

## ARTICLE

# Population PBPK modeling using parametric and nonparametric methods of the Simcyp Simulator, and Bayesian samplers

Janak R. Wedagedera | Anthonia Afuape | Siri Kalyan Chirumamilla  | Hiroshi Momiji  | Robert Leary | Mike Dunlavy | Richard Matthews | Khaled Abduljalil | Masoud Jamei  | Frederic Y. Bois 

CERTARA UK Limited, Simcyp Division, Sheffield, UK

**Correspondence**

Janak R. Wedagedera and Frederic Y. Bois, Certara UK Ltd, Simcyp Division, 1 Concourse Way, Level 2-Acero, Sheffield, S1 2BJ, UK  
Emails: [janak.wedagedera@certara.com](mailto:janak.wedagedera@certara.com) (J.R.W.); [frederic.bois@certara.com](mailto:frederic.bois@certara.com) (F.Y.B.)

**Funding information**

No funding was received for this work.

**Abstract**

Physiologically-based pharmacokinetic (PBPK) models usually include a large number of parameters whose values are obtained using in vitro to in vivo extrapolation. However, such extrapolations can be uncertain and may benefit from inclusion of evidence from clinical observations via parametric inference. When clinical interindividual variability is high, or the data sparse, it is essential to use a population pharmacokinetics inferential framework to estimate unknown or uncertain parameters. Several approaches are available for that purpose, but their relative advantages for PBPK modeling are unclear. We compare the results obtained using a minimal PBPK model of a canonical theophylline dataset with quasi-random parametric expectation maximization (QRPEM), nonparametric adaptive grid estimation (NPAG), Bayesian Metropolis-Hastings (MH), and Hamiltonian Markov Chain Monte Carlo sampling. QRPEM and NPAG gave consistent population and individual parameter estimates, mostly agreeing with Bayesian estimates. MH simulations ran faster than the others methods, which together had similar performance.

**Study Highlights****WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?**

No single software platform allows practitioners to compare parametric and nonparametric estimates of calibrated physiologically-based pharmacokinetic (PBPK) model parameters.

**WHAT QUESTION DID THIS STUDY ADDRESS?**

How do a maximum likelihood parametric (quasi-random parametric expectation maximization) and nonparametric (nonparametric adaptive grid estimation) algorithms, and two Bayesian numerical methods (Hamiltonian Markov Chain Monte Carlo and MH) compare in results and timing for calibration of a PBPK model?

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

### WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

We demonstrate, for the theophylline dataset studied, a strong coherence of the four approaches tested. Computation times differ largely, with MH being fastest, but this is partly implementation dependent.

### HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

Practitioners will be able to test parametric assumptions with a nonparametric method when calibrating PBPK model parameters and will have greater confidence when making important decisions in a drug development process.

## INTRODUCTION

Physiologically-based pharmacokinetic (PBPK) models usually integrate a large number of parameters because of their mechanistic nature.<sup>1</sup> The common practice workflow for PBPK modeling in clinical studies involves in vitro to in vivo extrapolation (IVIVE) of parameter values.<sup>2</sup> However, such extrapolations can be uncertain and may benefit from inclusion of evidence from clinical observations, via parametric inference on some model parameters.<sup>3</sup> When clinical interindividual variability is high, or the data sparse, it is essential to use a population pharmacokinetics framework when conducting such parametric inference.<sup>4</sup> Several approaches and algorithms are available for that purpose, and their relative advantages for PBPK modeling are unclear. A recent article<sup>5</sup> compared the application of two Markov Chain Monte Carlo (MCMC) numerical samplers to a diazepam PBPK model: the Metropolis-Hastings sampler (MH),<sup>6</sup> and the Hamiltonian MCMC sampler (HMCMC).<sup>7</sup> Both produced very good fits at the individual and population levels. HMCMC outperformed MH in sampling efficiency, due to its almost uncorrelated sampling. Yet, MH was much faster per iteration, and because of that it converged faster and produced more samples per unit time. A previous paper<sup>8</sup> compared NONMEM<sup>9</sup> (using a frequentist approach) and Winbugs<sup>10</sup> (using a Bayesian approach) on the same data set.

Physiologically-based pharmacokinetic models include species-specific physiological parameters, drug-specific parameters, and parameters defining the trial design. It is essential to document distributions or best values for this vast number of parameters (up to several hundreds), and their covariances. The Simcyp Population-Based PBPK Simulator compound database holds more than 100 drugs for which mean parameter values and, in most cases, interindividual variance (determined via meta-analysis of the available data<sup>1,11-13</sup>) are available. Similarly, distributional estimates for physiological parameters are available for more than 20 populations (e.g., Chinese, Japanese,

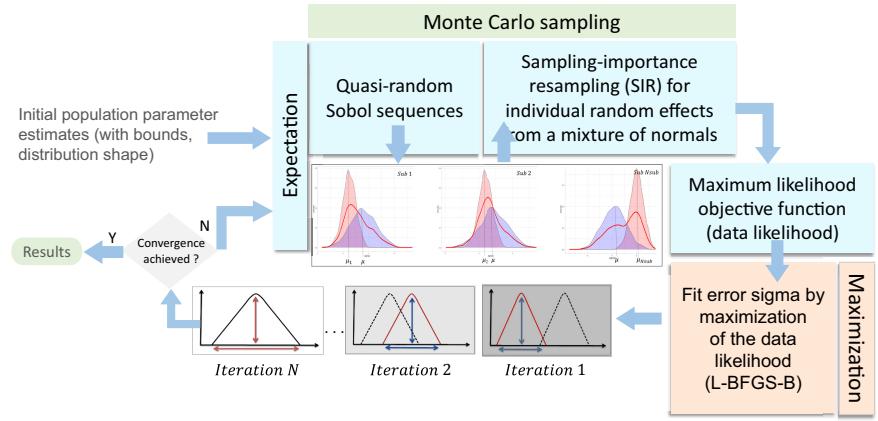
North European, pediatric, and the obese). In the Simcyp Simulator, the covariance between parameters is typically modeled via covariate submodels<sup>14</sup> when prior information is available. For several compounds, parameter values are often extrapolated from in vitro measurements. In such cases, simulations may not accurately match the available clinical data, and parametric inference based on the latter may yield better parameter estimates. Population estimation methods are computationally intensive, so unless the number of model parameters is already small, it may be necessary to confine population fitting to a small subset of them. For the rest of the parameters, which have a mechanistic definition, literature values supported by IVIVE have been used. When the number of uncertain parameters is large, sensitivity analysis is recommended to decide which ones are the most influential and should be estimated.

In this work, we compared results obtained using a minimal PBPK model of a classical theophylline dataset<sup>15</sup> with two frequentist asymptotic methods—quasi-random parametric expectation maximization (QRPEM),<sup>16,17</sup> nonparametric adaptive grid estimation (NPAG),<sup>18</sup> both implemented in the Simcyp Simulator—and two Bayesian numerical sampler, MH,<sup>19</sup> and Hamiltonian HMCMC.<sup>20</sup>

## INFERENCE METHODS AND ALGORITHMS

Statistical parameter estimation uses frequentist point-estimation or Bayesian distribution-based methods. Frequentist methods produce confidence intervals around point-estimates, which should lead to approximately the same decisions as Bayesian methods, given enough data. An independent classification differentiates parametric (assuming fixed distributions) from nonparametric methods. The methods investigated here cover the parametric and nonparametric frequentist cases, as well as the parametric Bayesian case.

