ACCP
American College of Clinical Pharmacology
Advancing Clinical Care through Pharmacology

Review

Parametric and Nonparametric Methods in Population Pharmacokinetics: Experts' Discussion on Use, Strengths, and Limitations

The Journal of Clinical Pharmacology
2022, 62(2) 158–170

DOI: 10.1002/jcph.1993

Sylvain Goutelle, PharmD, PhD[1,2,3], Jean-Baptiste Woillard, PharmD, PhD[4,5,6], Thierry Buclin, MD[7], Laurent Bourguignon, PharmD, PhD[1,2,3], Walter Yamada, PhD[8,9], Chantal Csajka, PharmD, PhD[10,11], Michael Neely, MD[8,9], and Monia Guidi, PhD[7,10]

Abstract

Population pharmacokinetics consists of analyzing pharmacokinetic (PK) data collected in groups of individuals. Population PK is widely used to guide drug development and to inform dose adjustment via therapeutic drug monitoring and model-informed precision dosing. There are 2 main types of population PK methods: parametric (P) and nonparametric (NP). The characteristics of P and NP population methods have been previously reviewed. The aim of this article is to answer some frequently asked questions that are often raised by scholars, clinicians, and researchers about P and NP population PK methods. The strengths and limitations of both approaches are explained, and the characteristics of the main software programs are presented. We also review the results of studies that compared the results of both approaches in the analysis of real data. This opinion article may be informative for potential users of population methods in PK and guide them in the selection and use of those tools. It also provides insights on future research in this area.

Keywords

MIDD (model-informed drug development), modeling and simulation, pharmacometrics, population pharmacokinetics, therapeutic drug monitoring

In 2 separate articles, we have reviewed the characteristics of parametric (P) and nonparametric (NP) approaches in population pharmacokinetics (popPK).[1,2] The aim of this third article is to answer some frequently asked questions that are raised by scholars, clinicians, and researchers interested in popPK and model-informed precision dosing (MIPD). The comparative strengths and limitations of both approaches will be highlighted to inform potential users and guide them in the use of those tools.

Briefly, both P and NP methods share in common their usual reliance on compartmental models to describe the pharmacokinetics (PK) of a drug. In both approaches, the structural model, that is, the group

[1]Hospices Civils de Lyon, Groupement Hospitalier Nord, Service de Pharmacie, Lyon, France
[2]CNRS, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, Villeurbanne, France
[3]Univ Lyon, Université Claude Bernard Lyon 1, ISPB—Faculté de Pharmacie de Lyon, Lyon, France
[4]Univ. Limoges, IPPRITT, Limoges, France
[5]INSERM, IPPRITT, U1248, Limoges, France
[6]Department of Pharmacology and Toxicology, CHU Limoges, Limoges, France
[7]Service of Clinical Pharmacology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[8]Keck School of Medicine, University of Southern California, Los Angeles, California, USA
[9]Laboratory of Applied Pharmacokinetics and Bioinformatics at the Saban Research Institute, Children's Hospital Los Angeles, Los Angeles, California, USA
[10]Center for Research and Innovation in Clinical Pharmaceutical Sciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[11]Institute of Pharmaceutical Sciences of Western Switzerland, University of GenevaUniversity of Lausanne, Geneva, Lausanne, Switzerland

Submitted for publication 14 October 2021; accepted 25 October 2021.

Corresponding Author:
Sylvain Goutelle, PharmD, PhD, Hospices Civils de Lyon, Groupement Hospitalier Nord, Hôpital Pierre Garraud, Service de Pharmacie, 136 rue du Commandant Charcot, 69005, Lyon, France.
Email: sylvain.goutelle@chu-lyon.fr

of equations describing the relationship between drug dosage and drug concentration as a function of time, is assumed to be similar across all individuals. Differences in the concentration-time profiles among subjects receiving the same drug at the same dosage regimen, that is, interindividual variability (IIV), are assumed to arise from each individual's specific set of PK parameters. A major goal of a popPK analysis is to estimate the variability of those PK parameters by estimating their distribution within the population of subjects.

The P and NP approaches differ mainly in the statistical properties of the distribution of population PK parameters. A formal statistical distribution (eg, Gaussian, log-normal) is assumed in P methods. Most of the time, the variability is described as a mixed-effects model, with a distinction between fixed effects (average or typical PK value in the population) and random effects (IIV around the average).[3] In contrast, NP methods consider a discrete distribution, without specific shape. There is no mixed-effects model for the PK parameters, and the entire discrete joint distribution is computed as a set of discrete "support points," with each point comprising values for every parameter in the model and an associated probability of those values based on their "ability" to predict the observed data.

Modern P and NP popPK approaches use maximum likelihood principles to estimate the population PK parameters. Both approaches involve a residual error term capturing how the model individual predictions deviate from the actual observations.

General Questions on Both Approaches

How "Different" Are Results Obtained for a Data Set When Using Parametric vs Nonparametric Analysis? Are the Results Substantially Different?

The main difference lies in the way that PK parameters are reported.

As mentioned previously, P methods provide separate estimates of a fixed effect (typical value, often denoted as θ) and a random effect (a measure of variability, often denoted as η) for each PK parameter included in the structural model. This means that 2 values should be available and examined in the results for each parameter (eg, θ_{CL} and η_{CL} as the fixed and random population values, respectively, for the clearance [CL] parameter). The P methods also provide a standard error (SE) or relative standard error (RSE) for both the fixed and random effects, which quantifies parameter uncertainty in the estimation. Sometimes, the modeler may decide not to estimate variability for a given parameter by fixing $\eta = 0$, for example, when the SE of the corresponding random effect is too high. In such a case, only a fixed effect is reported. The coefficients describing the residual error are also provided along with their confidence interval.

With P methods, the estimates of fixed and random effects strongly depend on prior assumptions about the statistical distribution of the popPK parameters (eg normal, log-normal, mixture of log-normals). Such assumptions cannot be directly verified from the concentration data. However, different distributions can be compared in their ability to fit the data.

It should be noted that the assumed distribution is not explicitly provided in the output of a run with P methods. Also, the estimates of random effects from a P method should not be confused with the variability of the PK parameter themselves.

In P methods, random effects are assumed to have a Gaussian distribution: $\eta \sim N(0,\omega^2)$. Programs such as NONMEM (ICON, Dublin, Ireland) or Monolix (Lixoft, Antony, France) provide the standard deviation (SD; ω) or the variance (ω^2) of random effects. Under the assumption of a log-normal distribution, population variance and SD of the clearance, for example, should be derived as follows:

$$\mathrm{Var(CL)} = \theta_{CL}^{2*}\omega_{CL}^2$$

$$\mathrm{SD\,(CL)} = \theta_{CL} * \omega_{CL}$$

The IIV on CL is often estimated by its coefficient of variation (CV%), that is, the ratio between the population estimate and its SD—ω_{CL} in the case of $\omega_{CL} \leq 0.1$. The P algorithms can compute the covariance between random effects; however, it is the modeler's decision to include no covariance or covariance between some or all PK parameters. This decision can influence the model fit and parameter estimates.

