Population Pharmacokinetic Modeling: The Importance of Informative Graphics[1]

Ene I. Ette[2,3] and Thomas M. Ludden[2]

Received July 5, 1995; accepted July 25, 1995

Purpose. The usefulness of several modelling methods were examined in the development of a population pharmacokinetics model for cefepime.

Methods. The analysis was done in six steps: (1) exploratory data analysis to examine distributions and correlations among covariates, (2) determination of a basic pharmacokinetic model using the NONMEM program and obtaining Bayesian individual parameter estimates, (3) examination of the distribution of parameter estimates, (4) multiple linear regression (MLR) with case deletion diagnostics, generalized additive modelling (GAM), and tree-based modelling (TBM) for the selection of covariates and revealing structure in the data, (5) final NONMEM modelling to determine the population PK model, and (6) the evaluation of final parameter estimates.

Results. An examination of the distribution of individual clearance (CL) estimates suggested bimodality. Thus, the mixture model feature in NONMEM was used for the separation of subpopulations. MLR and GAM selected creatinine clearance (CRCL) and age, while TBM selected both of these covariates and weight as predictors of CL. The final NONMEM model for CL included only a linear relationship with CRCL. However, two subpopulations were identified that differed in slope and intercept.

Conclusions. The findings suggest that using informative graphical and statistical techniques enhance the understanding of the data structure and lead to an efficient analysis of the data.

KEY WORDS: graphics; MLR; GAM; TBM; diagnostics; jackknife; population pharmacokinetics; subpopulations.

INTRODUCTION

Finding a model that adequately describes a given population pharmacokinetic data can be a complicated and time consuming task. It is not enough to find the covariates that are significantly associated with the pharmacokinetic parameters. The determination of the form of the relationship between covariates and parameters is also of importance.

A well chosen graph or graphical technique is a very powerful tool that can be used to obtain information about the structure of the data. Graphical techniques enable one to explore data thoroughly, to look for patterns and relationships, to confirm or disprove the expected, and to discover new phenomena. Data analysis procedures such as NONMEM (1) are based explicitly on assumptions about the data, and the validity of analyses depends upon the validity of the assumptions. Graphical displays provide powerful diagnostic tools for confirming assumptions, or when the assumptions are not met, for suggesting corrective actions. Without such tools, confirmation of assumptions can only be replaced by hope.

The eye-brain system is the most sophisticated information processor ever known to man, and through graphical displays this system can be put to good use. In this paper we discuss the modelling of population pharmacokinetics data using a combination of graphical displays and statistical techniques. In addition, population parameter estimates were evaluated using the jackknife technique in conjunction with a comparison of the NONMEM population parameter estimates with those obtained using the standard two stage approach.

METHODS

Data

Data were pooled from studies involving healthy subjects, cystic fibrosis patients, renal impaired patients, and patients with liver impairment which were submitted as part of the Bristol-Myers Squibb new drug application (NDA) package for cefepime, a cephalosporin antiinfective agent. 1000 or 2000 mg single or multiple doses of the drug were administered as either a 5 or 30 min intravenous (IV) infusion. A validated high performance liquid chromatographic assay with intra-assay and inter-day coefficient of variation of less than 9% was used to determine cefepime in plasma samples (2).

A total of 138 individuals comprising both sexes with age ranging from 5 to 81 years and weighing between 17 and 96 kg supplied 2084 plasma concentrations. The number of samples/subject ranged from 7 to 20, and there were 45 females. Demographic data collected on each individual (patient or normal healthy subject) included age, weight, creatinine clearance, and medical status (major disease state, if any). Classification of hepatic failure was based on clinical diagnosis and abnormal liver function tests (other than transaminases alone). The degree of renal dysfunction was assessed using routine serum creatinine measurements to calculate creatine clearance (CRCL) with the Cockroft and Gault equation (3) for adults and the Dechaux *et al.* equation (4) for children under 20 years.

Twenty three (16.5%) had a calculated CRCL < 3 L/h (<50 ml/min). Of these, 14 had end stage renal disease (ESRD) with a CRCL less than 2.3 L/h (30 ml/min). Eight of

[1] The views expressed in this article are personal opinions of the authors and not those of the U.S. Food and Drug Administration.

[2] Division of Biopharmaceutics, Center for Drug Evaluation and Research, Food and Drug Administration, 5600 Fishers Lane, Rockville, Maryland 20857.

[3] To whom correspondence should be addressed.

Abbreviations: MLR = Multiple Linear Regression; GAM = Generalized Additive Modelling; TBM = Tree Based Modelling; NONMEM = Nonlinear Mixed Effects Modeling program; CRCL = Creatinine Clearance; CL = Clearance; V1 = Apparent volume of the central compartment; V2 = Apparent volume of the peripheral compartment; Q = Intercompartmental clearance; Vss = Volume of distribution at steady state; NDA = New Drug Application; IV = Intravenous; ESRD = End Stage Renal Disease; PPK = Population Pharmacokinetics; PK = Pharmacokinetics; GLM = General Linear Model; C_i = Cook's Distance; AIC = Akaike Information Criterion; LLD = Loglikelihood Difference; MIX = Mixture Model Subroutine in NONMEM; RSE = Relative Standard Error; JKK = Jackknife; STS = Standard Two Stage; Est = Estimate; MAE = Mean Absolute Error.

0724-8741/95/1200-1845$07.50/0

these received hemodialysis, while the others were undialysed. There was no concomitant drug therapy. Each subject was sampled intensively, with blood samples taken predose and at specified time intervals.

Data Analysis

The methods used in the analysis of the data emphasize the combination of statistical techniques and graphical displays to discern the structure of the multivariable data. A stepwise approach was used in the analysis of the data: (1) exploratory data analysis to examine distributions and correlations between covariates, (2) determination of a basic pharmacokinetic model using the NONMEM program and obtaining Bayesian individual parameter estimates, (3) examination of distributions of estimates for parameters, (4) multiple linear regression (MLR) with case deletion diagnostics, generalized additive modelling (GAM), and tree-based modelling (TBM) for the selection of covariates and revealing structure in the data, (5) final NONMEM modelling to determine the population pharmacokinetic (PPK) model, and (6) evaluation of final PPK estimates using the jackknife technique.

Essentially, model building using steps (2), part of (4), and (5) was originally suggested by Maitre (5), subsequently elaborated upon by Mandema *et al.* (6), and Davidian and Gallant (7), and later used, for example, by Burtin *et al.* (8) in published analyses.

Step 1: Examination of Distributions and Correlations

The sampling distribution of each covariate was examined graphically by means of histograms. To reduce dimensionality of the covariate vector, graphic inspection of bivariate scatterplots and pairwise analysis based on the calculation of correlation coefficients were used (9).

