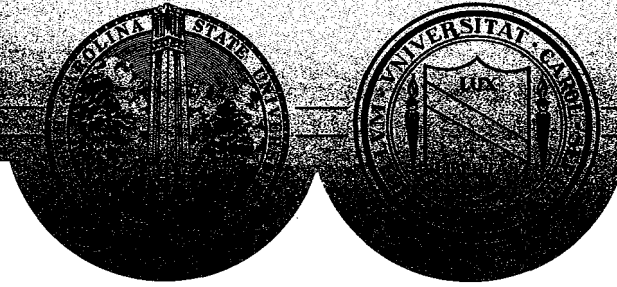


# THE INSTITUTE OF STATISTICS

UNIVERSITY OF NORTH CAROLINA SYSTEM



SMOOTH NONPARAMETRIC MAXIMUM LIKELIHOOD  
ESTIMATION FOR POPULATION PHARMACOKINETICS, WITH  
APPLICATION TO QUINIDINE

by

Marie Davidian and A. Ronald Gallant

Institution of Statistics Mimeograph Series No. 2221

May, 1992

NORTH CAROLINA STATE UNIVERSITY  
Raleigh, North Carolina

MIMEO Marie Davidian and  
SERIES A. Ronald Gallant  
#2221 SMOOTH NONPARAMETRIC  
MAXIMUM LIKELIHOOD ESTIMATION  
ESTIMATION FOR POPULATION PHARMA-  
COKINETICS, WITH APPLICATION TO  
QUINIDINE.....

NAME	DATE
------	------

The Library of the Department of Statistics  
North Carolina State University

# Smooth Nonparametric Maximum Likelihood Estimation for Population Pharmacokinetics, with Application to Quinidine <sup>1</sup>

Marie Davidian  
North Carolina State University

A. Ronald Gallant  
North Carolina State University

March 1992

Revised April 1992

<sup>1</sup>Corresponding author: Marie Davidian, Department of Statistics, North Carolina State University, Campus Box 8203, Raleigh NC 27695-8203 USA, 1-919-515-2532. This material is based upon work supported by the National Science Foundation and the North Carolina Agricultural Experiment Station. We should like to thank Lewis B. Sheiner and Thomas M. Ludden for use of data and helpful discussions.

## ABSTRACT

The SNP method, popular in the econometrics literature, is proposed for use in population pharmacokinetic analysis. For data that can be described by the nonlinear mixed effects model, the method produces smooth nonparametric estimates of the entire random effects density and simultaneous estimates of fixed effects by maximum likelihood. A new graphical model building strategy based on the SNP method is introduced. The methods are illustrated by a population analysis of plasma levels in 136 patients undergoing oral quinidine therapy.

**KEY WORDS:** Population pharmacokinetics; nonlinear mixed effects models; density estimation; nonparametric estimation, maximum likelihood, quinidine.

# INTRODUCTION

The statistical nonlinear mixed effects model has been used in population pharmacokinetic analysis since the pioneering work of Beal and Sheiner[1]; see also [2, 3, 4]. These models account for inter-individual variability in pharmacokinetic parameters and characterize the distribution of these parameters. The parameters vary from individual to individual due to variation in observable individual attributes and sampling variation in unobservable random effects that follow a distribution. The objectives of an analysis based on these models include estimation of population characteristics (mean, variance, etc.) of the pharmacokinetic parameters, assessment of the effect of individual attributes (weight, age, etc.) on the population characteristics, and computation of empirical Bayes estimates of an individual's pharmacokinetic parameters, which can be used in individual dosage adjustments [5]. The work of Mallet [6], who proposed a nonparametric maximum likelihood (NPML) estimator, and his co-workers Mentré, Steimer, and Lokiec [7], has generated an interest in estimation of the entire distribution of the random effects and demonstrated the importance of nonparametric estimation of the distribution when there is cause to expect departures from standard specifications such as the normal or log normal distribution. One important departure is multi-modality, which is often caused by omission of a relevant individual attribute from the description of pharmacokinetic parameters in the model.

Here we discuss a nonparametric method called SNP that is particularly relevant to population pharmacokinetic applications using the nonlinear mixed effects model. It is based on a presumption that the distribution of the random effects has a smooth density. The method was originally developed for the models of inter-individual heterogeneity that occur in labor economics [8] and has also been used in time series analysis [9, 10, 11]. It was adapted to nonlinear mixed effects models in [12], which contains an application of the method to the population pharmacokinetics of phenobarbital and simulations assessing its ability to fit multi-modal random effects densities. SNP is an acronym for seminonparametric which is a usage that predates its application to nonlinear mixed effects models [13, 14]. With respect to nonlinear mixed effects models, smooth nonparametric would be a more meaningful interpretation of the acronym.

The SNP method simultaneously estimates the fixed effects and the entire distribution of the random effects of the nonlinear mixed effects model. Subsequent computations based on SNP estimates are convenient: statistical tests of the significance of individual attributes, tests for normality of the random effects density, empirical Bayes estimation of individual random effects, and simulation from the estimated density. The ability to simulate greatly facilitates computation of the population characteristics of pharmacokinetic parameters that are affected nonlinearly by the random effects.

In this article, we relate the SNP method to population pharmacokinetic applications, introduce a new graphical model building strategy that exploits the tendency for omitted individual attributes to produce disparate empirical Bayes estimates of individual effects, and illustrate by application to clinical data on quinidine concentration.

## MAXIMUM LIKELIHOOD ESTIMATION

The SNP method is similar to the NPML method of [6, 7] in that it invokes the likelihood principle as the basis for estimation. Mallet et al. [7] provide an excellent discussion of the underlying ideas, relating them specifically to population pharmacokinetics. The derivation of the likelihood involves two steps: specification of an individual likelihood function in terms of an individual's pharmacokinetic parameters and specification of the population likelihood in terms of a distribution for the pharmacokinetic parameters. In both methods this distribution is not restricted to belong to a parametric family but is estimated nonparametrically in its entirety; the methods differ in choice of nonparametric estimation method.

### Individual Likelihood Function

Let  $y_{ij}$ ,  $1 \leq j \leq J_i$ , be the observed concentration measurements on individual  $i$ ,  $1 \leq i \leq n$ , at settings  $x_{ij}$  of a  $k$ -vector of independent variables such as time, dose, rate of administration, etc. These are assumed to follow the intra-individual regression model

$$y_{ij} = f(x_{ij}, \beta_i) + e_{ij}.$$

The function  $f$  describes the pharmacokinetics in terms of the independent variables  $x_{ij}$  and the  $p_\beta$ -vector of pharmacokinetic parameters  $\beta_i$  specific to the individual;  $e_{ij}$  denotes the

intra-individual error associated with the  $j$ th observation on individual  $i$ . Note that  $e_{ij}$  is unrestricted so that this assumption subsumes specifications such as  $y_{ij} = f(x_{ij}, \beta_i)(1 + \eta_{ij})$  with  $\eta_{ij}$  either iid or serially correlated. The total number of observations is  $N = \sum_{i=1}^n J_i$ .

Measurement error, primarily due to variation in the assay used to process samples, is the largest component of  $e_{ij}$ . Specification of a distribution for this error completes the description of the individual likelihood. Denote the joint density of the errors  $e_{ij}$  by

$$p_e(e_{i1}, \dots, e_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, \sigma, \beta_i)$$

where  $\beta_i$  is as above and  $\sigma$  is a  $p_\sigma$ -vector denoting parameters specific to the density such as parameters that characterize the second and higher moments of the density. For example, a common assumption is that the errors are independently and normally distributed with standard deviation proportional to level [15] which is written as

$$p_e(e_{i1}, \dots, e_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, \sigma, \beta_i) = \prod_{j=1}^{J_i} n\{e_{ij} | 0, [\sigma f(x_{ij}, \beta_i)]^2\},$$

where  $n(\cdot | \mu, \omega^2)$  denotes the normal density with mean  $\mu$  and variance  $\omega^2$ . Specifications other than the constant coefficient of variation variance function  $[\sigma f(x_{ij}, \beta_i)]^2$  are possible; for example, see [7]. One can specify more general forms for  $p_e$  that, for example, account for serial correlation, in which case the density would not be a product of individual marginal densities.

## Population Likelihood Function

The pharmacokinetic parameters  $\beta_i$  vary from individual to individual. Part of this inter-individual variation may be explained by systematic dependence of pharmacokinetic parameters on demographic and other individual attributes, also called covariates. The unexplained portion is assumed to be random and is characterized by a probability density  $h$ . These dependencies are usually represented by an inter-individual regression model of the form

$$\beta_i = g(w_i, \gamma, z_i).$$

The inter-individual regression function  $g$  describes the interaction of systematic and random sources of variation. In this function,  $w_i$  is a vector of covariates,  $\gamma$  is a  $p_\gamma$ -dimensional vector

of unknown fixed effects, and  $z_i$  is an  $M$ -dimensional vector of inter-individual random effects with density  $h(z)$ . An example is a log-additive specification of clearance  $Cl_i$  for individual  $i$  as a function of body weight  $w_{1i}$ , an indicator variable  $w_{2i}$  for ethanol abuse, and a random effect  $z_{1i}$ :

$$Cl_i = \exp(\gamma_0 + \gamma_1 w_{1i} + \gamma_2 w_{2i} + z_{1i}),$$

which would be one equation from the system of equations  $g(w, \gamma, z)$ . Note that the dimension of  $g$  is  $p_\beta$ , the dimension of  $z$  is  $M$ , and they need not be equal, as in the case where a pharmacokinetic parameter is assumed to be fixed (not random) across the population.

Collecting the parameters  $\gamma$  and  $\sigma$  together into the single  $p_\tau$ -vector

$$\tau = (\gamma, \sigma), \quad p_\tau = p_\gamma + p_\sigma,$$

the population likelihood can be written as

$$\ell(\tau, h) = \prod_{i=1}^n \int p_y(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_i, \tau, z_i) h(z) dz.$$

Here,  $p_y(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_i, \tau, z_i)$  is the joint density of the observations on individual  $i$  implied by the pharmacokinetic model  $f$  and the error density  $p_e$ . It is obtained by substituting the equation  $e_{ij} = y_{ij} - f[x_{ij}, g(w_i, \gamma, z_i)]$  into the error density  $p_e[e_{i1}, \dots, e_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, \sigma, g(w_i, \gamma, z_i)]$  (because additive errors,  $y_{ij} = f(x_{ij}, \beta_i) + e_{ij}$ , imply that the Jacobian of  $(e_{i1}, \dots, e_{iJ_i})$  with respect to  $(y_{i1}, \dots, y_{iJ_i})$  is the identity matrix of order  $J_i$ ).

