A Population Pharmacokinetic Model with Time-Dependent Covariates Measured with Errors

Lang Li,[1,*] Xihong Lin,[2] Morton B. Brown,[2] Suneel Gupta,[3] and Kyung-Hoon Lee[4]

[1]Division of Biostatistics, Indiana University, Indianapolis, Indiana 46254, U.S.A.
[2]Department of Biostatistics, The University of Michigan, Ann Arbor, Michigan 48109, U.S.A.
[3]ALZA Corporation, Clinical Pharmacology, Mountain View, California 94039, U.S.A.
[4]Department of Pharmacology, Sungkyunkwan University, Suwon, Hyunggi do, 440-746, South Korea
[*]*email:* lali@iupui.edu

SUMMARY. We propose a population pharmacokinetic (PK) model with time-dependent covariates measured with errors. This model is used to model S-oxybutynin's kinetics following an oral administration of Ditropan, and allows the distribution rate to depend on time-dependent covariates blood pressure and heart rate, which are measured with errors. We propose two two-step estimation methods: the second-order two-step method with numerical solutions of differential equations (2orderND), and the second-order two-step method with closed form approximate solutions of differential equations (2orderAD). The proposed methods are computationally easy and require fitting a linear mixed model at the first step and a nonlinear mixed model at the second step. We apply the proposed methods to the analysis of the Ditropan data, and evaluate their performance using a simulation study. Our results show that the 2orderND method performs well, while the 2orderAD method can yield PK parameter estimators that are subject to considerable biases.

KEY WORDS: Differential equations; Laplace approximation; Measurement error; Nonlinear mixed models; Pharmacokinetics; Two-compartment model.

1. Introduction

1.1 *The Ditropan Study*

S-oxybutynin, a major active component of Ditropan (ALZA Corp.), is commonly prescribed as an antispasmodic for patients with uninhibited neurogenic or reflex neurogenic bladder. It is used to relieve symptoms associated with voiding, such as urge incontinence, urgency, and frequency of urination arising from overactivity of the detrusor muscle (AHFS Drug Information, 2000).

In a Ditropan drug study sponsored by ALZA Corp., the controlled released drug formulation (once a day) was compared to the regular formulation (three doses a day) to establish the bioequivalence. In this article, we focus on the pharmacokinetics of S-oxybutynin based on the three-dose data. This study involved 40 healthy subjects, and the three doses were given 8 hours apart. Each subject had 25 plasma measures over 48 hours. Figure 1d shows individual S-oxybutynin plasma concentration time profiles. In addition to the plasma concentration of S-oxybutynin, diastolic blood pressure (DBP), systolic blood pressure (SBP), and heart rate (HR) were measured nine times repeatedly during the 48-hour window for each subject, and the times these covariates were measured were different from those of the plasma measures. The time profiles of DBP, SBP, and HR are given in Figure 1a–1c. An objective of this study was to develop a pharmacokinetic (PK) model for the S-oxybutynin plasma concentration by modeling the PK parameters as functions

of the time-varying covariates blood pressure and heart rate, which were measured with error. We present in Section 1.2 some preliminary-analysis results and provide biological motivations for such a model.

1.2 *Preliminary Analysis of the Ditropan Data and Motivation for the Proposed Research*

Following an oral administration of S-oxybutynin, absorption from the gut is rapid, and it is metabolized extensively by cytochrome P450 3A4 in the liver and the gut wall. Douchamps et al. (1988) conducted a one-oral-dose study and found that S-oxybutynin followed a two-compartment model with linear pharmacokinetics. Li et al. (2002) analyzed the three-dose Ditropan data, and found that both the dose-dependent maximum concentration (*Cmax*) and the dose-dependent area under the concentration (AUC) of S-oxybutynin were nondecreasing functions of the three consecutive doses, suggesting that some pharmacokinetic parameters were dose dependent. Therefore, we started our pharmacokinetic analysis from a three separate two-compartment model corresponding to the three doses. Specifically, we assumed that the PK parameters were constant during each dosing period and different among the three dosing periods.

Denote by t_l the time the dose l is given ($l = 1$–3). For the Ditropan study, $(t_1, t_2, t_3) = (0, 8, 16)$. Let $\{A_{1,il}(t - t_l)_+, A_{2,il}(t - t_l)_+, A_{3,il}(t - t_l)_+\}$ be the amount of drug in gut, blood, and tissue at time t for subject i ($i = 1, \ldots, 40$) from
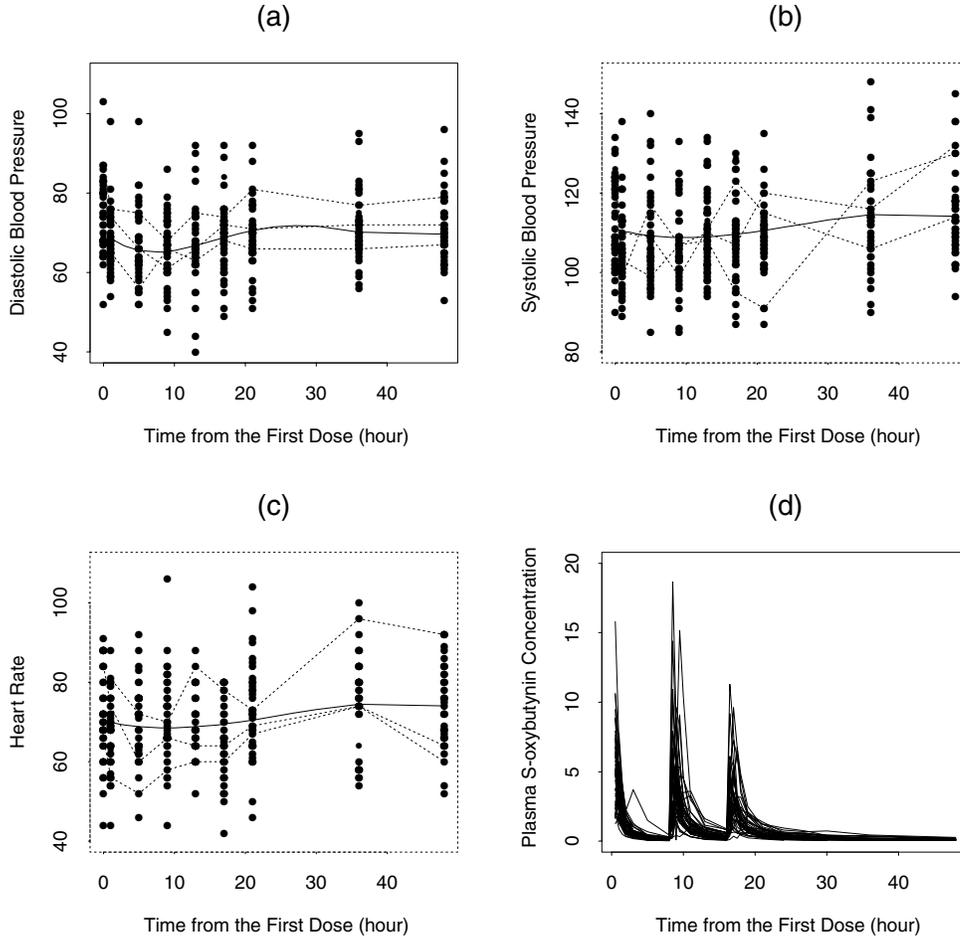
Figure 1. Raw data (**a**) diastolic blood pressure (DBP), (**b**) systolic blood pressure (SBP), (**c**) heart rate (HR), and (**d**) S-oxybutynin plasma concentration. The solid lines in (a), (b), and (c) are the fitted polynomial curves corresponding to the lowest AICs using the linear mixed model (10).

