Pharmacokinetics

Population Pharmacokinetics II: Estimation Methods

Ene I Ette and Paul J Williams

OBJECTIVE: To present, compare, and contrast the various approaches to estimating population pharmacokinetic (PPK) models with respect to the mathematical foundation, statistical aspects, software programs for implementation, and underlying assumptions.

DATA SOURCES: Information on PPK was retrieved from a MEDLINE search (1977–August 2004) of literature and a bibliographic review of review articles and books. This information is used in conjunction with experience to explain the various methodologic approaches to PPK.

STUDY SELECTION AND DATA EXTRACTION: All articles indentified from data sources were evaluated and relevant information was included in this review.

DATA SYNTHESIS: Over 80 articles dealing with PPK estimation methods and/or their implementation were identified and reviewed. Sixty-four of these were chosen for their direct relevance to the subject of this article. Different estimation methods ranging from the naïve averaging and naïve pooled approaches through the standard two-stage approach to the nonlinear mixed-effects modeling approaches for estimating PPK are reviewed with their advantages and limitations.

CONCLUSIONS: PPK estimation methods that rely on the characterizing of mixed (fixed and random) effects are known to produce PPK parameter estimates that are less biased than those obtained using the naïve and standard two-stage approaches. The NONMEM software is the most widely used software for the characterization of PPK.

KEY WORDS: estimation, pharmacokinetics, population.

Ann Pharmacother 2004;38:xxxx.

Published Online, 14 Sept 2004, *www.theannals.com*, DOI 10.1345/aph.1E259

Population pharmacokinetic (PPK) models have a wide range of application both to direct patient care and drug development. There are several advantages to employing PPK models compared with traditional pharmacokinetic model development. The previous installment of this series addressed the types of conceptual models necessary for an understanding of PPK.[1] The current installment explains the various methods used to estimate PPK models.

Over the past two and half decades, a variety of methods have been proposed for the characterization of the PPK of drugs. A discussion of some of the methods follows and is the focus of this section. The goals of a PPK analysis and the data type will determine the method selected for the analysis.

Author information provided at the end of the text.

See also Part I (2004;38:1702-6, DOI 10.1345/aph.1D374)

Methods Applied to Population Pharmacokinetic Modeling

NAÏVE AVERAGE DATA APPROACH

It is common practice in preclinical and clinical pharmacokinetics to perform studies in which the drug administration and sampling schedules are identical for all subjects. For this type of analysis, there are as many data points as there are individuals at each sampling time. Analysis of such data using the naïve averaging of data (NAD) approach consists of the following procedure.

(1) Computing the average value of the data for each sampling time:

$$\bar{y}_i = 1/N \sum_{j=1}^{N} y_{ij} \qquad \text{Eq. 1}$$

for $i = 1,\ldots, n$ where n is the standard number of individual data. The averaging of data across individuals makes sense, because all y_{ij} for $j = 1, \ldots, N$ have been measured under identical conditions.

(2) A model $y^m = f(\phi)$ is fitted to the mean data n-vector $\bar{y} = (\bar{y}_1,\ldots\bar{y}_n)^t$ and estimates the best-fit parameter values ϕ^*. The latter notation (ϕ^*) is used to distinguish it from individual estimates, denoted $\hat{\phi}$.

The NAD approach is attractive because of its simplicity. One unique fitting is sufficient for obtaining estimates of parameters describing the mean response. ϕ^* Components are quite often interpreted as "mean" parameter values. Correspondingly, $\hat{\mu}_{NAD}$ will be used for ϕ^* in the latter. The method is widely applicable in experimental data (EP) studies with standardized designs, including bioavailability, bioequivalence, and dose proportionality studies. Because of the smoothing effect of averaging, mean data generally look nicer than individual data, and better fitting often results when compared with individual data.

However, the NAD approach provides an estimate of the $\hat{\mu}_{NAD}$ sample mean. In this regard, several drawbacks of this approach must be pointed out. The use of NAD to establish a pharmacokinetic model may be misleading. Data averaging can, quite often, produce a distorted picture of the response. Averaging of monoexponential data from 2 subjects with very different half-lives has been shown to produce a mean curve that exhibits an apparent biexponential decay.[2] Sometimes the opposite situation is the case. The smoothing effect of the averaging will tend to obscure peculiarities that can be seen in individual data. The existence of secondary peaks in the plasma level–time course of individuals may be undetectable in the average curve if the rebounds occur at different time points.

NAD also performs poorly in terms of parameter estimation. The reference to individual data disappears after data averaging. All sources of variability are confounded. Because of this, important information on drug disposition is obscured. The average concentration curve derived with the NAD approach does not necessarily follow the individual model function. A wrong model may be obtained.[3] Undefined statistical uncertainties and large "unknown" subject variations might smooth the average response curve in an unpredictable manner. Thus, the NAD estimate $\hat{\mu}_{NAD}$ should not, as a general rule, be regarded as a valuable estimate of the expected value of pharmacokinetic parameters. This rule holds even if the true model, that is, the one that adequately describes the individual data, has been used for the fitting. The essential parametric nonlinearity of pharmacokinetic models is responsible for this.

Exceptions to this rule occur when the signal-to-noise ratio is small. This is the case when variability contributes less to the spread in observations than other sources of fluctuation (inter-occasion variability, measurement error, model misspecification). This situation might be seen when concentrations are measured in standardized laboratory animals. The quality of estimates may be improved by using averaging methods other than the straightforward arithmetic mean.[4] These ad hoc solutions do not fundamentally solve the problem. Moreover, no estimate of pure interindividual variability can be obtained with the NAD approach because it masks variability rather than reveals it. Thus, the NAD approach is not a reliable method for pharmacokinetic data analysis.