**FIGURE 1** Key components of the quasi-random parametric expectation maximization (QRPEM) algorithm in Simcyp version 19/20



## Quasi-random parametric expectation maximization

The QRPEM algorithm (Figure 1) is a population pharmacokinetic likelihood maximizer first implemented in the Phoenix NLME software<sup>21</sup> and available in Simcyp version 19. In QRPEM, importance sampling of the posteriors is based on quasi-random (“low discrepancy” or Sobol) sequences<sup>22</sup> with increased sampling efficiency over pseudo-random sequences. QRPEM also uses sampling importance resampling to draw individual random effects from their conditional empirical Bayes distribution (EBD). The overall marginal log-likelihood function evaluated in QRPEM is:

$$LL(\theta, \Omega, \sigma^2) = \sum_{i=1}^N \log \left( \int_{-\infty}^{+\infty} p(Y_i | x_i, \sigma^2) p(x_i | \theta, \Omega) dx \right) \quad (1)$$

where  $Y_i$  represents the observations for individual  $i$ ,  $x_i$  the structural parameter vector for  $i^{\text{th}}$  individual ( $i = 1, \dots, N$ ). The individual parameters  $x_i$  are assumed to be normally or log-normally distributed. The  $\theta$  is a vector of fixed effects (population means),  $\Omega$  is the population parameters’ variance–covariance matrix, and  $\sigma^2$  is the residual error variance. The integrand of Equation (1) is usually written as an expectation  $\mathbb{E}_g \left[ p(Y_i | x_i, \sigma^2) \frac{p(x_i | \theta, \Omega)}{g(x_i | \theta, \Omega)} \right]$  with respect to proposal distribution  $g$ . The EBD is defined as:

$$EBD(x) = \frac{1}{\alpha} p(Y_i | x, \sigma^2) \frac{p(x | \theta, \Omega)}{g(x | \theta, \Omega)} \quad (2)$$

where  $\alpha$  is a normalization factor defined by the integral in (1),  $p(Y_i | x_i, \sigma^2)$  is the individual likelihood of the observation  $Y_i$ , and  $p(x | \theta, \Omega)$  is the conditional distribution of the individual parameters  $x_i$  with population mean and covariance.  $EBD(x)$  is approximated by a discrete distribution on  $M_i$  support points ( $x_{i1}, \dots, x_{iM}$ ) where the probability on each support point is:

$$p_{i,k} = P(x_i = x_{ik}) = w(x_{i,k}) / \sum_{k=1}^{M_i} w(x_{i,k}) \quad (3)$$

with

$$w := P(Y_i | x, \sigma^2) \frac{p(x | \theta, \Omega)}{g(x | \theta, \Omega)} \quad (4)$$

In Simcyp version 20,  $g$  is a mixture of two multivariate normal distributions. The first distribution is  $g_1(x_i | \mu_i, \lambda \Omega_i)$  where  $\mu_i$  and  $\Omega_i$  are the individual  $i$ ’s Empirical Bayes estimates of the mean and the covariance of current iteration of the EM algorithm and  $\lambda$  is a dilation factor. Distribution  $g_2(x_i | \mu, \Omega)$  is wider,  $\mu = \text{mean}(\mu_i)$  and  $\Omega$  being respectively the current population mean and covariance matrix estimates. Equal numbers of quasi-random samples for individual parameters are drawn from these two distributions to calculate a current (discrete) EBD via:

$$w(x_i) = p(Y_i | x_i, \sigma^2) \frac{p(x_i | \mu, \Omega)}{0.5(g_1(x_i | \mu_i, \lambda \Omega_i) + g_2(x_i | \mu, \Omega))} \quad (5)$$

The QRPEM maximization step is done via optimization of the residual error model parameters with a Newton method or L-BFGS-B.<sup>23</sup> A fast bootstrap method<sup>24</sup> is used to calculate the confidence regions for estimated parameters and predictions. Simcyp QRPEM calculates individual estimates as means (eventually in log-space):

$$x_i = \sum_{k=1}^{M_i} p_{i,k} x_{i,k} \quad (6)$$

Individual covariance matrices using:

$$\omega_i = \sum_{k=1}^{M_i} p_{i,k} (x_{i,k} - x_i)(x_{i,k} - x_i)' \quad (7)$$

Population mean:

$$\theta = \frac{1}{N} \sum_{i=1}^N x_i \quad (8)$$

The population variance–covariance matrix is calculated as:

$$\Omega = \frac{1}{N-1} \sum_{i=1}^N (x_i - \theta)(x_i - \theta)' + \frac{1}{N} \sum_{i=1}^N \omega_i \quad (9)$$

## Nonparametric adaptive grid estimation

The NPAG algorithm is a likelihood maximizer that does not rely on distributional assumptions of individual values. The key problem in the nonparametric maximum likelihood (NPML) method is to estimate the optimal distribution function,  $F^{\text{ML}}$ , that maximizes the likelihood of the observations among the set of discrete distributions with no more support points than the number of subjects  $N$  in the population. The fact that the number of support points must be limited to  $N$  is proven by the Carathéodory Theorem.<sup>25</sup>

In the approach proposed by Leary and Burke,<sup>26</sup> the original NPML algorithm<sup>27</sup> is applied to a modestly sized grid  $H^0$  to produce an  $F^{\text{ML}}$  solution. A primal-dual interior point algorithm with near-quadratic convergence is used to solve the NPML estimation problem.<sup>18</sup> Support points with very small probability are deleted from  $H^0$ ; new support points are added around each of these remaining support points in  $H^0$  leading to a new grid  $H^1$  and an improved likelihood (Figure 2). The process is repeated until the convergence of likelihood is achieved.

NPAG results in a conditional probability matrix  $\psi$  with one row per subject  $i$  and one column per final support point  $j$ , where  $\psi_{ij}$  is the likelihood of subject  $i$ 's data given the parameter values of the support point. Given  $\psi$ , the primal dual algorithm estimates an optimal probability vector  $\pi$  with one element per support point. The population estimate of a parameter is calculated as the expectation over the final support points,  $\sum_j (\pi_j x_j)$ , where  $x_j$  is the value of the parameter vector at the  $j^{\text{th}}$  grid point. Parameter estimates for subject  $i$  are computed using the discrete empirical

Bayes distribution for that subject. The probabilities in that distribution form a vector  $p_i$  with elements equal to:

$$p_{i,j} = \frac{\psi_{ij} \cdot \pi_j}{\sum_j (\psi_{ij} \cdot \pi_j)}. \quad (10)$$

The parameter vector estimate for subject  $i$  is then given by the expectation  $\sum_j (p_{i,j} x_j)$ .