In contrast, NP methods do not provide values of θ and ω with their SE but, instead, provide a collection of support points, as previously mentioned. This collection can be viewed as a matrix of population parameter values. Each line of the matrix is a support point of the discrete distribution, that is, a set of PK parameter values associated with a given probability in the population. A statistical summary of this matrix is provided with mean, median, and variability measures. However, it is the matrix that fully describes the population distribution, not the statistical moments or summaries directly estimated by the P methods. The NP outputs also include the estimate of the residual error SD, which includes an additive or multiplicative coefficient that inflates the fixed assay error polynomial model in NPAG/Pmetrics (Laboratory of Applied Pharmacokinetics and Bioinformatics, USC, Los Angeles, California). By default, no parameter SE/RSE is provided, since the algorithms estimate support points, not means and variances. However, SE around support

Table 1. Population PK Parameter Estimates With a Parametric Method

Parameter	Value	Relative Standard Error (%)
θ_V	17.63	6.32
θ_{CL}	14.27	5.68
ω_V	0.061	108.68
ω_{CL}	0.115	27.78
b	0.312	9.24

CL, clearance; PK, pharmacokinetic; V, volume.
θ_V and θ_{CL} are the population volume (in L) and CL (in L/h), respectively.
ω_V and ω_{CL} are the standard deviations of the random effects η_V and η_{CL}
b is the coefficient of the residual error model.
The standard deviation of V and CL can be derived by calculation:
$SD(V) = \theta_V{*}\omega_V = 1.08$ L and $SD(CL) = \theta_{CL}{*}\omega_{CL} = 1.64$ L/h.
The data set included 84 plasma concentrations of cloxacillin from 11 patients.[4] This was a preliminary run with a 1-compartment model parameterized with V and CL and a proportional residual error model. Log-normal distribution was assumed, without covariance. The Monolix software was used.

Table 2. Population PK Parameters Estimates With a Nonparametric Method

Support Point			
Support Point Number	CL (L/h)	V (L)	Probability
1	11.185	16.296	0.182
2	22.003	23.852	0.104
3	18.689	20.455	0.186
4	23.367	39.667	0.088
5	15.181	19.635	0.107
6	18.494	15.066	0.086
7	13.719	10.673	0.098
8	12.939	13.074	0.102
9	12.939	13.016	0.047
Population Parameter Value Summaries			
	CL (L/h)	V (L)	
Mean	16.34	19.12	
SD	4.06	7.43	
CV%	24.83	38.86	
Variance	16.47	55.19	
Median	15.18	16.30	
Covariance and correlation			
Covariance	23.15		
Correlation	0.768		
Residual error			
Gamma	1.15		

CL, clearance; CV, coefficient of variation; PK, pharmacokinetic; SD, standard deviation; V, volume.
The data set included 84 plasma concentrations of cloxacillin from 11 patients.[4] This was a preliminary run with a 1-compartment model parameterized with V and CL. The Pmetrics program was used.

points can be obtained with some advanced techniques based on Monte Carlo sampling.[2]

As the entire joint distribution of PK parameter is estimated, covariance between parameters is always computed and provided in the outputs.

As an example, Tables 1 and 2 show the typical results obtained after the analysis with a P and an NP program, for the same real data set from a previously published study.[4]

Only a few studies have reported comparative analysis of results provided by P and NP methods. Table S1 summarizes the findings of selected studies that compared P and NP analysis of real data sets.[5–14] To our knowledge, no study reported an independent, blinded analysis of a data set with the 2 approaches. Mentré and Mallet[5] as well as Carlsson et al[7] used the 2 approaches sequentially in the model building, with results from NONMEM (P) being used for the NP analysis (NPML or NPAG algorithm) and vice-versa. De Velde et al[14] stated that their analysis with NONMEM (P method) and NPAG (NP method) were independent. In that study, the structural and covariate models identified with NONMEM and NPAG were the same, which supports the consistency of both approaches. Regarding PK parameter values, the estimates of mean or median parameter values were overall similar in all studies reported in Table S1. However, larger IIV was consistently reported with NP methods. This was especially striking in the study from de Velde et al, where IIV was not set on PK parameters except the elimination rate constant (k_e) in the NONMEM analysis, while variability was set on all parameters on the NPAG analysis.[14] This means that k_e was the only parameter that could explain overall PK variability in the NONMEM analysis. Interestingly, the CV% for k_e was still greater with NPAG than with NONMEM (34% vs 19%), and high CV% values were found for the other parameters with NPAG. Despite this large difference in variance estimation, the model predictions from both approaches were very similar. Studies from the Limoges group compared the performances of an in-house iterative 2-stage P method versus NPAG in the analysis of rich PK data sets of immunosuppressant drugs from transplanted patients.[11,13] P and NP models were built in parallel, using 2 gamma distributions to describe the absorption phase and a 1 compartment model with first-order elimination for mycophenolic acid and tacrolimus. Interestingly, estimates of parameters that depended on volume of distribution were very similar between the approaches. By contrast, estimates of the shape and scale of the 2 gamma distributions describing the absorption phase were quite different between the P and NP models. Parameter variability was also larger with the NPAG algorithm. However, the performance in terms of exposure (area under the curve [AUC]) estimation was overall very similar. In another study on cyclosporine, the same group compared 3 popPK approaches including 2 P approaches (IT2B and NONMEM) and NPAG.[9] The typical value of

central volume of distribution in the population were very similar with NONMEM and IT2B (222 L and 204 L, respectively) while it was lower with NPAG (97L). The typical values of the absorption rate constant and CL were similar among the 3 approaches with absorption rate constant values of 5.72/h, 4.06/h, and 6.95/h and CL values of 41.2, 59.5 and 61.1 L/h for NONMEM, IT2B, and NPAG respectively. Finally, the predictive performances were similar but slightly better for the NPAG model in comparison to NONMEM or IT2B.

To summarize, previously published data reported some differences in popPK parameter estimates between P and NP methods in the analysis of a given data set. Most of the time, variability reported as SD or CV% was larger with NP methods. However, there is an inherent problem comparing studies by mean and SD or CV%. P or NP techniques will generally equally describe and summarize parameters that have a real Gaussian distribution and those that are not normally distributed but have a nonetheless reasonably symmetric distribution (eg, a bimodal distribution). However, the bimodal nature of the latter may not be identified by the P methods.[15] Parameter distributions that are highly skewed or include outliers (eg, clearance by polymorphic enzyme pathways) are likely to be described differently by P and NP approaches. With the latter, support points describing the outliers with extreme parameter values but low probabilities will inflate the overall variability described as SD or CV%.

If results are substantially different, how do we apply this information to the dosing of drugs in patients, such that patients get safe and effective therapy? In terms of therapeutic drug monitoring (TDM), will P and NP analysis give substantially different results in terms of dosing?

Limited information exists on this question. In their study on gentamicin in neonates, Mallet and Mentré[5] showed that the models built with NONMEM and NPML could lead to different dosages in some individuals. This was due to difference in the way covariates were coded and handled in the 2 programs.

