_{ijk}) \tag{5}$$

where the random variables, ε_{ij}, and ε_{ijk} are independent with variances σ_c^2 and σ_r^2, respectively. A value of 0.1 is used throughout for σ_r, whereas σ_c varies, but has a basic value of zero. Each population data set consists of samples obtained at 16 time points (0.3, 0.6, 1, 1.5, 2, 2.5, 3, 4, 5, 6, 8, 10, 12, 16, 20, and 24 hr) after a single extravascular dose, from each of 16 subjects. With each individual, only one duplicate observation is generated. The time at which this duplicate is obtained differs between individuals. For each of the nine simulation conditions, 100 data sets are generated.

What follows is a specification of the difference between each of the nine simulations and the basic conditions. Whereas CONT1 is generated entirely under the basic conditions, CONT2 and CONT3 have values of σ_c of 0.1 and 0.2, respectively. In simulation DUPL, σ_c is 0.2 and duplicate

observations are generated at each sampling time, rather than at just one time per individual. Simulation CORR is the only simulation where a correlation according to Eq. (2) is included. The values of σ_c and t_{corr} are 0.2 and 2.86, respectively. The latter value was chosen as to yield a high (0.9) correlation between the errors of the early adjacent observations. The simulation TIME differs from the basic conditions in that the magnitude of σ_c varies over time between 0.6 at early times to 0.2 at late times (the exact shape of this time-varying function is given in Fig. 3).

In the last three simulations, ε_{ij} errors as such are not generated, i.e., $\sigma_c = 0$. In the simulation SAMPL, data are generated in the following manner: (i) to each of the sampling times given above (nominal times) a uniformly distributed random number, in the range -0.15 and $+0.15$ is added, and (ii) predictions are generated using the new times. In the generated data set, the nominal, rather than the actual times are given as covariate values. The input model in the simulation BIEXP is not a first-order process, but the sum of two parallel first-order processes (each accounting for half the dose), with values for $\tilde{k}_a$ of 0.4 and 2 hr^{-1}. The value for η^{ka} is the same for the two processes, to ensure that the input profile in each subject is truly biexponential rather than near-monoexponential as may occasionally occur otherwise. In simulation ERRAT the value of k_a is varied within an individual over time to mimic erratic absorption. For each new subject and each time interval between two observations, a new k_a is generated according to

$$k_a = \tilde{k}_a \exp(\eta^{ka} + ran) \tag{6}$$

where *ran* is a normally distributed random variable with mean 0 and variance 1 (different for each subject and each time interval).

The data, as described above, are all analyzed using four models that differ only in the way ε_{ij} is modeled. The structural model and the models for interindividual error and ε_{ijk} are identical to those of the basic simulation model. The different ε_{ij} models for the analyses are (i) the traditional error model where σ_c is set to zero, (ii) a model where σ_c is estimated, (iii) a model where the ε_{ij} are autocorrelated according to Eq. (3), and (iv) a model where two values for σ_c are estimated, each value governing the error on one side of the breakpoint of a step function in time. In this model, the breakpoint occurs after absorption is virtually complete (at 5 ln $2/\tilde{k}_a$). Data sets from the simulations TIME and ERRAT are also analyzed using a model in which σ_c is given by a natural cubic spline (7) in time. Splines are defined by the number and location of breakpoints and by the height of the function at these breakpoints. The number of breakpoints is determined based on the Akaike criterion (8), and the location of internal breakpoints is set so that an equal number of observation times lie between two breakpoints (but note that the observations themselves may not be spaced

uniformly over time). The heights of the spline function are estimated simultaneously with the other population parameters.

For all analyses, the first-order method as implemented in the program NONMEM (9), version V, is used. With version V, autocorrelated errors can be handled. The spline calculations in NONMEM were made using the B-spline package, based on PPPACK (7) (available from D. Verotta, Box 0446, University of California, San Francisco, CA 94123, upon request). The performance of parameter estimation is reported with the means and SDs of the estimates. For some simulations, the means of the SEs, as provided by the program, are presented. Individual parameter estimates, used to make individual-specific predictions for residual plots, are obtained as empirical Bayes estimates, which are available in NONMEM, versions IV and later.

Real Data Set

An illustration of the use of the different residual error models is given using a real data set, previously described (10). The drug is an antihypertensive agent administered to 64 patients orally twice daily. After at least 4 weeks of continuous therapy a concentration–time profile (samples at 1, 2, 3, 4, 5, 6, 8, 10, and 12 hr postdose) was obtained from each patient. A second profile was obtained 4 weeks or more after the first one from each of 35 patients. The population model used to describe the data is a one-compartment model with first-order input and with some covariates (height, age, race, and presence/absence of concomitant therapy with hydrochlorothiazide) influencing the expected values for clearance and/or volume of distribution. Proportional error models are used to describe both the random interindividual and residual variability. The population model developed by Mandema *et al.* (10) is taken to be the basic model in this study, to which we add extensions to the residual error model. Since replicates are not available in this data set, residual error cannot be separated into two components. For the same reason, the autocorrelation and the time-dependent error models are used for total residual error. In addition, a model that takes into account variability in pharmacokinetic parameter values between different studies within a subject is incorporated into the model (11). In this model the value of a parameter P for subject i during study j can be described by

$$P_{ij} = \tilde{P} \exp(\eta_i^P + \kappa_{ij}^P) \tag{7}$$

where η_i^P is as before, and κ_{ij}^P is a random variable with variance π_P^2.