NAÏVE POOLED DATA ANALYSIS

Sheiner and Beal[5] proposed the naïve pooled data (NPD) approach for the method in which all data from all individuals are considered as arising from one unique individual.

This reference subject is characterized by a set of parameters $\bar{\phi}$. With least-squares fitting, $\bar{\phi}$ will be the parameter vector minimizing the global objective function.

$$O_{NPD}(\phi) = \sum_{j=1}^{N} \{ \sum_{i=1}^{nj} [y_{ij} - f_{ij}(\phi)]^2 \} \qquad \text{Eq. 2}$$

where $\{f_{ij}, i = 1,\ldots,n_j\}$ is the set of components of f_j, and the summation is over all individuals and all measurements for a given individual.

Unlike the NAD approach, the NPD approach is far more general. It can easily deal with experimental data, nonstandard data, and routine pharmacokinetic data. After a unique fitting of all data at once, parameter estimates are obtainable. It may perform well when variations between subjects are small. This is occasionally the case in a group of homogeneous laboratory animals from a given strain, but it is rarely true for humans. The drawbacks of NPD are the same as those of NAD, as has been repeatedly pointed out.[6-8] The NPD approach tends to confound individual differences and diverse sources of variability in a manner different from the NAD approach, but with similar negative consequences. The NPD estimate for the reference individual $\bar{\phi}$ should be considered as a rough approximation ($\hat{\mu}_{NPD}$) of the population expectation μ, although the consequences of the omission can be minor.[9] In addition, estimates of the dispersion of parameters in the population are not provided. Extrapolation of mean outcomes on the basis of the set of estimates $\hat{\mu}_{NPD}$ should be done with caution.

These problems notwithstanding, it has been shown that, for several drugs used in anesthesia, a pooled analysis approach provided population mean parameters that, when prospectively tested, accurately predicted drug concentrations after drug administration by a computer-controlled infusion pump.[10-12] The data in all circumstances originated from well-controlled experiments with extensive sampling. That is, the data were of the EP type (see the installment of articles in this series of tutorials). Moreover, the NPD analysis provided similar population mean parameter estimates compared with estimates obtained using several other population analysis methods.[13,14] These findings are in contrast with an earlier simulation study which showed that the NPD approach provided biased estimates of the population

mean parameters even when a well-balanced experimental study design was used.[6] The discrepancy may be due to the large amount of interindividual variability present or inappropriate weighting scheme used in the latter study.

Imbalance and confounding correlations present in a data set pose serious problems for the NPD approach. These features are prevalent in observational data and make the NPD approach inappropriate for this type of data. Data imbalance occurs when there are many more observations taken from some individuals than others. An example would be a case where 6 samples are taken from some individuals, 4 from some others, and one from others.

When the design of the study correlates with the outcome, confounding correlations occur. That is, the presence or absence of an observation is dependent on the subject's pharmacokinetics. Confounding correlations are usually prevented with randomization. This, however, is not guaranteed with observational data. A case in point would be a pharmacokinetic study in which concentrations fall below the limit of quantitation during the study. Only individuals with the smallest clearance or largest volume of distribution would contribute measurable concentrations toward the end of the study. Biased estimate of the terminal half-life will result and may be wrongly interpreted as an additional phase of the pharmacokinetic profile. Clearly, the NPD approach should not be used in this setting.

THE TWO-STAGE APPROACH

With this approach, individual parameters are estimated in the first stage by separately fitting each subject's data and then, in the second stage, obtaining parameters across individuals, thus obtaining population parameter estimates. The data are summarized in the set $[(\hat{\phi}_j, M_j), j = 1,\ldots, N]$. $\hat{\phi}_j$ Is the p-vector of the parameter estimates and the $p \times p$ symmetric variance–covariance matrix of the corresponding individual estimate. To derive values for population characteristics according to a given strategy, the individual parameter estimates are combined. The salient features of the methods that constitute the two-stage approach are discussed briefly.

Standard Two-Stage Approach

The standard two-stage (STS) approach refers to a well-known and widely used procedure. Population characteristics of each parameter are estimated as the empirical mean (arithmetic or geometric) and variance of the individual estimates $\hat{\phi}_j$ according to the following equations:

$$\hat{\mu}_{STS} = 1/N \sum_{j=1}^{N} \hat{\phi}_j \qquad \text{Eq. 3}$$

$$\hat{\Omega}_{STS} = 1/N \sum_{j=1}^{N} (\hat{\phi}_j - \hat{\mu}_{STS})^2 \qquad \text{Eq. 4}$$

The estimate of the standard deviation ($\hat{s}$) is easily obtained by taking the square root of $\hat{\Omega}$. N — p can be used instead of N in the denominator of the variance estimate.

With the STS approach, estimates of individual parameters are combined as if the set of estimates were a true N-sample from a multivariate distribution. It has been recommended as a very simple and valuable approach for pooling individual estimates of pharmacokinetic parameters derived from experimental pharmacokinetic studies.[15] The advantage of the STS approach is its simplicity, but the validity of its results should not be overemphasized. However, it has been shown from simulation studies that the STS approach tends to overestimate parameter dispersion (the variance–covariance matrix).[6,16]