De Velde et al[14] mentioned that the difference in parameter variability between the P and NP models of imipenem could result in differences in simulations based on those models, such as probability of target attainment. However, this was not investigated in that study. Woillard et al[9] proposed combining the estimation of the different approaches for cyclosporine because AUC estimates outside the ±20% interval in the validation data set were not the same with the 3 approaches (NONMEM, IT2B, and NPAG). They compared 2 strategies for combining the dose recommendations: (1) the arithmetic mean of all 3 proposed doses; and (2) exclusion of 1 of the 3 values when it was too different from the 2 others (difference >15%) and arithmetic mean of the remaining 2. Overall, both strategies were similar and better than using a single method.[9] These authors also used a similar approach for mycophenolate mofetil in heart transplant recipients for whom 2 independent models were developed (IT2B and NPAG).[10] They showed again that mispredicted AUCs were not the same with the 2 approaches and that arithmetic mean of AUC from the 2 approaches provided the best dose suggestion, consistently with the reference method based on the linear trapezoidal rule applied to the full PK profiles.[10]

However, further research should be conducted to assess the potential consequences of differences between P and NP models on dosage design for both populations and individuals.

Regarding TDM applications, as mentioned in the paper on the NP methods, programs based on P and NP models do not use the same Bayesian framework to estimate individual PK parameters and compute the dosage based on TDM results.[2] To our knowledge, all programs based on P models use a single parameter estimate, the maximum a posterior (MAP) to compute the dosage. By contrast, the multiple-model method (also known as stochastic control) uses the entire discrete NP Bayesian posterior distribution in dosage design.

There are arguments supporting the superiority of the stochastic control approach in control theory.[16] This was confirmed in a simulation study with vancomycin from Jelliffe et al.[17] They showed that the multiple-model dosing method based on NP population models was better than the traditional MAP based on P models for reducing IIV and achieving the concentration target after feedback from TDM results.

However, there is a lack of comparative studies with real data. As shown in Table S1, Premaud et al performed external validation of the mycophenolic acid models built with NONMEM and NPAG by computing AUC and predicted concentrations in a data set different from that used in model building.[8] They reported better predictive performance of the NPAG model on this occasion. One can assume that this might lead to difference in dose adjustment based on the estimated AUC in some patients, but this was not formally assessed in the study.

We are currently performing comparative studies on this matter using BestDose (Laboratory of Applied Pharmacokinetics and Bioinformatics, USC) and Tucuxi (HEIG-VD, Yverdon-les-Bains, Switzerland), a recently developed software based on the parametric approach.[18] For busulfan in children, the results in terms of predicted exposure and dosages were very similar, and the differences were not clinically relevant.[19]

Why do regulatory agencies ask for and accept parametric popPK for submissions? It seems that agencies do not accept NP analysis. Is this true? If so, is it a function of misunderstanding by regulatory agencies or less confidence in nonparametric analysis?

Both the US Food and Drug Administration and the European Medicines Agency provide a guidance for industry on popPK analysis for drug submission.[20,21] They both require a detailed description on the modeling development steps, that is, from base model building to final model simulations and applications, so that the assessor can understand the modeler choices, without establishing strict rules on how to make the analysis. In particular, the methods section of the European Medicines Agency document states, "The choice of analysis (eg, *parametric* maximum likelihood, *non-parametric* maximum likelihood, Bayesian) and the choice of estimation method (eg, FO, FOCE, FOCE INTER) should be stated and justified."[20] The Food and Drug Administration guidance[21] does not explicitly mention the allowed type of analysis, but cites the paper of Tatarinova et al[22] on NP methodologies among those to consider for popPK investigations. Based on the American and European agencies' official documents on popPK analyses, both approaches appear to be acceptable for submissions.

On the other hand, it is true that the P models, as implemented in NONMEM,[23] MONOLIX,[24] and Phoenix (Certara, Princeton, New Jersey) are considered as the industry standards and are the most widely employed software programs in this setting. We believe that the main cause of the larger use of P methods in drug development is historical. The development of NONMEM by the San Francisco group led by Lewis Sheiner was contemporary to that of popPK modeling itself. Thus, popPK is often described in terms of nonlinear mixed-effects modeling and associated with P methods only. We also think that the NP methodology suffers from the same preconceived opinion as NP statistics in general. The NP approach is indeed probably harder to conceptualize and understand. The P methods are based on well-known statistical distributions (eg, Gaussian, log-normal) and are probably easier to handle for beginners. Finally, the NP approach has been more associated with individualized drug therapy and dosage adjustment than drug development by its supporters.[25]

Do P and NP Analyses Require a Different Type and Level of Knowledge to Use and Interpret?

As previously explained, concepts, output results, and interpretation of P and NP approaches are a bit different. Yet we do not think that the level of knowledge required is fundamentally different. We have trained many master and PhD students to work with both P and NP methods for analyzing the same data sets and they did not report a significant difference in skills requirements.

What Are the Strengths and Limitations of Each of These Methods?

Compared to the classic methods, population approaches allow for PK characterization using fewer or unequal numbers of observations per patient, irregularly measured concentrations, and concentrations without regard to steady-state conditions, all opening up the possibility of gathering information, especially for populations who are challenging to study, like pregnant women or neonates.[26,27] Both P and NP methods can be used in both industrial drug development and optimization of clinical patient care and constitute the basis of model-based TDM software programs.

Common limitations of both P and NP approaches are those related to the use of mathematical models to describe a phenomenon, that is, data reliability along with problem oversimplification. The latter is a limitation in the sense that the information of all the underlying physiological phenomena is averaged and thus lost. However, this can also be considered as a strength. Indeed, few parameters allow for the essential description of a system and inform on its important properties (eg, absorption, distribution, metabolism, and excretion in the case of popPK models). Whichever the choice between P and NP approaches, the quality of the data is crucial for popPK analysis, as they are the foundations of the model, and poor data lead to a poor model useless for the intended application.[28] It is thus strongly recommended to limit inaccuracies and biases in data collection by establishing and successively following clear protocols and procedures.

A strength of P approaches is the use of summary statistics to characterize the probability distribution of the parameters, that is, typical values and variabilities, allowing for a huge range of statistical calculations and tests based on the central limit theorem. The method also allows for easy interpretation of covariate effects, as they are integrated as fixed parameters in the models, and for simulation from the naturally continuous probability function describing the joint population parameter value densities. Another strength of the P methods is the larger community of users, which facilitates training and collaborations.

The most important criticism of the P approaches is the assumption of a priori distributions of the parameter value distributions in the population and that these distributions can be entirely described by their statistical moments or summaries such as means and covariances. Any violations of the assumed distribution, for example, nonnormal data or unexpected subpopulations and extreme individuals, are likely to be

poorly characterized.[15] Of course, as the drug development process is focused on dosing for the majority, this approach often works adequately.

The main strength of the NP methods in population PK is the flexibility associated with the lack of statistical assumptions about parameter distribution. Therefore, those methods are likely to better identify subpopulations and outliers. When applied to TDM and drug dosing in individuals, NP models are especially suited to compute optimal dosage with maximum precision. The BestDose software includes various fitting methods based on NP methods, which is also a practical advantage in this task (see below).[29]

Although not a theoretical limitation of NP methods, current NP software programs do not model some specific pharmacodynamic data (eg, count data, time to event) and provides only a global estimation of variance without distinction between interindividual and interoccasion variability. Finally, the discrete prior distributions must be converted into sums of normal for simulations.

More details on the strengths and limitations of each method are provided in the 2 previously mentioned papers on P and NP methods.[1,2]

Is There a Way to Decide Which Method Gives the Most Precise and Accurate Information As It Applies to Drug Use in Patients?