dose l (l = 1, 2, 3), respectively, where $A_{k,il}(t - t_l)_+ = A_{k,il}(t - t_l)$ ($k = 1$–3) if $t > t_l$ and 0 if $t \leq t_l$ except $A_{1,il}(0) = 1$. The two-compartment model for dose l can be written as the following differential equations:

$$\frac{dA_{1,il}(t - t_l)_+}{dt} = -k_{a,il}A_{1,il}(t - t_l)_+$$

$$\frac{dA_{2,il}(t - t_l)_+}{dt} = k_{a,il}A_{1,il}(t - t_l)_+$$
$$- (k_{e,il} + k_{23,il})A_{2,il}(t - t_l)_+$$
$$+ k_{32,il}A_{3,il}(t - t_l)_+$$

$$\frac{dA_{3,il}(t - t_l)_+}{dt} = k_{23,il}A_{2,il}(t - t_l)_+ - k_{32,il}A_{3,il}(t - t_l)_+, \quad (1)$$

where $k_{a,il}$ is the absorption rate, $k_{e,il}$ is the elimination rate, and $k_{23,il}$ and $k_{32,il}$ are the distribution rates for subject i given dose l. Note that we here assume that the PK parameters are constant within each dose period and different among the three doses. One can show that under this assumption, the

above system of differential equations (1) has a closed form solution. Specifically, we have

$$A_{2,il}(t - t_l)_+ = \frac{k_{a,il}(k_{32,il} - k_{a,il})e^{-k_{a,il} \times (t - t_l)_+}}{(\lambda_{1,il} - k_{a,il})(\lambda_{2,il} - k_{a,il})}$$
$$+ \frac{k_{a,il}(k_{32,il} - \lambda_{1,il})e^{-\lambda_{1,il} \times (t - t_l)_+}}{(\lambda_{2,il} - \lambda_{1,il})(k_{a,il} - \lambda_{1,il})}$$
$$+ \frac{k_{a,il}(k_{23,il} - \lambda_{2,il})e^{-\lambda_{2,il} \times (t - t_l)_+}}{(\lambda_{1,il} - \lambda_{2,il})(k_{a,il} - \lambda_{2,il})}, \quad (2)$$

where

$$\lambda_{1,il} = \Big[k_{e,il} + k_{23,il} + k_{32,il}$$
$$+ \big\{ (k_{e,il} + k_{23,il} + k_{32,il})^2 - 4k_{e,il}k_{32,il} \big\}^{1/2} \Big] / 2$$
$$\lambda_{2,il} = \Big[(k_{e,il} + k_{23,il} + k_{32,il}$$
$$- \big\{ (k_{e,il} + k_{23,il} + k_{32,il})^2 - 4k_{e,il}k_{32,il} \big\}^{1/2} \Big] / 2.$$

The log-transformed plasma drug concentration given the three consecutive doses is

$$f_i(V_{il}, k_{a,il}, k_{e,il}, k_{32,il}, k_{23,il}, t)$$
$$= \log\Big[\big\{A_{2,i1}(t-t_1)_+ + A_{2,i2}(t-t_2)_+$$
$$+ A_{2,i3}(t-t_3)_+\big\}\big/V_{il}\Big], \qquad (3)$$

where V_{il} is the volume of the distribution for subject i from dose l. Let Y_{ij} be the plasma drug concentration for subject i at the jth time point ($j = 1, \ldots, 25$). We assume Y_{ij} follows the nonlinear mixed model

$$Y_{ij} = f_i(V_{il}, k_{a,il}, k_{e,il}, k_{32,il}, k_{23,il}, t_{ij}) + \epsilon_{ij}, \qquad (4)$$

where $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, and

$$\log(V_{il}) = \log(V_l) + b_{1i},$$
$$\log(k_{a,il}) = \log(k_{a,l}) + b_{2i}, \quad \log(k_{e,il}) = \log(k_{e,l}) + b_{3i},$$
$$\log(k_{23,il}) = \log(k_{23,l}) + b_{4i}, \quad \log(k_{32,il}) = \log(k_{32,l}),$$

and $\mathbf{b}_i = (b_{1i}, b_{2i}, b_{3i}, b_{4i})^T \sim N(\mathbf{0}, \mathbf{D})$, $\mathbf{D} = \text{diag}(\sigma_{11}^2, \ldots, \sigma_{44}^2)$, $\mathbf{b}_i$ and ϵ_{ij} are independent. In particular, $k_{32,il}$ is assumed to be only dose dependent, because the algorithm did not converge during our initial subject-specific and dose-dependent model. The failure to convergence might be due to the fact that variance of the random effects in $k_{32,il}$ was close to zero. It was also probably because there was not enough information for $k_{32,il}$ from some subjects.

We fit the two-compartment nonlinear mixed models (2)–(4) using the conditional first-order method (Lindstrom and Bates, 1990). Table 1 gives the parameter estimates and their standard errors, and the last column gives the p values testing for the equivalence of the PK parameters among the three doses. The results show that the absorption rate $k_{a,l}$ and the distribution rate $k_{23,l}$ are statistically significantly different among the three doses. The absorption rate decreases with the dose level. This might be explained by the saturability of transporting passages across gastrointestinal membranes. The distribution rate appears to have a bell shape, with the lowest rate given by the second dose. It is of scientific interest to investigate the factors that drive the bell-shaped relationship between the distribution rate and the dose.

The distribution rate is usually either blood flow sensitive or binding sensitive. It is known in vitro that oxybutynin demonstrates high binding specificity for M3/m3 receptors. If oxybutynin were binding sensitive due to saturation with repeated doses, one would expect the distribution rate decreases monotonically with the doses instead of a bell shape. In view of our results in Table 1, we have no evidence that oxybutynin is binding sensitive. We hence focus on investigating the factors that might drive the distribution rate to be blood flow sensitive. It is known that blood pressure is a driving force for blood to flow through the arterial system (Eugene, 2001). If the distribution rate is flow sensitive, it could be correlated with blood pressure. In addition, since blood pressure usually has a circadian pattern, it might explain the observed bell-shaped relationship between the distribution rate and the dose. Therefore, we are interested in extending models (2)–(4) to allow for the distribution rate to be a function of blood pressure and heart rate. In the Ditropan study, this investigation is further challenged by the fact that blood pressure is measured with considerable error and its measurement times are different from those of plasma-concentration measures. We hence need to take these features into account in our statistical model.

1.3 *The Statistical Background*

Nonlinear mixed models have been widely used in modeling pharmacokinetic data (Lindstrom and Bates, 1990; Davidian and Giltinan, 1995), where the PK parameters are modeled as parametric functions of time-independent covariates. Li et al. (2002) extended such models to allow for some time-dependent PK parameters to be nonparametric functions of time using splines. Although there is an extensive literature on measurement error in nonlinear models (Carroll, Ruppert, and Stefanski, 1995), the literature on measurement error in nonlinear mixed effects PK models is rather limited. Higgins, Davidian, and Giltinan (1997) proposed a simple two-step regression calibration method. They assumed the PK parameters to be time dependent. Under this assumption, the solution of the differential equation does not have an analytic form. They assumed in their estimation procedure an approximate analytic solution. We will show in our simulation study that such a closed-form approximation can be subject to considerable bias. Their two-step method also does not consider the contribution of the variation of the estimator obtained from the first step to the estimation at the second step. Their

Table 1
Parameter estimates under the dose-specific two-compartment models (2)–(4) *for the Ditropan study*

Fixed effects	Dose 1	Dose 2	Dose 3	p value
V	0.163 (0.021)	0.175 (0.020)	0.195 (0.013)	0.687
k_a	6.394 (0.882)	3.553 (0.391)	2.574 (0.207)	0.001
k_e	0.652 (0.027)	0.605 (0.026)	0.632 (0.022)	0.330
k_{23}	0.312 (0.060)	0.209 (0.051)	0.324 (0.015)	0.041
k_{32}	0.122 (0.104)	0.078 (0.082)	0.091 (0.025)	0.914
Variance components				
σ_ϵ^2	0.135 (0.011)			
σ_{11}	0.247 (0.039)			
σ_{22}	0.162 (0.046)			
σ_{33}	0.121 (0.094)			
σ_{44}	0.204 (0.095)			

simulation results show that the model-based standard errors seriously underestimate the true SEs.

We consider in this article nonlinear mixed effects PK models where the PK parameters are functions of time-dependent covariates which are measured with errors. We describe the model in Section 2, and propose in Section 3 two two-step estimation methods: the second-order two-step method with numerical solutions of differential equations (2orderND), and the second-order two-step method with closed form approximate solutions of differential equations (2orderAD). Both methods yield approximate MLEs where the likelihood function is approximated using the Solomon–Cox–Laplace approximation. They are both computationally easy and stable, and require fitting a linear-mixed model at the first step and a nonlinear-mixed model at the second step. The proposed methods are applied to the Ditropan study in Section 4, and their performance is evaluated using a simulation study in Section 5. Conclusions are given in Section 6.