Global Two-Stage Approach

The $\hat{\phi}$ can be viewed as observations of the individual parameters. The estimate for a subject may be biased and imprecise because of poor experimental design, poor study execution, or a high level of measurement error. The global two-stage (GTS) approach makes extensive use of the matrices $|M_j, j = 1,\ldots, N|$, which reflect the deviations (bias), together with the estimates $|\hat{\phi}_j, j = 1,\ldots, N|$. The expectation E(.) and the variance–covariance Var(.) of each (random) $\hat{\phi}_j$ can be calculated as:

$$E(\hat{\phi}_j) = \mu \quad \text{for } j = 1, \ldots, N \qquad \text{Eq. 5}$$

$$\text{Var}(\hat{\phi}_j) = M_j + \Omega \quad \text{for } j = 1, \ldots, N \qquad \text{Eq. 6}$$

where μ is the true population expectation and Ω is the true population variance–covariance. An extensive description of the method is provided by Steimer et al.[16] The GTS approach provides a maximum likelihood estimate of μ and Ω by an iterative method. It assumes that the estimates of individual parameters are normally distributed around the true parameters with variance Var_j. The population parameters θ are the p components of the vector μ and the $p(p + 1)/2$ independent components of the symmetric matrix Ω. The objective function to be minimized is as follows:

$$O_{GTS}(\mu, \Omega) = \sum_{j=1}^{N} [(\hat{\phi}_j - \mu)^t (M_j + \Omega)^{-1}(\hat{\phi}_j - \mu) + \ln \det(M_j + \Omega)] \qquad \text{Eq. 7}$$

The first term in the right side of Equation 7 is the summation (over individuals) of the weighted squared deviations of individual estimates from the expected value μ. The weighting matrix is dependent on the quality of the estimate through the factor $(M_j + \Omega)^{-1}$. The last term in the equation is the logarithm of the determinant of the $(M_j + \Omega)$ matrix. It prevents the variance–covariance matrix from going to zero through its determinant.

The GTS approach has been shown, through simulation, to provide unbiased estimates of the population mean parameters and their variance–covariances, whereas the estimates of the variances were upwardly biased if the STS approach was used.[16] These simulations were done under the ideal situation that the residual error was normally distributed with a known variance. However, it is a well-known fact that the asymptotic covariance matrix used in the calculations is approximate and, under less ideal conditions, that the approximation can be poor.[17,18]

The Iterative Two-Stage Approach

A computationally "heavier" two-stage method that relies on repeated fittings of individual data, the iterative two-stage (IT2S) approach, has been described.[16,19,20] The IT2S approach can be implemented with rich data, sparse data, or a mixture of both. An approximate a priori population model is required to initiate the procedure. Provided that considerable informative data are available, the population values may be obtained from the literature, the NPD approach performed with the current study data and a reasonable choice of parameter variability, or the STS approach.[16] As the name implies, the IT2S approach is implemented in 2 stages. In the first stage, the population model is used as the set of prior distributions for Bayesian estimation of the individual parameters for all patients, irrespective of the number of samples supplied by each individual.

In the second stage, the population parameters are recalculated with these new individual parameters in order to form the new set of prior distributions. The estimation process (ie, parameters from the second stage are used for a repeat of the first stage and the results are used for a repeat of the second stage) is repeated until the difference between the new and old prior distributions is essentially zero. The method may be implemented with programs supporting Bayesian estimation and least-squares regression or with the IT2S routine,[20] which has been implemented with the USC*PACK collection of programs.[21]

A method close to the IT2S procedure is the expectation-maximization–like (EM) method presented by Mentre and Geomeni.[22] It can be viewed as an extension of IT2S when both random and fixed effects are included in the model and for heteroschedastic errors known to a proportionality coefficient. This algorithm is implemented with the software P-PHARM.[23]

Bayesian Two-Stage Approach

A method that is Bayesian in nature is that proposed by Racine-Poon.[24] The method uses the estimates of the individual parameters ϕ_j and asymptotic variance matrix V_j obtained from the individual fits, with very weak assumptions about the prior distribution of the population parameters to calculate a posterior density function from which ϕ and Ω can be obtained. In an iterative method suggested by Dempster et al.,[25] the EM algorithm is used to calculate the posterior density function. Simulation studies in which several varying and realistic conditions were assumed have shown that the Bayesian two-stage approach provides good estimates of PPK and pharmacodynamic parameters.[24,26]

THE NONLINEAR MIXED-EFFECTS MODEL APPROACH

The first attempt at estimating interindividual pharmacokinetic variability without neglecting the difficulties (eg, data imbalance, sparse data, subject-specific dosing history) associated with data from patients undergoing drug therapy was made by Sheiner et al.[27] using the nonlinear mixed-effects model approach. The vector θ of population characteristics is composed of all quantities of the first 2 moments of the distribution of the parameters: the mean values (fixed effects) and the elements of the variance–covariance matrix that characterize random effects.[5,6,8,28-30]

The number of samples per subject used for this approach is typically small, ranging from 1 to 6. The difficulties associated with this type of data preclude the use of the STS approach because there are not enough data to separately estimate the pharmacokinetic parameters for each subject. There are too few measurements to estimate the parameters accurately or the model may be unidentifiable in a specific individual. As does the pooled analysis technique, nonlinear mixed-effects modeling approaches analyze the data of all individuals at once, but take the interindividual random effects structure into account. This ensures that confounding correlations and imbalance that may occur in observational data are properly accounted for.

Most of the nonlinear mixed-effects modeling methods estimate the parameters by the maximum likelihood approach. The probability of the data under the model is written as a function of the model parameters, and parameter estimates are chosen to maximize this probability. This amounts to asserting that the best parameter estimates are those that render the observed data more probable than they would be under any other set of parameters.

It is difficult to calculate the likelihood of the data for most pharmacokinetic models because of the nonlinear dependence of the observations on the random parameters η_i and possibly ε_{ij}. To deal with these problems, several approximate methods have been proposed. These methods, apart from the approximation, differ widely in their representation of the probability distribution of interindividual random effects.

First-Order (NONMEM)

The first nonlinear mixed-effects modeling program introduced for the analysis of large amounts of pharmacokinetic data was NONMEM.[31] In the NONMEM program, linearization of the model in the random effects is effected by using the first-order (FO) Taylor series expansion with respect to the random effect variables η_i and ε_{ij}. This software is the only program in which this type of linearization is used.