Some individual attributes may change over the period of observation. This situation is accommodated by permitting the individual pharmacokinetic parameters to depend on  $j$  as well as  $i$  and by writing the inter-individual regression model as  $\beta_{ij} = g(w_{ij}, \gamma, z_i)$ . For instance, in the example above write  $Cl_{ij} = \exp(\gamma_0 + \gamma_1 w_{1ij} + \gamma_2 w_{2i} + z_{1i})$  to indicate that measured body weight may change over the period of observation. In this case the joint density of the observations on individual  $i$  has the form  $p_y(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_{i1}, \dots, w_{iJ_i}, \tau, z_i)$ .

## SNP Estimation

In this framework, an important objective is determination of an appropriate inter-individual regression model  $g(w, \gamma, z)$  from the data. For each  $g$ , estimation of and inference regarding



the parameters  $\tau = (\gamma, \sigma)$  is of interest. Another important objective is the determination of  $h$ , as  $h$  characterizes the population of individuals.

Nonparametric estimation of  $h$  allows one to detect unusual features of the population such as multi-modality, which often indicates the presence of systematic inter-individual variability and the need for a more refined inter-individual regression model  $g$ . It also affords protection against incorrect assumptions regarding  $h$  that can bias estimates of  $\tau$  and lead to erroneous inferences.

Nonparametric estimation of  $h$  can be accomplished within the likelihood framework by maximizing the likelihood  $\ell(\tau, h)$ , either simultaneously or sequentially, in the fixed parameters  $\tau$  and the density  $h$ . One maximizes with respect to  $h$  over a wide class of distributions. The NPML estimator is obtained when the class of distributions is completely unrestricted. A consequence is that NPML estimates are discrete densities that assign probability to a finite number of points.

As the true density of the random effects associated with pharmacokinetic parameters is likely to be smooth, estimates ought have this characteristic as well. One can either obtain smooth estimates by smoothing the NPML estimator *ex post* or by imposing smoothness on the estimator *ex ante*. The SNP estimator is obtained by adopting the latter approach.

To eliminate ambiguity in the sequel, let  $\tau^\circ$  denote the true values of the fixed parameters, let  $h^\circ$  denote the true random effects density, let  $\beta_i^\circ$ ,  $1 \leq i \leq n$ , denote the true realized values of the individual pharmacokinetic parameters, and let  $z_i^\circ$ ,  $1 \leq i \leq n$ , denote the true realized values of the random effects.

SNP estimates  $(\hat{\tau}, \hat{h})$  maximize the likelihood

$$\ell(\tau, h) = \prod_{i=1}^n \int p_y(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_i, \tau, z) h(z) dz,$$

where  $p(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_i, \tau, z_i)$  is the joint density of the observations on individual  $i$ . The approach is based on the assumption that  $h^\circ$  belongs to a class of smooth densities  $\mathcal{H}$  described below. This assumption allows  $h$  to be written such that maximization of  $\ell(\tau, h)$  over  $\tau$  and  $h$  becomes a standard nonlinear optimization problem. The method is summarized in this section; for the theory see [8, 12].

The maximum likelihood estimates  $(\hat{\tau}, \hat{h})$  of  $(\tau^\circ, h^\circ)$  may be equivalently computed by

minimizing

$$s_n(\tau, h) = -\frac{1}{N} \sum_{i=1}^n \log \int p_y(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_i, \tau, z) h(z) dz$$

over  $\tau$  and  $h \in \mathcal{H}$ . Once the estimates  $(\hat{\tau}, \hat{h})$  are obtained, empirical Bayes estimates of the  $z_i^o$ ,  $1 \leq i \leq n$ , are computed as the values  $\hat{z}_i$  that maximize

$$p(y_{i1}, \dots, y_{iJ_i} | x_{i1}, \dots, x_{iJ_i}, w_i, \hat{\tau}, z) \hat{h}(z).$$

From these, the empirical Bayes estimate of an individual's pharmacokinetic parameter  $\beta_i^o$  is obtained by evaluating  $\hat{\beta}_i = g(w_i, \hat{\gamma}, \hat{z}_i)$  [or  $\hat{\beta}_{ij} = g(w_{ij}, \hat{\gamma}, \hat{z}_i)$ ].

The true density  $h^o$  is presumed to be in a class  $\mathcal{H}$  of smooth densities. The smoothness restriction is that  $h$  must be at least  $M/2$  times differentiable. This restriction rules out unusual behavior such as kinks, jumps, or oscillation but does permit  $h$  to be skewed, multi-modal, and fat- or thin-tailed relative to the  $M$ -variate normal density. The class  $\mathcal{H}$  thus contains densities that accommodate a wide range of behavior. For a mathematical description of  $\mathcal{H}$  see [8].

As a practical matter, it is the ability to represent multi-modal densities and densities that are more spread-out than the normal density (fat-tailed) that is important. Multimodality is often a consequence of systematic dependence of a pharmacokinetic parameter on an individual attribute that has been omitted from the inter-individual regression function  $g$ . Distributions that exhibit more heterogeneity than permitted by the normal density (fat-tails) are not unusual in observational data on human subjects. Failure to track these characteristics can distort empirical Bayes estimates of individual random effects and estimates of population characteristics that are nonlinear in the random effects.

Gallant and Nychka [8] have shown that a density from  $\mathcal{H}$  can be written as the infinite series expansion

$$h(z) = \left[ \sum_{|\lambda| < \infty} a_\lambda (R^{-1}z)^\lambda \right]^2 n_M(z | 0, RR'),$$

where: The  $M$ -dimensional vector  $\lambda = (\lambda_1, \dots, \lambda_M)$  has nonnegative integers as elements;  $u^\lambda$  is the monomial  $u^\lambda = u_1^{\lambda_1} \cdots u_M^{\lambda_M}$  of order  $|\lambda| = \sum_{k=1}^M \lambda_k$ ;  $n_M(\cdot | \mu, \Sigma)$  denotes the multivariate normal density of dimension  $M$ , with mean  $\mu$ , and with variance-covariance matrix  $\Sigma$ ; and  $R$  is an upper-triangular matrix.

One can approximate  $h$  by truncating the infinite expansion, retaining only the leading terms  $[P_K(R^{-1}z)]^2 n_M(z|0, RR')$ , where  $P_K(u) = \sum_{|\lambda| < K} a_\lambda u^\lambda$  denotes a polynomial of degree  $K$ . The truncated expansion will be a density if the coefficients  $\{a_\lambda : 0 \leq |\lambda| \leq K\}$  are chosen so that it integrates to one, that is,  $\int \{P_K(u)\}^2 n_M(u|0, I) du = 1$ . The easiest way to impose this condition is to put  $a_0 \equiv 1$  and write the truncated expansion as

$$h_K(z) = \frac{[P_K(R^{-1}z)]^2 n_M(z|0, RR')}{\int [P_K(u)]^2 n_M(u|0, I) du}.$$

The denominator is an easily computed weighted sum of products of the moments of the standard normal distribution [16, p. 47].

Let  $\theta_{(1)}$  be a vector whose elements are the coefficients  $\{a_\lambda : 0 \leq |\lambda| \leq K\}$ , let  $\theta_{(2)} = (r_{11}, r_{12}, r_{22}, r_{13}, r_{23}, r_{33}, \dots, r_{MM})$ , let  $\theta = (\theta_{(1)}, \theta_{(2)})$ , and let  $p_\theta$  denote the dimension of the vector  $\theta$ , which is determined solely by the degree  $K$  of  $P_K$ . The vector  $\theta$  completely describes the truncated expansion.

As an example of the truncated expansion, consider  $M = 2$ ,  $K = 2$ . In this case

$$\{\lambda : 0 \leq |\lambda| \leq K\} = \{(0, 0), (1, 0), (0, 1), (2, 0), (0, 2), (1, 1)\},$$

and

$$P_2(u) = a_{00} + a_{10}u_1 + a_{01}u_2 + a_{20}u_1^2 + a_{02}u_2^2 + a_{11}u_1u_2.$$

The denominator of  $h_2(z)$  is a weighted sum of products of moments to the 4th order of the standard normal distribution, and  $\theta_{(1)} = (a_{00}, a_{10}, a_{01}, a_{20}, a_{02}, a_{11})$ ,  $\theta_{(2)} = (r_{11}, r_{12}, r_{22})$ ,  $\theta = (\theta_{(1)}, \theta_{(2)})$ , so that  $p_\theta = 9$ .

If  $h$  is represented by the truncated expansion  $h_K(z)$ , estimation of  $\tau^\circ$  and  $h^\circ$  becomes a standard problem in nonlinear optimization. One minimizes  $s_n[\tau, h_K(\cdot|\theta)]$  in the variables  $\tau$  and  $\theta$  to get  $\hat{\tau}$  and  $\hat{\theta}$ . The estimate of  $h^\circ$  is then  $\hat{h}_K(\cdot) = h_K(\cdot|\hat{\theta})$ .

Gallant and Nychka [8] show that  $\hat{h}_K$  is a nonparametric estimator by proving that if the degree  $K$  of the polynomial increases with the sample size  $N$  then the estimates  $(\hat{\tau}, \hat{h})$  obtained in this way converge to the true values  $(\tau^\circ, h^\circ)$ . Moments of  $\hat{h}$  converge to their true values as well. This convergence property is called consistency. Confidence intervals can be computed for the elements of  $\tau$  and characteristics of  $h$  using maximum likelihood formulæ [17].

If  $h^\circ$  is assumed to have mean zero, that is,  $\int z h^\circ(z) dz = 0$ , then the constraint may be imposed when computing  $\hat{h}$  without altering the consistency result. When the constraint is imposed, the estimators with  $K = 0$  and  $K = 1$  are the same so that  $K = 2$  is next in the progression. For  $K > 0$ , the off-diagonal elements of  $R$  can be constrained to be zero which attenuates estimated correlations but does not affect the consistency result.