2. The Population PK Model with Time-Dependent Covariates Measured with Errors

The preliminary results of the Ditropan data in Table 1 provide no evidence that V_l, $k_{e,l}$, and $k_{32,l}$ are dose dependent, whereas there is a strong statistical evidence that $k_{a,l}$ and $k_{23,l}$ are dose dependent. Little between-subject variation of k_e was found. We hence assume V_l, $k_{e,l}$, and $k_{32,l}$ are dose independent and (k_e, k_{32}) are fixed effects in our subsequent measurement error PK model. Denote these PK parameters by $\boldsymbol{\beta}_i(t) = \{V_i, k_{a,i}(t), k_e, k_{23,i}(t), k_{32}\}^T$. We consider a two-compartment model similar to (1) except that the PK parameters $(V_{i,l}, k_{e,il}, k_{32,il}, k_{a,il}, k_{23,il})$ are replaced by $\{V_i, k_e, k_{32}, k_{a,i}(t), k_{23,i}(t)\}$, where $k_{a,i}(t)$ and $k_{23,i}(t)$ are time dependent,

$$\frac{dA_{1,il}(t - t_l)_+}{dt} = -k_{a,i}(t)A_{1,il}(t - t_l)_+$$

$$\frac{dA_{2,il}(t - t_l)_+}{dt} = k_{a,i}(t)A_{1,il}(t - t_l)_+$$
$$- \{k_e + k_{23,i}(t)\}A_{2,il}(t - t_l)_+$$
$$+ k_{32}A_{3,il}(t - t_l)_+$$

$$\frac{dA_{3,il}(t - t_l)_+}{dt} = k_{23,i}(t)A_{2,il}(t - t_l)_+ - k_{32}A_{3,il}(t - t_l)_+. \quad (5)$$

It follows that the differential equations (5) do not have an analytic closed-form solution. We solve it numerically using an IMSL Fortran77 function (DIVPRK), which uses the Runge–Kutta–Verner fifth-order method (IMSL USER's MANUAL STAT/LIBRARY Version 1.0, 1987).

We describe below a three-stage nonlinear mixed model to account for measurement errors in covariates for the two-compartment model. Suppose the data are composed of m subjects. The jth observation ($j = 1, \ldots, J_i$) of the ith subject ($i = 1, \ldots, m$) is measured at time t_{ij} and consists of an outcome variable Y_{ij}, accurately measured time-varying covariates $\boldsymbol{X}_{ij} = \boldsymbol{X}(t_{ij})$, and an unobserved time-varying covariate $U_{ij} = U_i(t_{ij})$, e.g., U_{ij} is the true diastolic blood pressure measured at time t_{ij} for subject i. Suppose that W is an observed error-prone measure of U and is measured at time points t_{ik}^* ($k = 1, \ldots, K_i$), where t_{ik}^* might differ from t_{ij} and J_i might differ from K_i. For example, $W_{ik} = W_i(t_{ik}^*)$ is the

observed diastolic blood pressure measured at time t_{ik}^*. In the Ditropan data, the times diastolic blood pressure was taken (t_{ik}^*) differed from the times plasma concentration was measured (t_{ij}), and $J_i = 25$ and $K_i = 9$.

The first-stage model describes the observational level pharmacokinetics. Specifically, we rewrite the nonlinear mixed model (4) for the plasma concentration Y_{ij} ($i = 1, \ldots, m, j = 1, \ldots, J_i$) as

$$Y_{ij} = f_i(\boldsymbol{\beta}_{ij}, t_{ij}) + \epsilon_{ij}, \quad (6)$$

where $f_i(\boldsymbol{\beta}_{ij}, t_{ij}) = \log[\{A_{2,i1}(\boldsymbol{\beta}_{ij}, t_{ij} - t_1)_+ + A_{2,i2}(\boldsymbol{\beta}_{ij}, t_{ij} - t_2)_+ + A_{2,i3}(\boldsymbol{\beta}_{ij}, t_{ij} - t_3)_+\}/V_i]$, $A_{2,il}(\cdot)$ solves (5), $\boldsymbol{\beta}_{ij} = \boldsymbol{\beta}_i(t_{ij})$, the ϵ_{ij} are independent and follow $N(0, \sigma_\epsilon^2)$, and are independent of $\boldsymbol{\beta}_{ij}$.

The second stage model describes the association between time-dependent PK parameter vector $\boldsymbol{\beta}_{ij}$ and the accurately measured covariates $\boldsymbol{X}_{ij}$ and the unobserved covariate U_{ij}, e.g., U_{ij} is the true DBP at t_{ij}, as follows:

$$\boldsymbol{\beta}_{ij} = \mathbf{d}(\boldsymbol{X}_{ij}, U_{ij}, \boldsymbol{\alpha}, \mathbf{b}_i), \quad (7)$$

where $\boldsymbol{\beta}_{ij} = \{V_i, k_{a,i}(t_{ij}), k_e, k_{23,i}(t_{ij}), k_{32}\}^T$, $V_i = \exp(\alpha_1 + b_{1i})$, $k_{a,i}(t_{ij}) = \exp\{I(t_{ij} \leq 8)\alpha_2 + I(8 < t_{ij} \leq 16)\alpha_3 + I(16 \leq t_{ij})\alpha_4 + b_{2i}\}$, $I(\cdot)$ is an indicator function, $k_e = \exp(\alpha_5)$, $k_{23,i}(t_{ij}) = \{\alpha_6 + \alpha_7 U_{ij}\} \exp(b_{3i})$, $k_{32} = \exp(\alpha_8)$, $\mathbf{d}$ is a five-dimensional function, $\mathbf{b}_i = (b_{1i}, b_{2i}, b_{3i})^T \sim N(0, \mathbf{D})$, and $\mathbf{D} = \text{diag}(\sigma_{11}^2, \sigma_{22}^2, \sigma_{33}^2)$. In particular, a linear relationship between log-transformed k_{32} and U_{ij} was initially investigated; however, the association (result not reported) was not as significant as that in raw scale.

The unobserved covariate U_{ij} can be regarded as a realization of the underlying true continuous process $U_i(t)$ at time t_{ij}, i.e., $U_{ij} = U_i(t_{ij})$. Suppose $U_i(t)$ is measured with error by the error-prone process $W_i(t)$, and W_{ik} is the value of $W_i(t)$ at time t_{ik}^* ($k = 1, \ldots, K_i$). The third stage relates the error-prone covariate $W_{ik} = W_i(t_{ik}^*)$ to the true unobserved values $U_i(t_{ik}^*)$ using an additive measurement error model (Carroll, Ruppert, and Stefanski, 1995)

$$W_{ik} = U_i(t_{ik}^*) + e_{ik}, \quad (8)$$

where $e_{ik} \sim N(0, \sigma_e^2)$ is an independent measurement error and is independent of ϵ_{ij} and $\boldsymbol{b}_i$. Following Higgins, Davidian, and Giltinan (1997), we assume a polynomial model for the underlying process $U_i(t_{ik}^*)$

$$U_i(t_{ik}^*) = \mathbf{T}_{ik}^{*T}\boldsymbol{\gamma} + \mathbf{B}_{ik}^T\mathbf{a}_i, \quad (9)$$

where $\mathbf{T}_{ik}^* = (1, \ldots, \{t_{ik}^*\}^r)^T$, $\boldsymbol{\gamma}$ is a $(r + 1) \times 1$ vector, $\mathbf{B}_{ik} = (1, \ldots, \{t_{ik}^*\}^h)^T$, $\mathbf{a}_i$ is an $(h + 1) \times 1$ vector of random effects following $N(\mathbf{0}, \boldsymbol{\Omega})$, and $\boldsymbol{\Omega}$ is diagonal.

It follows from (8) and (9) that

$$W_{ik} = \mathbf{T}_{ik}^{*T}\boldsymbol{\gamma} + \mathbf{B}_{ik}^T\mathbf{a}_i + \mathbf{e}_{ik}. \quad (10)$$

Denote by $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{iJ_i})^T$, $\boldsymbol{Y} = (\boldsymbol{Y}_i^T, \ldots, \boldsymbol{Y}_m^T)^T$, and $\mathbf{f}$, $\mathbf{t}$, $\boldsymbol{\epsilon}$, $\boldsymbol{\beta}$, $\mathbf{X}$, $\mathbf{t}$, $\mathbf{U}$, $\mathbf{t}^*$, and $\mathbf{e}$ similarly. Let $\mathbf{B} = \text{diag}(\boldsymbol{B}_i)$. Equations (6)–(10) can be rewritten in a matrix form

$$\mathbf{Y} = \mathbf{f}(\boldsymbol{\beta}, \mathbf{t}) + \boldsymbol{\epsilon}, \quad (11)$$

$$\boldsymbol{\beta} = \mathbf{d}\{\mathbf{X}, \mathbf{U}(\mathbf{t}), \boldsymbol{\alpha}, \mathbf{b}\} = \mathbf{d}(\mathbf{X}, \boldsymbol{\gamma}, \mathbf{a}, \boldsymbol{\alpha}, \mathbf{b}), \quad (12)$$

$$\mathbf{W} = \mathbf{U}(\mathbf{t}^*) + \mathbf{e}, \tag{13}$$

$$\mathbf{U}(\mathbf{t}^*) = \mathbf{T}\boldsymbol{\gamma} + \mathbf{B}\mathbf{a}, \tag{14}$$

where $\boldsymbol{b} \sim N\{0, \text{diag}(\boldsymbol{D})\}$, $\boldsymbol{a} \sim N\{0, \text{diag}(\boldsymbol{\Omega})\}$, $\boldsymbol{\epsilon} \sim N(0, \sigma_\epsilon^2 \boldsymbol{I})$, and $\boldsymbol{e} \sim N(0, \sigma_e^2 \boldsymbol{I})$, and $\boldsymbol{b}$, $\boldsymbol{a}$, $\boldsymbol{\epsilon}$, and $\boldsymbol{e}$ are independent.