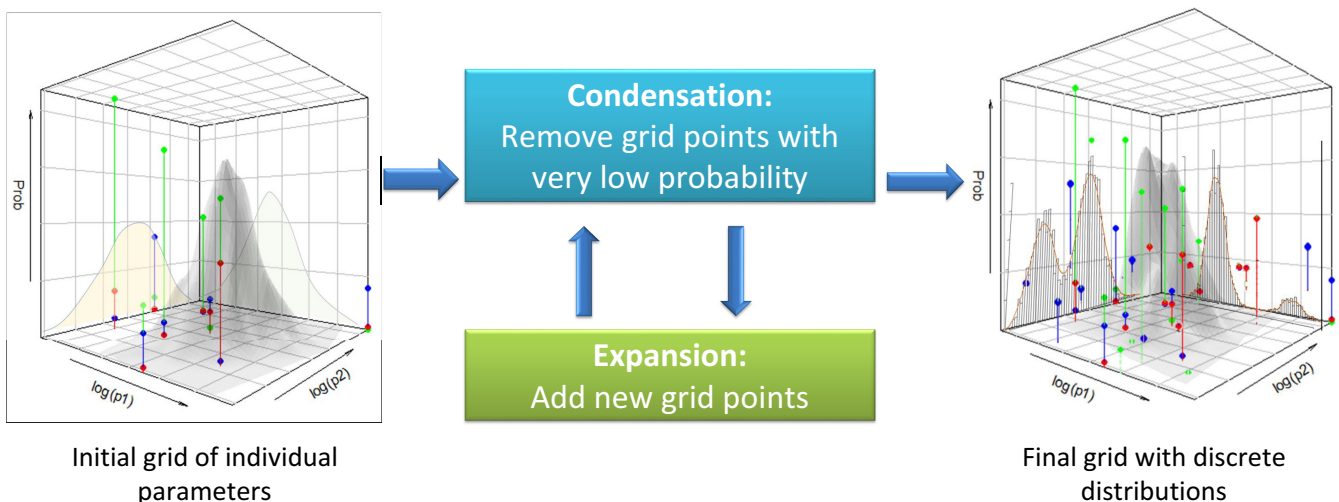
In Simcyp version 19, NPAG<sup>18</sup> was added as a population estimation algorithm. It uses the QRPEM to generate the NPAG initial grid. The population covariance matrix, calculated using individual estimates, and the probability distribution of the final grid are reported with other numerical and graphical performance measures. The coefficients  $b_0$  and  $b_1$  of the combination error model-estimated within the initial QRPEM run (see Equation (15) below) are input to the NPAG algorithm and remain fixed.<sup>18</sup>

## Metropolis-Hastings sampling

The MH algorithm<sup>6,28</sup> is an MCMC sampling method. It generates a sequence of random draws from any distribution for which the density can be computed up to a constant. Let  $\theta$  be a set of parameters of interest. According to Bayes' theorem, the joint posterior distribution of  $\theta$  given data  $\mathbf{y}$  is proportional to the product of its prior distribution by the data likelihood:

$$[\theta | \mathbf{y}] \propto [\theta] [\mathbf{y} | \theta] \quad (11)$$

To simplify the notation, let  $f(\theta)$  be the product  $[\theta] [\mathbf{y} | \theta]$ . At iteration zero of the sampler, the MH algorithm starts from



**FIGURE 2** The NPAG algorithm starts from a set of initial grid points obtained with only a few iterations of QRPEM. The grid is updated (removing and adding points) according to probabilities re-calculated in every iteration of the algorithm, as explained in the text. NPAG, nonparametric adaptive grid estimation; QRPEM, quasi-random parametric expectation maximization

an arbitrary point  $\theta^0$  (the exponent is used here for indexing). It then samples a “candidate” point  $\theta'$  using a so-called “instrumental” (i.e., arbitrary but judiciously chosen) conditional distribution  $J$ . The conditioning is, in general, on the previously sampled value of  $\theta$  (in this case  $\theta^0$ ), hence the appellation “Markov chain.”  $J$  is often a multivariate normal distribution centered on the previous draw of  $\theta$ . The candidate  $\theta'$  is selected or not, according to the following procedure:

- a. Compute the ratio

$$r_i = \frac{f(\theta')}{f(\theta^i)} \cdot \frac{J(\theta^i | \theta')}{J(\theta' | \theta^i)} \text{ with } i = 0 \quad (12)$$

- b. If  $r \geq 1$ , accept  $\theta'$ ; otherwise, accept it with probability  $r$  by sampling a uniform random number  $u$  between zero and one and accepting  $\theta'$  if  $u \leq r$ ;
- c. If  $\theta'$  is accepted, keep and call it  $\theta^1$ ; otherwise discard it and make  $\theta^1 = \theta^0$ .

The algorithm continues sampling proposed values and accepting or rejecting them, according to the value of  $r_j$ , as long as needed. It can be shown that after an infinite number of draws, the retained values of  $\theta$  form a random sample from the desired posterior distribution. In practice, it is enough to run several chains from different starting points  $\theta^0$  and check that they approximately converge to the same distribution.<sup>29</sup> Note that if  $J(\theta|\bullet)$  is symmetrical and centered on the previous value, the ratio  $J(\theta^i|\theta')/J(\theta'|\theta^i)$  in Equation 12 is always equal to 1 and does not need to be evaluated.

Instead of sampling the whole parameter vector  $\theta$  at once, it is common practice to split it into multiple components using conditional independence arguments and update these components one by one. In the simplest case, each component is a scalar, and its proposal distribution  $J$  is univariate. We selected that sampling scheme for our MH computations.

## Hamiltonian Markov Chain Monte Carlo sampling

This MCMC sampler avoids random walk. It uses Hamiltonian dynamics to generate proposal values. Imagine a particle with some position and momentum, moving over time in a frictionless space. The model parameters correspond to the position vector and an auxiliary vector represents the momentum vector. The potential and kinetic energy of the particle constitute the

total energy (equal to the Hamiltonian in the HMC context), which remains constant over time due to lack of friction. The position and momentum of the particle can be calculated by solving Hamilton's equations (a system of partial differential equations).<sup>9,30</sup>

An advantage of HMC is that it informs the proposal about the target distribution through numerical approximation of its gradient. Therefore, HMC performs better in high dimensional problems where the posterior has high curvature compared to other MCMC methods. Yet, per step, HMC sampling is much heavier computationally.<sup>31</sup>

## MATERIALS AND METHODS

### Experimental data

A canonical theophylline data set<sup>15,32</sup> was used (see Table S1). Eighteen healthy male subjects were administered an immediate release formulation of theophylline. For subjects A–K, 320 mg, for subject L, 268 mg, and for subjects M–R, 190 mg oral doses of theophylline were given, respectively. The concentration (in mg/L) of theophylline in patient's blood plasma samples was measured. The concentrations at time zero were not used (the model estimates would always be zero). Body weight was recorded for only 12 of the subjects (codes A to L) and was missing for the others.