Standard statistical model selection criteria can be used to choose the truncation point  $K$ . Most criteria pick the value of  $K$  that minimizes an expression of the form  $s_n(\hat{\tau}, \hat{h}_K) + c(N)(p_{net}/N)$ , where  $p_{net} = p_r + p_\theta - 1$  if the constraint  $\int z h(z) dz = 0$  is not imposed and  $p_{net} = p_r + p_\theta - M - 1$  if it is. The term  $c(N)(p_{net}/N)$  is a penalty factor designed to compensate for small  $s_n(\hat{\tau}, \hat{h}_K)$  achieved by fitting an over-parameterized model. The standard criteria are: the Schwarz or BIC with  $c(N) = (1/2) \log N$ , the Hannan-Quinn (HQ) with  $c(N) = \log \log N$ , and the Akaike or AIC with  $c(N) = 1$ .

These criteria have been extensively studied when  $(-N)s_n(\hat{\tau}, \hat{h}_K)$  is replaced in the expression above by the optimized log likelihood of a linear regression model [18, 19, 17]. The BIC criterion has the largest penalty that will not underfit in large samples with little restriction on regressors. The HQ criterion has the smallest penalty that will not overfit under more stringent conditions on the regressors. The AIC criterion adds regressors at an appropriate rate when the regression is a series expansion. For the formula that relates this work to the present context see [20, p. 366]. Note that the penalty increases as one goes from BIC to HQ to AIC. Thus,  $K$  may, but need not, increase from BIC to HQ to AIC.

Our recommendation is to inspect plots such as Figure 1 for all models between those chosen by the BIC and AIC criteria inclusively and make a visual selection. We cannot state the case for visual inspection better than Silverman:

A natural method for choosing the smoothing parameter  $[K]$  is to plot out several curves and choose the estimate that is most in accordance with one's prior ideas about the density. For many applications this approach will be perfectly satisfactory. Indeed, the process of examining several plots of the data, all smoothed by different amounts, may well give more insight into the data than merely considering a single automatically produced curve. [21, p. 44]

If one insists upon an automatic selection rule we recommend the HQ criterion because, upon checking several published applications of the SNP method in the econometrics literature, we found that the HQ criterion usually selected the same model that the authors of these articles had chosen after extensive diagnostic testing. The BIC criterion nearly always selected a smaller model than the authors chose, and the AIC model nearly always selected a larger model.

One structural aspect of the truncation estimator  $h_K$  deserves comment. If  $K = 0$  then  $h_K$  is the normal density; that is, the normal density is the leading term in the expansion of  $h^\circ$ . This is an advantage in applications when the normal distribution is a reasonable first approximation and one only expects modest departures from the normal such as an extra mode. Moreover, the fact that the leading term of the series is the normal density provides a convenient means to test the hypothesis that  $h^\circ$  is normal. One can compare the optimized likelihood for  $K > 0$  with that for  $K = 0$  using, say, either the model selection criteria above or the asymptotic  $\chi^2$  test. The asymptotic  $\chi^2$  statistic for a choice between specifications  $K_H < K_A$  having  $p_{net} = p_H$  and  $p_A$ , respectively, is  $2N[s_n(\hat{\tau}_H, \hat{h}_{K_H}) - s_n(\hat{\tau}_A, \hat{h}_{K_A})]$  on  $p_A - p_H$  degrees of freedom. In the econometric literature it has been noted that the asymptotic  $\chi^2$  tends to select unnecessarily large  $K$  and its use has been largely abandoned in favor of the model selection criteria. Thus we prefer the model selection criteria for determining  $K$ .

For given  $K$ , the model selection criteria can also be used determine if the current specification of  $g$  should be augmented by additional covariates. The model selection criteria are evaluated for both the current and augmented models. If the three model selection criteria select the larger model, one has rather persuasive statistical evidence in favor of the augmentation.

Imposing the constraint  $\int z \hat{h}(z) dz = 0$  usually has little effect on estimates and can be convenient when reporting results. Sometimes, however, the constraint increases the value of  $K$  required to obtain an adequate fit. We recommend not imposing it unless it leaves the estimates of  $\tau$ , the selected value of  $K$ , and the visual appearance of the fitted density essentially unchanged. When  $K = 1$ ,  $\int z \hat{h}(z) dz = 0$  imposes normality.

Putting the off-diagonal elements of  $R$  to zero improves numerical stability, especially when  $M$  is large. We recommend that it be imposed if estimates of  $\tau$  and the visual appear-

ance of the fitted density are little changed.

A Fortran program implementing the SNP method is in the public domain. It is called `nlmix` and is available together with a User's Guide as a PostScript file either via ftp anonymous at `ccvr1.cc.ncsu.edu` (128.109.212.20) in directory `pub/arg/nlmix` or from the Carnegie-Mellon University e-mail server by sending the one-line e-mail message "send `nlmix` from general" to `statlib@lib.stat.cmu.edu`. `Nlmix` computes parameter estimates, empirical Bayes estimates of the random effects, data for plotting, and simulations from the estimated density. Its use is illustrated in the next section.

## PHARMACOKINETICS OF QUINIDINE

### Kinetic Model

In their review article, Ochs, Greenblatt, and Woo [22] summarize the literature, mostly experimental studies, on the pharmacokinetics of quinidine:

Typical ranges for kinetic properties of quinidine in healthy persons weighing 80 kg are: apparent volume of distribution,  $V$ , 160 L to 280 L; elimination rate constant,  $k_e$ ,  $0.06 \text{ hr}^{-1}$  to  $0.14 \text{ hr}^{-1}$ ; and clearance,  $Cl$ , 12 L/hr to 24 L/hr. Quinidine clearance is reduced in the elderly, in patients with cirrhosis, and in those with congestive heart failure. Oral quinidine is available either as relatively rapidly absorbed conventional tablets (usually quinidine sulphate) or as a variety of slowly absorbed sustained release preparations. The fraction available,  $F$ , is generally 0.7 or greater. Values of the first order absorption rate constant,  $k_a$ , range from  $0.63 \text{ hr}^{-1}$  to  $2.97 \text{ hr}^{-1}$ . Evidence of a dependence of  $F$  or  $k_a$  on dosage form is conflicting. Quinidine is 70 to 90% bound to plasma protein, primarily to albumin but also to a number of other plasma constituents such as  $\alpha_1$ -acid glycoprotein. Binding is reduced in patients with cirrhosis, partly because of hypoalbuminaemia, but is not influenced by renal insufficiency.

Recently, Fattinger et al. [23] obtained estimates within the ranges above from routine clinical data on 60 patients treated with quinidine for arrhythmias. They found that weight,

age, and mild or moderate heart failure had no effect on clearance but that renal insufficiency, severe heart failure, and severe liver failure reduced clearance.

In the literature, a one-compartment open model with first-order absorption and a two-compartment open model with zero order absorption are the two most common characterizations of quinidine disposition [22, 24, 23]. Because the data used here have been analyzed using several statistical methodologies at an American Statistical Association Invited Paper Session [25] using a one-compartment open model with first-order absorption, that model is used here to permit comparison:

For the non-steady state at a dosage time,  $t = t_\ell$

$$\begin{aligned} X_a(t_\ell) &= X_a(t_{\ell-1})e^{-k_a(t_\ell-t_{\ell-1})} + FD_\ell/V \\ X(t_\ell) &= X(t_{\ell-1})e^{-k_e(t_\ell-t_{\ell-1})} + X_a(t_{\ell-1})\frac{k_a}{k_a - k_e} \left[ e^{-k_e(t_\ell-t_{\ell-1})} - e^{-k_a(t_\ell-t_{\ell-1})} \right] \end{aligned}$$

For the steady state at a dosage time,  $t = t_\ell$

$$\begin{aligned} X_a(t_\ell) &= (FD_\ell/V)(1 - e^{-k_a\Delta})^{-1} \\ X(t_\ell) &= (FD_\ell/V)\frac{k_a}{k_a - k_e} \left[ (1 - e^{-k_e\Delta})^{-1} - (1 - e^{-k_a\Delta})^{-1} \right] \end{aligned}$$

Between dosage times,  $t_\ell < t < t_{\ell+1}$

$$X(t) = X(t_\ell)e^{-k_e(t-t_\ell)} + X_a(t_\ell)\frac{k_a}{k_a - k_e} \left[ e^{-k_e(t-t_\ell)} - e^{-k_a(t-t_\ell)} \right]$$

where  $t_\ell, \ell = 0, 1, \dots$  are the times at which doses  $D_\ell$  are administered,  $X(t)$  is the concentration of quinidine at time  $t$ ,  $X_a(t_\ell)$  is the concentration of quinidine in the absorption depot at time  $t_\ell$ ,  $X_a(t_0) = FD_0/V$ ,  $X(t_0) = 0$ ,  $F$  is the fraction of dose available,  $k_a$  is the absorption rate constant,  $k_e = Cl/V$  is the elimination rate constant,  $Cl$  is the clearance,  $V$  is the apparent volume of distribution, and  $\Delta$  is the steady state dosing interval.

## Data

The data, from [26], consist of quinidine concentration (mg/L) measurements for 136 hospitalized men treated for either atrial fibrillation or ventricular arrhythmias with oral quinidine therapy. A total of 361 quinidine measurements ranging from 1 to 11 observations per patient were obtained by enzyme immunoassay during the course of routine clinical treatment.

Concentration measurements ranged from 0.4 mg/L to 9.4 mg/L with mean 2.45 mg/L and standard deviation 1.22 mg/L. These measurements were taken within a range of 0.08 hr to 70.5 hr after dose, with mean 6.10 hr after dose and standard deviation 5.78 hr.