Table II. Estimation of Residual Error Components

	Simulation		
	CONT1	CONT2	CONT3
σ_r^2			
True value	0.01	0.01	0.01
Mean of estimates	0.01	0.01	0.01
SD of estimates	0.0029	0.0045	0.0054
σ_c^2			
True value	0.0	0.01	0.04
Mean of estimates	0.0022	0.014	0.048
SD of estimates	0.0026	0.0054	0.010

RESULTS

Replication Error

Results of using the different models to analyze the simulated data sets are shown in Tables II and III. When σ_c^2/σ_r^2 is small, σ_c^2 is imprecisely estimated. This is illustrated by the results presented in Table II. On the other hand, σ_r^2 can be estimated even with as few as one replicate per subject, and the magnitude of σ_c^2 does not have a major influence on the precision of the estimate. It can be noted that in the control simulations the estimated misspecification error (0.0022, 0.014, and 0.048) is slightly larger than the comparable values used in the generation of data (0, 0.01, and 0.04). These differences are larger than are explainable by sampling error, and they may

Table III. Parameter Estimates for Simulations CONT3 and DUPL

Parameter	True value	σ_r		$\sigma_r+\sigma_c$	
		$\bar{x}$	SD	$\bar{x}$	SD
		CONT3			
CL	10	9.7	0.8	9.7	0.8
V	100	101	9	101	9
ka	1	1.1	0.2	1.1	0.2
ω_{CL}^2	0.09	0.08	0.04	0.08	0.04
ω_V^2	0.09	0.07	0.03	0.07	0.03
ω_{ka}^2	0.09	0.15	0.17	0.15	0.17
σ_r^2	0.01	0.057	0.008	0.01	0.005
σ_c^2	0.04			0.048	0.010
		DUPL			
CL	10	9.7	0.8	9.6	0.8
V	100	100	9	101	9
ka	1	1.1	0.2	1.1	0.2
ω_{CL}^2	0.09	0.09	0.04	0.08	0.04
ω_V^2	0.09	0.08	0.03	0.07	0.03
ω_{ka}^2	0.09	0.20	0.19	0.14	0.14
σ_r^2	0.01	0.055	0.008	0.011	0.001
σ_c^2	0.04			0.047	0.008

be due to the presence of approximations with the first-order method used in these analyses.

Even when a 20% misspecification error is included in the data (simulations CONT3 and DUPL) there is still no appreciable difference in the parameter estimates for fixed or interindividual random effects (Table III), nor in their estimated SEs (not shown), between those obtained with the simple model and those obtained with the more elaborate model involving both σ_c^2 and σ_r^2.

Autocorrelated Errors

Simulation BIEXP explores the case that a simpler structural model is used to analyze the data than the true one, while simulation CORR explores the case that autocorrelation is actually present in the data. In neither case does the inclusion of an autocorrelation term in the analysis model result in markedly different estimates of the fixed effects parameters. In contrast, the estimates of the random effects parameters, especially ω_{ka} and σ_c, depend considerably on the error model uses (Table IV). The correlation model used in the CORR simulation specifies less correlation between neighboring observations the greater the distance between them. This could explain the

Table IV. Parameter Estimates for Simulations CORR and BIEXP

Parameter	True Value	$\sigma_r+\sigma_c$			σ_r+autocorrelated σ_c		
		$\bar{x}$	SD	$\bar{x}$ SE	$\bar{x}$	SD	$\bar{x}$ SE
				CORR			
CL	10	9.4	0.84	0.72	9.7	0.83	0.73
V	100	105	11	9.2	101	10	8.6
ka	1	0.92	0.17	0.13	1.15	0.21	0.15
t_{corr}	2.86				4	2.6	1.9
ω^2_{CL}	0.09	0.090	0.038	0.031	0.076	0.039	0.031
ω^2_V	0.09	0.10	0.048	0.037	0.062	0.045	0.034
ω^2_{ka}	0.09	0.45	0.37	0.26	0.13	0.18	0.11
σ^2_c	0.04	0.026	0.008	0.006	0.060	0.023	0.018
σ^2_r	0.01	0.011	0.006	0.004	0.011	0.004	0.003
				BIEXP			
CL	10	10.0	1.1	0.93	10.4	0.93	0.85
V	100	121	12	11	122	10	9
ka	1	1.5	0.12	0.10	1.7	0.15	0.14
t_{corr}					15	7.4	5.6
ω^2_{CL}	0.09	0.10	0.049	0.041	0.086	0.053	0.039
ω^2_V	0.09	0.089	0.050	0.034	0.060	0.037	0.030
ω^2_{ka}	0.09	0.051	0.042	0.029	0.024	0.024	0.014
σ^2_c		0.0055	0.0038	0.0029	0.036	0.018	0.014
σ^2_r	0.01	0.010	0.0038	0.0030	0.0090	0.0016	0.0015