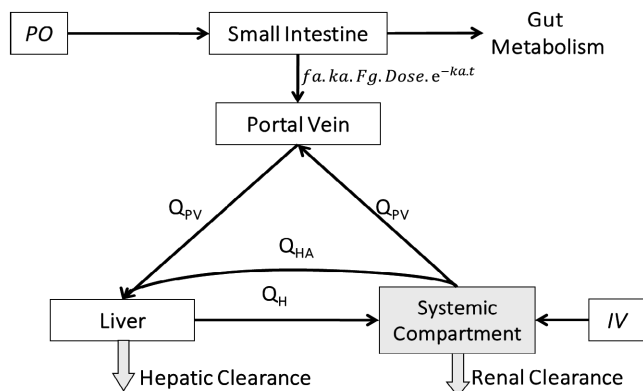
### Physiologically-based pharmacokinetic model

#### Physiologically-based pharmacokinetic model equations

The Simcyp minimal PBPK model consists of three well-stirred compartments, predicting the systemic, portal vein, and liver concentrations (Figure 3). Using a first order absorption model, input from the gut lumen can be obtained analytically. Because the volume of the portal vein blood is very small compared to the others, we assumed that the portal vein drug concentration is always at equilibrium. For theophylline, we also assumed negligible gut wall metabolism. The model equations are provided in Section S2.

#### Individuals' physiological covariates and parameter values

The subject-specific values of most physiological parameters were scaled using subjects' age.<sup>14</sup> Because Trembath



**FIGURE 3** The minimal physiologically-based pharmacokinetic (PBPK) model of the Simcyp Simulator

and Boobis only give the age range (21- to 31-year-old) of the subjects, we computed their expected age by randomly sampling the age of 1000 White men in that age range from a Weibull distribution, using Simcyp version 20. The average age of the simulated subjects (27.6 years) was then imputed to the 18 subjects studied here. When body mass was given as data, it was used, rather than computed as a function of age.

The absorption rate constant ( $k_a$ ; 1/h), the volume of distribution of the drug at steady state ( $V_{ss}$ ; L/kg) and CYP1A2 liver enzyme mediated maximum rate of metabolism ( $V_{max}$ ) of theophylline (pmol/min/pmol) were estimated for each subject, because they are the most uncertain or variable and sensitive parameters of the model.

## Statistical model and estimation procedures settings

The statistical model we consider assume that the  $j^{\text{th}}$  theophylline plasma concentration measurement for the  $i^{\text{th}}$  subject was normally distributed with a mean a time  $t_{ij}$  given by the structural PBPK model,  $f$ , and a residual error variance  $\sigma^2$ :

$$y_{ij} \sim N(f(t_{ij}, \theta_i, \varphi), \sigma^2(f(t_{ij}, \theta_i, \varphi), b_0, b_1)), 1 \leq i \leq N, 1 \leq j \leq n_i \quad (13)$$

where  $\theta_i$  was a vector of three estimated subject-specific parameters ( $k_a$ ,  $V_{ss}$  and  $V_{max}$ ),  $\varphi$  was a vector of fixed parameters (given in Table S2),  $b_0$  and  $b_1$  being residual error model parameters (also estimated).

At the population level, the subject-specific parameters were assumed to be distributed according to a multivariate log-normal distribution, with population geometric means  $\mu$  (estimated) and population covariance matrix  $\Omega$  in log-space:

$$\log(\theta_i) \sim N(\log(\mu), \Omega) \quad (14)$$

The residual error variance,  $\sigma^2$ , describes modeling and measurement errors and was chosen to be proportional to the prediction with an added constant error:

$$\sigma^2(f(t_{ij}, \theta_i, \varphi), b_0, b_1) = (b_0 + b_1 \times f(t_{ij}, \theta_i, \varphi))^2 \quad (15)$$

## Quasi-random parametric expectation maximization estimation

Table S3 gives the initial values of the population means  $\mu$  and residual error parameters, the bounds applied to  $\mu$  and  $\theta$  components, and the coefficients of variation specifying the initial population variances. The covariances (off-diagonal elements of  $\Omega$ ) were initially set to zero. The initial value of  $b_0$  corresponds to 10% of the average data values.

The QRPEM-specific control parameters' settings are given in Table S5. The number of QRPEM iterations was set at 100. For the expectation step, 500 values per subject were randomly drawn from the population distribution. An acceptance ratio ( $\alpha$ ) was used to set the dilation factor ( $\lambda$ ; see QRPEM description above) as:

$$\lambda = 1 / \alpha^{2/N_{\text{reff}}} \quad (16)$$

where  $N_{\text{reff}}$  is the number of random effects in the model. L-BFGS-B optimization was used to estimate the residual error model parameters in the maximization step. The SUNDIALS CVODES<sup>33</sup> solver, with the settings given in Table S7, was used to integrate the model's differential equations.

## Nonparametric adaptive grid estimation estimation

For NPAG, parameter initial values and bounds were the same as for QRPEM (Table S4). The NPAG-specific control parameters' settings are given in Table S6. The number of iterations was set to 100. The initial NPAG grid needs only to hold approximate solutions, so 10 iterations of initial QRPEM were run, with 200 Monte Carlo samples. The SUNDIALS CVODES solver was used for integration, as above (Table S7).

## Metropolis-Hastings estimation

Because of software limitations, a diagonal population covariance matrix was used instead of a full matrix with MH. Bayesian methods require the specification of prior distributions. At the population level, the priors for the population means of ( $K_a$ ,  $V_{ss}$ , and  $V_{max}$ ) were uniform within the same bounds as QRPEM and NPAG (Table S3). The priors for the population geometric SDs of  $K_a$ ,  $V_{ss}$ , and  $V_{max}$  were set to normal distributions with mean 1.1, SD

0.3 and truncation bounds 1 and 4.0 (corresponding to a half-normal for the population SDs in log-space). The prior for  $b_0$  was half-normal with a SD of 5, and the prior for  $b_1$  half-normal with a SD of 0.1.

Four Markov chains of 10,000 iterations were simulated. The initial 5000 samples were discarded and only one in two of the last 5000 samples were used, forming a final sample of 10,000 posterior parameter vectors. The convergence of the chains was assessed visually and confirmed using Gelman and Rubin potential scale reduction factor  $\hat{R}^{34}$ ;  $\hat{R}$  lower than 1.05 indicates well converged chains.<sup>29</sup> The Livermore *Lsode*<sup>35</sup> implicit solver was used to integrate the structural model, with relative and absolute tolerances of  $10^{-7}$  and  $10^{-8}$ , respectively.

## Hamiltonian Markov Chain Monte Carlo estimation

The population means and error model parameters  $b_0$  and  $b_1$  were assigned the same priors as for MH (above). The full population covariance matrix prior was specified using a scale vector distributed like the populations SDs in MH, and a Lewandowski–Kurowicka–Joe prior on the correlation matrix, with parameter 3.