Time periods over which patients were observed ranged from 0.13 hr to 8095.0 hr. The drug was administered as quinidine sulphate to 53 patients, as quinidine gluconate to 57 patients, and in both forms to 26 patients. Doses were adjusted for differences in salt content between the two forms by conversion of both forms to mg of quinidine base; doses ranged from 83 mg to 603 mg. Under steady state conditions, the mean dosing interval was 6.25 hr for the sulfate form and 7.70 hr for the gluconate form. Initial body weights ranged from 41 kg to 119 kg with mean 79.58 kg and standard deviation 15.64 kg; initial ages ranged from 42 years to 92 years with mean 66.88 years and standard deviation 8.92 years; and heights ranged from 60 in to 79 in with mean 69.63 in and standard deviation 3.37 in. There were 91 Caucasians, 10 Blacks, and 35 Latins; 91 non-smokers and 45 smokers; 90 non- or social drinkers, 16 ethanol abusers, and 30 ex-abusers; and 40 with severe, 40 with moderate, and 56 with no or mild congestive heart failure. There were 84 patients with measured creatinine clearance greater than 50 ml/min throughout the observation period, 41 with creatinine clearance less than 50 ml/min, and 11 whose creatinine clearance varied about 50 ml/min. Albumin concentration (g/dl) measurements were available for some but not all patients so that this attribute cannot be incorporated into the inter-individual regression model but its potential importance can be assessed graphically as seen below.  $\alpha_1$ -acid glycoprotein concentration measurements (mg/dl) were taken periodically on all patients and varied considerably within each patient; these ranged from 39 mg/dl to 316 mg/dl overall; the initial measurements on each patient had mean 118.54 mg/dl and standard deviation 46.23 mg/dl.

## Analysis

The basic set of three pharmacokinetic parameters that we considered were the clearance  $Cl$ , the apparent volume of distribution  $V$ , and the absorption rate constant  $k_a$ . These three constitute the vector  $\beta_{i,j}$ . The  $j$  index indicates that some covariates that appear in the final specification of the inter-individual regression function change over the period of observation.

There were an insufficient number of samples collected during the absorption phase to



allow precise estimation of the  $k_a$  value. Common practice in this situation is to treat  $k_a$  as a fixed parameter that has a prespecified value based on previous results. A value of  $0.9 \text{ hr}^{-1}$  is reasonable with respect to the range of values  $0.63 \text{ hr}^{-1}$  to  $2.97 \text{ hr}^{-1}$  above. The alternatives to this approach are to treat  $k_a$  as either a fixed parameter to be estimated or as a random parameter. For completeness, we report results for all three cases:

Case 1.  $k_a$  fixed and prespecified at  $0.9 \text{ hr}^{-1}$ .

Case 2.  $k_a$  fixed and estimated.

Case 3.  $k_a$  random.

In all cases, the constraint  $\int zh(z)dz = 0$  was imposed so the progression of SNP specifications is  $K=0,2,3,\dots$

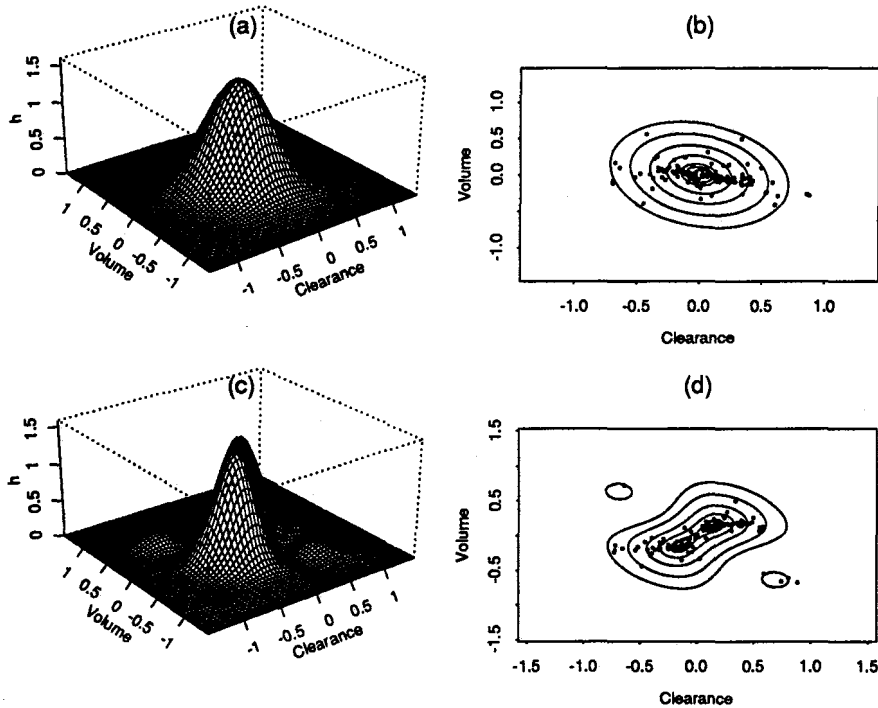
These data do not permit estimation of the fraction of dose available,  $F$ ; therefore we specified  $F = 1$ . As the lowest previous estimate of  $F$  reported by [22] was 0.7, the largest possible upward bias in our estimates of  $V$  and  $Cl$  is 40%.

The intra-individual errors associated with measured plasma concentration were taken as normally distributed with constant coefficient of variation.

The SNP methodology is suited to a graphical model building strategy which we shall illustrate here. The procedure relies on the tendency, noted earlier, for omission of influential factors from the inter-individual regression function  $g(w, \gamma, z)$  to induce multi-modality on nonparametric estimates of the inter-individual random effects density  $h(z)$ . Empirical Bayes estimates of the inter-individual random effects computed from a multi-modal estimate will be separated, a well known phenomenon in the nonparametric literature. The separation will be related to the omitted factors. This relationship can then be detected graphically.

We recommend fitting models with no covariates and increasing  $K$  until empirical Bayes estimates of the random effects separate. Next, plot these estimates against each potential covariate and look for a relationship.

Because the graphics for Case 3 ( $k_a$  random,  $M = 3$ ) take more space to report than those for Cases 1 and 2 ( $k_a$  fixed,  $M = 2$ ) and because we arrived at the same specification for the pharmacokinetic parameters in terms of the covariates in all three cases, we only illustrate the model building strategy for Case 2. In all three cases, however, we report parameter estimates.



**Fig. 1.** Estimated inter-individual random effects density and corresponding empirical Bayes estimates, Case 2, no covariates. Panel (a) is the perspective plot of the estimated joint normal ( $K = 0$ ) inter-individual random effects density and (b) is the contour plot at quantiles 10%, 25%, 50%, 75%, 90%, and 95% together with the empirical Bayes estimates of the random effects (dots). Panel (c) is the perspective plot of the estimated joint SNP ( $K = 2$ ) inter-individual random effects density and (d) is the contour plot at quantiles 10%, 25%, 50%, 75%, 90%, and 95% together with the empirical Bayes estimates of the random effects (dots).

We begin the Case 2 analysis by fitting the SNP models  $K=0,2,3$  using a log-linear inter-individual regression function [27] without covariates:

$$Cl = \exp(\gamma_1 + z_1)$$

$$V = \exp(\gamma_2 + z_2)$$

$$k_a = \exp(\gamma_3).$$

With these data it is essential to enforce positivity of the pharmacokinetic parameters during nonlinear optimization. In other data sets we have found that this complication does not arise and enforcement of positivity during computations is not as important. One would, however, be unlikely to tolerate a specification whose converged estimates were negative for any admissible setting of the covariates or inter-individual random effects, which is a tedious condition to verify. The log-linear specification enforces positivity during computations and eliminates the need to check for global positivity.

Figure 1 plots the estimated densities for  $K = 0$  and  $K = 2$  together with the corresponding empirical Bayes estimates  $\hat{z}_i$  for each specification. The tendency for omitted, influential

factors to induce multi-modal estimates is seen in the  $K = 2$  plots as is the separation of empirical Bayes estimates of the inter-individual random effects.

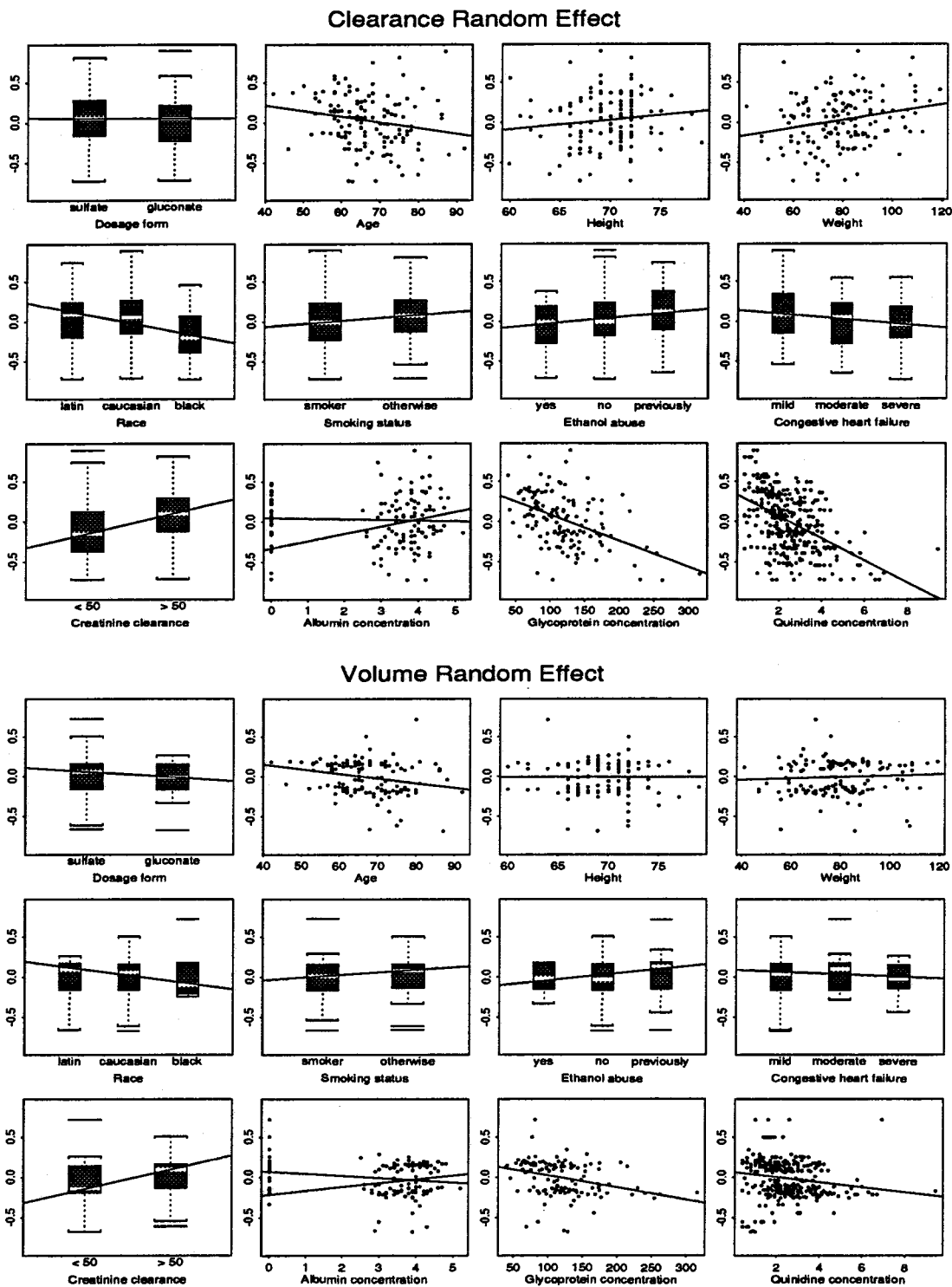
The separation is due to omitted covariates in the inter-individual regression. Thus, for each pharmacokinetic parameter, a univariate regression of each individual's estimated random effect on his demographic attribute suggests the importance of that attribute as a covariate in the inter-individual regression function. A regression against measured concentration reveals the importance of all omitted covariates taken as a group. It is easier to interpret a univariate regression when it is overlaid upon a scatter plot of the underlying data. Omitted curvilinearity that attenuates slope, influential outliers, and other characteristics can be immediately assessed by eye. When the data are categorical rather than measured along a continuum, boxplots take the place of scatter plots. Examples and a description of boxplots are in Figure 2.