greater estimation error with ω_{ka} than with ω_{CL} or ω_V when the correlation model is not used in the analysis, since the latter two parameters rely on information during the entire time span, while the former relies entirely on the early, more closely spaced samples. Interindividual variability is overestimated when autocorrelation is present but not accounted for. The underlying reason for this can be understood by considering the extreme case of no interindividual variability, but high autocorrelation. If such data were to be analyzed by a model without autocorrelation, consistent differences between profiles caused by highly serially correlated errors would be interpreted as real differences between individuals. Thus, overestimation of interindividual variability would result. When the correct model is used in the analysis of simulations CORR and BIEXP, some deviations from the true parameter values can still be observed, particularly in σ_c and t_{corr}. Again, these may be due to the first-order approximation which induces a misspecification error.

The patterns in NONMEM-computed SEs reflect those in the SDs of the parameter estimates (Table IV). The SEs consistently, but not greatly, underpredict the SDs of the parameter estimates. The accuracy of the SEs is similar in both models that incorporate autocorrelation and in models that do not. Thus, underprediction of SEs, which, as mentioned in the Theory section, can occur in the analyses of single-subject data when true autocorrelation is ignored, does not seem to be as great a problem in population analyses.

Time-Dependent Error Variance

The fit of the data from simulations TIME and ERRAT using the traditional error model, sometimes leads to grossly overestimated values for k_a and ω_{ka} (Table V). Although some overestimation remains when the two-step error model is used, this error model generally leads to more accurate and precise estimates of all parameters. An indication of the need for a model with time-dependent error variance can be obtained from residual analyses. The trends in the variance-adjusted residuals in one data set each from simulations TIME and ERRAT, analyzed with the traditional error model, are shown in Figs. 1 and 2, together with the relative absence of such trends when the two-step error model is used. When, instead of the two-step function, a spline is used to model the error variance over time, further improvement of the fit is obtained. For both the TIME and ERRAT simulations, the most appropriate number of breakpoints for the spline varies among data sets in the range of 3–5. The approximation by the two-step function, and by a 4-breakpoint spline, of the function used in the generation of σ_c for the TIME simulation is shown in Fig. 3. Despite the slightly better

Table V. Parameter Estimates for Simulations TIME, SAMPL, and ERRAT

Parameter	True value	σ_r		σ_r + time variant σ_c	
		$\bar{x}$	SD	$\bar{x}$	SD
		TIME			
CL	10	10.6	1.0	9.7	0.8
V	100	95	12	99	9
ka	1	90	>99	1.5	0.7
ω^2_{CL}	0.09	0.11	0.07	0.08	0.04
ω^2_V	0.09	0.10	0.04	0.08	0.04
ω^2_a	0.09	>99	>99	0.32	0.71
σ^2_r	0.01	0.11	0.03	0.010	0.006
σ^2_c	see Fig. 3			see Fig. 3	
		SAMPL			
CL	10	10.0	0.8	10.1	0.8
V	100	102	9	102	9
ka	1	1.3	0.2	1.1	0.2
ω^2_{CL}	0.09	0.09	0.04	0.09	0.04
ω^2_V	0.09	0.08	0.03	0.08	0.03
ω^2_{ka}	0.09	0.14	0.12	0.16	0.18
σ^2_r	0.01	0.015	0.002	0.008	0.002
σ^2_c				not shown	
		ERRAT			
CL	10.4	1.0	10.1	0.8	
V	101	9	102	9	
ka	6.6	21	1.8	0.4	
ω^2_{CL}	0.10	0.09	0.09	0.04	
ω^2_V	0.09	0.04	0.08	0.03	
ω^2_{ka}	31	>99	0.50	0.24	
σ^2_r	0.01	0.024	0.011	0.009	0.003
σ^2_c				not shown	

fit with the spline, the precision and accuracy of parameter estimates using either approximation, for both the TIME and ERRAT simulations, are the same (not shown).

In three simulations, SAMPL, BIEXP, and ERRAT, model misspecification is introduced in a manner designed to mimic model misspecification as it occurs in real data. The mean of the σ_r in these three simulations is 0.12, 0.12, and 0.15, respectively, when the traditional error model is used. The data are generated with $\sigma_r = 0.1$ and the mean of the σ_r in the corresponding control simulation (CONT1) is 0.11. Thus, the increases in residual error due to the various misspecifications are small compared to the differences normally seen between the assay error and the total residual error (see Table I).