The *j*th measurement in the *i*th subject of the population can be obtained from a variant of Equation 5 in the first tutorial as follows:

$$y_{ij} = f(\phi, x_{ij}, \eta_i) + \varepsilon_{ij} \qquad \text{Eq. 8}$$

The FO Taylor series expansion of the above model with respect to the random variables η_i (intersubject variability) and ε_{ij} (residual variability) around zero is given by

$$y_{ij} = f(\phi, x_{ij}) + G_{ij}(\phi, x_{ij})\eta_i + \varepsilon_{ij} \qquad \text{Eq. 9}$$

where

$$G_{ij}(\phi, x_{ij}) = \delta f(\theta, x_{ij}, \eta_i, \varepsilon_{ij})/\delta \eta_i^T \Big|_{\eta = 0} \qquad \text{Eq. 10}$$

$G_{ij}(\phi, x_{ij})$ is $1 \times p$ matrix of the first derivatives of $f(\theta, x_{ij}, \eta_i, \varepsilon_{ij})$ with respect to η_i, evaluated at η_i equals zero. In Equation 9, the model is linear in ε_{ij}; therefore, no ap-

proximation is made with respect to ε_{ij}. Logarithmic transformation of the data can be done to ensure linearity in ε_{ij}.

The random effect parameters η_i and ε_{ij} are independent (multivariate), normally distributed with zero means and variances Ω and σ^2, respectively. Ω Is the p × p covariance matrix of the p vector η_i. Based on the fact that η_i and ε_{ij} are independent and identically normally distributed, and the linearization of Equation 9, the expectation and variance–covariance of all observations for the *i*th individual (first 2 moments) are given by:

$$E_i = f(\theta,x_i) \qquad \text{Eq. 11}$$

and

$$C_i = G_i((\theta,x_i)\Omega G_i(\theta,x_i)^T + \sigma^2 I_{ni} \qquad \text{Eq. 12}$$

where $f(\theta,x_i)$ is the vector of model predictions of y_i, $G_i((\theta,x_i)$ represents the $n_i \times p$ matrix of first derivatives of $f(\theta,x_i,\eta_i,\varepsilon_i)$ with respect to η_i evaluated at η_i equals zero, and I_{ni} represents the identity matrix of size n_i. Maximum estimates of the population parameters θ, Ω, and σ^2 can be obtained by minimizing minus twice the logarithm of population likelihood as expressed below:

$$-2LL = \sum_{i=1}^{N} (\log(\det(C_i)) + (y_i - E_i)^T C_i^{-1}(y_i - E_i)) \qquad \text{Eq. 13}$$

This approach is called the FO method in NONMEM. This is the most widely used approach in PPK and pharmacodynamic data analysis and has been evaluated by simulation. The use of the FO Taylor series expansion to approximate the nonlinear model in η_i and possibly ε_{ij} by a linear model in these parameters is the greatest limitation of the FO approach.

The performance of the FO approach for the analysis of observational and experimental data has been evaluated by Sheiner and Beal with the Michaelis–Menten pharmacokinetic model[5] and the 1- and 2-compartment models.[6,7] In all instances, a comparison was made with the NPD and STS approaches for the analysis of the 2 types of data. The FO approach outperformed the NPD and the STS approaches on both data types. Despite the approximation, the FO approach provides good parameter estimates. When the residual error increases, the STS approach quickly deteriorates, especially with respect to variance parameters. However, the STS approach still performs reasonably well, but the bias and imprecision of the estimates tend to increase with increasing residual error.[7] Estimates of residual random effects have been shown to deteriorate with the FO approach when residual error increases.[32]

Deterioration in parameter estimation has been observed in simulation studies in which the value of the intersubject variability was >60% and the residual variability was set at 15%.[33] A series of studies in which observations were randomly deleted from a data-rich set to create a sparse data set and parameter estimation was done using the FO method showed good performance of the FO approach compared with the results obtained using the full data set.[34-38] The correspondence of the results in the 2 situations suggests that the FO approach can be used to estimate parameters using only a few observations per individual. Simulation studies have been performed to show that the FO approach can be used in the limiting case where only one sample is obtained per subject.[39] In this case, there is an upper limit of residual variability (not exceeding 20%) for the production of reliable parameter estimates.

The impact of the linearization approximation of the FO approach for a simple 1-compartment model was evaluated by Beal.[29] He compared the performance of this approach with the exact solution to the population likelihood. No difference was observed, which indicated that the approximation used in the FO method is not detrimental to the analysis under the conditions evaluated, which included an interindividual variability set at 25% (CV%). Other simulation studies, however, have shown that the FO approach has a potential for providing modestly biased estimates.[6,18,28,34,40-43]

For a 1-compartment multidose scenario, White et al.[18] showed that biased estimates are more likely when residual and intersubject variability are very high. Ette et al.[33] observed that the biased estimates are obtained at high levels of intersubject variability with a 2-compartment multidose situation, although the residual variability did not exceed 15%. The bias may be due to the fact that the FO Taylor series expansion is not a particularly good approximation of the underlying "real" (log-normal) distribution used to generate the simulated data in these studies. Also, it may be that the FO Taylor series expansion is evaluated at η_i equals zero (the population mean estimate of η_i). This may not be a good approximation depending on the magnitude of intersubject variability and the nonlinearity of the pharmacokinetic model. During data analysis, this can be compensated for, in part, by including explanatory covariates in the model to reduce the variance of η_i. With a 1-compartment model experimental data set, the GTS approach was shown to outperform the FO approach with respect to bias and precision of both the population mean and variance estimates. Similar results were obtained in a study in which the FO approach was compared with the Bayesian two-stage approach.[40]