First, we believe it is unlikely that a given method would be better for all drugs and situations. The modeling results may depend on many characteristics of the drug, the population, and study design. The only way to establish the superiority of one method for a given analysis would be to perform a comparative analysis of a common set of popPK data, and to validate both models using an external data set to figure out which method better predicts drug exposure and dosage requirements. This has scarcely been done. The studies from the Limoges group are good examples.[8–13] Indirect clues that may help to define a preference between P and NP models applied to a common data set can be derived from classic diagnostics of model fit, such as the concordance between observations and predictions and the distribution and absence of suspect trends in residuals.

Is One Method More "Labor" Intensive Than the Other?

Regarding the modeling labor, we see little difference between the methods. The preparation of data sets and analysis of results are very similar. A common characteristic of all the P and NP programs is the necessity to format the data according to their specific requirements before starting the analysis. Essential information for all of them is individual dosage regimen details, times of drug intake, time of sampling, and measured concentrations, together with any relevant factors susceptible to influence drug concentrations. Such data are organized chronologically for each subject in the data set, and minor modifications are required to switch from a P or NP software to another. Importantly, visual and statistical inspection of the data set to detect potential inaccuracies should be the initial step of any PK analysis. All the most widely used popPK software programs but NONMEM have a graphical interface and/or embedded code for visual or statistical analyses of popPK results. Nevertheless, this is possible for NONMEM users thanks to the development of Pirana, a workbench software tool that allows for organization and management of model runs and interpretation of results.[30] Simple model execution together with advanced modeling and simulation calculations with NONMEM are in addition possible with the Perl-speaks-NONMEM library.[31–33] Use of Pirana and Perl-speaks-NONMEM or other tools developed to ease NONMEM use (eg, Wings for NONMEM) is currently standard practice among NONMEM modelers, and, when not, should be strongly encouraged because these programs allow for analysis traceability and reproducibility. In addition, their employment removes several barriers in the learning process of NONMEM for young researchers. However, those programs necessary for pre- and postprocessing of NONMEM require additional training.

In terms of computation time, the NP methods have been described as slower than P methods, which was true for the first algorithms (NPML, NPEM) compared with the fast first-order approximation (FO; but inaccurate and, therefore, no more used in P analyses) and FO with conditional expectation (FOCE) methods in NONMEM. The most recent NP algorithm (NPAG) is much faster. In any case, the improvement in computer performances, including the possibility to parallelize calculations on multi-core processors, have sped up the running time for both NP and P methods. Indeed, computation duration rarely represents a barrier for popPK calculations nowadays, and very complex models can be executed in acceptable running times. However, there is a lack of modern comparative studies on this question.

How do methods of validation of each method work? It seems if one has either a very large data set, which can be broken randomly into 2 parts (model building and validation) or 2 separate data sets, this would be good. Using data that is randomly chosen from the data set that was analyzed to validate the model may raise concerns.

The validation methods are similar for P and NP methods. Those include internal validation (goodness of fit and simulation-based diagnostics such as

prediction-corrected visual predictive checks) and external validation methods. External validation can be performed with data splitting or by using a true external "independent" data set, which, whenever possible, should be the method of choice. It is most important to perform external validation for models to be used for TDM, as this emulates the real conditions of model use.

Questions on P Methods

There are several programs supporting population P methods (NONMEM, MONOLIX, Phoenix, R). There are differences in the algorithm available within each one, as well as in the overall environment. For example, users of NONMEM often have to handle a collection of programs for postprocessing of data. Could the author summarize the mains characteristics, strengths, and limitations of the main programs available?

The authors are most familiar with NONMEM and MONOLIX. NONMEM is the oldest program implementing P methods,[23] developed by the popPK pioneers, and it has no graphical user interface to directly retrieve the models and explore runs' results in contrast to the more recent Phoenix and MONOLIX[24] programs. NONMEM was originally conceived to provide model parameter estimations and simulations from a command line. Therefore, a modeler needs to compile the control streams to perform the analyses in NONMEM by writing the code directly, and to perform data management and results exploration by means of self-developed tools. However, several programs have been developed to facilitate NONMEM use[33] and fill the gap in terms of ease of use between it and the newer software programs for P analyses.

MONOLIX and Phoenix NLME are more recent programs. They provide an integrated environment for modeling and simulation as desktop software. Both programs have a user-friendly graphical user interface that provide a visual workflow for modeling project. Phoenix NLME embeds a higher variety of models than the templates offered by Pirana, in addition to modeling features to handle other type of data (eg, data below the limit of quantification). However, this is also possible in NONMEM by writing the appropriate code in the model scripts. Advanced self-written models can be implemented in the commercially available computer programs, so that they are all highly flexible. It is worth noting that NONMEM is regularly updated with the addition of the newest algorithms for popPK analyses, and therefore it offers a large choice of minimization techniques for parameter estimations[34] that might be hard to handle in other software programs.

Of course, the algorithms used by differing software packages also differ. NONMEM includes various options for FO (rarely used due to mathematical inconsistencies), FOCE, and FOCE with interaction (FOCEI). Recent versions of NONMEM also includes modern algorithm including Monte Carlo importance sampling and stochastic approximation expectation maximization (SAEM), as well as a semi-NP method. Phoenix NLME also provide several P methods (FOCE, Laplacian, QRPEM) as well as a NP engine. However, popPK analysis with the NP methods included in NONMEM and Phoenix have been rarely reported. By contrast, MONOLIX uses only the SAEM algorithm, with strongly proven convergence and efficacy.[35]

The open-source R software,[36] typically used for pre- and postprocessing of data and results, now implements a specific package, nlmixr, for P nonlinear mixed-effect modeling.[37] Parameters estimates and precisions obtained with the nlmixr algorithms were comparable to those estimated using MONOLIX and NONMEM.[38] This package requires the knowledge of the R environment and language but represents a valid free and open-source alternative to the other commercially available P software programs. On the other hand, all-purpose statistical software such as Stata (StataCorp, College Station, Texas) or SAS (SAS Institute, Cary, North Carolina) in addition to R increasingly include nonlinear mixed-effect regression routines that perfectly allow nowadays to treat popPK problems of a reasonable level of complexity.

Despite the availability of modern algorithms computing exact likelihoods (eg, SAEM), it seems that the FOCE(I) methods, which compute approximate likelihoods, remains widely used. What are the reasons for this? Are some data of expert opinion available about which algorithm should be used depending on the data set and model parameterization?

In a recent tutorial on the estimation methods implemented in NONMEM, Robert Bauer delineates general guidelines for the use of classic linearization and expectation-maximization (EM) algorithms depending on the available data and model parametrization.[34] While stating that the modern EM methods can solve all the possible problems and are more suited for very sparse data than FOCE, he points out their considerable overlap in application and usefulness with the classic methodology and suggests to compare the results obtained using a variety of methodologies. Indeed, the performances of the EM algorithms against FOCE(I) on parameters estimations do not markedly differ in complex but also simple scenarios.[34,39,40] However, the FOCE(I) runtime has been reported to be shorter

than for SAEM, as implemented in NONMEM, in several settings including sparse or rich sampling study design.[39,40] We believe that the modeler's habits play a great role in the choice of an algorithm. It is indeed important to realize that FOCE(I) is among the first algorithms implemented in NONMEM, so that modelers got familiar with it since the beginning and trained the following generation of researchers to use this algorithm in favor of those computing exact likelihoods.