Real Data Example

The basic population model used for this data set is a one-compartment model with first-order absorption. Several covariates influence the typical values of CL and V, but as these relationships are of little interest in the context of this paper, only the estimated values for the typical covariate

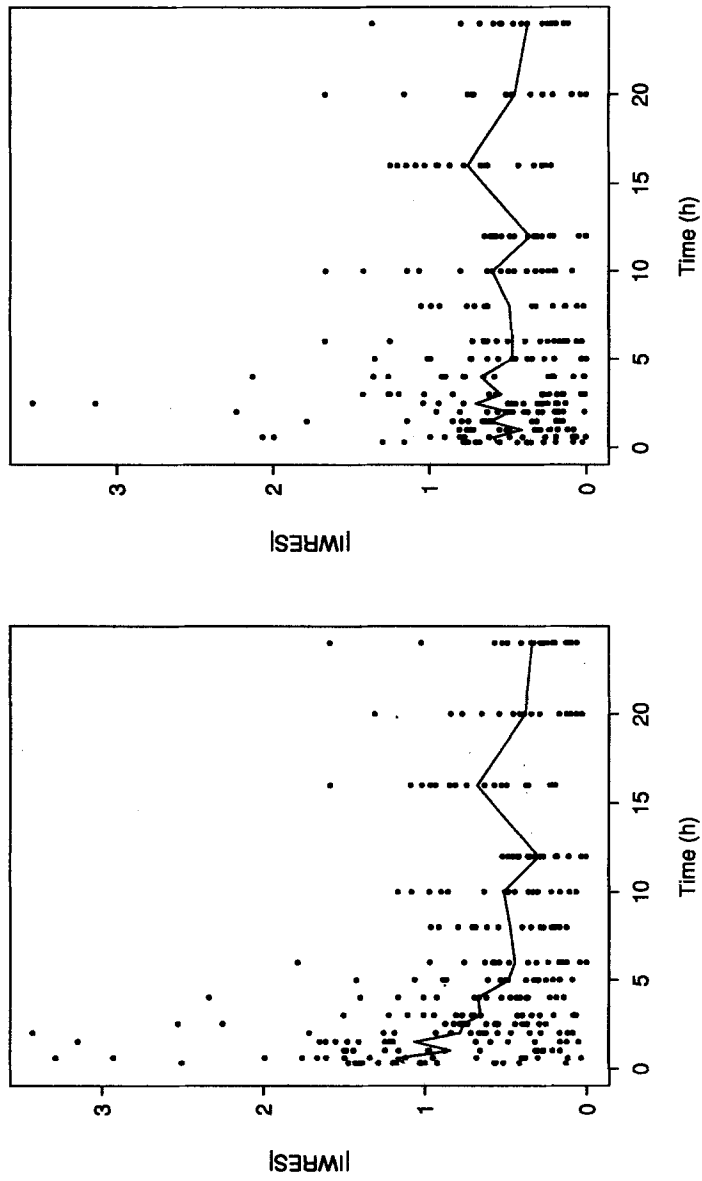

Fig. 1. Residual plots for a randomly selected simulated data set of type TIME. Left: analysis with traditional error mode ($\sigma_c = 0$). Right: analysis with an error model in which the standard deviation of ε_{ij} is specified by a two-step function. The abscissa (|IWRES|) represents the absolute difference between subject-specific predictions and the observation, divided by the estimated standard deviation of the observation. The continuous line is a smooth ("supsmu"-function in S-PLUS (12) with default conditions) through the points.