Denote by $\boldsymbol{\zeta} = (\boldsymbol{\alpha}, \mathbf{D}, \sigma_\epsilon^2)$. The integrated likelihood function of the observed data $\boldsymbol{Y} \,|\, (\boldsymbol{W}, \boldsymbol{X})$ is

$$L\big(\mathbf{Y} \,|\, \mathbf{W}, \boldsymbol{X}; \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2\big)$$

$$= \int L_1(\mathbf{Y} \,|\, \mathbf{b}, \mathbf{a}; \boldsymbol{\gamma}, \boldsymbol{\zeta}) L_2(\mathbf{b}; \mathbf{D}) L_3\big(\mathbf{a} \,|\, \mathbf{W}; \boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2\big) \, d\mathbf{b} \, d\mathbf{a}, \tag{15}$$

where $L_1(\mathbf{Y} \,|\, \mathbf{b}, \mathbf{a}; \boldsymbol{\gamma}, \boldsymbol{\zeta})$ is the conditional likelihood of $\boldsymbol{Y}$ under the first-stage model (11), L_2 $(\mathbf{b}; \mathbf{D})$ is the likelihood of $\boldsymbol{b}$ under the second-stage model (12), and L_3 $(\mathbf{a} \,|\, \mathbf{W}; \boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2)$ is the likelihood of $\boldsymbol{a}$ under the third-stage models (13) and (14). Since the function $f(\cdot)$ is nonlinear, the likelihood (15) does not have a closed form and its evaluation involves a high dimensional integral.

3. The Two Estimation Procedures

3.1 *The Solomon–Cox–Laplace Approximation of the Likelihood*

In view of the high dimensional integral involved in the likelihood (15), we consider a two-step estimation procedure by approximating the likelihood (15) using the Solomon–Cox–Laplace approximation. Since the $\boldsymbol{W}$ data contain the majority of the information about $(\boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2)$, at the first step, we use the $\boldsymbol{W}$ data to estimate $(\boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2)$ by fitting the linear-mixed model (10) and obtain the BLUP estimator of $\boldsymbol{\xi} = (\boldsymbol{\gamma}^T, \boldsymbol{a}^T)$ denoted by $\tilde{\boldsymbol{\xi}}$ and its covariance $\tilde{\boldsymbol{V}}_{\boldsymbol{\xi}}$. It can be easily shown that the resulting estimators $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\Omega}}, \hat{\sigma_e}^2)$ are consistent. At the second step, we treat $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{V}}_{\boldsymbol{\xi}}$ to be known and to maximize the Solomon–Cox–Laplace approximation of the integrated likelihood (15) as a function of $\boldsymbol{\zeta}$ to obtain an approximate MLE of $\boldsymbol{\zeta}$.

Specifically, denote by $\boldsymbol{U} = \boldsymbol{U}(\boldsymbol{t}) = \boldsymbol{T}\boldsymbol{\gamma} + \boldsymbol{B}\boldsymbol{a} = \boldsymbol{S}\boldsymbol{\xi}$, where $\boldsymbol{S} = (\boldsymbol{T}, \boldsymbol{B})$ and $\boldsymbol{\xi} = (\boldsymbol{\gamma}^T, \boldsymbol{a}^T)^T$. Given $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\Omega}}, \hat{\sigma_e}^2)$, the distribution of $(\boldsymbol{U} \,|\, \boldsymbol{W})$ can be estimated using the distribution of $(\boldsymbol{\xi} \,|\, \boldsymbol{W})$, i.e., the distribution of the BLUP estimator of $\boldsymbol{\xi}$. Denote by $\boldsymbol{H} = \boldsymbol{S}\boldsymbol{S}^T + \text{diag}(\mathbf{0}, \widehat{\sigma_e^2}\hat{\boldsymbol{\Omega}}^{-1})$. Then $\boldsymbol{\xi} \,|\, \boldsymbol{W}$ follows $N(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{V}}_\xi)$, where

$$\tilde{\boldsymbol{\xi}} = \boldsymbol{H}^{-1}\boldsymbol{S}\boldsymbol{W}, \quad \tilde{\boldsymbol{V}}_{\boldsymbol{\xi}} = \widehat{\sigma_e^2}\boldsymbol{H}^{-1}. \tag{16}$$

It follows that the distribution of $\boldsymbol{U} \,|\, \boldsymbol{W}$ can be estimated as $N(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}_{\boldsymbol{u}})$, where $\tilde{\boldsymbol{U}} = \boldsymbol{S}\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{V}}_{\boldsymbol{u}} = \boldsymbol{S}\tilde{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{S}^T$. Note that in fact $\tilde{\boldsymbol{U}} = E(\boldsymbol{U} \,|\, \boldsymbol{W})$ evaluated at $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\Omega}}, \widehat{\sigma_e^2})$, and $\tilde{\boldsymbol{U}}$ corresponds to the regression calibration estimator of $\boldsymbol{U}$.

To estimate $\boldsymbol{\zeta}$, we approximate the doubly integrated likelihood (15) using the Solomon–Cox and Laplace approximations at the second stage. Specifically, some calculations show that the likelihood (15) can be rewritten by integrating out $\boldsymbol{U}$ and $\boldsymbol{b}$ as

$$L\big(\mathbf{Y} \,|\, \mathbf{W}, \boldsymbol{X}; \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2\big)$$

$$= \int L_1(\mathbf{Y} \,|\, \mathbf{b}, \mathbf{U}; \boldsymbol{\zeta}) L_2(\mathbf{b}; \mathbf{D}) L_3\big(\mathbf{U} \,|\, \mathbf{W}; \boldsymbol{\gamma}, \boldsymbol{\Omega}, \sigma_e^2\big) \, d\mathbf{U} \, d\mathbf{b}, \tag{17}$$

where $L_3(\boldsymbol{U} \,|\, \boldsymbol{W})$ is the likelihood of $\boldsymbol{U} \,|\, \boldsymbol{W}$ and can be estimated as $\boldsymbol{U} \,|\, \boldsymbol{W} \sim N(\tilde{U}, \tilde{V}_u)$, and $\tilde{V}_u$ is the conventional variance of the BLUP of $\boldsymbol{U}$. To approximate the double integrals in (17), we first approximate the integral that involves $\boldsymbol{U}$ using the Solomon–Cox approximation (Solomon and Cox, 1992) by treating the integrand as a function of $\boldsymbol{U}$ and taking a quadratic expansion of the integrand about $\tilde{\boldsymbol{U}}$, the mean of $\boldsymbol{U}$. We next approximate the integral that involves $\boldsymbol{b}$ using the Laplace approximation by treating the integrand as a function of $\boldsymbol{b}$ and taking a quadratic expansion of the integrand about its mode $\tilde{\boldsymbol{b}}$. The resulting Solomon–Cox–Laplace approximation can be regarded as the second-order approximation of the integrated likelihood (15).