The standard HMC no-U-turn sampler was used. Four Markov chains were simulated with 500 warmup iterations and 1000 sampling iterations per chain, and no thinning. Convergence was assessed visually, with the  $\hat{R}$  criterion, and with the effective number of samples. Backward differentiation formula method (an implicit integrator, like for the other methods) was used for solving differential equations, with relative and absolute tolerances both set at  $10^{-6}$ .

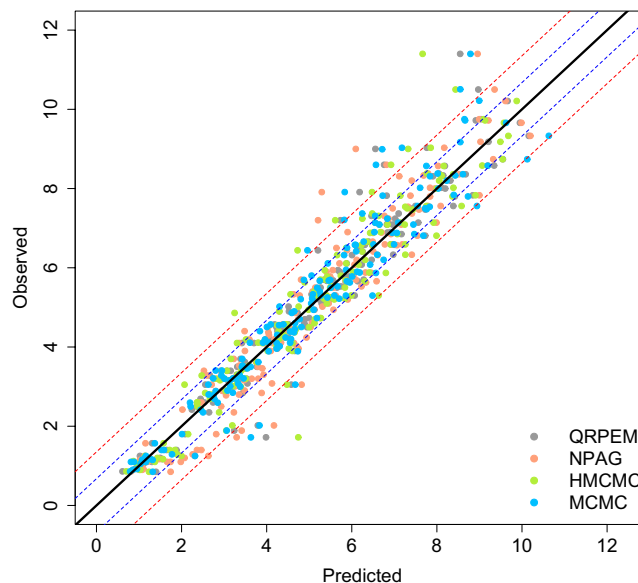
## Software used

We used the QRPEM and NPAG algorithms implemented in *Simcyp* version 20. MH simulations were performed with *GNU MCSim*<sup>19,36</sup> available at [www.gnu.org/software/mcsim](http://www.gnu.org/software/mcsim). HMCMC simulations were performed with the *R* package *rstan*<sup>20</sup> under *R* version 4.0.5.<sup>37</sup> Graphs and additional statistical treatment were performed in *R*. Software was run on a Dell Precision 5540 Notebook computer, with 16 GB RAM and 6 Intel Core i7-9850H processors clocked at 2.6 GHz.

## RESULTS

### Model fits to the data

Figure 4 shows the best (maximum likelihood or maximum posterior) predictions versus observations for each

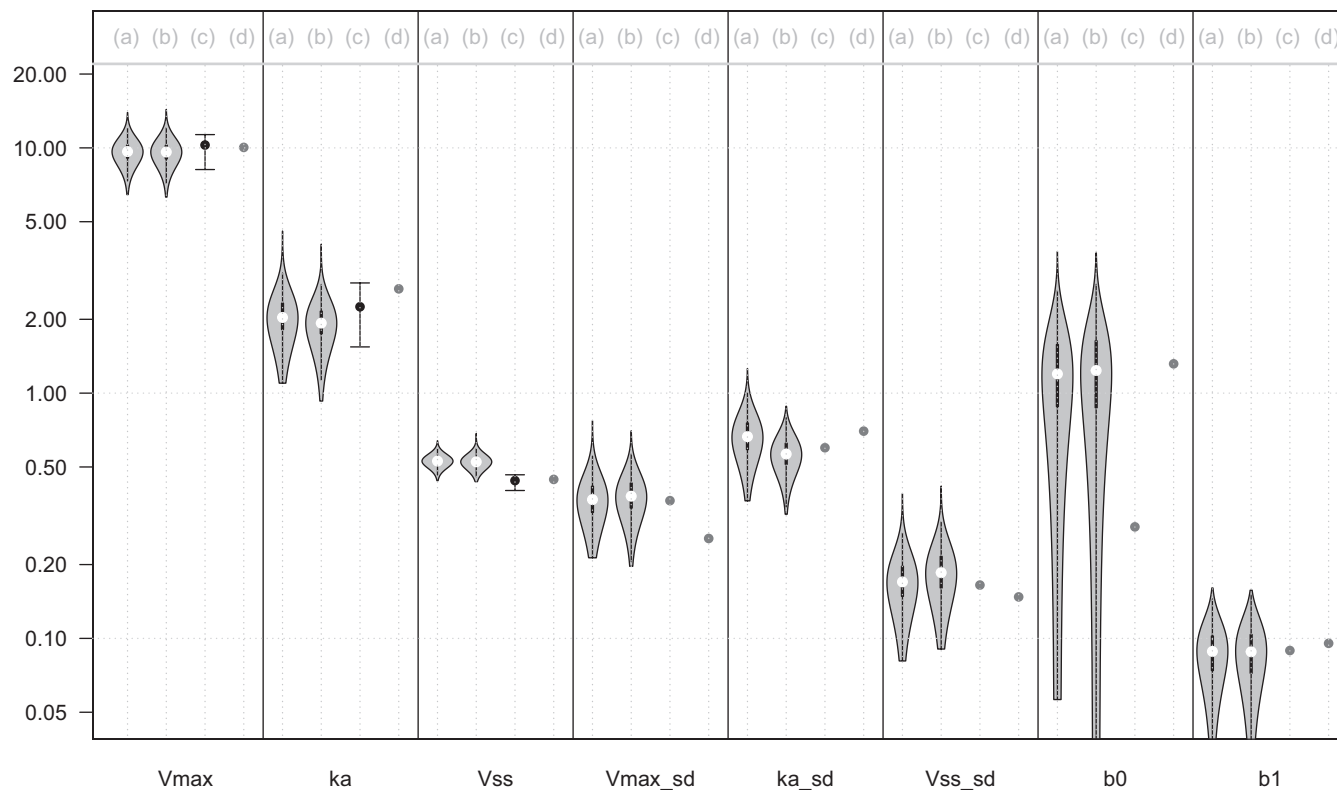


**FIGURE 4** Individual observations vs. predictions by all four parameter estimation methods. The solid line is the unity line; the blue and red dashed lines mark, respectively, 1 and 2 standard deviations above and below line of unity. HMCMC, Hamiltonian Markov chain Monte Carlo; MCMC, Markov Chain Monte Carlo; NPAG, nonparametric adaptive grid estimation; QRPEM, quasi-random parametric expectation maximization

of the four parameter estimation methods used. The predicted plasma concentration versus time curves, overlaid with individual observations, are shown in [Figures S1–S4](#) of the Supplementary Material from QRPEM, NPAG, MH, and HMCMC, respectively. All four methods gave reasonable fits to the data and the residual error is probably close to the data measurement error. There is a significant, even if not very high, intersubject variability, as can be seen by the deviations from the population average time curve.