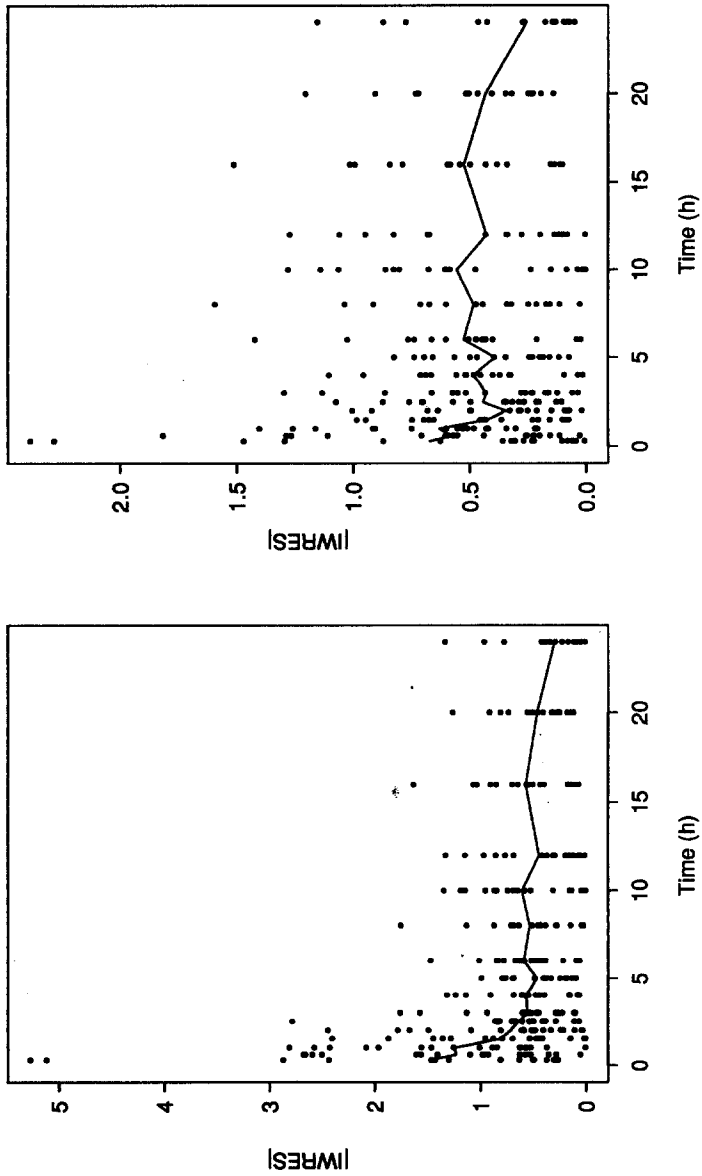

Fig. 2. Residual plots for a randomly selected simulated data set of type ERRAT. For details see legend to Fig. 1.

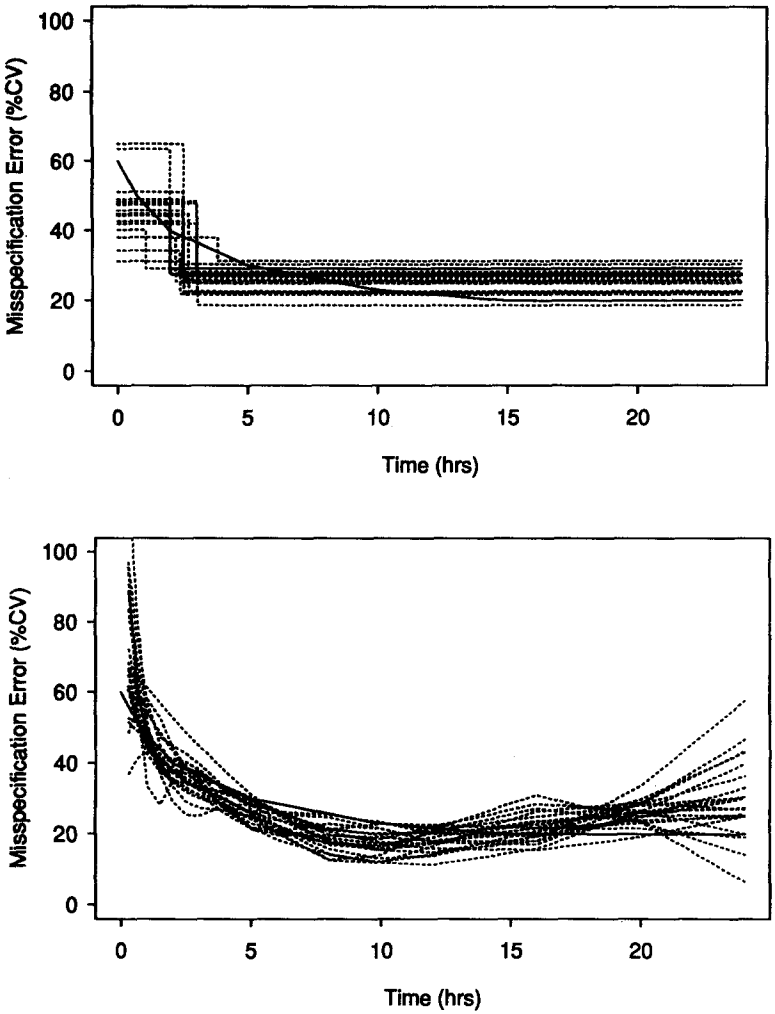

Fig. 3. The standard deviation of ε_{ij} versus time from the analyses of 20 randomly chosen data sets from the TIME simulation is presented. The broken lines in the top graph shows σ_c versus time, as estimated using the two-step error model. The corresponding lines in the lower graph show estimated σ_c versus time, using a spline function with four breakpoints. In both graphs the continuous line represents the function actually used for σ_c in the simulation.

values are presented. The runs described below are meant to exemplify the application of the autocorrelation model and the model for time-dependent error magnitude as diagnostics and for obtaining final parameter estimates.

Residual analyses of the fit using the basic model (Run 1) do not display any obvious model misspecification, but when the residual error model is

extended to include autocorrelation (Run 2) a significant drop in the objective function occurs (Table VI). The magnitude of the residual correlation coefficient between nearby observations (those separated by 1 hr) is 0.6. A type of model misspecification that can cause such a high correlation, without being evident from traditional residual analyses, is random variability in one or several pharmacokinetic parameters over time. When, as with the real data set, we use the observations that are collected on distinct occasions, separated by relatively long periods with no observations, it is reasonable to believe that most of the random variability over time within an individual may be attributable to differences in his pharmacokinetic parameters between the study occasions. When a model that accounts for random variability in CL and V between study occasions in included in the analysis (Run 3), a better fit results. Not only is the objective function value markedly lower (Table VI), but σ_r decreases from 0.20 to 0.13, and the correlation between nearby observations drops from 0.60 to 0.18. In fact, omitting the autocorrelation part of the model after the inclusion of interoccasion variability (Run 4) results in little change in either the parameter estimates or the objective function value.