Specifically, following Wolfinger and Lin (1997, Section 3), one can show that the Solomon–Cox approximation for the integral involving $\boldsymbol{U}$ corresponds to the loglikelihood of $\boldsymbol{Y}$ as a function of $\boldsymbol{\zeta}$ under the nonlinear mixed model

$$\boldsymbol{Y} = f(\boldsymbol{\beta}, \boldsymbol{t}) + \boldsymbol{e}^*, \tag{18}$$

where $\boldsymbol{\beta} = \boldsymbol{d}(\boldsymbol{X}, \tilde{U}, \boldsymbol{\alpha}, \boldsymbol{b}) = \boldsymbol{d}(\boldsymbol{X}, \tilde{\boldsymbol{\xi}}, \boldsymbol{\alpha}, \boldsymbol{b}), \boldsymbol{b} \sim N(0, \boldsymbol{D}), \boldsymbol{e}^* \sim N(0, \boldsymbol{R}^*), \boldsymbol{R}^* = \sigma_\epsilon^2 \boldsymbol{I} + \boldsymbol{F}_{\boldsymbol{u}}\tilde{\boldsymbol{V}}_{\boldsymbol{u}}\boldsymbol{F}_{\boldsymbol{u}}^T = \sigma_\epsilon^2 \boldsymbol{I} + \boldsymbol{F}_{\boldsymbol{\xi}}\boldsymbol{V}_{\boldsymbol{\xi}}^{-1}\boldsymbol{F}_{\boldsymbol{\xi}}^T$, and $\boldsymbol{F}_{\boldsymbol{u}} = \partial\boldsymbol{f}/\partial\boldsymbol{U}^T$ and $\boldsymbol{F}_\xi = \partial\boldsymbol{f}/\partial\boldsymbol{\xi}^T$. Note that here we assume $(\tilde{U}, \tilde{\boldsymbol{V}}_{\boldsymbol{\xi}})$, equivalently $(\tilde{\boldsymbol{\xi}}, \tilde{\boldsymbol{V}}_{\boldsymbol{\xi}})$, which are obtained from the first step, are fixed and known parameters.

Following Lindstrom and Bates (1990) and Wolfinger and Lin (1997, Section 4), we can show that a further application of the Laplace approximation for the integral involving $\boldsymbol{b}$ in (17) gives, apart from a constant,

$$\ell_{\text{SCL}}(\boldsymbol{Y} \,|\, \boldsymbol{W}; \boldsymbol{\zeta}) = -\tfrac{1}{2}\ln|\tilde{\boldsymbol{V}}| - \tfrac{1}{2}\{\boldsymbol{Y} - f(\tilde{\boldsymbol{\beta}}, \boldsymbol{t}) - \tilde{\boldsymbol{Z}}\tilde{\boldsymbol{b}}\}\tilde{\boldsymbol{V}}^{-1}$$

$$\times \{\boldsymbol{Y} - f(\tilde{\boldsymbol{\beta}}, \boldsymbol{t}) - \tilde{\boldsymbol{Z}}\tilde{\boldsymbol{b}}\}, \tag{19}$$

where $\tilde{\boldsymbol{\beta}} = \boldsymbol{d}(\boldsymbol{X}, \tilde{U}, \boldsymbol{\alpha}, \tilde{\boldsymbol{b}}), \tilde{\boldsymbol{V}} = \tilde{\boldsymbol{Z}}\boldsymbol{D}\tilde{\boldsymbol{Z}}^T + \boldsymbol{R}^*, \tilde{\boldsymbol{Z}} = \partial\boldsymbol{f}/\partial\boldsymbol{b}^T|_{\boldsymbol{b}=\tilde{\boldsymbol{b}}}$ and $\tilde{\boldsymbol{b}}$ solves

$$\frac{1}{\sigma_\epsilon^2}\tilde{\boldsymbol{Z}}^T\{\boldsymbol{Y} - f(\tilde{\boldsymbol{\beta}}, \boldsymbol{t})\} - \boldsymbol{D}^{-1}\tilde{\boldsymbol{b}} = 0.$$

We term (19) the Solomon–Cox–Laplace approximation of the likelihood (15). The resulting approximate maximum likelihood estimator of $\boldsymbol{\zeta}$ that maximizes (19) correponds to fitting the nonlinear mixed model (18) using the Lindstrom and Bates (1990) algorithm.

3.2 *The Two-Stage Estimation Methods*

The above discussions suggest that our two-step estimation procedure is hence simple. At the first step, one fits the linear-mixed model (10) using the W data to obtain the MLE $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\Omega}}, \widehat{\sigma_e^2})$ and the BLUP estimator $\tilde{\boldsymbol{\xi}}$ and its covariance $\tilde{\boldsymbol{V}}_{\boldsymbol{\xi}}$. At the second step, one treats $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{V}}_{\boldsymbol{\xi}}$ as known, and fits the nonlinear-mixed model (18) with $\boldsymbol{b}$ as the random effects using the Lindstrom and Bates algorithm.

Following Lindstrom and Bates (1990) and Wolfinger and Lin (1997), one can show that fitting the nonlinear-mixed model at the second stage can proceed by iteratively fitting the linear-mixed model

$$\boldsymbol{Y}^* = \boldsymbol{X}^*\boldsymbol{\alpha} + \boldsymbol{Z}^*\boldsymbol{b} + \boldsymbol{e}^*, \tag{20}$$

where denoting by $\boldsymbol{\alpha}^*$ and $\boldsymbol{b}^*$ the estimates at the previous iteration, $\boldsymbol{Y}^*$ is a working vector,

$$\boldsymbol{Y}^* = \boldsymbol{Y} - \boldsymbol{f}^* + \boldsymbol{X}^*\boldsymbol{\alpha} + \boldsymbol{Z}^*\boldsymbol{b}^*, \quad \boldsymbol{e}^* \sim N\left(0, \boldsymbol{R}^* = \boldsymbol{F_\xi}\tilde{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{F_\xi}^T\right)$$

$$\boldsymbol{f}^* = \boldsymbol{f}_{\boldsymbol{\alpha}^*,\boldsymbol{b}^*,\tilde{\boldsymbol{\xi}}}, \qquad\qquad \boldsymbol{X}^* = \partial\boldsymbol{f}\big/\partial\boldsymbol{\alpha}^T|_{\boldsymbol{\alpha}^*,\boldsymbol{b}^*,\tilde{\boldsymbol{\xi}}}$$

$$\boldsymbol{Z}^* = \partial\boldsymbol{f}\big/\partial\boldsymbol{b}^T|_{\boldsymbol{\alpha}^*,\boldsymbol{b}^*,\tilde{\boldsymbol{\xi}}}, \qquad \boldsymbol{F_\xi} = \partial\boldsymbol{f}\big/\partial\boldsymbol{\xi}^T|_{\boldsymbol{\alpha}^*,\boldsymbol{b}^*,\tilde{\boldsymbol{\xi}}}.$$

The variance components $\boldsymbol{D}$ and σ_ϵ^2 are estimated iteratively using REML under (20).

Note that the regression calibration method (Higgins, Davidian, and Giltinan, 1997) ignores the variability of $\boldsymbol{U} \mid \boldsymbol{W}$ by simply plugging $\tilde{\boldsymbol{U}}$ into $\boldsymbol{\beta}(\cdot)$ and setting $\boldsymbol{U} = \tilde{\boldsymbol{U}}$. Our procedure accounts for the variability in the regression calibration estimator $\tilde{\boldsymbol{U}}$.

The nonlinear mixed model (18) and the linear-mixed model (20) involve evaluating the nonlinear function $f(\cdot)$. We propose two two-step estimation procedures in the next two sections. Both procedures are the same at the first step. They differ by the second step. The first procedure evaluates $f(\cdot)$ by numerically solving the differential equation (5) at the second step, while the second one approximates the solution of (5) analytically.

3.3 *The Second-Order Two-Step Method with Numerical Solution of Differential Equations*

Our first two-step estimation procedure fits the linear mixed model (10) using the W data at the first step and fits the nonlinear mixed model (18) or iteratively fits the linear mixed model (20), where $f(\cdot)$ is evaluated by numerically solving the differential equation (5). We call this method the second-order two-step method with numerical solution of differential equations (2orderND).

