A Stochastic Optimization Approach for the Design of Individualized Dosage Regimens

José M. Laínez-Aguirre and Gintaras V. Reklaitis
School of Chemical Engineering, Purdue University, 480 Stadium Mall Dr., West Lafayette, IN 47907

DOI 10.1002/aic.14100
Published online April 12, 2013 in Wiley Online Library (wileyonlinelibrary.com)

Quantitative methods for individualizing and optimizing the dosage regimen and clinically monitoring each patient are desirable to insure that each patient can obtain effective therapeutic benefit while minimizing undesirable side effects. This is of special concern for medicines that are expensive or whose toxic side effects are severe (e.g., oncological agents). The optimal dosage regimen for an individual is a combination of dose amount and/or dosing interval (i.e., time between doses) which minimizes the risk that the drug exposure deviates from the desired therapeutic window. The therapeutic window is defined as the range of drug exposure (e.g., blood concentration, area under the curve concentration-time) which is below a threshold defining an acceptable toxic side effect and above a threshold defining a minimum acceptable level of therapeutic efficacy. In this work, the dosage regimen optimization problem defined in terms of general pharmacometric models (i.e., described by differential-algebraic equations) is presented and a solution approach outlined which uses a scenario-based stochastic optimization formulation that minimizes a risk metric. The scenarios are derived from the posterior joint probability distribution of the individual's pharmacometric parameters which is obtained following an approximate Bayesian inference approach. A Smolyak rule is used for the selection of the scenarios (i.e., combination of pharmacometric parameters) to be considered and for computing the approximation to the risk metric. Two case studies, gabapentin and cyclophosphamide, are presented to elucidate the advantages and limitations of the proposed approach. The numerical results demonstrate that low risk optimal solutions can be generated via the proposed stochastic optimization; while significantly reducing the computational burden in comparison with the conventional Markov chain Monte Carlo—grid search approach. This partially alleviates implementation issues preventing the deployment of dosage regimen individualization in clinical practice. Since stochastic optimization has been extensively used in other domains, the approach for uncertainty characterization proposed in this work may have general relevance beyond the pharmacometrics domain. © 2013 American Institute of Chemical Engineers *AIChE J*, 59: 3296-3307, 2013
Keywords: pharmacometrics, dosage regimen individualization, uncertainty characterization, stochastic optimization, NLP, MINLP

Introduction

Researchers in process systems engineering (PSE) are actively contributing to the solution of problems arising in physiology and clinical medicine that can be addressed by model and optimization-based methodologies that are native to PSE.[1] One such area of application is the growing research directed at providing improved approaches to individualized drug treatment using pharmacometric models. Briefly, pharmacometrics deals with the interpretation of data generated during clinical trials and on-going drug administration using mathematical and statistical methods so as to support decision making in drug development and patient care. Pharmacometrics can be divided into two major areas, pharmacokinetics (PK) and pharmacodynamics (PD). The former studies how the body handles a drug by describing the time course of drug concentration in the plasma; while the latter describes the relationship between drug concentration and therapeutic effect.

Practicing clinicians are aware of the wide variability among patients in therapeutic responses to a given dose of any particular drug and recognize that this variability may be explained by population pharmacokinetic and/or pharmacodynamic differences.[2] The study and incorporation of these differences in dosage regimen design is especially relevant for those drugs that have tight therapeutic windows and highly toxic side effects. The standard "one fits all" dosing regimen for such drugs is likely to result in (1) ineffective use of the drug and no therapeutic response or (2) adverse drug reactions with additional costs due to overmedication and in some cases even substantially diminished patient quality of life. A systems approach can provide suitable tools to support drugs prescription and ultimately improve the therapeutic outcomes for individual patients. A sample of some relevant research focused on dosage regimen individualization is presented next.

In order to build an individualized patient-specific model, Jellife et al.[3] estimated PK parameters for individuals by