The residual error models in most parametric programs is not based on measurement error, which is surprising. It is not rare to see published models with constant additive error, which is unrealistic considering the precision pattern of most drug assays. When data are available from laboratories regarding the assay error, those should be taken into account in the modeling of such data.

In P analyses, a combined additional and proportional error model most frequently depicts the residual unexplained variability, which arises from physiological intraindividual variation, model misspecification, errors in independent variables along with assay error. Sometimes, models with only the proportional components are preferred, while models with only additive components are best suited for drugs having flat concentration-time profiles, that is, negligible differences in concentration values all over the dosing interval. It is true that information on assay measurement error is not explicitly integrated in the residual error for P methods with which we are familiar (but see NP methods below). However, in 1995 Karlsson et al[41] reported that the analytical error was usually much lower than the proportional error term of the residual unexplained variability, meaning that model misspecification, physiological intraindividual variability, or errors in sampling and/or dose information would be of a greater impact. We then believe that the addition of the analytical assay data would marginally influence the results, taking into account the important improvement in the quantification methods made during the past years. On the other hand, even if the effect is small, neglecting a quantifiable, known source of variability seems counter to the philosophy of modeling. At the least, a good practice for a P modeler is to verify that the additive error component is comparable to the lowest quantifiable concentrations, and the proportional component is higher than the reported analytical method CV%. Of note, the only information about bioanalytical assay explicitly considered in parametric analyses is its limit of quantification, as several methods exist to handle concentrations below this value when they are available.[42,43]

Quite often, simplification of the statistical model is done with P models, such as ignoring the covariance between parameters or setting parameters with only fixed effect and no random effect (eg, absorption rate, volume, or intercompartment clearance). This is often justified by the inability to precisely estimate those random effects, despite being unrealistic from a biological perspective. In addition, this is likely to increase the chance to identify significant covariates on the remaining random parameters, possibly false positive.

Inclusion of interindividual variabilities on the PK parameters and covariance terms between them is clearly data driven in P analyses. It is historically true that covariance has often been neglected, even when it deserved to be taken into account. Still, these random effects must be added in the models using a stepwise strategy and kept in it only if a significant statistical improvement of the data description is observed. The inclusion of variabilities not supported by the data would, first, not be statistically significant and, second, give unsuccessful results, with poor estimations of several fixed and/or random-effects parameters. Although unrealistic from a biological perspective, the model with less but well-estimated parameters will produce an adequate fit of the data. The principle of parsimony is largely accepted in the P approach: The simplest model should be chosen among all those providing appropriate fitting. In addition, models are judged for their fitness of purpose, as increased model complexity does not always correspond to better data description. To limit the detection of false-positive PK parameter–covariate relationships and their integration in the model, however, it is strongly recommended to verify their biological plausibility before starting the covariate search. It is worth commenting that when modelers are forced to fix model parameters to have zero (co)variance, this represents a failure of study design and suboptimal sampling.

A strength of parametric methods is the ability to include intraindividual also denoted as interoccasion variability in the modeling. However, is it possible to handle such variability in model-based TDM software?

Parametric model-based TDM software relies on Bayesian inference, which allows the computation of the individual PK parameters by informing the population model with patients' characteristics and concentration measurements. Such parameters are then used for therapy individualization by maximizing the probability of target achievement between several possible dosage regimens. This strategy is now referred to as MIPD,[44] and several dedicated software tools implement it.[45,46] The interoccasion variability (IOV) in P models allows for different parameter values within an individual when PK observations are obtained on

several occasions. Its incorporation for future treatment decision is not trivial.[44,45,47] Abrantes et al[45] proposed and evaluated different strategies to handle IOV in Bayesian forecasting: estimation of the dose and the individual parameters with the original P model containing both IIV and IOV vs dose predicted neglecting IOV but employing individual parameters computed with IIV and IOV or ignoring IOV, with a new IIV calculated as the square root of the sum of the variances of the original IIV and IOV or reestimated from the original data set neglecting the IOV term. This simulation study showed that the best approach for model-based TDM is to include IOV in parameter estimations while neglecting it to calculate the next occasion dose, even when a unique measurement is considered for MIPD. It also pointed out that the proposed strategies perform similarly when IOV is lower than IIV, and the reestimation of the IIV, possible only if the original data are available, might be a valuable alternative in case of multiple occasions. These findings are also supported by the work of Wicha and Hennig.[47] Recently, Keutzer and Simonsson[48] confirmed that MIPD for drugs with high IOV allows for appropriate dosing individualization when this type of variability is included in the parameter estimation but not in dose forecasting. Of note, several of the existing MIPD software tools based on P approaches integrate IOV for future dose computation.[46]

Questions on the NP Methods

Could you please comment on the ability of NP methods to distinguish covariates? In my experience, NP methods yield robust parameter estimates such that very few covariates ever improve the model compared to using SAEM or FOCEI. Is this an expectation given that one is more exact than the other?

NP methods can distinguish covariate effects just as P methods can. However, in our experience, we found that covariates are often less relevant in NP models than in P models. A first reason may be related to covariance. As explained above, NP methods always estimate the entire joint distribution of popPK parameters, which maximizes the likelihood. Parameter covariances are not optimized to be minimal but are derived from the maximally likely parameter distribution. Therefore, covariates that are modeled as variance descriptors may be more frequently identified with P models in an attempt to minimize the covariance in random effects, that is, by allowing for explained variability in the fixed parameters through covariate relationships. Quite often, P models include parameters with only fixed effect and no random effect (eg, volume or intercompartmental clearance). This is likely to increase the chance to identify significant covariates to reduce variance in the remaining random parameters. The underlying distributional assumptions might further increase the ease of testing the effects of covariates, with a recognized risk of including some clinically futile covariates (overfitting resulting from multiple testing) that a P modeler should minimize by an a priori screening of biologically relevant factors as well as a postprocessing investigation of covariate pertinence. All this seems to be less common in the NP community, probably due to both cultural and technical factors. NP models used for simulation may be more likely to include covariates to introduce appropriate variability or strata in the simulated population(s). NP models used for estimation of individual patient exposures, particularly for MIPD, tend to have fewer covariates because the regions of the model parameter space are amplified as necessary given the patient's data, without being "held back" by a centralized probability weighting of most P distributions, for example, normal. However, to the best of our knowledge, this perceived difference has not been formally evaluated.

Another feature that is not described is modeler time. The time that it takes to run large data sets with multiple covariates with NPAG is not as efficient as SAEM. To get similar levels of efficiency, what hardware needs are necessary? Are we there yet, or is this for institutions with high computing cores?

There has been no recent comparison of running times between SAEM and NPAG, and this would be of interest. However, there is no need for supercomputers to run NP algorithms anymore. This time is over. There have been large gains in speed with the replacement of NPEM by NPAG. In addition, NPAG in Pmetrics can be parallelized in multicore computers. It can be run on personal machines for the most common applications in PK. For large models, in our experience, we can converge models with differential equations, 5 outputs, 25 parameters, and 16 subjects in a single day. The time spent on modeling depends little on the population method, but rather on the quality of data and the complexity of the model.

Since initial conditions are so important/influential on final parameter estimates, when dealing with evaluation of new chemical entities, would it be better to use NPAG first to get a better handle on initial parameter estimates, that is, can the use of NPAG be introduced into the regular workflow to help with this and then rely on SAEM for covariate model building? This is going back to efficiency considerations and integration into the current Pharma workflow.