Examination of equation (20) suggests that we need to calculate both $\mathbf{f}\{\boldsymbol{\beta}(t), t\}$ and its first derivative $\partial\mathbf{f}\{\boldsymbol{\beta}(t), t\}/\partial\boldsymbol{\beta}$. For the two-compartment model, equation (5) indicates that we need to calculate $\boldsymbol{A}_{il}(t - t_l)_+$ and $\boldsymbol{A}_{\boldsymbol{\beta},il}(t - t_l)_+ = \partial\boldsymbol{A}_{il}(t - t_l)_+/\partial\boldsymbol{\beta}$, where $\boldsymbol{A}_{il}(t - t_l)_+ = \{A_{1,il}(t - t_l)_+, A_{2,il}(t - t_l)_+, A_{3,il}(t - t_l)_+\}^T$. Write (5) as $\mathbf{A}_{il}(t - t_l)_+/dt = \mathbf{G}_i\{\boldsymbol{\beta}(t)\}\mathbf{A}_{il}(t - t_l)_+$ and take a derivative with respect to $\boldsymbol{\beta}$ on both sides, and we have

$$\frac{d\mathbf{A}_{\boldsymbol{\beta},il}(t - t_l)_+}{dt} = \mathbf{G}_{\boldsymbol{\beta},i}\{\boldsymbol{\beta}(t)\}\mathbf{A}_{il}(t - t_l)_+$$
$$+ \mathbf{G}_i\{\boldsymbol{\beta}(t)\}\mathbf{A}_{\boldsymbol{\beta},il}(t - t_l)_+, \qquad (21)$$

where $\mathbf{G}(\cdot)_{\boldsymbol{\beta},i} = \partial\mathbf{G}_i(\cdot)/\partial\boldsymbol{\beta}$. Hence both $\mathbf{A}_{il}(\cdot)$ and $\mathbf{A}_{\boldsymbol{\beta},il}(\cdot)$ can be obtained by solving the differential equations (1) and (21). Note that the initial value $\mathbf{A}_{\beta l}, (0)$ is always 0, because changing the PK parameters $\boldsymbol{\beta}$ will not affect the initial value $\mathbf{A}_{il}(0)$. More details of the numerical solution of the differential equations can be found in Li et al. (2002).

Wu (2002) applied an EM algorithm to maximize a marginal likelihood function similar to (17). He treated both the random effects $\mathbf{b}$ and the unobserved variable $\mathbf{U}$ as missing data. At the E-step, $\mathbf{b}$ and $\mathbf{U}$ were estimated using the observed data $(\mathbf{Y}, \boldsymbol{W})$ by assuming the conditional distribution of $(\boldsymbol{b}, \boldsymbol{U} \mid \boldsymbol{Y}, \boldsymbol{W})$ to be normal. At the M-step, given the estimated $\mathbf{b}$ and $\mathbf{U}$, he maximized the joint likelihood function of $(\mathbf{Y}, \mathbf{W}, \mathbf{b}, \mathbf{U})$ over all the other parameters.

Our two-step method assumes that almost all of the information about $\mathbf{U}$ is contained in $\mathbf{W}$. Hence, $\mathbf{U}$ is estimated from $\mathbf{W}$ exclusively at step one. At step two, the parameters of the PK model and the random effects $\boldsymbol{b}$ are estimated jointly from the Solomon–Cox–Laplace approximation (19) of the likelihood function of $\mathbf{Y}$. Since he assumed normality for the conditional distribution of $(\boldsymbol{b}, \boldsymbol{U} \mid \boldsymbol{Y}, \boldsymbol{W})$, his EM algorithm can be viewed as a second-order approximation of the marginal likelihood function (15), similar to ours. However, unlike his method, we estimate $\mathbf{U}$ only once, and therefore the computational burden required by numerically solving the differential equations is considerably reduced.

3.4 *Second-Order Two-Step Method with Approximately Closed Form Solutions of Differential Equations*

Unless the PK parameters are time independent, there is no closed form solution for the differential equations (5) in general. However, if the PK parameters change only slightly over time, the closed form solution based on the constant PK parameter assumption would provide a good approximation to the exact solution. In this section, we make such an assumption and solve the differential equations analytically in a closed form. The resulting estimation procedure is hence computationally much faster. Specifically, under this assumption, the solution of (5) is

$$A_{2,il}(t - t_l)_+ = \frac{k_{a,i}(t)\{k_{32} - k_{a,i}(t)\}e^{-k_{a,i}(t)\times(t-t_l)_+}}{\{\lambda_{1,i}(t) - k_{a,i}(t)\}\{\lambda_{2,i}(t) - k_{a,i}(t)\}}$$
$$+ \frac{k_{a,i}(t)\{k_{32} - \lambda_{1,i}(t)\}e^{-\lambda_{1,i}(t)\times(t-t_l)_+}}{\{\lambda_{2,i}(t) - \lambda_{1,i}(t)\}\{k_{a,i}(t) - \lambda_{1,i}(t)\}}$$
$$+ \frac{k_{a,i}(t)\{k_{32} - \lambda_{2,i}(t)\}e^{-\lambda_{2,i}(t)\times(t-t_l)_+}}{\{\lambda_{1,i}(t) - \lambda_{2,i}(t)\}\{k_{a,i}(t) - \lambda_{2,i}(t)\}}, \qquad (22)$$

where

$$\lambda_{1,i}(t) = \big[k_e + k_{23,i}(t) + k_{32}$$
$$+ \big\{(k_e + k_{23,i}(t) + k_{32})^2 - 4k_e k_{32}\big\}^{1/2}\big]\big/2,$$
$$\lambda_{2,i}(t) = \big[k_e + k_{23,i}(t) + k_{32}$$
$$- \big\{(k_e + k_{23,i}(t) + k_{32})^2 - 4k_e k_{32}\big\}^{1/2}\big]\big/2.$$

We call this method the second-order two-step method with approximately closed form solutions of differential equations (2orderAD). It is the same as the 2orderND method described in Section 3 except that a closed form expression $f(\cdot)$ is used in the estimation at the second step when fitting (18).

4. Data Analysis

We applied the proposed measurement error PK model (6)–(9) to the analysis of the Ditropan data introduced in Section 1 to study the effect of DBP on the distribution rate $k_{23}(t)$. We also considered the effects of SBP and HR on $k_{23}(t)$. Since DBP, SBP, and HR are highly correlated, we analyzed each of them separately by assuming W to be either of these three variables.

At step one, the linear mixed model (9) with a random intercept was fitted to DBP, SBP, and HR separately using the Splus function LME. We modeled the time effects using quadratic, cubic, and quartic regression, and compared their fits using the AICs. The quartic model for DBP, and the

Table 2
AICs and parameter estimates of the polynomial mixed models (9) *for DBP, SBP, and HR for the Ditropan study*

		Diastolic blood pressure	Systolic blood pressure	Heart rate
AIC	Linear	1844.2	1874.0	1857.9
	Quadratic	1839.5	1871.4	1854.2
	Cubic	1827.1	1868.4	1850.4
	Quartic	1826.8	1879.9	1855.5
Parameter estimate[a]	Intercept	66.76	107.97	68.79
	$(t-13)$	0.33	0.11	0.13
	$(t-13)^2$	0.02	0.02	0.01
	$(t-13)^3$	-0.002	-0.0004	-0.0004
	$(t-13)^4$	0.00004		
	σ_a^2	6.43	8.08	6.88
	σ_e^2	5.90	7.27	7.86

[a]The parameter estimates of the model with the smallest AIC.

cubic models for SBP and HR gave the lowest AIC. The AIC values of all models and the parameter estimates associated with the model with the smallest AICs for each variable are given in Table 2. Higher-order random effects models, e.g., random slope models, were also tried. They all had higher AIC values. The solid lines in Figure 1a–1c give the polynomial fits with the lowest AICs for these three time-dependent covariates. The curves all first decreased, then increased, and then decreased again. This indicates that BP and HR react to the three doses.

At step two, the plasma concentration was modeled using the nonlinear mixed model (18), where $f_i(\cdot) = f[\{A_{2,i1}(t-t_1)_+ + A_{2,i2}(t-t_2)_+ + A_{2,i3}(t-t_3)_+\}/V_i]$, $A_{2,il}(t)$ is given (5), and

$$k_{23,i}(t_{ij}) = \{\alpha_6 + \alpha_7 U(t_{ij})\}\exp(b_{3i}), \quad (23)$$

and $U(t)$ is set to be DBP, SBP, or HR. Parameter estimators were obtained using the 2orderND method and the 2orderAD method described in Section 3 to account for measurement error in DBP, SBP, and HR. The results are given in Table 3. Using the 2orderND method, we found that $k_{23,i}(t)$ was significantly positively associated with true DBP(t) ($\alpha_7 = 0.0042$, p value $= 0.022$), and there was some evidence of a positive association between $k_{23,i}(t)$ and true *SBP(t)* ($\alpha_7 = 0.0023$, p value $= 0.125$) and true *HR(t)* ($\alpha_7 = 0.0032$, p value $= 0.075$). These results partially support our hypothesis that the distribution process could be blood flow sensitive.