### Comparison of population parameter estimates

Figure 5 shows the population means and SDs for  $V_{\max}$ ,  $k_a$ , and  $V_{ss}$ , and the error model parameter estimates for each method. The populations SDs are given on the log-scale, because the three parametric methods assume a log-normal distribution of individual values in the population. This is not relevant for NPAG, but we computed the SD estimates after log-transformation of the parameter grid points coordinates for easy comparison with the results of the other methods. MH and HMCMC yield posterior distributions, QRPEM computes confidence intervals, and NPAG does not provide uncertainty estimates. Yet, overall, the four methods gave similar estimates for population means and SDs, and for the proportional error component



**FIGURE 5** Population (mean, SD) and error model ( $b_0$ ,  $b_1$ ) parameters posterior distributions (a) HMC; (b) MH or maximum likelihood estimates (c) QRPEM, with 95% confidence intervals when available; (d) NPAG. HMC, Hamiltonian Markov chain Monte Carlo;  $k_a$ , absorption rate constant; MCMC, Markov Chain Monte Carlo; NPAG, nonparametric adaptive grid estimation; QRPEM, quasi-random parametric expectation maximization;  $V_{\max}$ , maximum rate of metabolism;  $V_{ss}$ , volume of distribution of the drug at steady state

( $b_1$ ) of the measurement error model. The additive error component ( $b_0$ ) has a long left-tail as it is poorly identifiable from the data, and the best QRPEM estimate is lower than the others, whereas still falling in a feasible region of the Bayesian posteriors. Note that the prior bounds imposed on the population means were not reached by any of the methods used and the estimates are practically only constrained by the data in all cases. Histograms of the posterior population parameter samples are shown in [Figures S5 and S6](#) for HMC and MH, respectively.

[Figure 6](#) shows the (best) population distributions densities according to QRPEM, MH, and HMC for  $V_{\max}$ ,  $k_a$ , and  $V_{ss}$ , overlaid with the NPAG individual estimates. The NPAG estimates do not point to any significant deviation from lognormality of the population distribution, but there are only 18 subjects to check that. Note that the final NPAG grid of support points can have at most 18 points, but in fact only eight have a probability greater than  $10^{-5}$ . The population distributions are consistent across the four methods, even though QRPEM and NPAG population mean estimates for  $V_{ss}$  are a bit lower than those of MH and HMC. Detailed population parameter estimates are given in [Tables S8–S11](#) for QRPEM, NPAG, HMC, and MH, respectively. Detailed individual parameter estimates are given in [Tables S12–S15](#). The last two tables also

give the  $\hat{R}$  convergence values obtained with HMC and MH, respectively. Approximate convergence of the simulated Markov chains was always achieved.

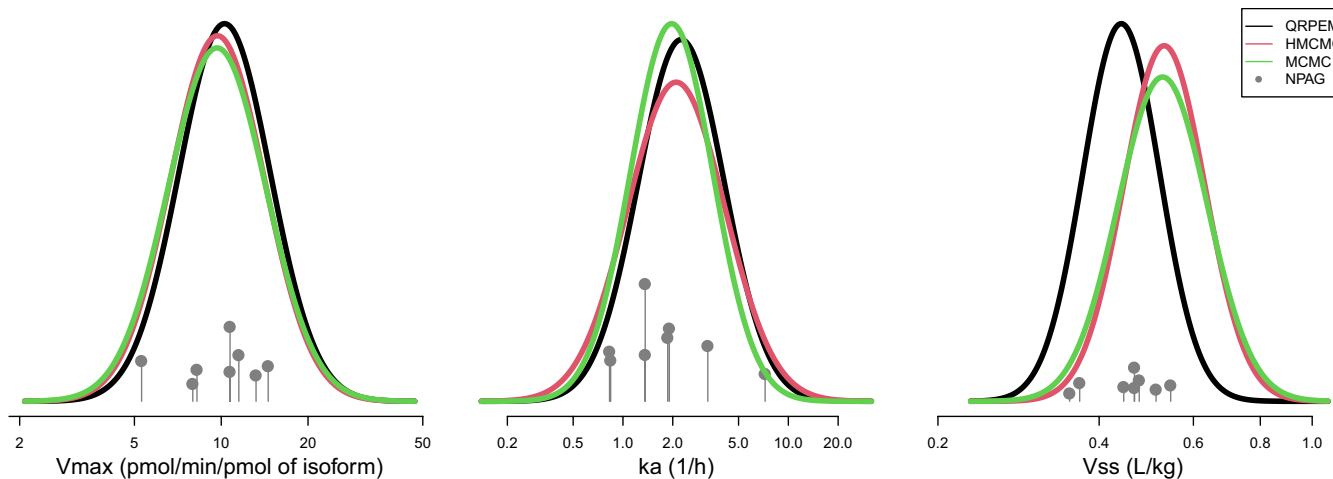
## Computation time

Run times of QRPEM and NPAG were 4.5 h and 1 h 20 min respectively, and 1 h 15 min and 1.5 min for HMC and MH, respectively. To check whether the much shorter run time of MH was partly due to a simpler (diagonal) matrix estimation, we ran separately an HMC estimation with a diagonal matrix. The HMC run time was the same as with a full covariance matrix and the results were the same as those of MH (data not shown), so the better performance of MH is not due to simpler matrix estimation.

## DISCUSSION

In this work, we used the parametric QRPEM and the nonparametric NPAG methods, both implemented in Simcyp version 20, together with two Bayesian numerical methods (HMC and MH) to calibrate a PBPK model using population data on theophylline plasma concentration following





**FIGURE 6** Best population distributions density estimates according to QRPEM, MH, and HMCMC for  $V_{\max}$ ,  $k_a$ , and  $V_{ss}$ , overlaid with the NPAG individual estimates (points). HMCMC, Hamiltonian Markov chain Monte Carlo;  $k_a$ , absorption rate constant; MCMC, Markov Chain Monte Carlo; NPAG, nonparametric adaptive grid estimation; QRPEM, quasi-random parametric expectation maximization;  $V_{\max}$ , maximum rate of metabolism;  $V_{ss}$ , volume of distribution of the drug at steady state

oral administration of an immediate release form. We used a minimal PBPK model consisting of only two ordinary differential equations and a first-order absorption rate but complemented with a rich covariate model.

We could have used a more complex PBPK model, but such a model was not needed to describe theophylline pharmacokinetics. Theophylline undergoes negligible intestinal metabolism and exhibits almost 100% bioavailability, with a time to peak plasma level of about 45 min in humans given theophylline orally as a solution.<sup>38,39</sup> Those justify using a first-order model for theophylline absorption and estimating an absorption rate. Theophylline distributes mostly into body water. Its apparent volume of distribution is variable across subjects (which justifies estimating it, as we did here) with a mean of about 0.48 L/kg.<sup>39</sup> This low volume of distribution justifies using a minimal PBPK model. The baseline values of our theophylline model parameters were extracted from the Simcyp compound library file for that drug. Metabolism of theophylline by hepatic CYP1A2 accounts for 70–80% of its total metabolism. Out of the three theophylline metabolic pathways mediated by CYP1A2, we selected the main N3-demethylation pathway<sup>40</sup> and  $V_{\max}$  for that pathway was fitted, given its relatively high expected interindividual variability. Its value is only apparent and includes contributions from all pathways and enzymes.