Decades ago, we recommended the use of a P algorithm first to establish appropriate boundaries for the NP search space.[49] However, by restricting parameter

values in NP models to physiologic values, this strategy is usually not necessary. Therefore, it may be a reasonable workflow to reverse the process and start with the NP algorithm to define the distribution of parameter values without a priori specifying its nature. This could aid the modeler to identify subpopulations and/or subjects with extreme parameter values and visually check the assumption of Gaussian or log-normal distribution before running a P program. Additionally, the modeler should select a tool according to the available data and objectives of the modeling work. For example, one might prefer to use NPAG to make PK models that will be used for Bayesian adaptive control of drug therapy, as NP models are especially suited for this goal.[17] It may also be better to use NP models for popPK studies in small groups of patients, where the assumptions about normal or log-normal distribution of random effects cannot be verified with parametric methods. Although NP methods can estimate individual interoccasion variability, P methods may be preferred when it is desirable to distinguish population level between subjects and interoccasion PK variability. In addition, they are especially suited for the PK/PD analysis of count data or time-to-event data.

In general, we believe that it would be better to use and compare more often several methods in the analysis of a given data set to get more robust conclusions about the structural and covariate models, as those are strongly interconnected.[50] A nice example has been recently published.[14] We support more systematic comparisons of this kind.

What Are the Steps for NP Model Building?

Model building with NP methods is not different from that of P methods. After data formatting, model building includes the evaluation of various structural and residual error models. Then, covariate modeling is performed. The final model is selected based on standard criteria (eg, goodness of fit, simulation-based diagnostics, external validation if possible). In Pmetrics, a special feature is the tuning of parameter bounds. The user has to define the range of possible values for each parameter before running the model. Those bounds define the initial multidimensional grid where the algorithm will search for population values. The final results may vary with different bounds for a given model. Therefore, it is important to run the model with various parameter bounds and ensure that the population parameter distributions lie within the predefined bound.

How do you decide if a PK parameter is constant or is variable in the study population? Do you start by assuming that all the parameters have an associated IIV? How is the issue of covariance between parameters treated by NP methods?

By default, all parameters are assumed to be random and to have a discrete distribution, even coefficients of covariate/parameter relationships. However, in Pmetrics, it is possible to set the parameter value to a fixed value (known population mean, zero variance) or to a single value to be estimated (unknown population mean, zero variance).

Covariance is not an issue for NP methods. The distributions of each random parameter are not estimated separately. NP methods compute the entire discrete joint distribution of random parameters, so a full covariance matrix is always calculated and provided in the results.

When fitting an NP model, is the residual (intraindividual) variability regarded nonparametrically as well, or do NP methods actually rely on distributional assumptions for the residual errors no less than P methods?

In Pmetrics, the residual error model follows a special equation based on the assay error that is described by a polynomial equation and a multiplicative or additive terms for extra noise to be estimated. However, it is assumed that the residual errors are Gaussian, so the individual weighted residuals should be normally distributed, as for P methods.

Characterization of the variability in the population is one of the aims of population approaches. However, in the parametric methodology, the distribution of a PK parameter is assumed to be a priori known, and the associated random effect is related to the parameter IIV. Conversely, in NP methodology, no constraints exist on the parameter distribution, and it is not clear how the interindividual variabilities reported in the papers are computed. How are they quantified?

NP methods compute the discrete joint distribution of all model parameters, providing a grid of support points as shown in Table 1. These parameters are all random effects, so the variance of this distribution is the interindividual variance. Note that it is similar, but not numerically the same, to the variance for the IIV term in P methods, unless the number of support points in the NP model is equal to the number of subjects. Average (mean, median) as well as variability measures (variance, standard deviation, coefficient of variation) are just descriptive statistics computed from the grid of support points corresponding to the entire distribution of individual parameters.

When facing the empirical matrix of support points describing the variability of PK parameters, are there approved methods to determine whether their distribution is compatible with a given classical statistical law (eg, log-normal, normal, gamma distributions)?

First, as shown by Davidian and Gallant,[51] it is theoretically possible to estimate a smooth nonparametric distribution of random effects that allows testing for normality of the parameter distribution. Similarly, Claret and Iliadis[52] used a kernel method to estimate the nonparametric probability density function.

When the final population PK parameter distribution is discrete, as with NPAG, post hoc smoothing is also possible with similar kernel methods.[53] Then, one can statistically assess if the obtained continuous distribution is consistent with a given law.

What Are the Advantages of NP Over P for Bayesian TDM?

The strengths of the NP approach in model-based TDM have been presented in details in the article dedicated to the NP methods.[2]

A first advantage comes from the use of the entire distribution of the Bayesian posterior distribution of PK parameters in dosage calculation, unlike P methods that use only a single set of PK parameter values, most often the Bayesian MAP. The BestDose program uses the multiple-model (MM) approach, an optimal control theory also known as stochastic control. With this approach, the dosage calculation is based on the optimization of a precision criterion, which cannot be done with a MAP approach. A study has shown that this MM method was better than the MAP approach in reducing IIV in drug exposure and keeping patients in the therapeutic target.[54]

A second advantage of NP methods for model-based TDM is the ability to target an exposure target interval, not only a single value. This is also due to the use of the entire Bayesian posterior distribution in dosage design. With a MAP Bayesian approach, it is only possible to target a single value within a range. Examples of exposure interval targeting have been published elsewhere.[55–57]

Finally, the BestDose software features 3 different methods of Bayesian fitting: the previously mentioned MM, the interacting multiple model and hybrid multiple model, for any NP PK drug model. These 3 methods have been recently compared in their ability to fit past concentrations and predict future concentrations in real data sets.[29] Each method has its strengths and limitations. The user can try each one and select the most appropriate for a given patient. To our knowledge, TDM programs using P models provide only a single method of Bayesian fitting (MAP). The hybrid MM approach is a promising method that combines the flexibility of MAP and the precision of the MM method. It allows the identification of the individual PK parameter values outside the discrete NP prior. This is similar to the "flattened priors" that has been more recently considered for MAP estimation with P models.[58]

What Are the Tools Available for NP TDM?

To our knowledge, the BestDose software is the only tool available for TDM that uses NP PK models. BestDose is available as a stand-alone Windows version, an R version, and a web version from www.lapk.org. It is also embedded in the InsightRx platform for commercial purchase, although its use in InsightRx is experimental currently.

What are the barriers or what would be the facilitators to widen the use of NP approaches? And why is this approach still a minority in the landscape of pharmacometrics?

Across all fields of statistics, P approaches dominate over NP methods. The simplicity of summarizing a distribution with 1 or 2 parameters such as mean and variance is appealing. Equally important are the statistical inferences possible with P analysis, such as confidence intervals around point estimates. These characteristics make statistical tests based on P assumptions, such as *t*-tests, much more widely known than the corresponding NP Wilcoxon rank-sum test. This same familiarity and useful properties extend to the greater use of P approaches in population modeling. Nevertheless, P assumptions can often be untrue, making NP methods important additional tools in the hands of a modeler.