To investigate the effect of not incorporating the time-varying covariate, e.g., DBP, in the $k_{23,i}(t)$ model, we compared the residuals of the model assuming $k_{23,i}(t)$ was a function of DBP given in equation (23) with $U(t) = $ true *DBP(t)*, with the model assuming $k_{23,i}(t)$ was free of DBP as $k_{23,i}(t_{ij}) = \alpha_6\exp(b_{3i})$. These two residual plots are given in Figure 2a and 2b. Figure 2c shows the average differences between these two sets of residuals, and Figure 2d shows the estimate curve of $k_{23}(t)$ as a function of t and its 95% confidence interval.

Figure 2c suggests that if the effect of DBP is not modeled, the residuals are too high during the second dosing period,

Table 3
Parameter estimates under the PK model with the time-dependent covariate measured with error (6)–(9) *for the Ditropan study*

Parameters	2orderND (SE) (DBP)	2orderND (SE) (SBP)	2orderND (SE) (HR)	2orderAD (SE) (DBP)
$V = e^{\alpha_1}$	0.171 (0.010)	0.170 (0.010)	0.171 (0.010)	0.187 (0.009)
$k_{a1} = e^{\alpha_2}$	6.536 (0.839)	6.547 (0.854)	6.541 (0.861)	7.232 (1.112)
$k_{a2} = e^{\alpha_3}$	3.591 (0.327)	3.668 (0.344)	3.625 (0.342)	4.654 (0.400)
$k_{a3} = e^{\alpha_4}$	2.612 (0.169)	2.673 (0.173)	2.650 (0.166)	3.276 (0.287)
$k_e = e^{\alpha_5}$	0.603 (0.016)	0.611 (0.017)	0.607 (0.016)	0.455 (0.020)
α_6	-0.040 (0.088)	-0.013 (0.072)	-0.025 (0.077)	-0.079 (0.111)
α_7	0.0042 (0.002)	0.0023 (0.0015)	0.0032 (0.0018)	0.0059 (0.0017)
$k_{32} = e^{\alpha_8}$	0.091 (0.008)	0.084 (0.008)	0.085 (0.008)	0.075 (0.006)
σ_ϵ^2	0.162 (0.012)	0.164 (0.012)	0.164 (0.012)	0.166 (0.013)
σ_{11}	0.265 (0.037)	0.269 (0.038)	0.266 (0.037)	0.255 (0.044)
σ_{22}	0.167 (0.042)	0.166 (0.042)	0.167 (0.042)	0.170 (0.039)
σ_{33}	0.256 (0.111)	0.273 (0.122)	0.265 (0.118)	0.268 (0.121)

2orderND: second-order two-step method with the numerical solutions of the differential equations.
2orderAD: second-order two-step method with the approximately closed form solution of the differential equations.
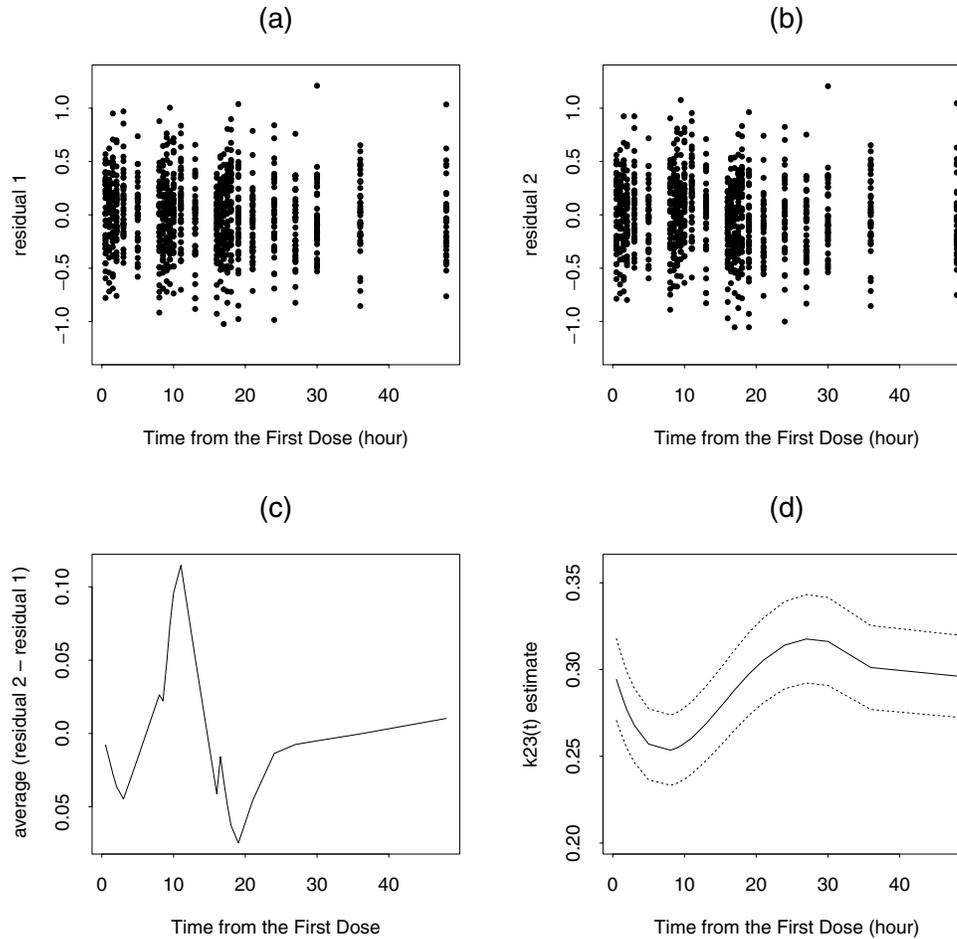
Figure 2. (a) The residual plot when $k_{23,i}(t)$ was modeled as $\{\alpha_6 + \alpha_7 \times \text{true}DBP(t)\}\exp(b_{3i})$; (b) the residual plot when $k_{23,i}(t)$ was modeled as $\{\alpha_6\}\exp(b_{3i})$; (c) the average residual differences between (a) and (b); and (d) the estimate of $k_{23,i}(t)$ as a function of t under (a). Solid line is the fitted curve of $k_{23,i}(t)$; dotted lines are its 95% confidence interval.

and too low during the first and the third dosing periods. In other words, there was a clear lack of fit if the mean of $k_{23,i}(t)$ was assumed to be a constant and free of DBP. This lack of fit was caused by the fact that the second dose had a lower k_{23} than the other two doses, and the lower k_{23} in the second dosing period might be due to the fact that a higher DBP was associated with the second dose.

We also analyzed the data using the 2orderAD method, where model (23) was used in calculating the mean function $f(\cdot)$ in (3). The results in Table 3 show that the PK parameters were overestimated using the 2orderAD method. Our simulation study confirms this finding. We present 2orderAD analysis with only DBP in Table 3; the other two time-dependent covariates have similar results. The biases of the 2orderAD estimates were likely to result from the poor approximation of the closed form solution of the differential equations. Both methods, however, provided comparable variance component estimates.

5. Simulation

We performed a simulation study to evaluate the performance of the two-step 2orderND and 2orderAD procedures. The de-

sign of the simulation study was identical to that of the original Ditropan data. Specifically, we generated the data from model (6)–(9) with $f(\cdot)$ given in (3), and set the true parameter values equal to the estimates from the analysis of the real DBP data given in Tables 2 and 3. In each simulated data set, there were 40 subjects and each subject had 25 observations over time. In addition, every subject's DBP was simulated at nine different fixed sampling time points during 48 hours. We used 500 Monte Carlo data sets.

We analyzed each simulated data set using the 2orderND and 2orderAD methods. At step one, the AIC was used to estimate the degree of the polynomial of $DBP(t)$. Hence both methods had the same estimate of true $DBP(t)$. At step two, PK parameters were estimated from the mean function which was calculated using the differential equations (2orderND) or the approximate analytic solution of the differential equations (2orderAD). Table 4 summarizes the average point estimates, model-based SEs (MOD-SE), empirical SEs (EMP-SE), relative biases of the parameter estimates, and the ratios of the model-based SEs and the empirical SEs.