Absorption is relatively rapid in the real data, and early observations are made at hourly intervals. A plot of the error magnitude versus time (Fig. 4) gives some guidance as to whether a time-dependent model might be appropriate. Three different two-step models for time-varying residual error were tried, with the breakpoints after the first, second, and third sampling time. Use of any of the three models results in a significant decrease of the objective function, the largest decrease being associated with the model with a breakpoint after the first sample (Run 5, Table VI). With this error model, a further trend in error magnitude with time is no longer observed (Fig. 4). As noted previously with the simulations, the introduction of a time-dependent residual error results in marked decreases in the estimates of k_a, $\omega_{k\alpha}$, and π_{ka}.

The variation of the error magnitude over time was further explored using a spline (Fig. 5). The parameter estimates using the most flexible spline the data would support (Run 6) were similar to those obtained using the best two-step model (Table VI). Further, the shape of the spline confirms that the two-step model is appropriate.

As a final check on the adequacy of the two-step model, correlation between residuals was allowed for this model (Run 7). This extention to the model resulted in no decrease in the objective function, and the correlation between neighboring observations is a trivial 0.06 (see Table VI).

The use of extended models for the residual error has, apart from aiding the identification of time-varying parameters, led to revised values of some parameters (most notably k_a, ω_{ka}, and π_{ka}) and a strong indication that the

Table VI. Parameter Estimates for Real Data Set

Model feature	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7
Autocorrelation	—	Yes	Yes	—	—	—	Yes
Interoccasion variability	—	—	Yes	Yes	Yes	Yes	Yes
Time-dependent error	—	—	—	—	Two-step function	Spline function	Two-step function
ΔOFV[a]	0	−38	−384	−378	−453	−464	−453
Parameter							
CL/F ($L \cdot hr^{-1}$)	21	20	20	20	20	20	20
V/F (L)	99	99	100	100	103	98	103
ka (hr^{-1})	2.4	2.7	1.8	1.8	1.1	1.1	1.1
$\omega_{CL/F}$	0.32	0.31	0.28	0.29	0.30	0.29	0.30
$\omega_{V/F}$	0.39	0.37	0.34	0.34	0.33	0.34	0.33
ω_{ka}	1.50	2.10	0.74	0.69	0.47	0.38	0.47
$\pi_{CL/F}$	—	—	0.19	0.20	0.18	0.17	0.18
$\pi_{V/F}$	—	—	0.21	0.21	0.20	0.20	0.20
π_{ka}	—	—	0.99	0.96	0.71	0.62	0.70
σ	0.20	0.22	0.13	0.13	0.10[b]	Fig. 5	0.10[b]
t_{corr} (hr)	—	1.9	0.58	—	—	—	0.36

[a] Difference in objective function value (OFV) compared to basic model (Run 1).
[b] This figure refers to the error during the period 2–12 hr postdose. For samples obtained at 1 hr postdose the estimated error was 43% for both Run 5 and 7.

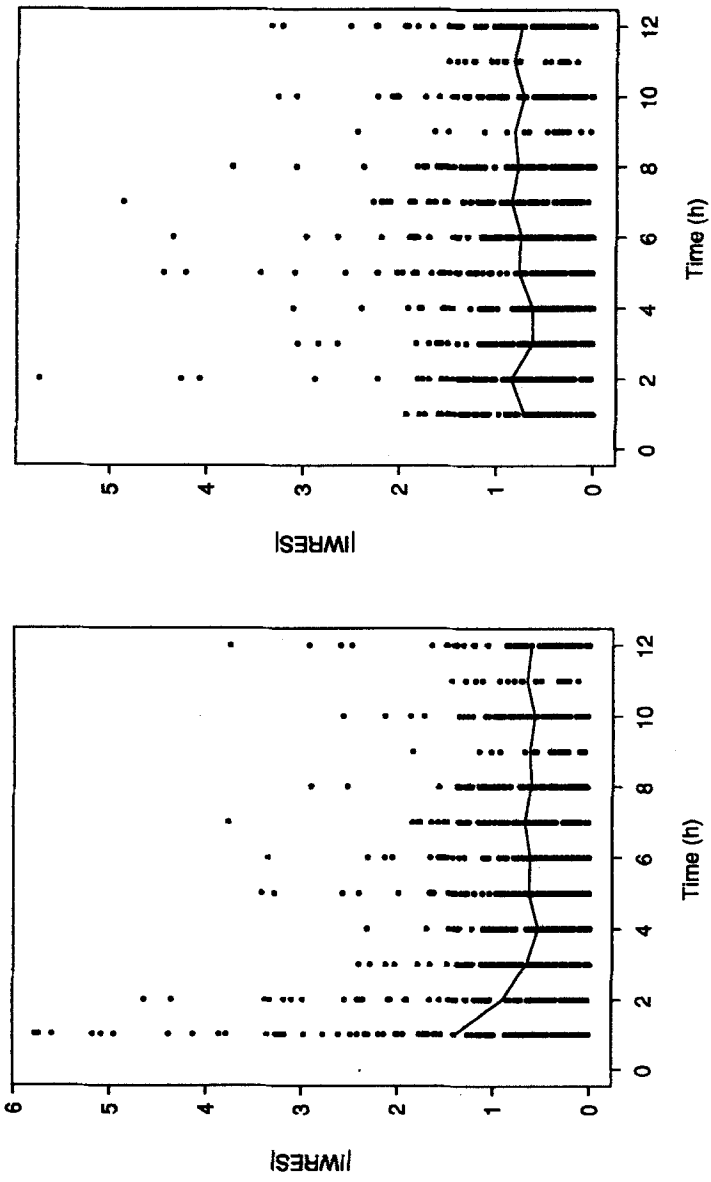

Fig. 4. Residual plots for real data set. For details see legend to Fig. 1.