Step 2: Determination of Basic Pharmacokinetic Model

The cefepime concentration-time data were fit to a two compartment open model parameterized in terms of CL, V1, Q, and V2 with input and elimination into and from the central compartment using NONMEM (Version IV, Level 2 of NONMEM and PREDPP version 3, level 1 (1)). Interindividual variability in CL was initially modeled using an exponential error model as shown in equation (1).

$$CL_j = CL(\exp \eta_j^{CL}); \qquad \eta_j^{CL} \text{ i.i.d. } \sim N(0,\omega^2_{CL}) \quad (1)$$

where CL_j is the hypothetical true total body clearance (CL) for the jth individual as predicted by the regression model. CL is the typical population value of CL; the η_j^{CL} represents the persistent difference between the jth individual's CL value and that predicted by the regression model; η_j^{CL}s are independent, identically distributed random variables. Interindividual variability in V1, Q, and V2 were similarly modelled. The exponential error model for interindividual variability was significantly better than the additive error model and was used for all subsequent analyses.

The residual intraindividual variability represents uncertainty in the relationship between the plasma concentrations predicted by the model and the observed concentrations. This uncertainty results from model misspecification, assay variability, etc. It was initially modelled using the proportional error model as shown in equation (2).

$$C_{ij} = C_{mij}(1 + \epsilon_{ij}); \qquad \epsilon_{ij} \text{ i.i.d. } \sim N(0,\sigma^2) \quad (2)$$

where C_{ij} is the ith observed concentration for the jth individual; C_{mij} is the ith concentration predicted by the model at the ith observation time for the jth individual. ϵ_{ij} are independent, identically distributed statistical errors of mean 0 and variance equal σ^2.

A combination of additive and constant coefficient of variation error models (eq. (3)) were found to describe the error in the data better than the proportional error model.

$$C_{ij} = C_{mij} + C_{mij}*\epsilon_1 + \epsilon_2 \quad (3)$$

In determining the basic PK model no covariance was assumed between elements of η at this stage. This was to ensure that each covariate had the opportunity to appear to be related to each η. With the fixed and random effects models chosen, empirical Bayes estimates of PK parameters were subsequently obtained using the POSTHOC option within the NONMEM program (1). The parameter estimates obtained from the initial analysis using the regression models typified by equations (1 & 3) were taken to be the population priors. With the empirical Bayes estimates of individual PK parameters obtained, the distributions of these parameter estimates and relationships between PK parameters and covariates were examined in subsequent steps.

Step 3: Examination of Distributions of Parameter Estimates

Density and normal scores plots were used to examine the sampling distributions of parameter estimates. The "straightness" of probability plots were measured using the modified Shapiro-Wilk test (10) which is based on the correlation coefficient. A very high correlation is consistent with normality.

Step 4: Selection of Covariates

Exploratory data analysis was performed on the empirical Bayesian parameter estimates from Step 2 treated as "data" to examine distributions, shapes, and relationships between covariates and individual PK parameter estimates.

MLR with Case Deletion Diagnostics

MLR. The "data" were subjected to stepwise (single terms addition/deletion) MLR using the general linear modelling (GLM) procedure in the SPLUS statistical program (Version 3.1, Statistical Sciences, Inc.) (9).

The relationship between parameters and covariates using MLR can be described with equation (4).

$$P_{kj} = \alpha_{ko} + \sum_{i=1}^{n} g_{ki}(x_{ij}) + \epsilon \quad (4)$$

where g_{ki} is a linear regression coefficient, α_{ko} is a constant, ϵ is a normally distributed with zero mean and constant variance, and P_{kj} is the predicted kth PK parameter in the jth subject.

Case Deletion Diagnostics. MLR is widely used to determine which variables are important predictors, and to find a reduced set of predictors. Thus, it is undesirable for the final model to depend strongly on only a few observations (individuals in this case). Conclusions drawn from the model could be misleading if those individuals (observations) were inconsistent with the bulk of the data. Measures of influence are thus very important for model building. They determine the influence an individual observation has on the fitted model, and what would happen to various aspects of the model if individuals (observations) were omitted. Cook's distance (C_i) (11) is one of such measures. The Cook's distance diagnostic was implemented with Splus (9) for each of the predictors of CL obtained from MLR.

A reference value is often useful for determining which values of C_i are actually "large". Although C_i does not have an F distribution (12) it is often compared to the 50th percentile of a standard F distribution on z parameters (i.e. 3 in this case) and n − z (i.e. 135, and n is the number of observations (individuals)) degrees of freedom, as a heuristic means of determining observations for which it is unusually large (13). This critical value is typically close to one, and this amounts to a cutoff of one for C_i which was used in this analysis.

GAM

The model in equation (4) for MLR makes a strong assumption about the linear dependence of P_k on each of the predictors, x_i. This may not always be the case. For many types of data a change in the mean of P is accompanied by a change in its variance. The GAM (14) approach presents a more general perspective for the handling of covariates in the multiple regression setting. This is a group of models that is as tractable as the linear model, but does not force the data into unnatural scales. Separate functions are introduced to allow for nonlinearity and heterogeneous variances. This is closer to a reparameterization of the model than to a reexpression of the response.

The use of the GAM approach for covariate selection in pharmacokinetics was recently described by Mandema *et al.* (6). With the GAM approach the function $g_{ki}(x_{ij})$ in equation (4) can be represented by any function, and smoothing spline functions were used in this study for their modelling flexibility. The GAM approach provides for straightforward interpretation of results by assuming an additive structure. It also allows the contributions of various covariates to be displayed graphically, making comparison with results obtained with straight forward MLR possible.

The building of the general additive model is done using the stepwise procedure described for MLR above. Each covariate is allowed to enter the model in any of several functional representations. The Akaike information criterion (AIC) is used as the model selection criterion (15). Single terms addition/deletion are carried out in a current model which reduces the AIC selection criterion the most. It stops when it hits a specified model boundary, or when no step will decrease the criterion any further.

TBM

TBM is an exploratory technique for uncovering structure in data, and assessing the adequacy of linear models (16). The TBM approach is especially effective when there are significant interactions among predictors (16).

Consider splitting the data into two parts, along any of the X predictors, so that the resulting groups are most homogeneous with respect to the response. Specifically, all splitting points are examined along all predictors, and the one that produces the smallest total within-group variance in the two groups is chosen. The split at a node is that split on the X variables which most successfully separates the high response values from the low ones (16). The data are then split into two parts, and the process is repeated on each part. At each stage all split points along all predictors are considered, so that the predictor can be used for splitting more than once. The splitting process can be terminated when no further splits can be found to significantly improve the homogeneity of the subgroups. Only the ranks of numeric predictors are used to define splits and not their values. This aspect of TBM for numeric predictors renders them invariant under monotone transformations of X. TBM as implemented in SPLUS (9) was used for the construction of the regression tree for CL.