The shallow slope of the regression of the empirical Bayes estimate of the  $V$  random effect on measured concentration in the lower panel of Figure 2 suggests that there are no omitted covariates in the  $V$  equation. The steep slope for the  $Cl$  effect in the upper panel does suggest omitted covariates in the  $Cl$  equation.

To determine these omitted covariates, we inspect the remaining plots in the upper panel of Figure 2. Slopes for the regressions for categorical variables are not comparable to the slopes for continuous variables; comparisons must be made within variable types. On this basis, creatinine clearance is the most important categorical variable. Next in importance is race. But the race relationship is due entirely to Blacks, as seen from the graph, of which there are only ten in the sample; thus, we do not include race as a covariate. No other categorical variables seem as important.  $\alpha_1$ -acid glycoprotein concentration is the most important continuous variable. It is hard to judge between age, height, and weight as next in importance. These are correlated variables in these data, any one of them could be expected to proxy for the other two; we selected weight. No other continuous variables seem as important.

The graphical analysis implies the inter-individual regression model

$$Cl = \exp[\gamma_1 + \gamma_4(\text{weight}) + \gamma_5(\text{glyco.}) + \gamma_6(\text{creat.}) + z_1]$$



**Fig. 2.** Inter-individual regression graphics, Case 2, no covariates. Empirical Bayes estimates of inter-individual random effects for clearance (upper panel) and volume (lower panel) plotted against initial values of candidate covariates and quinidine concentration. Plots against continuous variables such as age are scatter plots with least squares lines superimposed; the extra line for albumin concentration has the unmeasured (zero) values excluded. Plots against categorical variables such as dosage form are boxplots: A horizontal line is drawn in the box at the median, the upper and lower ends of the box are at the upper and lower quartiles, vertical lines go up and down from the median to 1.5 times the interquartile range, extreme points are plotted by themselves, and least squares lines through the medians are superimposed.

**Table I.** Optimization Results, Case 2, Covariates: Weight,  $\alpha_1$ -acid Glycoprotein Concentration, Creatinine Clearance

Specification	$K^a$	$p_{net}^b$	$s_n(\hat{\tau}, \hat{h}_K)^c$	BIC <sup>d</sup>	HQ <sup>e</sup>	AIC <sup>f</sup>
Without covariates	0	7	1.24547	1.30256	1.27985	1.25486
With covariates	0	10	1.10133	1.18289	1.15045	1.12903
Without covariates	2	10	1.23415	1.31571	1.28327	1.26185
With covariates	2	13	1.08493	1.19096	1.14878	1.12094
Without covariates	3	14	1.22912	1.34331	1.29788	1.26790
With covariates	3	17	1.07520	1.21386	1.15870	1.12229
With covariates	0	10	1.10133	1.18289	1.15045	1.12903
With covariates	2	13	1.08493	1.19096	1.14878	1.12094
With covariates	3	17	1.07520	1.21386	1.15870	1.12229

<sup>a</sup> $K$  is the degree of the polynomial part of  $h_K$ .

<sup>b</sup> $p_{net}$  is the effective number of parameters.

<sup>c</sup> $s_n(\hat{\tau}, \hat{h}_K)$  is the negative of the optimized log-likelihood divided by the total number,  $N = 361$ , of measured concentrations.

<sup>d</sup>BIC is the Schwarz model selection criterion.

<sup>e</sup>HQ is the Hannan-Quinn model selection criterion.

<sup>f</sup>AIC is the Akaike model selection criterion.

$$V = \exp(\gamma_2 + z_2)$$

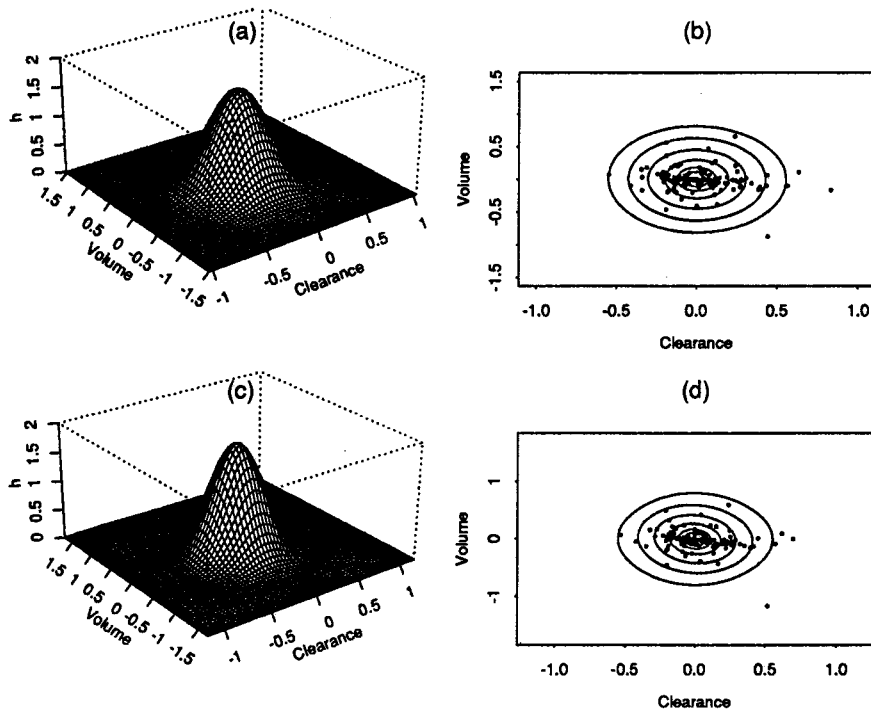
$$k_a = \exp(\gamma_3).$$

For this model, we repeated the computations and graphics, reporting them as Table I, Figure 3, and Figure 4.

In Table I, the inclusion of weight, creatinine clearance, and  $\alpha_1$ -acid glycoprotein concentration in the  $Cl$  equation is strongly supported by all criteria in all specifications,  $K = 0, 2, 3$ . For the models with covariates, the conservative BIC criterion selects the normal ( $K = 0$ ), and HQ and AIC criteria select the  $K = 2$  specification.

As seen from Figure 3, the covariates have removed the multimodality in the SNP ( $K = 2$ ) estimate which suggests that there may no longer be omitted covariates.

The lower panel of Figure 4 indicates that there are no omitted covariates in the  $V$  equation; the slope of the regression of the  $V$  effect against measured concentration is flat



**Fig. 3.** Estimated inter-individual random effects density and corresponding empirical Bayes estimates, Case 2, covariates: weight,  $\alpha_1$ -acid glycoprotein concentration, creatinine clearance. Panel (a) is the perspective plot of the estimated joint normal ( $K = 0$ ) inter-individual random effects density and (b) is the contour plot at quantiles 10%, 25%, 50%, 75%, 90%, and 95% together with the empirical Bayes estimates of the random effects (dots). Panel (c) is the perspective plot of the estimated joint SNP ( $K = 2$ ) inter-individual random effects density and (d) is the contour plot at quantiles 10%, 25%, 50%, 75%, 90%, and 95% together with the empirical Bayes estimates of the random effects (dots).