Given the unavoidable uncertainty affecting the results, the four methods gave comparable estimates of all model parameters at the population and subject levels. The posterior fit to the data were uniformly very good. Because we used uniform prior distributions for population means in Bayesian estimation, it is no surprise that the results are similar to those of maximum likelihood methods. It is reassuring that the choice of method does not strongly condition the data analysis results. The population estimates

of  $k_a$ , are close to  $2 \text{ h}^{-1}$ , which correspond to an absorption half-time of about 30 min. The population mean  $V_{ss}$  estimate is very close to the published value 0.48 L/kg, and the population mean for  $V_{\max}$  is close to 10 pmol/min/pmol of CYP1A2, which is compatible with the value used in the Simcyp compound library file (6 pmol/min/pmol) because, as mentioned above, our estimate includes contributions from other metabolic pathways which are modeled separately in Simcyp. Population variability estimates correspond to about 80% coefficient of variation (CV) for  $k_a$ , 20% CV for  $V_{ss}$ , and 50% CV for  $V_{\max}$ . A correlation at the population level between  $V_{ss}$  and  $V_{\max}$  seems to exist, but it is quite uncertain (according to HMCMC, which gives a measure of its uncertainty), and MH (which neglect those correlations) gave estimates consistent with those of the other methods. We suggest that the amount of information and covariance modeling that enters the PBPK model leaves their parameters quite uncorrelated. This has already been noticed by others.<sup>5,41</sup> This feature allows the use of diagonal population covariance matrices, simpler to estimate.

In terms of computation time, MH was the fastest, followed by NPAG and HMCMC, and then by QRPEM. However, QRPEM and NPAG are implemented in the Simcyp Simulator and we used only a subset of its minimal PBPK model functionalities. It is likely that some overhead in QRPEM is due to unused capabilities of the model.

Population estimation of PBPK model parameters has been previously reported using various Bayesian or maximum likelihood methods implemented in Gnu MCSim, Matlab, Stan, PK-Sim, and NONMEM.<sup>5,42,43,44,45</sup> The current study is the first to compare the performance of parametric Bayesian to those of the parametric QRPEM and non-parametric NPAG frequentist methods in the context of PBPK models.

The population parameter estimation module in Simcyp offers both a parametric method (QRPEM) and a nonparametric one (NPAG). This allows the usual (log)normality assumptions made by QRPEM to be immediately checked by NPAG. For theophylline, with 18 subjects, NPAG did not point to any flaw in that assumption. This could indeed be different with other drugs and populations.

Population inference can be used with complex models if they are supported by sufficient individual data. Obviously, computing power should be scaled up accordingly. Complex models often include mechanistic parameters for which informative priors can be defined. In a Bayesian context, that helps the estimation by increasing identifiability. Still, only the parameters to which the data predictions are sensitive should be included to limit identifiability problems. That is partly why global sensitivity analysis was recently implemented in Simcyp.

The possibility to use population inference for complex PBPK model opens interesting perspectives for improved simulations of clinical trials for drug–drug interactions, special populations, or bioequivalence assessments. The ability to estimate population variability in drug-specific PBPK parameters in Simcyp should help users in capturing the variability seen in available data when running clinical trial simulations. Population means and CV estimates given by QRPEM or NPAG can be input for the parameters of the compound of interest. However, one also needs to consider the variability affecting all the other PBPK model parameters that were kept fixed during the estimation process. For those, informative prior distributions determined by meta-analysis or IVIVE should be used when running predictive checks of the inference and further simulations. On a more fundamental level, better estimates of population variability for some parameters and processes can point to the need to better understand the biological determinants of that variability.

## ACKNOWLEDGEMENTS

The authors acknowledge the feedback and support from Prof. Amin Rostami, Dr. Ian Gardner, and Dr. Sibylle Neuhoff on various aspects. We are also thankful to the Simcyp Library team for helping with literature access.

## CONFLICT OF INTEREST

All authors are employees and potential shareholders of Certara.

## AUTHOR CONTRIBUTIONS

J.W., A.A., S.C., H.M., F.B., R.L., M.D., R.M., K.A., and M.J. designed the research. F.B., R.L., and M.D. contributed new analytical tools. J.W., A.A., S.C., H.M., F.B., R.L., and M.D. performed the research. J.W., A.A., S.C., H.M.,

and F.B. analyzed the data. J.W., F.B., A.A., S.C., and H.M. wrote the manuscript.

## ORCID

Siri Kalyan Chirumamilla  <https://orcid.org/0000-0002-5026-8369>

Hiroshi Momiji  <https://orcid.org/0000-0002-3015-5770>

Masoud Jamei  <https://orcid.org/0000-0002-3443-0194>

Frederic Y. Bois  <https://orcid.org/0000-0002-4154-0391>

## REFERENCES

1. Jamei M, Bajot F, Neuhoff S, et al. A mechanistic framework for in vitro–in vivo extrapolation of liver membrane transporters: prediction of drug–drug interaction between rosuvastatin and cyclosporine. *Clin Pharmacokinet.* 2014;53(1):73–87. doi:10.1007/s40262-013-0097-y
2. Rostami-Hodjegan A. A physiologically based pharmacokinetics joined with in vitro–in vivo extrapolation of ADME: a marriage under the arch of systems pharmacology. *Clin Pharmacol Ther.* 2012;92(1):50–61.
3. Emoto C, Johnson TN, Hahn D, et al. A theoretical physiologically-based pharmacokinetic approach to ascertain covariates explaining the large interpatient variability in tacrolimus disposition. *CPT: Pharm Syst Pharmacol.* 2019;8(5):273–284.
4. Calvier EAM, Nguyen TT, Johnson TN, et al. Can population modelling principles be used to identify key PBPK parameters for Paediatric clearance predictions? An innovative application of optimal design theory. *Pharm Res.* 2018;35(11):209.
5. Tsiros P, Bois FY, Dokoumetzidis A, Sarimveis H. Population pharmacokinetic reanalysis of a diazepam PBPK model – a comparison of Stan and GNU MCSim. *J Pharmacokinet Pharmacodyn.* 2019;4:173–192. doi:10.1007/s10928-019-09630-x
6. Hastings WK. Monte Carlo sampling methods using Markov chains and their application. *Biometrika.* 1970;57:97–109.
7. Girolami M, Calderhead B. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J Royal Stat Soc Ser B.* 2011;73:123–214.
8. Langdon G, Gueorguieva I, Aarons L, Karlsson M. Linking pre-clinical and clinical whole-body physiologically based pharmacokinetic models with prior distributions in NONMEM. *Eur J Clin Pharmacol.* 2007;63(5):485–498. doi:10.1007/s00228-007-0264-x
9. Betancourt M. A conceptual introduction to Hamiltonian Monte Carlo; 2017. <https://arxiv.org/abs/1701.02434v1>
10. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS — a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* 2000; 10:325–337.
11. Musther H, Harwood MD, Yang J, Turner D, Rostami-Hodjegan A, Jamei M. The constraints, construction, and verification of a strain-specific physiologically based pharmacokinetic rat model. *J Pharm Sci.* 2017;106:2826–2838. doi:10.1016/j.xphs.2017.05.003
12. Jamei M. Recent advances in development and application of physiologically-based pharmacokinetic (PBPK) models: a transition from academic curiosity to regulatory acceptance. *Curr Pharmacol Rep.* 2016;2:161–169. doi:10.1007/s40495-016-0059-9
13. Jamei M, Duncan Edwards SM, Wragg K, Feng K, Barnett A, Rostami HA. The Simcyp population based simulator: architecture, implementation, and quality assurance. *In Silico Pharmacol.* 2013;1(9):1–14. doi:10.1186/2193-9616-1-9