Step 5: Population Model Building Using NONMEM

For each NONMEM analysis the improvement in fit obtained upon addition of a factor into the regression model was assessed by the change in the NONMEM objective function. Minimization of the NONMEM objective function, equal to twice the negative log-likelihood of the data ($-2\log L_{max}$), is equivalent to maximizing the probability (likelihood) of the data. Thus, monitoring changes in the objective function serves as a statistical test showing which parameter values render the data most probable. The difference in the objective function values obtained for the full versus the restricted models is approximately chi-square distributed with degrees of freedom equal to the number of parameters which are set equal to a fixed value in the restricted model. This approach was used to estimate the significance of various patient characteristics, usually referred to as fixed effects or covariates, as predictors of pharmacokinetic parameters. Similarly, the appropriate statistical model was determined and the magnitude of inter- and intraindividual variability estimated (17).

The goodness of fit of each NONMEM analysis was also assessed by the examination of scatterplots of predicted versus measured cefepime concentrations and weighted residuals, the percent relative standard error of the mean (i.e., %RSE = (standard error estimate/parameter estimate) * 100%), and changes in the estimates of interindividual and residual intraindividual variability resulting from the addition or deletion of a parameter.

The NONMEM analysis was continued by testing covariates selected from Step 3 above, one at a time, in the regression model. During the process, a difference in the objective function value of ($3.8 = X^2$ critical value for 1 degree of freedom at $p = 0.05$) was considered statistically significant. The full model was then constructed to include all covariates which demonstrated a significant effect on CL, V1, Q and V2.

Testing for the significance of each covariate, weight

(either as a linear or nonlinear predictor) was found to be a predictor of V1, V2 and Q with a loglikelihood difference (LLD) of ≤ -20.87 between the full and reduced models. Sex was not a predictor of either V1 or V2. The significance of this covariate was not tested for Q because it was not selected by MLR or GAM. Age was found to be a predictor of V2 (LLD = −20.84) and not Q. Also, age was not a predictor of CL (LLD = 0 when compared with the base model of CL modeled without regard to any covariate), while CRCL modeled nonlinearly was found to be a significant predictor of CL (LLD = −327.09 when compared to the base model).

The GAM and TBM findings coupled with the nonnormality in the distribution of CL observed from Step 3 led to testing the appropriateness of using the MIX subroutine in NONMEM (1) for fitting the data. The MIX subroutine allows mixture modeling to be carried out within the context of mixed effects modeling. A mixture model assumes that the population consists of two or more subpopulations, each approximating a normal distribution and each subpopulation may have its own model. With two subpopulations it might be assumed that some fraction of the population has one set of typical values of pharmacokinetic parameters, and the remaining fraction has another set of typical values. The mixing fraction (θ) and both sets of typical values can be estimated, and NONMEM computes an estimate of the subpopulation to which an individual belongs (1). Thus, the estimation of CL for two subpopulations was tested. The basic model for CL with the mixture model (two subpopulations) was superior to basic model assuming only one population (LLD = −616.52). It was also better than the model in which CRCL was modeled as a nonlinear predictor of CL. Different models for CRCL as predictor of CL: (1) same intercept with different slopes, (2) same slope with different intercepts, and (3) different intercepts with different slopes were tested with the mixture model. The modified Shapiro-Wilk test showed that all other parameters approximated normal distributions.

To arrive at the minimum subset of covariates, the full model for each structural model parameter was tested against corresponding restricted models. A full model was developed for V2, while other parameters had only one covariate in their models. The models were:

$$CL_1\ (L/h) = \Theta_1 + \Theta_3 * CRCL \quad (5)$$

$$CL_2\ (L/h) = \Theta_1*\Theta_2 + \Theta_3*\Theta_4 * CRCL \quad (6)$$

$$V1\ (L) = \Theta_5 + WT^{\Theta 7} \quad (7)$$

$$Q\ (L/h) = \Theta_9 + \Theta_6*WT \quad (8)$$

$$V2\ (L) = \Theta_{10} + WT^{**}\Theta_7 + AGE^{**}\Theta_{11} \quad (9)$$

where CRCL is L/h, WT is in Kg, and AGE is in years. To partially compensate for the multiple comparisons, $p < 0.005$ was used. Thus, a change in the objective function value of 7.8 was necessary to show statistical significance between each proposed restricted model and the full model when the two models differed by 1 parameter. The final regression models included the fewest number of covariates in each parameter resulting in a model which was not significantly different from the full models.

Step 6a: Evaluation of Final Parameter Estimates

Numerical methods used to fit experimental data should, ideally, give estimates of both the primary and secondary parameters that are unbiased and of defined precision. In practice bias should be absent as long as the error in the data is of known distribution and variance. The preciseness of the primary parameters can be estimated from the final fit of the multiexponential function to the data, but they are of doubtful validity if the model is severely nonlinear (18). The preciseness of the secondary parameters (in this case variability) are likely to be even less reliable. Consequently, the results of statistical tests carried out with preciseness estimated from the final fit could easily be misleading. The first order method in NONMEM yield estimates of parameters which are sometimes biased. A possible way of reducing bias in parameter estimates and of calculating realistic variances for them is to subject the data to the jackknife technique (19,20). The technique requires little by way of assumption, or analysis.

Jackknifing for the homogeneous data involves one-at-a-time omission. However, the block-at-a-time (s-at-a-time) omission has been shown to have a similar efficiency to the one-at-a-time omission (21). A 10%-at-a-time omission was applied to the data set and reanalyzed with NONMEM. A naive Student t approximation for the standardized jackknife estimator (22) was used. The magnitude of bias reduction is the reciprocal of the total number of blocks omitted.

Step 6b: Performance of Mixture Modeling Within the Context of the Mixed Effects Model

Since mixture modeling was done within the frame work of mixed effects modeling, it was of interest to assess the performance of this approach. In doing this, parameter estimates obtained using the standard two stage (STS) approach were used as standards. Each parameter estimate was then transformed so that it could be expressed as a percentage of the corresponding STS value ($\%E_{(p)}$, p = 1,...,138) thus:

$$\%E_{(p)} = [(Est_{NONMEM} - Est_{STS})/Est_{STS}] * 100 \quad (10)$$

where Est_{NONMEM} and Est_{STS} are NONMEM and STS estimates of parameters, respectively. The dimensionless quantity ($\%E_P$) enabled us to evaluate the performance of the mixture modeling approach for parameter estimation within the mixed effects model setting. The mean of %E values for each parameter estimate obtained from 138 subjects provided a measure of accuracy with which the parameter had been estimated using NONMEM. Precision was computed using mean absolute error thus:

$$\%MAE_P = [1/N\ \Sigma(|Est_{NONMEM} - Est_{STS}|)/Est_{STS}] * 100 \quad (11)$$

where $\%MAE_P$ is the percent mean absolute error in the estimation of a parameter, P.