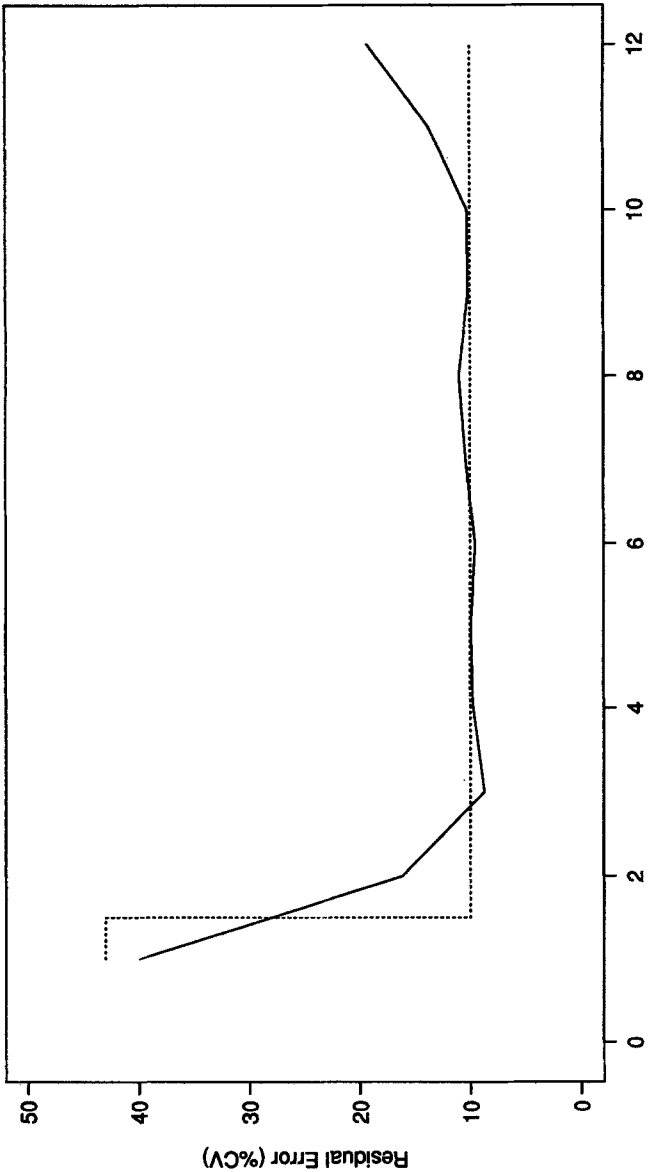

Fig. 5. Error magnitude as a function of time for the real data set. The broken line is the estimated σ_e as a two-step function of time. The continuous line is the estimated σ_e as given by a spline function with seven breakpoints.

model describing disposition is far more precise that the one describing absorption. The residual error in the postabsorption phase in the final model is 0.10, compared to 0.20 (overall) using the basic model. This low error, similar in magnitude to assay error, and the apparent absence of correlation between observations strongly suggest that little or no disposition model misspecification is present with the final model.

DISCUSSION

This paper suggests several residual error models for population analyses, dealing with the situation that sources other than assay error contribute significantly to total unexplained residual variability. Three error models were investigated using simulations. All identify the different properties of the residual error for which they are intended: (i) different residual error magnitudes for interreplicate and intersample error, (ii) serial correlation of errors, and (iii) time-dependent magnitude of error.

Correct specification of the magnitude and structure of residual error may be important if they, together with other population specifications, are to be used as prior information for subject-specific Bayesian estimation. In addition, the potential utility of the models used here is that they can provide (i) more detailed identification of sources of error, (ii) additional diagnostic tools and/or measures of modeling success, (iii) more realistic assessment of the appropriateness of models used to describe kinetic processes, and (iv) more accurate or precise estimates for structural parameters. The remainder of this section discusses the specific situations in which the three models can be useful.

The potential hazard of ignoring real differences in sources of error was mentioned in the Introduction and Theory sections of this paper. In the simulations performed here, it seems that little, if anything, is lost in terms of structural parameter estimation by using the traditional model in which all sources of error are lumped into one error term. However, the model with separate error components may be useful when one is particularly concerned about identifying different sources of error. The number of replicates necessary to detect significant replication error is small, and this added information is therefore available at little cost. The magnitudes of replication and assay error can be compared. If replication error is considerably higher than known assay error, this suggests problems elsewhere in the replication sequence.