Further reasons for the much smaller NP community are undoubtedly due to the almost exclusive use of P approaches by the large companies in the pharmaceutical industry with in-house modeling groups. It is well known that these companies are generally highly risk averse and unlikely to innovate, instead licensing new technologies or compounds from ideas developed in academic centers, or from smaller companies who have themselves spun out of academia. Because regulators are equally familiar with P approaches and accept them, the companies have little incentive to explore additional tools.

NP approaches can likely become more widespread through 2 means: (1) education and (2) increased ease of use to fit into existing workflows. Education, workshops, web-based tutorials, and publications using NP methods can introduce new students to these methods. For ease of use, an emphasis on universal or nearly universal data input formats will facilitate easy switching between P and NP software tools. Finally, as previously mentioned, at least 1 commercial precision dosing tool by InsightRx, incorporates both NP and P options, albeit experimental for the NP algorithm.

Conclusion

As in statistics in general, P and NP methods are available for population PK analysis. So far, P methods have been more widely used, especially in drug development, while NP methods have been especially developed in software used for drug TDM and MIPD. Each approach has its strengths and limitations that users should have in mind. A limited number of studies have thoroughly compared the results of both methods in the analysis of real data. Overall, those studies showed comparable goodness of fit and predictive performance between P and NP models, but reported discrepancies in parameter estimates, especially in the magnitude of IIV. The implications of such differences in terms of simulations results or dose suggestions remain an area for further research. Both approaches are equally suitable for the various applications of population PK.

Conflicts of Interest

MN and WY developed the BestDose software. The remaining authors declare no conflicts of interest.

Data Sharing

Data sharing is not applicable. This is a review article without original data to be shared.

References

1. Guidi M, Csajka C, Buclin T. Parametric approaches in population pharmacokinetics. *J Clin Pharmacol*. 2022;62(2):125-141.
2. Goutelle S, Woillard JB, Neely M, Yamada W, Bourguignon L. Nonparametric methods in population harmacokinetics. *J Clin Pharmacol*. 2022;62(2):142-157.
3. Lavielle M. *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools.* Boca Raton, FL: Chapman and Hall/CRC; 2014.
4. Courjon J, Garzaro M, Roger PM, et al. A population pharmacokinetic analysis of continuous infusion of cloxacillin during staphylococcus aureus bone and joint infections. *Antimicrob Agents Chemother*. 2020;64(12):e01562-20.
5. Mentre F, Mallet A. Handling covariates in population pharmacokinetics. *Int J Biomed Comput*. 1994;36(1-2):25-33.
6. Bustad A, Terziivanov D, Leary R, Port R, Schumitzky A, Jelliffe R. Parametric and nonparametric population methods: their comparative performance in analysing a clinical data set and two Monte Carlo simulation studies. *Clin Pharmacokinet*. 2006;45(4):365-383.
7. Carlsson KC, van de Schootbrugge M, Eriksen HO, Moberg ER, Karlsson MO, Hoem NO. A population pharmacokinetic model of gabapentin developed in nonparametric adaptive grid and nonlinear mixed effects modeling. *Ther Drug Monit*. 2009;31(1):86-94.
8. Premaud A, Weber LT, Tonshoff B, et al. Population pharmacokinetics of mycophenolic acid in pediatric renal transplant patients using parametric and nonparametric approaches. *Pharmacol Res*. 2011;63(3):216-224.
9. Woillard JB, Lebreton V, Neely M, et al. Pharmacokinetic tools for the dose adjustment of ciclosporin in haematopoietic stem cell transplant patients. *Br J Clin Pharmacol*. 2014;78(4):836-846.
10. Woillard JB, Saint-Marcoux F, Monchaud C, Youdarene R, Pouche L, Marquet P. Mycophenolic mofetil optimized pharmacokinetic modelling, and exposure-effect associations in adult heart transplant recipients. *Pharmacol Res*. 2015;99:308-315.
11. Woillard JB, Debord J, Monchaud C, Saint-Marcoux F, Marquet P. Population pharmacokinetics and Bayesian estimators for refined dose adjustment of a new tacrolimus formulation in kidney and liver transplant patients. *Clin Pharmacokinet*. 2017;56(12):1491-1498.
12. Riff C, Debord J, Monchaud C, Marquet P, Woillard JB. Population pharmacokinetic model and Bayesian estimator for 2 tacrolimus formulations in adult liver transplant patients. *Br J Clin Pharmacol*. 2019;85(8):1740-1750.
13. Labriffe M, Vaidie J, Monchaud C, et al. Population pharmacokinetics and Bayesian estimators for intravenous mycophenolate mofetil in haematopoietic stem cell transplant patients. *Br J Clin Pharmacol*. 2020;86(8):1550-1559.
14. de Velde F, de Winter BCM, Neely MN, et al. Population pharmacokinetics of imipenem in critically ill patients: a parametric and nonparametric model converge on CKD-EPI estimated glomerular filtration rate as an impactful covariate. *Clin Pharmacokinet*. 2020;59(7):885-898.
15. Neely MN, van Guilder MG, Yamada WM, Schumitzky A, Jelliffe RW. Accurate detection of outliers and subpopulations with Pmetrics, a nonparametric and parametric pharmacometric modeling and simulation package for R. *Ther Drug Monit*. 2012;34(4):467-476.
16. Schumitzky A. Application of stochastic control theory to optimal design of dosage regimens. In: D'Argenio DZ, ed. *Advanced Methods of Pharmacokinetic and Pharmacodynamic Systems Analysis*. Boston, MA: Springer US; 1991:137-152.
17. Jelliffe R, Bayard D, Milman M, Van Guilder M, Schumitzky A. Achieving target goals most precisely using nonparametric compartmental models and "multiple model" design of dosage regimens. *Ther Drug Monit*. 2000;22(3):346-353.
18. Drennan PG, Thoma Y, Barry L, Matthey J, Sivam S, van Hal SJ. Bayesian forecasting for intravenous tobramycin dosing in adults with cystic fibrosis using one versus two serum concentrations in a dosing interval. *Ther Drug Monit*. 2021;43(4):505-511.
19. Goutelle S, Buffet R, Philippe M, et al. Validation of a new Bayesian tool for busulfan therapeutic drug monitoring in children. *Special Issue: Abstracts of the Annual Meeting of French Society of Pharmacology and Therapeutics, and INSERM Clinical Research Centers (CIC) Meeting, 12–14 June 2019*, Lyon, France. 33; 2019:2347.
20. CHMP. *Guideline on Reporting the Results of Population Pharmacokinetic Analyses*. London: EMA; 2008.
21. CDER, CBER. *Population Pharmacokinetics: Guidance for Industry*. Silver Spring, MD: FDA; 2019.
22. Tatarinova T, Neely M, Bartroff J, et al. Two general methods for population pharmacokinetic modeling: non-parametric adaptive grid and non-parametric Bayesian. *J Pharmacokinet Pharmacodyn*. 2013;40(2):189-199.
23. Beal SL, Sheiner LB, Boeckmann A, Bauer RJ. NONMEM 7.4 users guides 1989-2018. https://nonmem.iconplc.com/nonmem743/guides. Accessed November 9, 2021.
24. MONOLIXSuite *[computer program]*. Version 2018R1. Antony, France: Lixoft SAS, 2019.
25. Jelliffe RW, Schumitzky A, Bayard D, et al. Model-based, goal-oriented, individualised drug therapy. Linkage of population modelling, new "multiple model" dosage design, bayesian