RESULTS

Distributions of Covariates, Parameter Estimates, and Correlations

The distributions of the various subject (patient) demographic variables are shown in Fig. 1, and the relationships between covariates are shown in the pairs plot in Fig. 2. There appears to be a slight bimodality in the distribution of CRCL. The non-normality in the distribution of CRCL was confirmed by formal testing. Although weight was significantly correlated with age ($p < .0001$) the correlation coefficient was 0.34. Creatinine clearance was similarly correlated with age ($p < .0007$), and the correlation coefficient was 0.29. Although these correlations were significant, the low values of the correlation coefficients did not allow a reduction in the dimensionality of the covariate vector to be achieved a priori.

A density plot of the empiric Bayes estimates of CL revealed a non-normality in the distribution of the estimates of this parameter (Fig. 3a) which may be due in part to non-normality of CRCL. However, density plots of V1, Q, and V2 did not reveal any significant departure from normality (Fig. 3 (b–d)), and this was confirmed by formal testing.

Exploratory Data Analysis

Exploratory analysis of the data with MLR and GAM was performed using empiric Bayes estimates of CL, V1, Q, and V2 against CRCL, age, weight, and sex as predictors of the parameters. The results of the MLR and GAM analyses are presented in detail below for CL, while those for Q, V1, and V2 are mentioned briefly.

MLR with Case Deletion Diagnostics

Fig. 4 (a & b) shows the relationship between CRCL, age and CL as obtained with MLR. However, CRCL as a linear predictor of CL was not satisfactory as observed from the partial residual plot (Fig. 4a). Subjects with CRCL less than 60 ml/min were over predicted. The Cook's distance diagnostic did not reveal any high leverage observations ($C_i < 1$) in the "data". The highest C_i value was 0.12 for CRCL and 0.125 for age.

Weight and sex were identified as predictors of V1, while age and weight were identified as predictors of Q (Table I). On the other hand, CRCL, weight, sex, and sex interacting with weight were identified as predictors of V2 by this GLM procedure (Table I).

GAM

Smoothing spline scatterplot smoother was used in the GAM approach for fitting. Setting interior knots at 33rd and 66th quantiles for the modelling of CL did not yield a result significantly different from using only one interior knot set at the median (the default setting). Consequently, the default setting was used for all models tested with the GAM approach. (Note that we were interested in the selection of covariates and the shapes of the fits only.) Like the MLR approach, CRCL and age were identified as predictors of CL using GAM (Fig. 4 (c & d)). GAM also identified a nonlinear relationship between these covariates and CL. It is worth noting from Fig. 4c that most subjects with low CRCL were poorly predicted with the suggested nonlinear relationship between CL and CRCL.

As with the MLR approach, GAM selected weight and

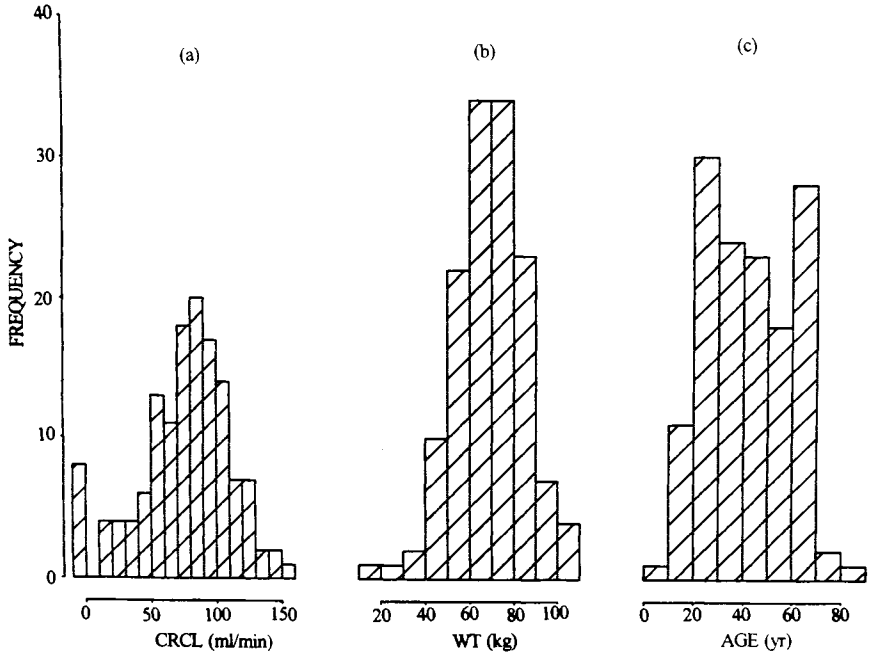

Fig. 1. Frequency distribution of demographic variables: (a) CRCL, (b) weight (WT), and (c) age.

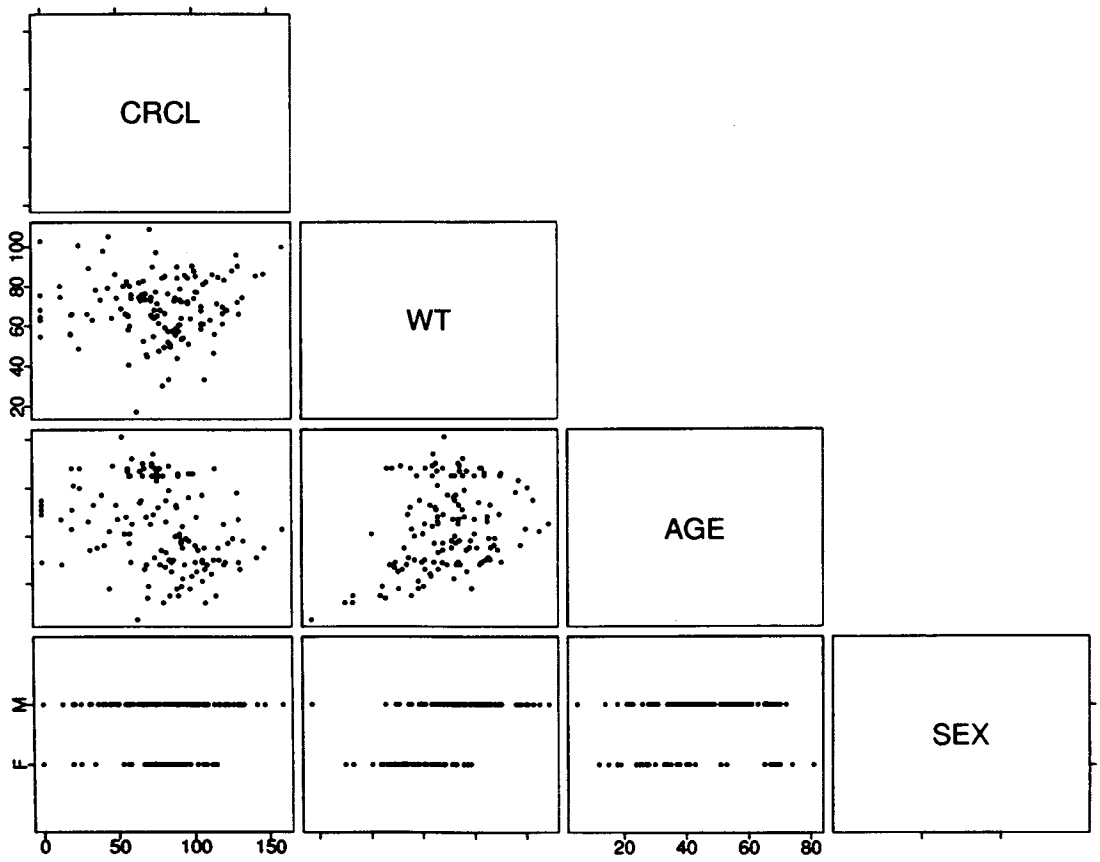

Fig. 2. Bivariate scatterplots showing relationships between covariates.

sex as predictors of V1, while weight and age were selected as predictors of Q. Moreover, the GAM approach identified weight and age as nonlinear predictors of V2, while sex was a linear predictor (Table I).