Correspondence concerning this article should be addressed to J. M. Laínez-Aguirre at jlainez@purdue.edu

minimizing the combination of deviations from population parameters and predicted plasma concentrations. They suggested that adaptive control could be then used to specify a dosage regimen which would achieve the desired therapeutic goal. Subsequently, Mehvar[4] showed how PK principles could be used to design dosage regimens so that the average of plasma concentration at steady state achieved an estimated target. Salinger and coworkers[5,6] determined individualized second doses of a cancer agent treatment to achieve a target area under the concentration-time curve. They proposed the use of a maximum a posteriori approach for this purpose. Laínez et al.[7] showed how a feasible range for a specific drug can be determined which would achieve an acceptable level of risk of over and under medication. This approach used the posterior distributions of parameters obtained from a Markov chain Monte Carlo (MCMC) computation and can be applied for those PK models for which the dose variable can be isolated from the administration interval effect. Research has also been reported on individualizing therapies which make use of control methods. For instance, Bayard et al.[8] presented a stochastic control framework for determining dosage regimens based on models with linear dynamics and quadratic cost assuming discrete distributions of the model parameters. Ji et al.[9] proposed a weighting function to penalize drug exposures that range outside of the therapeutic window. They also used a stochastic control framework with discrete probabilities for the parameters. A nonlinear model predictive control algorithm was developed to dose individualization of the chemotherapeutic agent tamoxifen by Florian Jr. et al.[10] The PK and PD parameters were fitted by minimizing the classical square of the differences between prediction and observations. Noble et al.[11] developed an adaptive model predictive control strategy for the dosing of the chemotherapeutic agent 6-mercaptopurine. The model parameters are fitted by minimizing the weighted squared differences between prediction and observations. More recently, Sopasakis and Sarimveis[12] proposed a finite impulse response model to fit experimental PK data and employed this model within an optimal control framework.

Most of the works to date for determining individualized dosage regimens are based on the use of point estimates or simplified approximations to the probability distributions of the pharmacometric parameters. The work described in the following sections proposes a dosage regimen individualization strategy which relies on the probability distributions of parameters for an individual which are generated by using approximate Bayesian inference. These distributions are used in a scenario-based stochastic optimization formulation which seeks to minimize a side risk metric. In addition, the framework is general enough to accommodate a wide variety of pharmacometric models. For those models described by a set of differential algebraic equations (DAE), orthogonal collocation on finite elements is employed. A Smolyak quadrature rule which has proved to be efficient for multivariate cases is used for the approximation of those model equations involving the posterior probability distribution.

The rest of the article is organized as follows. Section entitled The Proposed Bayesian Framework for Dosage Regimen Individualization describes the approach that has been recently suggested for dosage regimen design. We point out the contribution of this work to its further development. In Section entitled The Dosage Regimen Individualization Problem, the proposed stochastic optimization approach is mathematically posed. An approximation to this problem so that it can be solved as an Nonlinear program (NLP)/Mixed Integer Nonlinear program (MINLP) is presented in subsection entitled A discretized stochastic optimization. Next, two case studies are presented to highlight the capabilities of the proposed optimization framework. Finally, some conclusions are drawn and directions for future research suggested.

The Proposed Bayesian Framework for Dosage Regimen Individualization

In this article, we report on further advances in the development of the Bayesian framework for dosage regimen individualization that was proposed in our earlier work.[7] That framework is divided into three stages as depicted in Figure 1. The first stage comprises the use of clinical trial data for constructing the joint probability distribution of the pharmacometric model parameters for the population. Based on this distribution a procedure was proposed for determining a minimal sampling schedule (i.e., number of samples and their corresponding collection times) to obtain well characterized pharmacometric parameters for a new patient. This first stage constitutes an offline analysis procedure as it is performed just once to formalize the prior knowledge represented by the clinical trial data.

In the second stage, an individualized joint distribution is computed for a new patient using the population distribution and the data collected for the individual following the sampling schedule suggested in stage I. This individualized distribution of pharmacometric parameters is used in stage III to compute a dosage regimen that results in a drug exposure profile that falls within the desired therapeutic window. The second and third stages are repeated for each new patient for whom the dosage regimen is to be individualized. The objective of the research reported in this article is to pose the dosage individualization problem of stage III (cross-hatched area in Figure 1) as a constrained stochastic optimization formulation and to present and discuss approaches to its solution. Therefore, for purposes of this work, we assume that the individualized posterior distributions (i.e, the output of stage II) have already been determined.

A general representation of pharmacometric models

In general, a pharmacometric model can be described by a set of differential-algebraic equations of the general form

$$\begin{aligned} \hat{y}(u,\boldsymbol{\phi}) &= f(x(u,\phi,t)) \\ \frac{\mathrm{d}x(u,\boldsymbol{\phi},t)}{\mathrm{d}t} &= g(x(u,\boldsymbol{\phi},t),u,\boldsymbol{\phi}) \\ h(x(u,\boldsymbol{\phi},t),u,\boldsymbol{\phi}) &= 0 \\ x(t_{\mathrm{o}}) &= x_{\mathrm{o}} \end{aligned} \tag{1}$$

where $\hat{y}$ is the vector of predictions, while f is a vector function relating the state variables x to the predicted values. $\boldsymbol{\phi}$ represents the vector of uncertain pharmacometric parameters and u is the vector of control variables such as dose amount and administration interval. These control variables are the ones we are interested in optimizing. The vectors of differential and algebraic equations that define the predictive model are g and h, respectively. The vector of initial conditions is given by x_{o}.