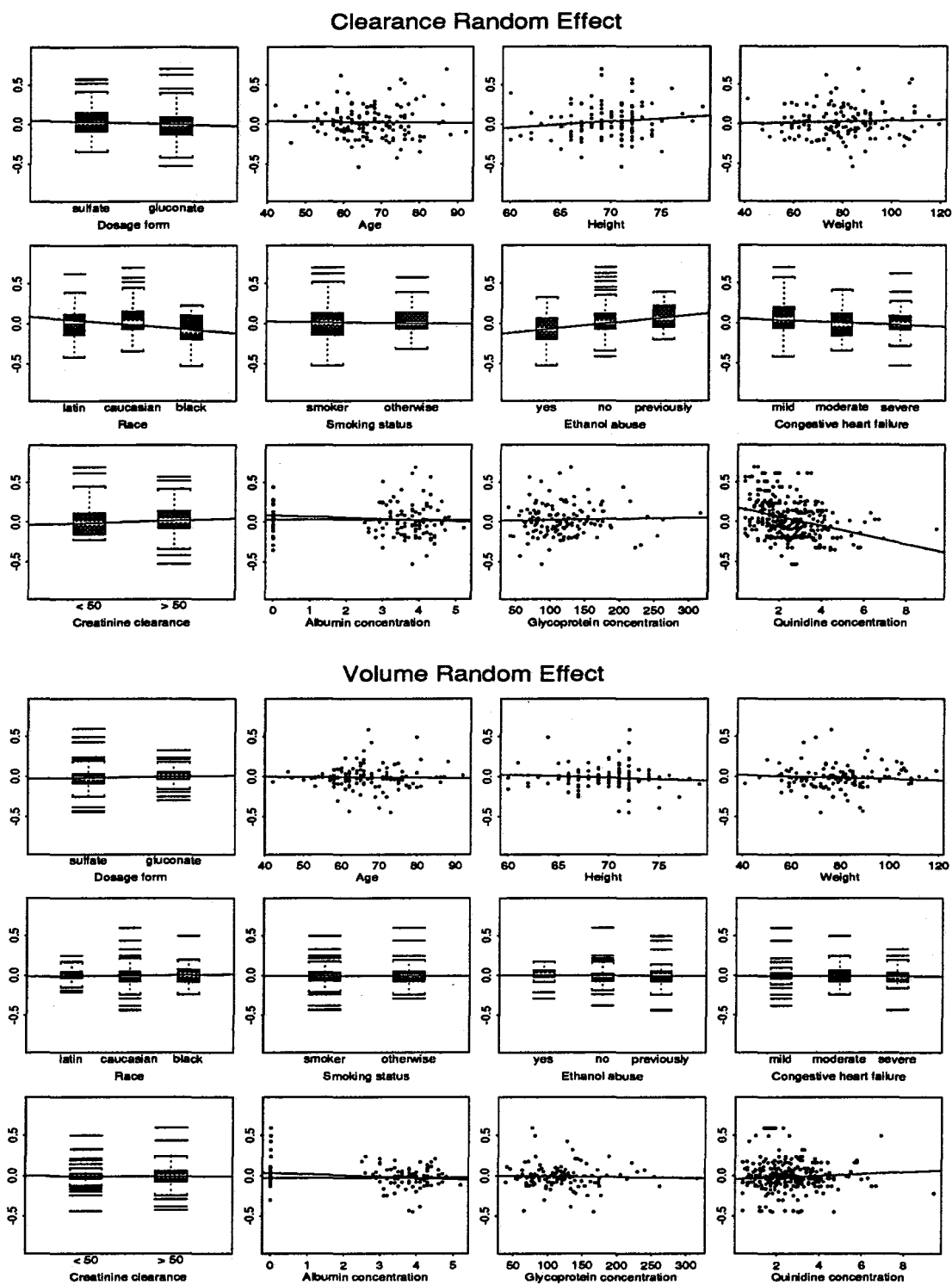
as are those for all variables.

In the upper panel, the slope of the regression of the  $Cl$  effect against quinidine concentration has been considerably attenuated from Figure 2. It is interesting to note that the plots for weight,  $\alpha_1$ -acid glycoprotein concentration, and creatinine clearance are flat. This attenuation and the shallow slopes of the regressions against the other variables suggests that there are no omitted covariates.

To confirm this impression, we computed the BIC, HQ, and AIC criteria for inclusion of each of dosage form, age, height, race, ethanol abuse, and congestive heart failure as an incremental variable in the equation

$$Cl = \exp[\gamma_1 + \gamma_4(\text{weight}) + \gamma_5(\text{glyco.}) + \gamma_6(\text{creat.}) + z_1]$$

for the  $K = 0$  and  $K = 2$  specifications. The BIC and HQ criteria do not support inclusion of these covariates for either the  $K = 0$  or  $K = 2$  specifications. The AIC criterion supports inclusion of congestive heart failure, ethanol abuse, race, and height for the  $K = 0$



**Fig. 4.** Inter-individual regression graphics, Case 2, covariates: weight, 1-acid glycoprotein concentration, creatinine clearance. Empirical Bayes estimates of inter-individual random effects for clearance (upper panel) and volume (lower panel) plotted against initial values of candidate covariates and quinidine concentration. Plots against continuous variables such as age are scatter plots with least squares lines superimposed; the extra line for albumin concentration has unmeasured (zero) values excluded. Plots against categorical variables such as dosage form are boxplots: A horizontal line is drawn in the box at the median, the upper and lower ends of the box are at the upper and lower quartiles, vertical lines go up and down from the median to 1.5 times the interquartile range, extreme points are plotted by themselves, and least squares lines through the medians are superimposed.

specification but does not support inclusion of any covariates when  $K = 2$ .

Because [23] reported lowered clearance for serious liver failure or serious congestive heart failure, we converted the tri-variate ethanol abuse and congestive heart failure variables to bi-variate categories by pooling non- or social drinkers with the ex-abusers and no or mild congestive heart failure patients with those experiencing moderate heart failure and repeated the above analysis, including both of the new variables simultaneously. For the  $K = 0$  specification both the HQ and AIC criteria support inclusion of these covariates. For  $K = 2$ , only the AIC supports inclusion.

The graphical analysis in Cases 1 and 3 leads to the same log-linear specification of the inter-individual regression equation for  $Cl$  and  $V$  as for Case 2. In Case 3, the equation  $k_a$  is log-linear without covariates,  $k_a = \exp(\gamma_3 + z_3)$ .

## Results

We found that no measured individual attributes affect the absorption rate  $k_a$  or the apparent volume of distribution  $V$ . We found that clearance  $Cl$  is positively related to weight, decreased in patients with impaired renal function (creatinine clearance  $< 50$ ), and decreased by increased levels of  $\alpha_1$ -acid glycoprotein concentration. We found weak statistical evidence that clearance may be reduced in patients with severe congestive heart failure or ethanol abuse. These results agree with previous results as summarized above.

Parameter estimates for Cases 1, 2, and 3 and specifications  $K = 0$  and  $K = 2$  are shown in Table II for the inter-individual regression

$$Cl = \exp[\gamma_1 + \gamma_4(\text{weight}) + \gamma_5(\text{glyco.}) + \gamma_6(\text{creat.})] \exp(z_1)$$

$$V = \exp(\gamma_2) \exp(z_2)$$

$$k_a = \exp(\gamma_3) \exp(z_3)$$

where  $z_3$  is interpreted as having zero variance in Cases 1 and 2. Estimates of the fixed effects  $\gamma$  and the intra-individual coefficient of variation  $\sigma$  are remarkably stable across cases and specifications as are the estimated variances of the random effects, with the exception of  $\text{var}(z_3)$ . Correlation estimates are small and unstable; in Case 3 with  $K = 0$  it was necessary to constrain them to zero to achieve numerical stability in the optimizations.



Table II. Parameter Estimates

Parameter <sup>a</sup>	Case 1		Case 2		Case 3	
	Estimate	Std.Dev. <sup>b</sup>	Estimate	Std.Dev.	Estimate	Std.Dev.
Normal Effects <sup>c</sup>						
$\gamma_1$ ln <i>Cl</i>	2.483840	0.161038	2.488578	0.160400	2.558929	0.162336
$\gamma_2$ ln <i>V</i>	5.424166	0.064146	5.360088	0.084944	5.242221	0.134361
$\gamma_3$ ln $k_a$	-0.105361	constr.	-0.889899	0.293296	-1.292581	0.365426
$\gamma_4$ Weight	0.005010	0.001697	0.005878	0.001636	0.005550	0.001711
$\gamma_5$ Glyco.	-0.004205	0.000426	-0.004592	0.000479	-0.004802	0.000458
$\gamma_6$ Creat.	0.182166	0.037989	0.165075	0.041611	0.154950	0.042073
$\sigma$ CV( <i>y</i> )	0.250014	0.011076	0.242044	0.010580	0.239513	0.010364
var( $z_1$ ) ln <i>Cl</i>	0.060156	0.009506	0.066420	0.010348	0.067296	0.010333
var( $z_2$ ) ln <i>V</i>	0.082108	0.051643	0.142595	0.066383	0.070784	0.089063
var( $z_3$ ) ln $k_a$	0.0	constr.	0.0	constr.	0.630126	0.497739
$\rho(z_1, z_2)$ ln <i>Cl</i> , <i>V</i>	0.137148	—	-0.013377	—	0.0	constr.
$\rho(z_1, z_3)$ ln <i>Cl</i> , $k_a$	0.0	constr.	0.0	constr.	0.0	constr.
$\rho(z_2, z_3)$ ln <i>V</i> , $k_a$	0.0	constr.	0.0	constr.	0.0	constr.
SNP Effects <sup>d</sup>						
$\gamma_1$ ln <i>Cl</i>	2.621325	0.153507	2.647315	0.157191	2.623371	0.151130
$\gamma_2$ ln <i>V</i>	5.429789	0.068826	5.356483	0.083635	5.222166	0.157562
$\gamma_3$ ln $k_a$	-0.105361	constr.	-1.065291	0.300597	-1.153077	0.428943
$\gamma_4$ Weight	0.003769	0.001667	0.004802	0.001671	0.004735	0.001599
$\gamma_5$ Glyco.	-0.004338	0.000410	-0.004947	0.000441	-0.004963	0.000417
$\gamma_6$ Creat.	0.121290	0.045138	0.112014	0.046321	0.169517	0.042232
$\sigma$ CV( <i>y</i> )	0.241634	0.010388	0.234707	0.009639	0.225575	0.009972
var( $z_1$ ) ln <i>Cl</i>	0.086285	0.025013	0.086321	0.025638	0.083067	0.019849
var( $z_2$ ) ln <i>V</i>	0.081184	0.048667	0.181529	0.101930	0.136504	0.107092
var( $z_3$ ) ln $k_a$	0.0	constr.	0.0	constr.	0.910501	0.897072
$\rho(z_1, z_2)$ ln <i>Cl</i> , <i>V</i>	-0.215559	—	-0.381538	—	0.010591	—
$\rho(z_1, z_3)$ ln <i>Cl</i> , $k_a$	0.0	constr.	0.0	constr.	-0.201680	—
$\rho(z_2, z_3)$ ln <i>V</i> , $k_a$	0.0	constr.	0.0	constr.	0.082857	—

<sup>a</sup>Model: The intra-individual pharmacokinetic model, represented as  $y = f(x, \beta) + e$ , is a one-compartment open model with first-order absorption where  $\beta = (Cl, V, k_a)$  are the pharmacokinetic parameters, *Cl* is clearance in L/hr, *V* is the apparent volume of distribution in L, and  $k_a$  is the absorption rate constant in hr<sup>-1</sup>. Quinidine concentration, *y*, is in mg/L, and the experimental variables  $x = (\text{dose}, \text{time})$  are in mg of quinidine base and hr respectively. The errors *e* are normal with mean zero and standard deviation  $\sigma f(x, \beta)$ . The inter-individual regression model for the pharmacokinetic parameters is  $\ln Cl = \gamma_1 + \gamma_4(\text{weight}) + \gamma_5(\text{glyco.}) + \gamma_6(\text{creat.}) + z_1$ ,  $\ln V = \gamma_2 + z_2$ , and  $\ln k_a = \gamma_3 + z_3$  where the *z*'s are the random effects. Weight is in kg,  $\alpha_1$ -acid glycoprotein concentration is in mg/dl, and creatinine clearance is in ml/min. Var( $\cdot$ ) denotes the variance of a random effect and  $\rho(\cdot, \cdot)$  a correlation. Cases differ with respect to which parameters are constrained.