14. Jamei M, Dickinson G, Rostami HA. A framework for assessing inter-individual variability in pharmacokinetics using virtual human populations and integrating general knowledge of physical chemistry, biology, anatomy, physiology and genetics: a tale of 'bottom-up' vs 'top-down' recognition of covariates. *Drug Metab Pharmacokinet*. 2009;24(1):53-75. doi:10.2133/dmpk.24.53
15. Trembath P, Boobis S. Pharmacokinetics of a sustained-release theophylline formulation. *Br J Clin Pharmacol*. 1980;9(4):365-369.
16. Leary RH, Dunlavey M. QRPEM – a new standard of accuracy, precision, and efficiency in NLME population PK/PD methods; 2020. <https://www.certara.com/white-paper/qrpem-a-new-standard-of-accuracy-precision-and-efficiency-in-nlme-population-pkpd-methods/>
17. Leary RH, Dunlavey M. A quasi-random parametric EM method; 2012. 2012 PAGE meeting abstract 2338. <https://www.page-meeting.org/default.asp?abstract=2338>
18. 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:E42. doi:10.3390/pharmaceutics13010042
19. Bois FY. GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models. *Bioinformatics*. 2009;25:1453-1454.
20. Stan Development Team. RStan: the R interface to Stan. R package version 2.21.2. 2020; Accessed June 29, 2021. <https://mc-stan.org/rstan/authors.html>
21. Phoenix® WinNonlin® User's Guide version 8.3. Certara USA, Inc.; 2020. Accessed June 29, 2021. [https://onlinehelp.certara.com/phoenix/8.3/responsive\\_html5\\_MasterPage!/WinNonlin%20User%27s%20Guide.pdf](https://onlinehelp.certara.com/phoenix/8.3/responsive_html5_MasterPage!/WinNonlin%20User%27s%20Guide.pdf)
22. Sobol IM. Uniformly distributed sequences with an additional uniform property. *Comput Maths Math Phys*. 1976;16:236-242.
23. Zhu C, Byrd RH, Nocedal J. L-BFGS-B: algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. *ACM Transact Math Soft*. 1997;23(4):550-560.
24. Leary RH, Dunlavey M, Chittenden J. A fast bootstrap method using EM posteriors. Presented at: PAGE meeting; 2013; Scotland, UK
25. Uchida D, Gabriele E. Carathéodory theorem. In: Floudas CA, Pardalos PM, eds. *Encyclopedia of Optimization*. Springer US; 2009:358-359.
26. Leary RH, Jelliffe R, Schumitzky A, Van Guilder M. An adaptive grid non-parametric approach to population pharmacokinetic/dynamic (PK/PD) population models. *Proc 14th IEEE Symp Computer Based Med Syst*. 2001;1(1):389-394.
27. Schumitzky A. Nonparametric EM. Algorithms for estimating prior distributions. *Appl Math Comput*. 1991;45(2):143-157.
28. Metropolis N, Rosenbluth M, Teller A, Teller E. Equation of state calculation by fast computing machines. *J Chem Phys*. 1953;21:1087-1092.
29. Gelman A, Rubin D. Inference from iterative simulation using multiple sequences. *Stat Sci*. 1992;7:457-511.
30. Brooks S, Gelman A, Jones GL, Meng XL. MCMC using hamiltonian dynamics. In: *Handbook of Markov Chain Monte Carlo*. 2nd ed. CRC Press; 2011.
31. Bois FY, Gao W, Chiu WA, Reissfeld B. Well-tempered MCMC simulations for population pharmacokinetic models. *J Pharmacokinet Pharmacodyn*. 2019;47(6):543-559. doi:10.1007/s10928-020-09705-0
32. Bates D, Maechler M, Bolker B. MEMSS: Data Sets from Mixed-Effects Models in S. R package version 0.9-3. R Foundation for Statistical Computing; 2019. Accessed June 29, 2021. <https://cran.r-project.org/web/packages/MEMSS/index.html>
33. Hindmarsh AC, Brown PN, Grant KE, et al. SUNDIALS: suite of nonlinear and differential/algebraic equation solvers. *ACM Transact Math Soft*. 2005;31(3):363-396.
34. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat*. 1998;7(4):434-455.
35. Radhakrishnan K. Description and use of LSODE, the Livemore solver for ordinary differential equations; NASA Reference Publication 1327, LLNL Report UCRL-ID-113855, 1993.
36. Bois FY, Maszle D. MCSim: a simulation program. *J Stat Softw*. 1997;2(9):1-60. doi:10.18637/jss.v002.i09
37. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2013. <https://cran.r-project.org/bin/windows/base/>
38. Upton RA, Sansom L, Guentert TW, et al. Evaluation of the absorption from 15 commercial theophylline products indicating deficiencies in currently applied bioavailability criteria. *J Pharmacokinet Biopharm*. 1980;8(3):229-242.
39. Trépanier EF, Nafziger AN, Amsden GW. Effect of terbinafine on theophylline pharmacokinetics in healthy volunteers. *Antimicrob Agents Chemother*. 1998;42(3):695-697.
40. Ha H, Chen J, Freiburghaus A, Follath F. Metabolism of theophylline by cDNA-expressed human cytochromes P-450. *Br J Clin Pharmacol*. 1995;39(3):321-326.
41. Gelman A, Bois FY, Jiang J. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J Am Stat Assoc*. 1996;91:1400-1412.
42. Banks HT, Ma Y, Potter LK. A simulation-based comparison between parametric and nonparametric estimation methods in PBPK models. *J Inv Ill-Posed Problems*. 2005;13(1):1-26.
43. Liu X, Wang Y. Comparing the performance of FOCE and different expectationmaximization methods in handling complex population physiologically-based pharmacokinetic models. *J Pharmacokinet Pharmacodyn*. 2016;43:359-370. doi:10.1007/s10928-016-9476-y
44. Krauss M, Burghaus R, Lippert J, et al. Using Bayesian-PBPK modeling for assessment of inter-individual variability and subgroup stratification. *In Silico Pharmacology*. 2013;1(6):1-11. doi:10.1186/2193-9616-1-6
45. Wendling T, Dumitras S, Ogungbenro K, Aarons L. Application of a Bayesian approach to physiological modelling of mavoglurant population pharmacokinetics. *J Pharmacokinet Pharmacodyn*. 2015;42:639-657. doi:10.1007/s10928-015-9430-4

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Wedagedera JR, Afuape A, Chirumamilla SK, et al. Population PBPK modeling using parametric and nonparametric methods of the Simcyp Simulator, and Bayesian samplers. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:755-765. doi: [10.1002/psp4.12787](https://doi.org/10.1002/psp4.12787)