TBM

The result of TBM which was carried out only on CL showed that CRCL was the most significant predictor of CL. The first node for splitting the data was based on CRCL (Fig. 5a). The histogram of the splits for CL and CRCL shows the splits for CL almost matching those of CRCL (Fig. 5 (b & c)). Two levels of response apparently associated with two CRCL subgroups is noticeable. No interactions were revealed.

Population Model Building

The NONMEM analysis carried out on the variables selected by GAM showed that the mixture model with CL modelled in two subpopulations was superior to modelling it without regard to the existence of two subpopulations in the data set. Although age was identified as a predictor of CL from the exploratory modelling methods, NONMEM did not identify age as a predictor of CL. Modelling CL with different intercepts and slopes with CRCL entering the model linearly (using the mixture model) was superior (LLD = −616.52 when compared with CL modeled without regard to any covariate) to other models tested for CL (assuming unimodality in CL distribution). Some examples of the "other" CL models were CRCL modeled as a nonlinear predictor of CL or a linear predictor of CL with a breakpoint at 50 ml/min. The LLD for these models when compared to the basic model (i.e. CL modeled without regard to any covariate) were −327.09 (for CRCL as nonlinear predictor of CL) and −300.02 (for CRCL modeled with a breakpoint at 50 ml/min). Thus, equations (5) and (6) describe the final model for CL. The distribution of individuals in the two apparent subpopulations are shown in Fig. 6.

Only weight was identified by NONMEM as a significant predictor of V1 and Q (Table I). Modelling Q linearly as a function of weight without an intercept yielded Q estimates which were more precise, and this was used as the final model for Q. On the other hand, V2 was similarly predicted by either age or weight. Although age and weight were included in the full model for V2, the addition of age to weight in the model did not significantly improve the goodness-of-fit. Thus in the final model, V2 was modelled as a nonlinear function of weight without the intercept which was infinitesimally small. Removing the intercept in the model for V2 did not alter the objective function.

Using the parameter estimates from equations (5) and (6), the equations for the estimation of CL are of the forms

CL_1(L/h) = 1.0(20.0%) + 0.25(44.0%) * CRCL for the

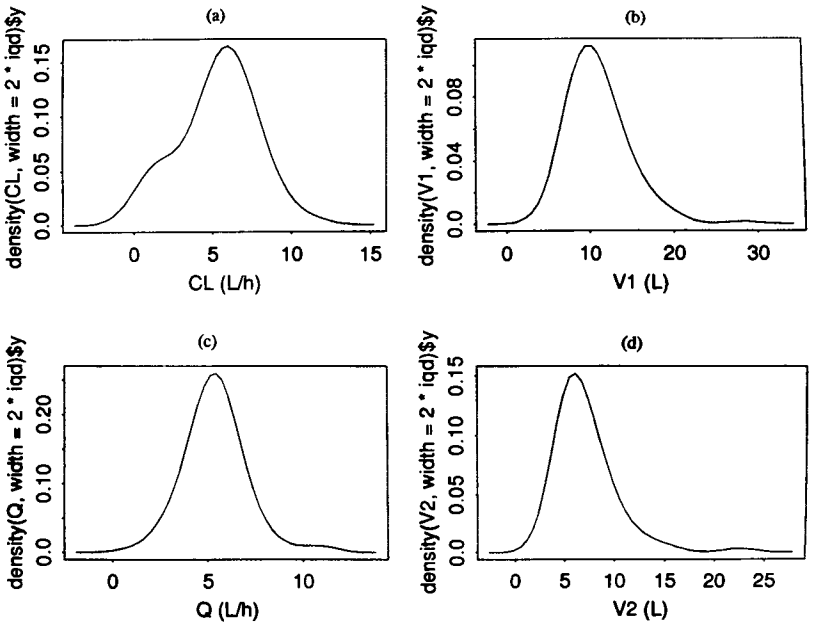

Fig. 3. Density plots showing the structure of empiric Bayesian estimates for: (a) CL, (b) V1, (c) Q, and (d) V2. Note that the extension of density plots beyond zero in the negative direction is a consequence of the smoothing function used in producing the plot; no parameter value was negative. The degree of smoothness of the density plots was determined by the width which was set to two times the interquartile distance.

group of individuals with a shallow slope of the CL regression model, and

CL_2(L/h) = 1.0(20.0%) * 0.27(35.0%) + 0.25(44.0%) * 4.43(45.0%) * CRCL for the group of individuals with a steep slope of the CL regression model. The numbers in parentheses are percent relative standard errors. In addition, the equations for the estimation of other structural model parameters are:

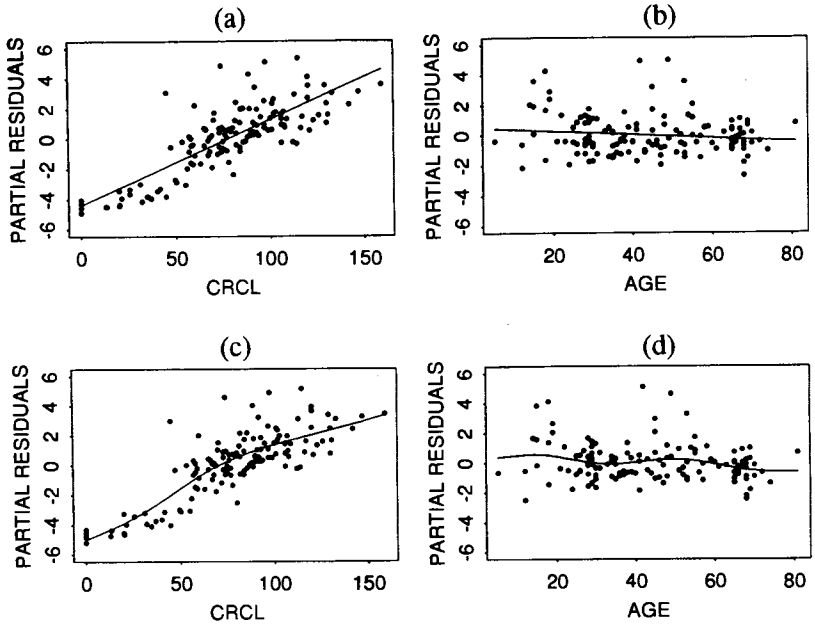

Fig. 4. MLR and GAM analyses results: scatterplots of partial residuals of CL (L/h) vs (a) CRCL (ml/min), and (b) age (yr) from MLR analysis; and CL (L/h) vs (c) CRCL (ml/min), and (d) age (yr) for GAM analysis. The ordinate represents the partial residuals, i.e, the individual empiric Bayesian estimates of CL minus the parameter estimate based on other subject covariates. The ordinate label is the expression used to specify the contribution of the covariate to the model formula in the SPLUS language. The same scale is used for the ordinate in both plots so that the relative importance of the covariates can be compared.