The NONMEM program implements 2 alternative estimation methods: the FO conditional estimation (FOCE) and the Laplacian methods.[31] The FOCE method uses an FO expansion about conditional estimates (empirical Bayesian estimates) of the interindividual random effects rather than about zero.[44] In this respect, it is like the conditional FO method of Lindstrom and Bates.[45] Unlike the latter, which is iterative, a single objective function is minimized, achieving a similar effect as with iteration. The Laplacian method uses second-order expansions about the conditional estimates of the random effects.[44]

Conditional First-Order (NLME)

The conditional FO method of Lindstrom and Bates[45] uses an FO Taylor series expansion about conditional estimates of interindividual random effects. Estimation involves an iterative generalized least-squares type algorithm. This estimation method is available in S-PLUS as the function NLME.[46]

Alternative First-Order (MIXNLIN)

This method, proposed by Vonesh and Carter,[47] also uses an FO series expansion of the interindividual random effects. They proposed the use of estimated generalized least squares and established the asymptotic properties of the resulting estimates. An alternative method is the use of the iteratively reweighted generalized least squares.[48] The MIXNLIN program also implements pseudo maximum likelihood (ML) and restricted maximum likelihood (REML) estimation by embedding the EM algorithm within an iteratively reweighted generalized least-squares routine. Expansion is either about zero or about the empirical best linear unbiased predictor (EBLUP) of the interindividual random effects. Only the fixed-effects and variance component estimates are updated after each call to the embedded EM algorithm (ie, the method uses the EBLUP estimates inherent within the EM algorithm only to update estimates of the variance components) when the expansion is about zero. ML estimation expanded about zero should result in estimates similar to those obtained using the NONMEM FO method, while expansion about the EBLUP should result in estimates similar to those obtained with the FOCE in NONMEM and the FO conditional method (NLME). These estimation methods are available in the SAS macro and MIXNLIN 3.0 version of Vonesh.[48]

Alternative First-Order (SAS)

This is an FO Taylor series expansion method, but the algorithm consists of iteratively fitting a set of generalized estimating equations until they stabilize.[49] The method uses a Taylor series expansion in the fixed-effects parameters, as well as one in the random effects; expansion is about the generalized least-squares estimates for the fixed-effects parameters and about zero for the random effects. It yields estimates similar to those obtained using the FO method of NONMEM. The method is implemented in the SAS macro NLINMIX. The NLINMIX program also implements expansion about the EBLUPs of the interindividual random effects as an alternative to expansion about zero, yielding estimates similar to those produced with the FOCE method in NONMEM.

Nonparametric Maximum Likelihood (NPML)

The NPML approach provides an estimate of the whole probability distribution of the pharmacokinetic parameters on a nonparametric basis.[50] The method relies on maximization of the likelihood of the set of observations of all individuals to estimate the distribution of the parameters. The basic conceptual framework is similar to that described above for NONMEM. The difference is that no specific model for the relationship between pharmacokinetic parameters and patient-specific covariates is specified. The individual parameters ϕ_i are assumed to be independent realizations of a given random variable Φ with probability distribution $F(\phi)$. The likelihood of all data is given by:

$$L(F) = \prod_{i=1}^{N} \int_D l_i(y_i|\phi) F(\phi) d\phi \qquad \text{Eq. 14}$$

where $l_i(y_i|\phi)$ is the likelihood of the observations y_i for ith individual, given ϕ. D is the domain in which the parameters lie. Maximization of this likelihood provides an estimate ^F of the probability distribution of the parameters. This distribution has been proven by Mallet to be discrete, involving N_p locations, where N_p is less than or equal to the number of individuals (N). To estimate the N_p locations q_k and their corresponding frequencies α_k, a specific algorithm was developed. The level of residual error and how well the parameters are known determines the number of locations. There will be N locations, each with a frequency of 1/N if the parameters are known very precisely for all N subjects. The set of locations q_k and frequencies α_k completely specify the estimate of the distribution of the parameters:

$$\hat{F} = \sum_{k=1}^{N_p} \alpha_k \bullet \delta(q_k) \qquad \text{Eq. 15}$$

where $\delta(x)$ denotes the Dirac probability distribution, which takes the value 1 at x and 0 elsewhere. With this method, a complete distribution of F with very soft assumptions, namely that F takes only positive values and that its integral over D domain, is equal to unity.[50,51] The NPML approach has been shown in a simulation study assuming a 1-compartment pharmacokinetic model with bimodal distribution to produce parameter estimates that accurately describe the distribution, even though only one measurement was available per individual.[52] Several summary statistics, such as mean or variance–covariance matrix, can easily be calculated from the distribution of F specified by Equation 15.

The method also allows for the inclusion of patient-specific covariates without the specifying a priori relationship between the pharmacokinetic parameters and covariates. The covariates are regarded as additional parameters, and the algorithm provides an estimate of the joint distribution of the pharmacokinetic parameters and the covariates.[53] The probability distribution of the parameters conditional on any value of the covariates can be computed and used for the initial dosage selection, given the distribution obtained. Thus, the shape of the relationship between parameters and covariates can be explored nonparametrically.

The major limitation of this approach is that the residual error must be known a priori. The method, therefore, is nonparametric with respect to the interindividual random effects, but requires the intraindividual error to be specified a priori. Pharmacokinetic analyses performed with the NPML approach and reported in the literature have used the residual error model based on drug concentration measurement assay variance.[52-55] This seems to be unrealistic. Intraindividual variability, inter-occasion variability, and model misspecification often will contribute significantly to the residual error.[56,57]

Also, the estimator of the distribution produced by the NPML approach is a point estimator, and no results on the accuracy of the estimation are obtained. Consequently, care should be taken in interpreting the results, especially when they are obtained from a small sample size. If the NPML approach is used primarily for exploratory analysis to improve the efficiency of subsequent parametric analysis, this may not be much of a problem. The NPML approach is a computationally expensive approach, which may limit the practicality of the approach when the dimension of the parameter space increases. An example of this would be the case of a complex pharmacokinetic model with numerous covariates.