When the 2orderND method was used, 6 out of 500 data sets did not converge. The results in Table 4 show that the

Table 4
Simulation results using the 2orderND method

Parameter	True value	Estimate	MOD-SE	EMP-SE	RB	MOD-SE/EMP-SE
$V = e^{\alpha_1}$	0.17	0.165	0.007	0.008	-0.029	0.875
$k_{a1} = e^{\alpha_2}$	6.50	5.877	0.649	0.711	-0.096	0.913
$k_{a2} = e^{\alpha_3}$	3.60	3.356	0.219	0.221	-0.065	0.990
$k_{a3} = e^{\alpha_4}$	2.50	2.457	0.182	0.212	-0.017	0.858
$k_e = e^{\alpha_5}$	0.60	0.612	0.021	0.022	0.020	0.954
α_6	-0.04	-0.035	0.081	0.091	-0.125	0.890
α_7	0.0042	0.0038	0.0016	0.0018	-0.095	0.888
$k_{32} = e^{\alpha_8}$	0.09	0.092	0.008	0.009	0.022	0.888
σ_ϵ^2	0.162	0.174	0.014	0.016	0.074	0.875
σ_{11}	0.265	0.247	0.042	0.046	-0.068	0.913
σ_{22}	0.167	0.149	0.049	0.059	-0.108	0.830
σ_{33}	0.256	0.207	0.152	0.178	-0.191	0.853

2orderND: second-order two-step method with the numerical solutions of the differential equations.
MOD-SE: model-based standard error.
EMP-SE: empirical standard error.
RB: relative bias.

Table 5
Simulation results using the 2orderAD method

Parameter	True value	Estimate	MOD-SE	EMP-SE	RB	MOD-SE/EMP-SE
$V = e^{\alpha_1}$	0.17	0.192	0.007	0.008	0.129	0.875
$k_{a1} = e^{\alpha_2}$	6.50	8.353	1.013	1.210	0.285	0.837
$k_{a2} = e^{\alpha_3}$	3.60	4.872	0.364	0.429	0.353	0.848
$k_{a3} = e^{\alpha_4}$	2.50	3.030	0.271	0.308	0.212	0.879
$k_e = e^{\alpha_5}$	0.60	0.490	0.018	0.019	0.183	0.947
α_6	-0.040	-0.0068	0.092	0.102	-0.830	0.902
α_7	0.0042	0.0057	0.0018	0.0020	0.357	0.900
$k_{32} = e^{\alpha_8}$	0.09	0.076	0.006	0.008	-0.156	0.750
σ^2	0.162	0.175	0.015	0.018	0.080	0.833
σ_{11}	0.265	0.242	0.045	0.049	-0.087	0.917
σ_{22}	0.167	0.143	0.054	0.067	-0.144	0.805
σ_{33}	0.256	0.201	0.151	0.181	-0.214	0.834

2orderAD: second-order two-step method with the approximately closed form solution of the differential equations.
MOD-SE: model-based standard error.
EMP-SE: empirical standard error.
RB: relative bias.

2orderND method performs well. The PK parameter estimates had small biases, with the largest 12% bias observed in the estimate of α_7, which describes the association between true $DBP(t)$ and $k_{23}(t)$. The variance component estimates also had small biases. The biases of the estimates of σ_ϵ^2 and σ_{11}^2 were smaller than those of σ_{22}^2 and σ_{33}^2. These results are expected, as ϵ_{ij} and b_{1i} are linearly related to y_{ij}, whereas b_{2i} and b_{3i} are nonlinearly related to y_{ij}. The model-based SEs slightly underestimated the true SEs.

The 2orderAD method was slightly less stable than the 2orderND method, and 9 out of 500 of the simulated data sets did not converge. The results in Table 5 show that the estimates are considerably biased, especially the estimates of α_6 and α_7, whose biases could be as high as 80%. This is expected, since the 2orderAD method uses an approximate analytic solution of the differential equations by assuming the PK parameters change slowly with t. If the PK parameters vary considerably with the time-varying covariate, the approxima-

tion is likely to be poor. Both methods, however, provided comparable variance component estimates.

6. Conclusion

We considered in this article a population PK model with the PK parameters modeled as functions of time-dependent covariates measured with errors. We proposed two two-step estimation methods. Both methods are computationally easy and stable, and only require fitting a linear mixed model at the first stage followed by fitting a nonlinear mixed model at the second stage. The 2orderND method accounts for the fact that the PK parameters depend on a time-varying covariate and change over time, and solves the differential equations numerically, while the 2orderAD method assumes the PK parameters vary slowly with time and approximates the solution of the differential equations in a closed form. Our simulation results show that the 2orderND method performs well, and has small biases in parameter estimators and the model-based

standards errors are close to the empirical standard errors. The 2orderAD method is subject to considerable bias. For simplicity, our presentation focused on the two-compartment model in this article. The proposed method is applicable to general nonlinear mixed models commonly used in pharmacokinetics.

We showed that the 2orderND method could be derived as the Solomon–Cox–Laplace approximation of the integrated likelihood. Our simulation study suggests that such a likelihood approximation works well in practice. In view of the computational advantage of the proposed method and the considerable additional computational burden of the maximum likelihood method, it would be of future research interest to investigate whether the magnitude of improvement of the maximum likelihood method over the 2orderND method is practically appreciable and is worth additional computational burden. Based on our proposed 2orderND method, we found using the Ditropan data that the true underlying processes of diastolic blood pressure, systolic blood pressure, and heart rate were all positively correlated with $k_{23}(t)$, although only diastolic blood pressure showed a statistically significant association. This finding suggests that drug-distribution rate might be flow sensitive as discussed at the end of Section 1.2. To have a better power to verify the association between $k_{23}(t)$ and blood pressure and HR, a larger range of systematically changed blood pressure and HR is needed. For example, it would be desirable to let subjects exercise during one dosing period, and rest in another period in a new study. On the other hand, in order to understand the association between blood pressure and S-oxybutynin's pharmacokinetics among hypertension and hypotension patients, more studies are necessary. As this study was performed on healthy subjects, it remains for further studies to verify our findings and determine its clinical significance in the patient population. If the decreasing k_a pattern is verified, the controlled release formulation may need to be modified to keep the drug concentration as stable as possible. If the flow-sensitive hypothesis of k_{23} is verified, more studies may need to be done carefully among hypertension and hypotension patients.

Acknowledgements

Résumé

Nous proposons un modèle de pharmococinétique de population (PK) avec des erreurs sur la détermination de covariables dépendant du temps. Nous utilisons ce modèle pour traiter la cinétique de la S-oxybutinine après administration orale de Ditropan en faisant varier la constante de distribution en fonction du rythme cardiaque et de la pression sanguine, quantités connues avec erreurs. Nous proposons deux méthodes du deuxième ordre et en deux étapes: une méthode avec résolution numérique des équations différentielles (2orderND) et une méthode avec des formes fermées de solutions approchées des équations différentielles (2orderAD). Ces deux méthodes donnent des calculs faciles et réclament l'ajustement d'un modèle linéaire mixte à la première étape et d'un modèle non linéaire mixte à la seconde. Nous appliquons les deux méthodes à l'analyse des données du Ditropan et leurs performances sont étudiées par simulation. Nos résultats montrent que la méthode 2orderND se comporte bien alors que la méthode 2orderAD peut donner des estimations des paramètres de PK sujettes à des biais considérables.

References

AHFS Drug Information (2000). Bethesda, Maryland: American Society of Hospital Pharmacists.

Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models.* London: Chapman and Hall.

Davidian, M. and Giltinan, D. G. (1995). *Nonlinear Models for Repeated Measurement Data.* London: Chapman and Hall.

Douchamps, J., Derenne, F., Stockis, A., Gangji, D., Juvent, M., and Herchuelz, A. (1988). The pharmacokinetics of oxybutynin in man. *European Journal of Clinical Pharmacology* **35,** 515–520.

Eugene, B. (2001). *Harrison's Principles of Internal Medicine.* New York: McGraw-Hill, Medical Publishing Division.

Higgins, K. M., Davidian, M., and Giltinan, D. M. (1997). A two-step approach to measurement error in time-dependent covariates in nonlinear mixed-effect models, with application to IGF-I pharmacokinetics. *Journal of the American Statistical Association* **92,** 436–448.

IMSL User's Manual STAT Library Version 1.0 (1987). Houston, Texas: IMSL, Inc.

Li, L., Brown, M., Lee, K., and Gupta, S. (2002). Estimation and inference for a spline-enhanced population pharmacokinetic model. *Biometrics* **58,** 601–611.

Lindstorm, J. M. and Bates, M. D. (1990). Nonlinear mixed effects models for repeated measurement data. *Biometrics* **46,** 673–687.

Solomon, P. J. and Cox, D. R. (1992). Nonlinear components of variance models. *Biometrika* **79,** 1–11.

Wolfinger, R. D. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika* **80,** 791–795.

Wolfinger, R. D. and Lin, X. (1997). Two Taylor-series approximations methods for nonlinear mixed models. *Journal of Computational Statistics and Data Analysis* **25,** 465–490.

Wu, L. (2002). A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *Journal of the American Statistical Association* **97,** 955–964.