<sup>b</sup>constr. means that the estimate is constrained to have the value shown.

<sup>c</sup>Normal Effects: maximum likelihood estimates with normal ( $K = 0$ ) random effects density.

<sup>d</sup>SNP Effects: maximum likelihood estimates with SNP ( $K = 2$ ) random effects density.

Table III. Population Characteristics of the Random Effects

		Case 1	Case 2	Case 3
Characteristic <sup>a</sup>		Estimate	Estimate	Estimate
Normal Effects <sup>b</sup>				
CV(exp $z_1$ )	$Cl$	0.248875	0.263182	0.269232
CV(exp $z_2$ )	$V$	0.299589	0.401253	0.277401
CV(exp $z_3$ )	$k_a$	0.0	0.0	0.931264
$\mathcal{E}$ (exp $z_1$ )	$Cl$	1.026680	1.030480	1.039373
$\mathcal{E}$ (exp $z_2$ )	$V$	1.039255	1.071143	1.037394
$\mathcal{E}$ (exp $z_3$ )	$k_a$	1.0	1.0	1.361435
var(exp $z_1$ )	$Cl$	0.065288	0.073551	0.078306
var(exp $z_2$ )	$V$	0.096939	0.184727	0.082814
var(exp $z_3$ )	$k_a$	0.0	0.0	1.607458
$\rho$ (exp $z_1$ , exp $z_2$ )	$Cl, V$	0.111926	-0.036051	0.0
$\rho$ (exp $z_1$ , exp $z_3$ )	$Cl, k_a$	0.0	0.0	0.0
$\rho$ (exp $z_2$ , exp $z_3$ )	$V, k_a$	0.0	0.0	0.0
SNP Effects <sup>c</sup>				
CV(exp $z_1$ )	$Cl$	0.321833	0.314340	0.327505
CV(exp $z_2$ )	$V$	0.288746	0.537890	0.387419
CV(exp $z_3$ )	$k_a$	0.0	0.0	1.684499
$\mathcal{E}$ (exp $z_1$ )	$Cl$	1.042353	1.040531	1.050885
$\mathcal{E}$ (exp $z_2$ )	$V$	1.035648	1.105573	1.075237
$\mathcal{E}$ (exp $z_3$ )	$k_a$	1.0	1.0	1.647813
var(exp $z_1$ )	$Cl$	0.112536	0.106982	0.118453
var(exp $z_2$ )	$V$	0.089424	0.353640	0.173528
var(exp $z_3$ )	$k_a$	0.0	0.0	7.704733
$\rho$ (exp $z_1$ , exp $z_2$ )	$Cl, V$	-0.193330	-0.319158	0.026915
$\rho$ (exp $z_1$ , exp $z_3$ )	$Cl, k_a$	0.0	0.0	-0.079568
$\rho$ (exp $z_2$ , exp $z_3$ )	$V, k_a$	0.0	0.0	0.026544

<sup>a</sup>Model: The inter-individual regression model for the pharmacokinetic parameters is  $Cl = \exp[\gamma_1 + \gamma_4(\text{weight}) + \gamma_5(\text{glyco.}) + \gamma_6(\text{creat.}) + z_1]$ ,  $V = \exp(\gamma_2 + z_2)$ , and  $k_a = \exp(\gamma_3 + z_3)$  where  $Cl$  is clearance in L/hr,  $V$  is the apparent volume of distribution in L,  $k_a$  is the absorption rate constant in  $\text{hr}^{-1}$ , and the  $z$ 's are the random effects. Weight is in kg,  $\alpha_1$ -acid glycoprotein concentration is in mg/dl, and creatinine clearance is in ml/min.  $CV(\cdot)$  denotes the coefficient of variation of a random effect,  $\mathcal{E}(\cdot)$  denotes the mean,  $\text{var}(\cdot)$  the variance, and  $\rho(\cdot, \cdot)$  a correlation. Cases differ with respect to which parameter estimates are constrained, see Table II. The moments are computed by Monte Carlo integration using 2000 repetitions.

<sup>b</sup>Normal Effects: maximum likelihood estimates with normal ( $K = 0$ ) random effects density.

<sup>c</sup>SNP Effects: maximum likelihood estimates with SNP ( $K = 2$ ) random effects density.

The population characteristics of  $[\exp(z_1), \exp(z_2), \exp(z_2)]$  are required in order to assess the population characteristics of the pharmacokinetic parameters as is seen from the inter-individual regression equations for  $\beta = (Cl, V, k_a)$  above. Since it is easy to simulate from the SNP density [12], as noted above, an estimate of a population characteristic such as a mean  $\mathcal{E} \exp(z_1) = \int \exp(z_1) h^\circ(z) dz$  can be computed by Monte Carlo integration by drawing a sample  $\{\tilde{z}_i\}_{i=1}^I$  from the estimated SNP density  $\hat{h}_K$  and averaging  $\hat{\mathcal{E}} \exp(z_1) \doteq (1/I) \sum_{i=1}^I \exp(\tilde{z}_{1i})$ . Similarly, a percentile would be computed as the corresponding percentile of  $\{\exp(\tilde{z}_{1i})\}_{i=1}^I$ . These computations can be made as accurate as desired by taking the number of Monte Carlo repetitions  $I$  large enough. Using this procedure, the coefficients of variation, the means, the variances, and correlations of  $[\exp(z_1), \exp(z_2), \exp(z_2)]$  were computed and are shown in Table III for Cases 1, 2, and 3 and specifications  $K = 0$  and 2.

The most notable feature of Table III is the high coefficient of variation of  $\exp(z_3)$ , which corresponds to absorption rate  $k_a$ , and its interaction with the coefficient of variation of  $\exp(z_2)$ , which corresponds to volume  $V$ . Treating  $k_a$  as a fixed parameter to be estimated, Case 2, increases the coefficient of variation of volume over both the case when  $k_a$  is fixed and specified, Case 1, and when treated as a random effect, Case 3. This is apparently due to an insufficient number of samples collected during the absorption phase. It is also interesting to note the stability of population characteristics of  $\exp(z_1)$ , which corresponds to quinidine clearance  $Cl$ , across Cases 1, 2, and 3.

Because the inter-individual regression model is multiplicative, estimates of the population characteristics of the pharmacokinetic parameters are obtained by rescaling the random effect estimates. For instance, at a given setting of weight,  $\alpha_1$ -acid glycoprotein concentration, and creatinine clearance

$$\hat{\mathcal{E}}Cl = \exp[\hat{\gamma}_1 + \hat{\gamma}_4(\text{weight}) + \hat{\gamma}_5(\text{glyco.}) + \hat{\gamma}_6(\text{creat.})] \hat{\mathcal{E}} \exp(z_1)$$

$$\widehat{SD}(Cl) = \exp[\hat{\gamma}_1 + \hat{\gamma}_4(\text{weight}) + \hat{\gamma}_5(\text{glyco.}) + \hat{\gamma}_6(\text{creat.})] \sqrt{\widehat{\text{var}}(\exp z_1)}$$

where  $\hat{\mathcal{E}} \exp(z_1)$  and  $\widehat{\text{var}}(\exp z_1)$  are the entries from Table III. Table IV displays estimates of population characteristics of the pharmacokinetic parameters  $\beta = (Cl, V, k_a)$  at some representative settings of the covariates weight,  $\alpha_1$ -acid glycoprotein concentration, and creatinine

Table IV. Population Characteristics of the Pharmacokinetic Parameters

Characteristic <sup>a</sup>		Case 1	Case 2	Case 3
		Estimate	Estimate	Estimate
Normal Effects <sup>b</sup>				
$\mathcal{E}Cl$ :	low weight, low glyco., creat.<50	12.374220	12.788633	13.378630
	low weight, low glyco., creat.>50	14.846755	15.083957	15.620883
	low weight, median glyco., creat.<50	10.262462	10.425062	10.804575
	low weight, median glyco., creat.>50	12.313039	12.296169	12.615417
	low weight, high glyco., creat.<50	7.957304	7.896319	8.080459
	low weight, high glyco., creat.>50	9.547280	9.313563	9.434740
	median weight, low glyco., creat.<50	13.678364	14.384003	14.949213
	median weight, low glyco., creat.>50	16.411483	16.965667	17.454695
	median weight, median glyco., creat.<50	11.344043	11.725579	12.072977
	median weight, median glyco., creat.>50	13.610735	13.830105	14.096403
	median weight, high glyco., creat.<50	8.795940	8.881378	9.029063
	median weight, high glyco., creat.>50	10.553487	10.475422	10.542331
	high weight, low glyco., creat.<50	15.426014	16.563289	17.079154
	high weight, low glyco., creat.>50	18.508337	19.536094	19.941613
	high weight, median glyco., creat.<50	12.793443	13.502093	13.793116
	high weight, median glyco., creat.>50	15.349744	15.925470	16.104837
high weight, high glyco., creat.<50	9.919776	10.226974	10.315511	
high weight, high glyco., creat.>50	11.901879	12.062527	12.044386	
$\mathcal{E}V$	235.725999	227.878890	196.160421	
$\mathcal{E}k_a$	0.9	0.410697	0.373797	
SNP Effects <sup>c</sup>				
$\mathcal{E}Cl$ :	low weight, low glyco., creat.<50	13.274630	13.860196	13.597986
	low weight, low glyco., creat.>50	14.986423	15.503025	16.109975
	low weight, median glyco., creat.<50	10.944244	11.121494	10.903329
	low weight, median glyco., creat.>50	12.355529	12.439709	12.917528
	low weight, high glyco., creat.<50	8.417936	8.244829	8.075273
	low weight, high glyco., creat.>50	9.503448	9.222076	9.567038
	median weight, low glyco., creat.<50	14.313951	15.257347	14.948660
	median weight, low glyco., creat.>50	16.159767	17.065778	17.710163
	median weight, median glyco., creat.<50	11.801111	12.242575	11.986346
	median weight, median glyco., creat.>50	13.322891	13.693670	14.200613
	median weight, high glyco., creat.<50	9.077008	9.075933	8.877382
	median weight, high glyco., creat.>50	10.247509	10.151690	10.517323
	high weight, low glyco., creat.<50	15.669101	17.121058	16.747712
	high weight, low glyco., creat.>50	17.689667	19.150392	19.841559
	high weight, median glyco., creat.<50	12.918361	13.738027	13.428887
	high weight, median glyco., creat.>50	14.584213	15.366375	15.909639
high weight, high glyco., creat.<50	9.936359	10.184573	9.945763	
high weight, high glyco., creat.>50	11.217675	11.391736	11.783069	
$\mathcal{E}V$	236.232459	234.357272	199.279249	
$\mathcal{E}k_a$	0.9	0.344628	0.520155	

<sup>a</sup>Weight: low, 10th percentile, 59.0 kg; median, 79.0 kg; high, 90th percentile, 103.0 kg.  $\alpha_1$ -acid glycoprotein concentration: low, 10th percentile, 69.0 mg/dl; median, 113.5 mg/dl; high, 90th percentile, 174.0 mg/dl. Creatinine clearance: creat. < 50 ml/min; creat. > 50 ml/min. Coefficients of variation and correlations for  $Cl$ ,  $V$ , and  $k_a$  are as for  $\exp(z_1)$ ,  $\exp(z_3)$ , and  $\exp(z_3)$ , respectively, in Table III. Cases differ with respect to which parameter estimates are constrained, see Table II.