Table I. Covariate Selection by Regression Method

Par.	Regression Method			
	MLR	GAM	TBM	NONMEM
CL	~CRCL + AGE	~s(CRCL) + s(AGE)	~CRCL + WT + AGE	~CRCL (MIX)[a]
V1	~WT + SEX	~WT + SEX	–	~WT (nl)[b]
Q	~WT + AGE	~WT + AGE	–	~WT
V2	~CRCL + WT + SEX + WT:SEX	~s(WT + s(AGE) + SEX	–	~WT (nl)[b]

[a] MIX—NONMEM mixture model.
[b] nl—entered the NONMEM model nonlinearly.

$$V1\ (L) = 3.81\ (11.1\%) + WT^{0.43\ (2.0\%)}$$

$$Q\ (L/h) = 0.08\ (12.4\%)*WT$$

$$V2\ (L) = WT^{0.43\ (2.0\%)}.$$

The NONMEM estimate of the fraction of individuals with shallow slope for the CL regression model (subpopulation I) was 0.16 (25.0%, percent relative standard error (%RSE)).

The unexplained interindividual differences in plasma concentration profile of cefepime could be described with the exponential error model (Eq. 1). The interindividual unexplained variability was found for CL to be 42.3 and 25.1% for subpopulations I (shallow slope group of the CL regression model) and II (steep slope group of the CL regression model), respectively. The unexplained interindividual variability in V1, V2, and Q were found to be 27.6, 28.0, 29.6%, respectively (Table II). The residual intrasubject variability was 14.5% at a concentration of 5 ug/ml and less than 11% when the concentration was greater than or equal to 30 ug/ml.

Evaluation of Final Parameter Estimates

Most of the JKK estimators for regression coefficients and variability were similar to the NONMEM estimates (Ta-

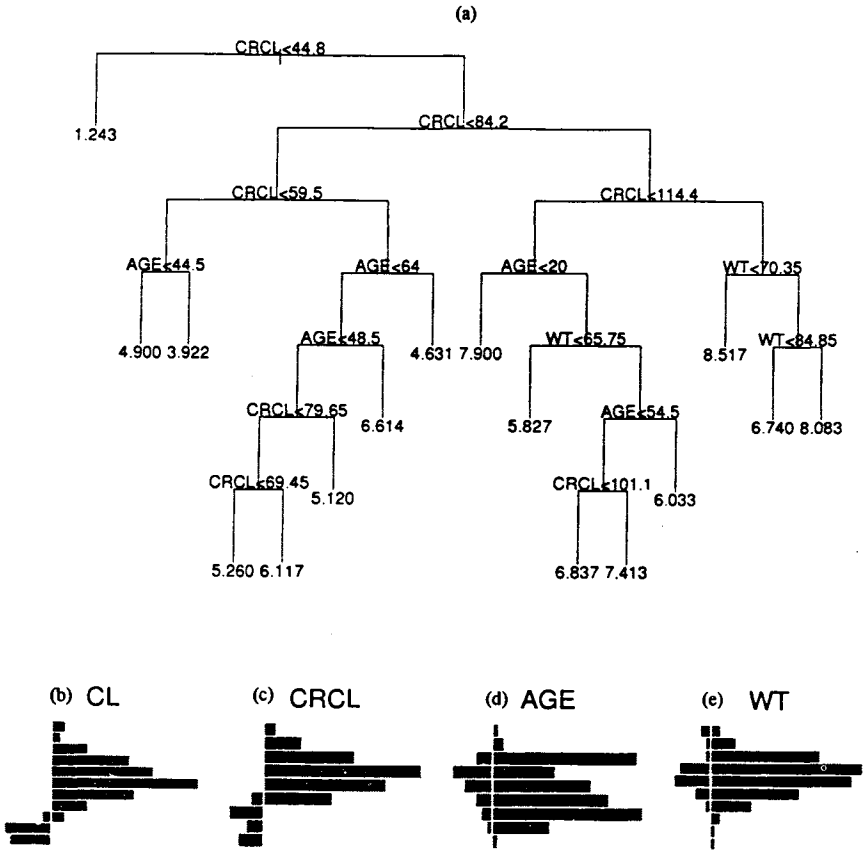

Fig. 5. (a) A tree-based model for CL. The top panel displays the labeled dendrogram. The lower panel displays a side-by-side histogram for: (b) CL, (c) CRCL CRCL (ml/min), (d) age (yr), and (e) WT (kg). The left-side histogram summarizes the observations following of the left split, and similarly for the right.

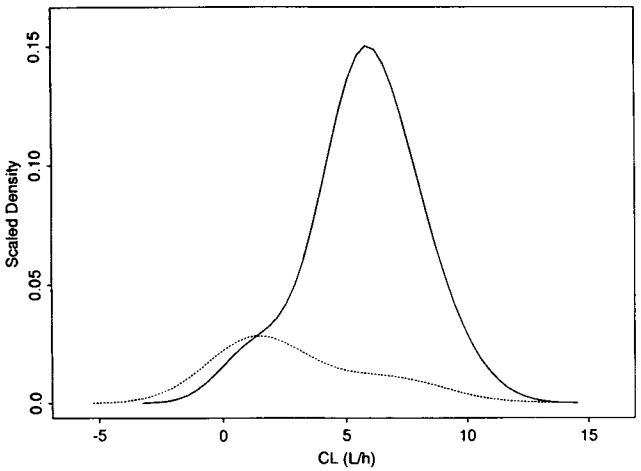

Fig. 6. The distribution of individuals into two subpopulations by NONMEM. (.....)—a majority of individuals with shallow slope of for the CL regression model. (____)—a majority of individuals with steep slope of for the CL regression model. The densities were rescaled to reflect the proportion of individuals in each subgroup. Note that the extension of density plots beyond zero in the negative direction is due to smoothing function used in producing the plot; no CL value was negative. The degree of smoothness of the density plots was determined by the width which was set to two times the interquartile distance.

bles III). In some cases, these were more precise than the NONMEM estimates.