The nonparametric expectation–maximization (NPEM) program of Schumitzky,[58] which is similar to the NPML program of Mallet,[50] computes the nonparametric ML using the nonparametric EM algorithm. NPEM has been developed as a segment of the USC*PACK collection of programs.[21] The results obtained using NPEM for PPK data analysis are similar to those of the NPML program. NPEM and STS give virtually identical estimates of PPK parameters in the same population when the results of NPEM indicate normal distribution for parameter estimates.[59,60]

Seminonparametric Maximum Likelihood (SNP)

Davidian and Gallant[61] introduced the SNP maximum likelihood from econometrics into pharmacokinetics. Like the NPML approach, the SNP approach provides an estimate of the entire distribution of the interindividual random effects. The SNP approach maximizes the likelihood over a class of distributions restricted to have a smooth density instead of maximizing the likelihood over all distribution functions, as does the NPML method. This assumption of smoothness is flexible enough to allow heavy-tailed, multimodal, and skewed distributions to be characterized, but prevents kinks, jumps, and oscillatory behavior.[62] Also, this method relies on maximizing the likelihood of the set of observations of all individuals to estimate the distribution of the random effects. The basic conceptual framework remains the same as that described for the population model in the "Models" subsection of the first installment in this series of tutorials.[1] The representation of the probability distribution and calculation of the likelihood are different from the NONMEM and NPML approaches. It has been shown by Gallant and Nychka[63] that the smooth distribution can be presented as an infinite series expansion, and they provide a full mathematical description. The SNP approach uses a finite number of leading terms resulting from an approximation of the infinite expansion. A single tuning parameter determines the number of terms retained. The density is multivariate normal if the value of this tuning parameter equals zero. The distribution becomes more flexible the larger the value of the tuning parameter. An important step in the modeling procedure is the selection of an appropriate value of this tuning parameter.[61] The density of the random effect parameters is represented by a multivariate normal distribution multiplied by a polynomial. The SNP approach computes the integral present in the population likelihood by quadrature. This is another useful feature of this approach. This obviates the use of the linearization approximation to the likelihood used in the NONMEM approach. Unlike the NPML approach, standard errors can be computed for the model parameters and used for inference.

The SNP approach is implemented in a public domain FORTRAN program called NLMIX. Experience with this approach is still very limited, and only a few simulations have evaluated the ability of the method to reveal multiple modes in the random effects density under conditions likely to be encountered in practice.

A method similar to the SNP approach was proposed by Fattinger et al.[64] to explore the complete distribution of interindividual effects using the FOCE approach in the NONMEM program. The method uses a monotone nondecreasing spline to transform the normally distributed interindividual random effects. The model for the interindividual random effect model is given as:

$$\phi_i = g(\theta, x_i) + sp(\eta_i) \qquad \text{Eq. 16}$$

where $sp(..)$ represents a monotone non-decreasing spline of which the parameters are estimated. Because splines are not multivariate, a different spline is used for each of the elements of η_i. The spline function transformation is very flexible and allows appropriate representations of skewed, heavily tailed, or multimodal distributions.

Summary

Thus far, the principles that serve as the foundation and the methods for PPK model estimation have been presented. These concepts are important so that the application of PPK will be executed in an informed manner. The current article serves as a bridge to the final PPK tutorial paper, which will address application of PPK modeling with informative examples.

Ene I Ette MSc PhD FCP FCCP, Senior Director of Clinical Pharmacology, Vertex Pharmaceuticals, Inc., Cambridge, MA

Paul J Williams PharmD MS FCP FCCP, Professor of Pharmacy, Department of Pharmacy Practice, School of Pharmacy, University of the Pacific; Trials by Design, LLC, Stockton, CA

Reprints: Ene I Ette MSc PhD FCP FCCP, Vertex Pharmaceuticals, Inc., 130 Waverly St., Cambridge, MA 02139-4242, fax 617/444-6713, Ene_Ette@vpharm.com

References

1. Ette EI, Williams PJ. Population pharmacokinetics I: background, concepts, and models. Ann Pharmacother 2004;38:1702-6. DOI 10.1345/aph.1D374
2. Levy G, Hollister LE. Inter- and intrasubject variations in drug absorption kinetics. J Pharm Sci 1964;53:1446-52.
3. Martin E, Moll W, Schmid P, Dettli L. Problems and pitfalls in estimating average pharmacokinetic parameters. Eur J Clin Pharmacol 1984;26: 595-602.
4. Cochetto DM, Wargin WA, Crow JW. Pitfalls and valid approaches to pharmacokinetic analysis of mean concentration data following intravenous administration. J Pharmacokinet Biopharm 1980;8:539-52.
5. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michealis–Menten model: routine clinical data. J Pharmacokinet Biopharm 1980;8:553-71.
6. Sheiner LB, Beal SL. Evaluation of methods for estimating population