<sup>b</sup>Normal Effects: maximum likelihood estimates with normal ( $K = 0$ ) random effects density.

<sup>c</sup>SNP Effects: maximum likelihood estimates with SNP ( $K = 2$ ) random effects density.

clearance.

## DISCUSSION

We propose the SNP method, taken from the econometrics literature, for the analysis of population pharmacokinetic data that can be described by a nonlinear mixed effects model. It produces smooth nonparametric estimates of the entire random effects density and simultaneous estimates of fixed effects by maximum likelihood. The benefits of estimating the entire distribution nonparametrically rather than a few leading moments have been discussed in [7]. We also introduce a new graphical model building strategy that exploits an inherent tendency of a nonparametric method to produce disparate empirical Bayes estimates when covariates have been omitted from the model. We illustrate by an analysis of the population pharmacokinetics of quinidine.

Because the SNP method is based on the principle of maximum likelihood and because the SNP density has a convenient representation, subsequent computations essential to a complete statistical analysis are straightforward: estimation of the precision of the estimates of fixed effects, statistical tests of the significance of covariates, tests for normality of the random effects density, empirical Bayes estimation of individual random effects, and estimation of population characteristics of pharmacokinetic parameters including those that are affected nonlinearly by the random effects. More elaborate, computationally intensive statistical analyses are also possible such as setting sup-norm confidence bands on a marginal density of a random effect or pharmacokinetic parameter [10] or placing a global restriction on the density such as unimodality [9].

The NPML [6, 7] method maximizes a likelihood over all distribution functions whereas the SNP method maximizes over a class of distributions restricted to have a smooth density. That is the difference between them. Both the NPML and SNP method require a complete specification of the intra-individual likelihood. The advantages the smoothness assumption confers upon the SNP method, relative to the NPML method, are as follows: Any unknown parameters that are fixed across individuals are estimated simultaneously with the distribution rather than requiring a two-step estimation procedure. Thus, for example, covariates in the inter-individual regression for the pharmacokinetic parameters are easily accommodated.

The SNP density estimate is inherently smooth; a subsequent smoothing computation is not required. Statistical inference is possible: one can test for normality and set confidence intervals on parameters.

If the true random effects distribution violates the smoothness assumption, however, the NPML estimator will be consistent and the SNP estimator will not. Empirical Bayes estimation of the random effects is convenient with both the SNP and NPML methods, and thus both allow use of the graphical model building strategy that we describe.

Mallet *et al.* [7] compare the features of nonparametric methods to other procedures for the analysis of the nonlinear mixed effects model: the popular first-order (FO) approximation with extended least squares (ELS) estimation [1, 28, 3], which provides estimates of fixed effects, allows empirical Bayes estimation based on normality, and is implemented in the powerful NONMEM package [29, 30]; and two-stage (TS) methods [31, 2]. Nonparametric methods are indicated when a knowledge of the entire distribution is required, not just first and second moments, and when there may be reason to doubt a parametric assumption, such as normality, for the purpose of computing empirical Bayes estimates, for example. Nonparametric estimation of the entire distribution may be based on the use of two-stage methods; these methods, however, require a sufficient number of observations on each individual to allow accurate estimation of individual pharmacokinetic parameters. Thus, two-stage methods are usually not applicable to clinical data, which are typically sparse.

In addition, as we have illustrated here, nonparametric methods are inherently well suited to model building. Even when there is no other reason to use nonparametric methods, their unique ability to fragment empirical Bayes estimates, thereby exaggerating their relation to influential individual attributes, can be of value in allowing graphical screening of a large number of attributes for their potential as covariates.

## References

- [1] S. L. Beal, and L. B. Sheiner. Estimating population kinetics. *CRC Crit. Rev. in Biomed. Eng.* 8:195–222 (1982).
- [2] J. L. Steimer, A. Mallet, J. L. Golmard and J. F. Boisvieux. Alternative approaches to

- estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed effect model. *Drug Metab. Rev.* **15**:265–292 (1984).
- [3] S. L. Beal and L. B. Sheiner. Methodology of population pharmacokinetics. In E. R. Garrett & J. L. Hirtz (eds.), *Drug Fate and Metabolism-Methods and Techniques*, Vol 5. Marcel Dekker, New York, 1985, pp. 135–83.
- [4] L. B. Sheiner and T. M. Ludden. Population pharmacokinetics/dynamics. *Annu. Rev. Pharmacol. Toxicol.* **32**:185–209 (1992).
- [5] S. Vozeh, R. Hillman, M. Wandell, T. M. Ludden, and L. B. Sheiner. Computer-assisted drug assay interpretation based on Bayesian estimation of individual pharmacokinetics: application to lidocaine. *Ther. Drug. Monit.* **7**:66–73 (1985).
- [6] A. Mallet. A maximum likelihood estimation method for random coefficient regression models. *Biometrika* **73**:645–656 (1986).
- [7] A. Mallet, F. Mentré, J-L Steimer, and F.Lokiec. Nonparametric maximum likelihood estimation for population pharmacokinetics, with application to cyclosporine. *J. Pharmacokin. Biopharm.* **16**:311–327 (1988).
- [8] A. R. Gallant and D. W. Nychka. Semiparametric maximum likelihood estimation. *Econometrica*, **55**:363–390 (1987).
- [9] A. R. Gallant and G. E. Tauchen. Semiparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications. *Econometrica* **57**:1091–1120 (1989).
- [10] A. R. Gallant, P. E. Rossi, and G. E. Tauchen. Stock prices and volume. *Rev. Financ. Stud.* **5**:in press (1992).
- [11] A. D. Brunner. Conditional asymmetries in real GNP: a semiparametric approach. *J. Bus. Econ. Statist.* **10**:65–72 (1992).

- [12] M. Davidian and A. R. Gallant. The nonlinear mixed effects model with a smooth random effects density. *Institute of Statistics Mimeo Series No. 2206*. North Carolina State University, Raleigh, 1992.
- [13] I. Elbadawi, A. R. Gallant, and G. Souza. An elasticity can be estimated consistently without a priori knowledge of functional form. *Econometrica* **51**:1731–1752 (1983).
- [14] A. R. Gallant. Identification et convergence en regression semi-nonparametrique. *Annals de l'INSEE* **59/60**:239–267 (1985).
- [15] S. L. Beal and L. B. Sheiner. Heteroscedastic nonlinear regression. *Technometrics* **30**:327–338 (1988).
- [16] N. L. Johnson and S Kotz. *Continuous Univariate Distributions*. Wiley, New York, 1970.
- [17] B. J. Eastwood. Asymptotic normality and consistency of semi-nonparametric regression estimators using an upward F test truncation rule. *J. Econometrics* **48**:151–182 (1991).
- [18] B. M. Potscher. Model selection under nonstationarity: autoregressive models and stochastic linear models. *Ann. Statist.* **17**:347–370 (1989).
- [19] E. J. Hannan. Rational transfer function approximation. *Statist. Sci.* **2**:1029–1054 (1987).
- [20] A. R. Gallant. *Nonlinear Statistical Models*. Wiley, New York, 1987.
- [21] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.
- [22] H. R. Ochs, D. J. Greenblatt, and E. Woo. Clinical pharmacokinetics of quinidine. *Clin. Pharmacokin.* **5**:150–168 (1980).
- [23] K. Fattinger, S. Vozech, H. R. Ha, M. Borner, and F. Follath. Population pharmacokinetics of quinidine. *Br. J. Clin Pharmac.* **31**:279–286 (1991).
- [24] W. A. Mahon, M. Mayerson, and T. Inaba. Disposition kinetics of two oral forms of quinidine. *Clin. Pharmacol. Ther.* **19**:566–575 (1976).



- [25] American Statistical Association. *Proceedings of the Biopharmaceutical Section, Annual Meetings, Boston, 1992* American Statistical Association, Alexandria VA, 1992.
- [26] C. N. Verme, T. M. Ludden, W. A. Clementi, and S. C. Harris. Pharmacokinetics of quinidine in male patients: A population analysis. *Clin. Pharmacokin.* **23**:in press (1992).
- [27] M. S. Driscoll, T. M. Ludden, D. T. Casto, and L. C. Littlefield. Evaluation of theophylline pharmacokinetics in a pediatric population using mixed effects models. *J. Pharmacokin. Biopharm.* **17**:141-168 (1989).
- [28] S. L. Beal. Population pharmacokinetic data and parameter estimation based on their first two statistical moments. *Drug Metab. Rev.* **15**:173-193 (1984).
- [29] S. L. Beal and L. B. Sheiner. The NONMEM system. *Am. Statist.* **34**:118-119 (1980).
- [30] A. J. Boeckmann, L. B. Sheiner, and S. L. Beal. *NONMEM User's Guide, Part V, Introductory Guide*. University of California, San Francisco, 1990.
- [31] L. B. Sheiner and S. L. Beal. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: Routine pharmacokinetic data. *J. Pharmacokin. Biopharm.* **9**:59-127 (1983).