Performance of Mixture Modeling

Relative to the STS approach the estimation of CL and Vss using MIX subroutine within NONMEM was associated with minimal to no "deviation", and good precision. The deviations associated with the estimation of CL and Vss were 0.09 and 10.09%, respectively. The %MAE in the estimation of these parameters were 0.19% (CL) and 18.06% (Vss).

DISCUSSION

The concepts of the exploratory data analysis (EDA) technique have been well defined for over a decade (23), but the application of EDA has been slow to gain acceptance in many fields of science. EDA is a powerful technique for interpreting data. The visualization of the data provided by the graphical techniques presented above has the potential to speed the analysis of PPK data.

The EDA approaches used in the analysis of this data set produced relatively consistent results. The non-normality in the distribution of CL and CRCL were confirmed by subsequent EDA approaches. This non-normality in CL distribution showed up as model inadequacy in the MLR results of CRCL as a linear predictor of CL (Fig. 4a). Case deletion diagnostics were used with MLR because of the ease with which this can be implemented within SPLUS. The GLM object returned by SPLUS for MLR fit was used for this purpose, and there were no high leverage points in the "data" set. Also, subjects with low CRCL values were poorly predicted with the GAM model. Although the MLR and GAM analysis selected CRCL and age as predictors of CL, ANOVA carried out on the residuals from the analyses indicated that age was not a significant predictor (MLR: $p < 0.115$, GAM: $p < 0.06$) of this parameter. This is amplified in Fig. 4b and 4d.

Based on the GAM analysis, there appeared to be a nonlinear relationship between cefepime CL and CRCL. The nonlinearity suggested by GAM in the relationship between CL and age was not significant. The relationships between the partial residuals and age was flat for the majority of the data. Only a few observations appeared to skew the values at the extremes of age. Smoothing splines are sensitive to outlying observations along the ordinate.

The MLR and GAM analysis suggested sex and weight as predictors of V1. ANOVA carried out on the residuals suggested that sex was not a significant predictor of V1. The GAM analysis suggested nonlinearity in the relationship between V1 and weight. Age and weight were shown to be significant linear predictors of Q by both modelling procedures. Nonlinearity in the relationship between age, weight and V2 was suggested by GAM in addition to sex. MLR also showed these factors to be predictors of V2 with interaction occurring between weight and sex.

The apparent existence of two subpopulations was confirmed with TBM (Fig. 6). The histogram of TBM at the first node of the tree shows the dichotomy in CL and CRCL values, with the latter being the most significant predictor of the former (Fig. 6 (a & b)). The first node for splitting of data occurred at 44.8 ml/min (CRCL) were the CL values in the right and left splits are most different; suggesting two levels of response by two different subpopulations. TBM was a powerful tool for revealing structure in the data, and confirmed the inadequacy of modeling CRCL as a linear predictor of CL.

The inadequacy of modeling CRCL as a nonlinear predictor of CL using GAM, coupled with the results of TBM and taking into consideration the non-normality in the distribution of CL estimates led to the use of the NONMEM MIX subroutine for testing models for CL. The results of models tested using mixture modeling were compared with the results obtained with CL models assuming unimodal distribution of CL estimates. Fits obtained with the mixture model were significantly better than any CL model without the MIX subroutine. Modelling CRCL as a linear predictor of CL (equations (5 & 6)) with different slopes and intercept was better than other models tested. The use of mixture modelling to account for the apparent existence of two sub-

Table II. Random Effect Parameters

Parameter	Intersubject Variability	95% CI
CL_1	42.3%	23.5–55.0%
CL_2	25.1%	19.1–30.0%
V1	27.6%	23.2–31.4%
Q	28.0%	12.2–40.0%
V2	29.6%	19.3–33.5%
	C_{mij}[a] (μg/ml) Residual Intrasubject Variability	
σ	5	14.5%
σ	30	10.8%

[a] C_{mij} is the ith concentration predicted by the model for the jth individual.

Table III. Comparison of Normal Theory Parameter and Variance Parameter Estimates with Jackknife Estimators

	Normal theory			Jackknife[a]		
Parameter	Estimate	SE	%RSE	Estimate	SE	%JKK SE
θ_1	0.990	0.200	20.200	0.780	0.075	9.615
θ_2	0.270	0.090	33.330	0.910	0.335	36.813
θ_3	0.250	0.110	44.000	0.260	0.055	21.150
θ_4	4.430	2.000	45.460	3.513	0.347	9.878
θ_5	3.810	0.420	11.024	3.830	0.073	1.906
θ_6	0.080	0.010	12.500	0.080	0.002	2.500
θ_7	0.430	0.010	2.326	0.430	0.006	1.395
θ_8	0.160	0.040	25.000	0.190	0.020	10.526
ω_1	0.180	0.060	33.333	0.190	0.025	13.158
ω_2	0.060	0.010	16.667	0.060	0.004	6.667
ω_3	0.080	0.010	12.500	0.080	0.002	2.500
ω_4	0.080	0.020	25.000	0.080	0.003	3.750
ω_5	0.090	0.040	44.444	0.080	0.007	8.750
ϵ_1	0.010	0.001	10.000	0.010	0.0003	3.000
ϵ_2	0.050	0.010	20.000	0.050	0.002	4.000

[a] 10% of subjects removed/run. θ_1 is the intercept for the regression of CRCL as a predictor of CL in subpopulation I. θ_2 is the intercept scaling factor for the regrssion of CL on CRCL in subpopulation II. θ_3 is the regression coefficient for CRCL as a predictor of CL in subpopulation I. θ_4 is the regression coefficient scaling factor for CRCL as predictor of CL in subpopulation II. θ_5 is the intercept for the regression of V1 on WT. θ_6 is the regression coefficient of WT as a predictor of Q. θ_7 is the regression coefficient of WT as a nonlinear predictor of V1 and V2. θ_8 is the fraction of individuals partitioned into subpopulation I. ω_1 is variance in CL_1. ω_2 is variance in CL_2. ω_3 is variance in V1. ω_4 is variance in Q. ω_5 is variance in V2. ϵ_1 is the constant coefficient of variation error variance, and ϵ_2 is the additive error variance.

populations of individuals in the sample provided a better fit for the data than using a nonlinear model for CRCL (in NONMEM) as suggested by GAM.

As suggested by GAM, NONMEM selected weight as a nonlinear predictor of V1. Gender had no influence in the estimation of this parameter by NONMEM (Table I). Also, Q was modelled as a function of body size (equation (8)).

The nonlinear relationship between age, weight and V2 suggested by GAM was confirmed in the NONMEM analysis during the model building phase. The final model for V2 included only weight.