feedback and individualised target goals. *Clin Pharmacokinet*. 1998;34(1):57-77.
26. Weisskopf E, Guidi M, Fischer CJ, et al. A population pharmacokinetic model for escitalopram and its major metabolite in depressive patients during the perinatal period: Prediction of infant drug exposure through breast milk. *Brit J Clin Pharmacol*. 2020;86(8):1642-1653.
27. Dao K, Guidi M, Andre P, et al. Optimisation of vancomycin exposure in neonates based on the best level of evidence. *Pharmacol Res*. 2020;154:104278.
28. *Pharmacokinetics in Drug Development: Clinical Study Design and Analysis*. Tucson, AZ: AAPS Press; 2004.
29. Goutelle S, Alloux C, Bourguignon L, Van Guilder M, Neely M, Maire P. To estimate or to forecast? Lessons from a comparative analysis of four Bayesian fitting methods based on nonparametric models. *Ther Drug Monit*. 2021;43(4):461-471.
30. Keizer RJ, van Benten M, Beijnen JH, Schellens JH, Huitema AD. Pirana and PCluster: a modeling environment and cluster infrastructure for NONMEM. *Comput Methods Programs Biomed*. 2011;101(1):72-79.
31. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)–a Perl module for NONMEM related programming. *Comput Methods Programs Biomed*. 2004;75(2):85-94.
32. Lindbom L, Pihlgren P, Jonsson EN. PsN-Toolkit–a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Meth Progr Bio*. 2005;79(3):241-257.
33. Keizer RJ, Karlsson MO, Hooker A. Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose. *CPT: Pharmacomet Syst Pharmacol*. 2013;2:e50.
34. Bauer RJ. NONMEM Tutorial Part II: Estimation Methods and Advanced Examples. *CPT: Pharmacomet Syst Pharmacol*. 2019;8(8):538-556.
35. Delyon B, Lavielle M, Moulines E. Convergence of a stochastic approximation version of the EM algorithm. *Ann Statist*. 1999;27(1):94-128.
36. R: A Language and Environment for Statistical Computing *[computer program]*. Version 3.4.1. Vienna, Austria: R Foundation for Statistical Computing; 2017.
37. Fidler M, Wilkins JJ, Hooijmaijers R, et al. Nonlinear mixed-effects model development and simulation using nlmixr and related R open-source packages. *CPT: Pharmacomet Syst Pharmacol*. 2019;8(9):621-633.
38. Schoemaker R, Fidler M, Laveille C, et al. Performance of the SAEM and FOCEI algorithms in the open-source, nonlinear mixed effect modeling tool nlmixr. *CPT: Pharmacomet Syst Pharmacol*. 2019;8(12):923-930.
39. Johansson ÅM, Ueckert S, Plan EL, Hooker AC, Karlsson MOJJoP. Evaluation of bias, precision, robustness and runtime for estimation methods in NONMEM 7. *J Pharmacokinet Pharmacodyn*. 2014;41(3):223-238.
40. Sukarnjanaset W, Wattanavijitkul T, Jarurattanasirikul S. Evaluation of FOCEI and SAEM estimation methods in population pharmacokinetic analysis using NONMEM® across rich, medium, and sparse sampling data. *Eur J Drug Metab Pharmacokinet*. 2018:43(6):729-736.
41. Karlsson MO, Beal SL, Sheiner LB. Three new residual error models for population PK/PD analyses. *J Pharmacokinet Biopharm*. 1995;23(6):651-672.
42. Ahn JE, Karlsson MO, Dunne A, Ludden TM. Likelihood based approaches to handling data below the quantification limit using NONMEM VI. *J Pharmacokinet Pharmacodyn*. 2008;35(4):401-421.
43. Bergstrand M, Karlsson MO. Handling data below the limit of quantification in mixed effect models. *AAPS J*. 2009;11(2):371-380.
44. Keizer RJ, Ter Heine R, Frymoyer A, Lesko LJ, Mangat R, Goswami S. Model-informed precision dosing at the bedside: scientific challenges and opportunities. *CPT: Pharmacomet Syst Pharmacol*. 2018;7(12):785-787.
45. Abrantes JA, Jonsson S, Karlsson MO, Nielsen EI. Handling interoccasion variability in model-based dose individualization using therapeutic drug monitoring data. *Br J Clin Pharmacol*. 2019;85(6):1326-1336.
46. Kantasiripitak W, Van Daele R, Gijsen M, Ferrante M, Spriet I, Dreesen E. Software tools for model-informed precision dosing: how well do they satisfy the needs? *Front Pharmacol*. 2020;11:620.
47. Wicha SG, Hennig S. *Handling Inter-Occasion Variability in Model Implementation for Bayesian Forecasting: A Comparison of Methods and Metrics*. Montreux, Switzerland: PAGE meeting; 2018.
48. Keutzer L, Simonsson USH. Individualized dosing with high inter-occasion variability is correctly handled with model-informed precision dosing-using rifampicin as an example. *Front Pharmacol*. 2020;11:794.
49. Jelliffe R, Schumitzky A, Van Guilder M. Population pharmacokinetics/pharmacodynamics modeling: parametric and nonparametric methods. *Ther Drug Monit*. 2000;22(3):354-365.
50. Wade JR, Beal SL, Sambol NC. Interaction between structural, statistical, and covariate models in population pharmacokinetic analysis. *J Pharmacokinet Biopharm*. 1994;22(2):165-177.
51. Davidian M, Gallant AR. Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. *J Pharmacokinet Biopharm*. 1992;20(5):529-556.
52. Claret L, Iliadis A. Nonparametric density estimation applied to population pharmacokinetics. *Math Biosci*. 1996;133(1):51-68.
53. Yamada WM, Neely MN, Bartroff J, et al. An algorithm for nonparametric estimation of a multivariate mixing distribution with applications to population pharmacokinetics. *Pharmaceutics*. 2020;13(1):42. https://pubmed.ncbi.nlm.nih.gov/33396749/
54. Bayard DS, Milman MH, Schumitzky A. Design of dosage regimens: a multiple model stochastic control approach. *Int J Biomed Comput*. 1994;36(1-2):103-115.
55. D'Argenio DZ, Rodman JH. Targeting the systemic exposure of teniposide in the population and the individual using a stochastic therapeutic objective. *J Pharmacokinet Biopharm*. 1993;21(2):223-251.
56. Philippe M, Neely M, Bertrand Y, Bleyzac N, Goutelle S. A nonparametric method to optimize initial drug dosing and attainment of a target exposure interval: concepts and application to busulfan in pediatrics. *Clin Pharmacokinet*. 2017;56(4):435-447.
57. Boidin C, Bourguignon L, Cohen S, et al. Amikacin initial dose in critically ill patients: a nonparametric approach to optimize a priori pharmacokinetic/pharmacodynamic target attainments in individual patients. *Antimicrob Agents Chemother*. 2019;63(11):e00993-19.
58. Hughes JH, Keizer RJ. A hybrid machine learning/pharmacokinetic approach outperforms maximum a posteriori Bayesian estimation by selectively flattening model priors. *CPT: Pharmacomet Syst Pharmacol*. 2021;10(10):1150–1160.

Supplemental Information

Additional supplemental information can be found by clicking the Supplements link in the PDF toolbar or the Supplemental Information section at the end of web-based version of this article.