pharmacokinetic parameters. I. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm 1981;9:635-51.
7. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. I. Monoexponential model and routine clinical data. J Pharmacokinet Biopharm 1983;11:303-19.
8. Sheiner LB, Beal SL. Estimation of pooled pharmacokinetic parameters describing populations. In: Endrenyi L, ed. Kinetic data analysis. New York: Plenum Press, 1981:271-84.
9. Fluhler H, Huber H, Widmer E, Brechbuhler S. Experiences in the application of NONMEM to pharmacokinetic data analysis. Drug Metab Rev 1984;15:317-39.
10. Dyck JB, Haack DL, Azarnoff L, Vuorilehto L, Shafer SL. Computer-controlled infusion of intravenous dexmedetomide hydrochloride in adult human volunteers. Anesthesiology 1993;78:821-8.
11. Shafer SL, Varvel JR, Aziz N, Scott JC. Pharmacokinetics of fentanyl administered by computer-controlled infusion pump. Anesthesiology 1990;73:1091-102.
12. Gustafsson LL, Ebling WF, Osaki E, Harapat S, Stanski DR, Shafer SL. Plasma concentration clamping in the rat using a computer-controlled infusion pump. Pharm Res 1992;9:800-7.
13. Kataria BK, Ved SA, Nicodemus HF, Lea D, Dubios MY, Mandema JW, et al. The pharmacokinetics of propofol in children using three different data analysis approaches. Anesthesiology 1994;80:104-22.
14. Egan TD, Lemmens HJ, Fiset P, Hermann DJ, Muir KT, Stanski DR, et al. The pharmacokinetics of the new short-acting opioid remifentanil (GI87084B) in healthy adult male volunteers. Anesthesiology 1993;79: 881-92.
15. Rodda BE. Analysis of sets of estimates from pharmacokinetic studies. In: Endrenyi L, ed. Kinetic data analysis. New York: Plenum Press, 1981:285-97.
16. Steimer JL, Mallet A, Golmard JL, Boisvieux JF. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with nonlinear mixed effects model. Drug Metab Rev 1984;15:265-92.
17. Sheiner LB, Beal SL. A note on confidence intervals with extended least squares parameter estimates. J Pharmacokinet Biopharm 1987;15:93-8.
18. White DB, Walawander CA, Tung Y, Grasela TH. An evaluation of point and interval estimates in population pharmacokinetics using NONMEM analysis. J Pharmacokinet Biopharm 1991;19:87-112.
19. Prevost G. Estimation of a normal probability density function from samples measured with non-negligible and non-constant dispersion, Internal Report 6-77, Anders-Gerbios, F-91120 Palaiseau, France, 1977.
20. Forrest A, Ballow CH, Nix DE, Birmingham MC, Schentag JJ. Development of a population pharmacokinetic model and optimal sampling strategy for intravenous ciprofloxacin in seriously ill patients. Antimicrob Agents Chemother 1993;37:1065-72.
21. Jellife RW, Schumitzky A, Van Guillder M. User manual for the nonparametric EM program for population modeling, version 2.17. Los Angeles: Laboratory for Applied Pharmacokinetics, USC School of Medicine, December 15, 1993.
22. Mentre F, Geomeni R. A two-step iterative algorithm for estimation in nonlinear mixed-effect models with an evaluation in population pharmacokinetics. J Biopharm Stat 1995;5:141-58.
23. P-PHARM user's guide, version 1.3, Créteil, France: SIMED, 1994.
24. Racine-Poon A. A Bayesian approach to nonlinear random effects models. Biometrics 1985;41:1015-23.
25. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via EM algorithm. J Roy Stat Soc B 1977;39:1-38.
26. Racine A, Grieve AP, Fluhler H, Smith AFM. Bayesian methods in practice: experiments in the pharmaceutical industry. Appl Stat 1986;35:93-100.
27. Sheiner LB, Rosenberg B, Melmon KL. Modeling of individual pharmacokinetics for computer-aided drug dosage. Comp Biomed Res 1972;5: 441-59.
28. Sheiner LB. The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. Drug Metab Rev 1984;15: 153-71.
29. Beal SL. Population pharmacokinetic data and parameter estimation based on their first two statistical moments. Drug Metab Rev 1984;15: 173-93.
30. Whiting B, Kelman AW, Grevel J. Population pharmacokinetics: theory and clinical application. Clin Pharmacokinet 1986;11:387-401.
31. Sheiner LB, Beal SL. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. J Pharm Sci 1982;71:1344-8.
32. Ette EI, Kelman AW, Howie CA, Whiting B. Analysis of animal pharmacokinetic data: performance of the one point per animal design. J Pharmacokinet Biopharm 1995;23:551-66.
33. Ette EI, Sun H, Ludden TM. Balanced designs and longitudinal population pharmacokinetic studies. J Clin Pharmacol 1998;38:417-23.
34. Grasela TH Jr, Antal EJ, Townsend RJ, Smith RB. An evaluation of population pharmacokinetics in therapeutic trials. Part I. Comparison of methodologies. Clin Pharmacol Ther 1986;39:605-12.
35. Collart L, Blashke TF, Boucher F, Prober CG. Potential of population pharmacokinetics to reduce the frequency of blood sampling required for estimating kinetic parameters in neonates. Dev Pharmacol Ther 1992;18: 71-80.
36. Aarons L, Mandema JW, Danhof M. A population analysis of the pharmacokinetics and pharmacodynamics of midazolam in the rat. J Pharmacokinet Biopharm 1991;19:485-96.
37. Kaniwa N, Aoyagi N, Ogata H, Ishi M. Application of the NONMEM method to evaluation of the bioavailability of drug products. J Pharm Sci 1990;79:1116-20.
38. Pai SM, Shukla UA, Grasela TH, Knupp CA, Dolin R, Valentine FT, et al. Population pharmacokinetic analysis of didanosine (2',3'-dideoxyinosine) plasma concentration obtained in phase I clinical trials in patients with AIDS or AIDS-related complex. J Clin Pharmacol 1992;32:242-7.
39. Jones CD, Sun H, Ette EI. Designing cross-sectional population pharmacokinetic studies: implications for pediatric and animal studies. Clin Res Reg Affairs 1996;13:133-65.
40. Racine A, Grieve AP, Fluhler H, Smith AFM. Bayesian methods in practice: experiences in pharmaceutical industry. Appl Stat 1986;35:1-38.
41. Ette EI, Kelman AW, Howie CA, Whiting B. Influence of interanimal variability on the estimation of population pharmacokinetic parameters in preclinical studies. Clin Res Reg Affairs 1994;11:121-39.
42. Ette EI, Howie CA, Kelman AW, Whiting B. Experimental design and efficient parameter estimation in preclinical pharmacokinetic studies. Pharm Res 1995;12:729-37.
43. Karlsson MO, Sheiner LB. The importance of modeling interoccasion variability in population pharmacokinetic analyses. J Pharmacokinet Biopharm 1993;21:735-50.
44. Beal SL, Sheiner LB. NONMEM users guide—part VII. Conditional estimation methods. San Francisco: University of California, 1992.
45. Lindstrom MJ, Bates DM. Nonlinear mixed effects models for repeated measures data. Biometrics 1990;46:673-87.
46. S-PLUS. Seattle: Insightful, 2002.
47. Vonesh EF, Carter RL. Mixed-effects nonlinear regression for unbalanced repeated measures. Biometrics 1992;46:673-87.
48. Vonesh EF. Nonlinear models for the analysis of longitudinal data. Stat Med 1992;11:1929-54.
49. Wolfinger RD. Laplace's approximation for nonlinear mixed models. Biometrika 1993;80:791-5.
50. Mallet A. A maximum likelihood estimation method for random coefficient regression models. Biometrika 1986;73:645-56.
51. Steimer JL, Mallet A, Mentre F. Estimating interindividual pharmacokinetic variability. In: Rowland M, Sheiner L, Steimer JL, eds. Variability in drug therapy: description, estimation, and control. New York: Raven Press, 1985:65-111.
52. Mentre F, Mallet A. Experiences with NPML—application to dosage individualisation of cyclosporine, gentamicin and zidovudine. In: Rowland M, Aarons L, eds. The population approach. Luxembourg: Commission of European Communities, 1992:75-90.
53. Mallet A, Mentre F, Gilles J, Kelman AW, Thomson AN, Bryson SM, et al. Handling covariates in population pharmacokinetics with an application to gentamicin. Biomed Meas Inform Contr 1988;2:673-83.
54. Mallet A, Mentre F, Steimer JL, Lokiec F. Nonparametric maximum likelihood estimation for population pharmacokinetics, with application to cyclosporine. J Pharmacokinet Biopharm 1988;16:529-56.
55. Mentre F, Escolano S, Diquet B, Golmard JL, Mallet A. Clinical pharmacokinetics of zidovudine: inter and intraindividual variability and relationship to long term efficacy and toxicity. Eur J Clin Pharmacol 1993;45(5):397-407.
56. Sidhu JS, Ashton M, Huong NV, Hai TN, Karlsson MO, Sy ND, et al. Artemisinin population pharmacokinetics in children and adults with uncomplicated falciparum malaria. Br J Clin Pharmacol 1998;45:347-54.
57. Hossain M, Wright E, Baweja R, Ludden TM, Miller R. Nonlinear mixed effects modeling of single dose and multiple dose data for an immediate release (IR) and controlled release (CR) dosage form of alprazo-