The modelling of CL as a function of CRCL led to a significant decrease in the objective function (LLD = −616.52), with a decrease in variability of 32.8% and 3.5% for subpopulations I (equation (5)) and II (equation (6)), respectively. The relatively lower precision associated with the estimation CL for subpopulation I is a consequence of a smaller proportion of individuals (twenty two) in this subgroup as compared to 116 individuals in subpopulation II. The partitioning of some individuals in subpopulation II into subpopulation I and vice versa was due to the fact that the tail of one distribution overlapped into the other (Fig. 6).

Significant improvements were also observed in the goodness-of-fit when V1, Q, and V2 were modelled as functions of weight. The intersubject variability in V1, Q, and V2 decreased by 4.2, 17.7, and 5.6%, respectively.

The finding that CRCL is a linear predictor of CL is in agreement with earlier reports by Barbhaiya *et al.* (24,25). The authors (24) also reported that gender was not a predictor of CL or any of the other pharmacokinetic parameters, and this was confirmed in this analysis. Barbhaiya *et al.* (24) observed statistically significant age-related effects for CL and volume of distribution at steady state, but concluded the magnitudes of these changes in the pharmacokinetics of cefepime was not significant enough to recommend dosage adjustment in the elderly. Our findings that age was not a predictor of CL or any of the volume terms supports the proposition that dosage adjustment for this drug should be based on renal function (25). In a recently published review article on cefepime it was also recommended that dosage adjustment should be based on renal function (26)

The similarity of the final NONMEM parameter estimates to JKK estimators indicated that the NONMEM estimates were relatively unbiased and precise. It is worth noting that in using the JKK technique, the PPK estimates were evaluated and not the population model. Also, NONMEM with the MIX subroutine produced accurate and precise parameter estimates relative to the STS estimates in a situation where the subjects were sampled intensively.

A rich data set was used for this analysis with an average of 15 data points per individual. Thus, the shrinking of empirical Bayes parameter estimates towards the population mean (which would have occurred with sparse sampling) was not a problem.

Although MLR and GAM may provide a selection of covariates which could be included in the NONMEM model for explaining variability, testing the contribution of these covariates by ANOVA (say, $p \leq 0.05$) may provide additional information about the significance of a covariate. If a covariate does not achieve significance at $p \leq 0.05$, it appears from these results that the covariate may not be selected by NONMEM.

The nature of the distribution of the empiric Bayes estimates for CL, model misspecification revealed in MLR and GAM analyses, and the results of the TBM technique were very useful in the regression model building process. EDA is an important first step for making a selection of essential covariates. Understanding the structure of a given data set is

of utmost importance in carrying out an informative analysis. EDA revealed patterns in the data that helped in the explanation of intersubject variability. Thus, understanding the structure of the data should guide the analysis carried out with NONMEM. The use of mixture modelling with NONMEM led to a better characterization of the population that was studied without violating the assumptions of normal theory.

The findings suggest that using informative graphical and statistical techniques enhance the understanding of the data structure and lead to an efficient analysis of the data. Although not universal, the approaches described herein have some level of general applicability.

ACKNOWLEDGMENTS

We are grateful to Bristol-Myers Squibb for allowing us to use the data from the cefepime NDA.

REFERENCES

1. Beal SL, Sheiner LB. NONMEM *Users Guide*, Parts I–VI, Division of Clinical Pharmacology, University of California, San Francisco, 1979–1992.
2. Barbhaiya RH, Forgue ST, Shyu WC, Papp EA, Pittman KA. High pressure liquid chromatography of BMY-28142 in plasma and urine. Antimicrob. Agents Chemother. **31**:55–59 (1987).
3. Cockroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron **16**:31–41 (1976).
4. Dechaux M, Gonzalez G, Broyer M. Creatinine plasmatique, clearance et excretion urinaire de la creatinine chez l'enfant. Arch. Fr. Ped. **35**:53–62 (1978).
5. Maitre PO, Buhrer M, Stanski DR, Thompson D. A three step population pharmacokinetic analysis using NONMEM and bayesian regression: an application to midazolam. J. Pharmacokinet. Biopharm. **19**:377–384 (1991).
6. Mandema JW, Verotta D, Sheiner LB. Building population pharmacodynamic models. I. Models for covariates. J. Pharmacokinet. Biopharm. **20**:511–528 (1992).
7. Davidian M, Gallant AR. Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. J. Pharmacokinet. Biopharm. **20**:529–556 (1992).
8. Burtin P, Jacqz-Aigrain E, Girard P, Lenclen R, Magny JF, Betremieux P, Tehiry C, Desplanques L, Mussat P. Population pharmacokinetics of midazolam in neonates. Clin. Pharmacol. Ther. **56**:615–625 (1994).
9. S-PLUS (version 3.1). Seattle, Washington: Statistical Sciences, Inc., 1992.
10. Shapiro SS, Francia RS. An approximate analysis of variance test for normality. J. Am. Stat. Assoc. **67**:215–216 (1972).
11. Cook RD. Detection of influential observations in linear regression. Technometrics **19**:15–18 (1977).
12. Obenchain RL. Letters to the editor. Technometrics **19**:348–351 (1977).
13. Cook RD, Weisberg S. Criticism and influence analysis in regression. In: Leinhardt S. (ed). *Sociological Methodology*. San Francisco: Jossey-Bass, 1982, pp 313–316.
14. Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. New York: Chapman and Hall, 1990, pp 82–103.
15. Hastie TJ. Generalized additive models. In: Chambers JM, Hastie TJ., eds. *Statistical Models in S*. Pacific Grove, California: Wadsworth & Brooks/Cole Advanced Books & Software, 1992, pp 249–307.
16. Brieman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. Belmont, California: Wadsworth International Group, 1984, pp 216–265.
17. Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J. Pharmacokinet. Biopharm. **5**:445–479 (1977).
18. Drapper NR, Smith H. *Applied Regression Analysis*. New York: Wiley, 1966, pp 263–304.
19. Quenouille MH. Notes on bias estimation. Biometrika **43**:353–360 (1956).
20. Miller RG. The jackknife—a review. Biometrika **61**:1–15 (1974).
21. Mosteller F, Tukey JW. Data analysis, including statistics. In: *Handbook of Social Psychology Vol 2*, Lindzey O, Aronson E, eds. Reading, MA.: Addison-Wesley, 1964, pp 144–145.
22. Hinkley DV. Jackknife confidence limits using Student t approximations. Biometrika **64**:21–28 (1977).
23. Tukey JW. *Exploratory Data Analysis*. Reading, MA.: Addison-Wesley, 1977.
24. Barbhaiya RH, Knupp CA, Pittman K. Effects of age and gender on pharmacokinetics of cefepime. Antimicrob. Agents Chemother. **36**:1181–1185 (1992).
25. Barbhaiya RH, Knupp CA, Forgue ST, Matzke GR, Guay RP, Pittman K. Pharmacokinetics of cefepime in subjects with renal insufficiency. Clin. Pharmcol. Ther. **48**:268–276 (1990).
26. Hardin TC, Jennings TS. Cefepime. Pharmacotherapy **14**:657–668 (1994).