Recognizing even significant serial correlation between residuals improves neither accuracy nor precision of the structural parameter estimates. Such improvement can only be obtained by using a better specified structural model. However, more accurate estimates of variance components

can be obtained through using an autocorrelated error structure when this is appropriate. Due to the nature of the inaccuracies in variance parameter estimates—overestimation of interindividual variability and underestimation of residual variability—the apparent value of therapeutic drug monitoring can be falsely high when serial correlation is neglected. Another situation in which accounting for serial correlation may be important is when conclusions regarding the population structural model are important. The presence of serial correlation may weaken mechanistic interpretations. It should be noted, though, that the inability to discover serial correlation, using the particular serial correlation model employed here, does not necessarily mean that serial correlation is absent.

The observation that SEs are not adversely affected when even relatively high autocorrelation is neglected must be put into perspective. That population analyses differ in this respect from single-subject analyses can be understood by considering the following example. With a constant infusion where clearance varies over time, there will be periods where the concentrations from a given subject reflect a higher or lower clearance than the average one for that subject. If many measurements are made during such a period, and serial correlation between them not considered, an analysis of these data would result in a seemingly precise estimate of the clearance value. In a population analysis with similar data from a number of subjects, the population average clearance is being estimated, and the precision of this estimate is influenced far more by interindividual variability (which is usually substantial) than it is by the (overoptimistically precise) subject-specific estimates of clearance.

The use of a time-dependent error model should be considered whenever residual error is considerably higher than assay or replication error. It may be possible to get an indication of the need for this kind of error model from residual analyses. It is not surprising that for the oral data in our real data example, residual error is larger during the absorption than elimination phase. Disposition models are better approximations to the underlying processes than are currently used absorption models. A time-dependent error model (was the only model that) resulted in structural parameter estimates different from those obtained with the traditional error model. The upward bias in k_a in the TIME simulation and in the real data example, using the traditional error model, is likely due to the use of a proportional error model. The only way to accommodate the larger residual errors during the absorption phase is to predict higher concentrations, since these, according to the proportional error model, are associated with (and therefore permit) larger errors. The only way to achieve higher predicted concentrations during the absorption, but not elimination phase, is to increase k_a above its true value; hence the bias noted.

Even when interest in the variance components of a population model is little, the error models presented herein can offer advantages. The real data set is a good example of this. The estimates of the structural parameters change somewhat as the variance components are more correctly specified, but more important, the unexplained error is partitioned rather differently. In the final model, during the disposition phase the residual error is only slightly larger than the assay error. This offers strong reassurance that no major misspecification remains.

The simulations SAMPL, BIEXP, and ERRAT were performed to mimic the errors introduced into the analysis when there are problems with sample collection practices or the structural model. Despite our judgment that the misspecifications introduced were large compared to what one would expect with real data situations, the estimated σ_r with traditional model was relatively low compared to the estimated σ_r which is often observed with real data sets. Therefore, caution should be exercised in interpreting our simulation results quantitatively.

In summary, we have presented and investigated extensions of the standard residual error model used in population analyses. The extensions try to accommodate properties of the residual error that can be anticipated from the nature of the data, but have not previously been incorporated into population analyses. The examples presented provide qualitative information on how different parameter estimates are influenced when different models for the residual error are used.

REFERENCES

1. G. A. F. Seber and C. J. Wold. *Nonlinear Regression*, J. Wiley, New York, 1989.
2. E. M. Chi and G. C. Reinsel. Models for longitudinal data with random effects and AR(1) errors. *J. Am. Statist. Assoc.* **84**:452–459 (1989).
3. M. Davidian and D. M. Giltinan. Some general estimation methods for nonlinear mixed-effects models. *J. Biopharm. Statist.* **3**:23–55 (1993).
4. J. Beran. Statistical method for data with long-range dependence. *Statist. Sci.* **7**:404–407 (1992).
5. C. A. Glasbey. Correlated residuals in non-linear regression applied to growth data. *Appl. Statist.* **28**:251–259 (1979).
6. D. Ruppert, N. Cressie, and R. J. Carroll. A transformation/weighting model for estimating Michaelis-Menten parameters. *Biometrics* **45**:637–656 (1989).
7. C. A. DeBoor. *A Practical Guide to Splines*. Springer-Verlag, New York, 1978.
8. H. Akaike. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**:716–723 (1974).
9. *NONMEM User's Guides*. NONMEM Project Group, University of California, San Francisco, 1992.
10. J. W. Mandema, D. Verotta, and L. B. Sheiner. Building population pharmacokinetic-pharmacodynamic models: I. Models for covariate effects. *J. Pharmacokin. Biopharm.* **20**:511–528 (1992).
11. M. O. Karlsson and L. B. Sheiner. The importance of modeling interoccasion variability in population pharmacokinetic analyses. *J. Pharmacokin. Biopharm.* **21**:735–750 (1993).
12. *S-PLUS, version 3.1*. Statistical Sciences, Inc., P.O. Box 85625, Seattle, WA, 1990.