lam. Pharm Res 1997;14:309-15.
58. Schumitzky A. Nonparametric EM algorithms for estimating prior distributions. Appl Math Comput 1991;45:141-57.
59. Dodge WF, Jellife RW, Richardson CJ, McCleery RA, Hokanson JA, Snodgrass WR. Gentamicin population pharmacokinetic models for low birth weight infants using a new nonparamaetric method. Clin Pharmacol Ther 1991;50:25-31.
60. Kisor D, Watling S, Zarowitz B, Jelliffe RW. Population pharmacokinetics of gentamicin in patients with indicators of malnutrition: the use of a new nonparametric expectation maximization (NPEM) algorithm. Clin Pharmacokinet 1992;23:62-8.
61. Davidian M, Gallant AR. Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. J Pharmacokinet Biopharm 1992;20:529-56.
62. Davidian M, Gallant AR. The nonlinear mixed effects model with a smooth random effects density, Institute of Statistics Mimeo Series No. 2206. Raleigh, NC: North Carolina State University, 1992.
63. Gallant AR, Nychka DW. Semi-nonparametric maximum likelihood estimation. Econometrica 1987;55:363-90.
64. Fattinger KE, Sheiner LB, Verotta D. A new method to explore the distribution of interindividual random effects in nonlinear mixed effects models. Biometrics 1995;51:1236-51.

RÉSUMÉ

OBJECTIF: Présenter, comparer, et mettre en relief les diverses approches permettant l'estimation des paramètres pharmacocinétiques populationnels en ce qui concerne les bases mathématiques, les aspects statistiques, les logiciels de mise en œuvre, et les revendications sous-jacentes.

MÉTHODE: L'information sur la pharmacocinétique de population a été repérée par une recherche sur MEDLINE (janvier 1979 à juin 2002) et une bibliographie d'articles de revues et d'ouvrages. Cette information est utilisée conjointement avec l'expérience pour expliquer les diverses approches méthodologiques de la pharmacocinétique de population.

RÉSUMÉ: Différentes méthodes d'estimation sont examinées, avec leurs avantages et leurs limites, depuis les approches par moyennage de données brutes et par analyse de données brutes poolées jusqu'aux approches de modèles non linéaires à effets mixtes en passant par les approches en 2 étapes.

CONCLUSIONS: Les méthodes d'estimation de pharmacocinétique de population qui reposent sur la caractérisation d'effets mixtes (fixes et aléatoires) sont connues pour produire des estimations des paramètres pharmacocinétiques d'une population moins biaisées que celles obtenues avec les approches utilisant les données brutes ou avec l'approche standard en 2 étapes. Le logiciel NONMEM est le plus largement employé pour la caractérisation des pharmacocinétiques de population.

Bruno Edouard