Bayesian inference

In accordance with Bayes' theorem (Eq. 2), the posterior belief about the model parameters involves considering the new data obtained from experiments (**D**), and the prior belief

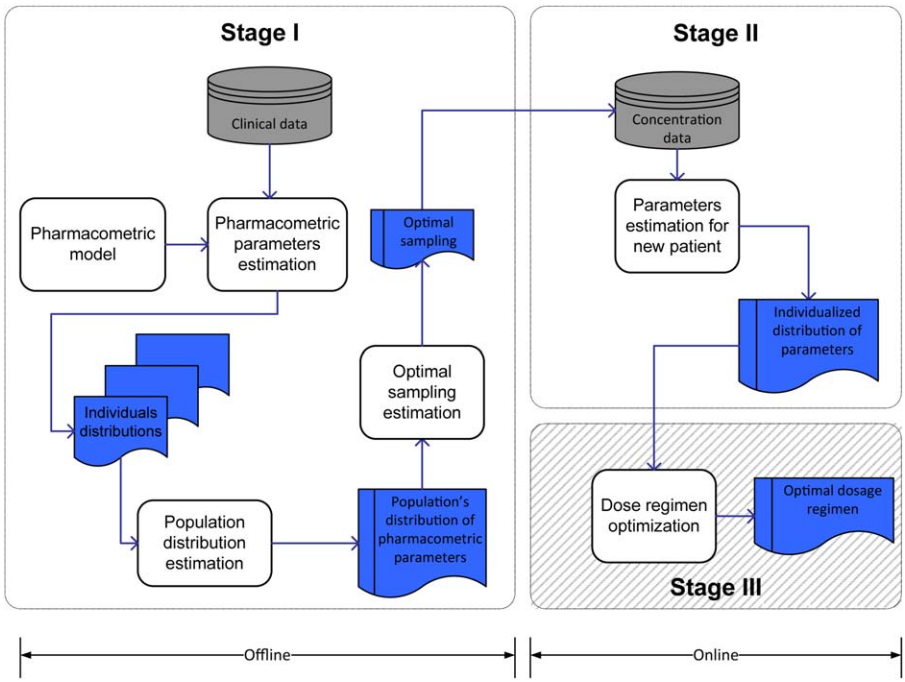


Figure 1. Schematic of the proposed Bayesian framework for dosage regimen individualization.[7]

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

represented by the *prior* distribution $p(\boldsymbol{\phi})$ (e.g., a distribution of pharmacometric parameters for the population). The probability that the experimental outcomes ($\mathbf{D}$) are explained by the model (Eq. 1) and a specific set of values for the pharmacometric parameters ($\boldsymbol{\phi}$) is captured by the likelihood function ($L(\mathbf{D}|\boldsymbol{\phi})$).

$$p(\boldsymbol{\phi}|\mathbf{D}) = \frac{L(\mathbf{D}|\boldsymbol{\phi})p(\boldsymbol{\phi})}{p(\mathbf{D})} \tag{2}$$

As data $\mathbf{D}$ becomes available for each new patient this information is processed to form an individualized posterior distribution $p(\boldsymbol{\phi}|\mathbf{D})$ for the pharmacometric parameters using Bayes' theorem. In this manner, the uncertainty associated with the pharmacometric parameters for that individual is characterized.

The most commonly used method for Bayesian inference is MCMC sampling.[13] One of the advantages of the MCMC approach is that once convergence is achieved the samples are drawn from $p(\boldsymbol{\phi}|\mathbf{D})$. However, MCMC can be very computationally demanding for large and complex models, for example, for those that involve a set of nonlinear differential-algebraic equations.[14] An alternative to MCMC are the family of variational Bayes'(VB) methods. They approximate the posterior distribution using a given structure and their accuracy depend on how appropriate this assumption is. Nonetheless, they have the potential to offer an attractive trade-off between computational cost and accuracy.

Variational Bayes' Approximation. VB methods translate the Bayesian inference into an optimization problem by approximating the posterior distribution using a known distribution form (e.g., a family of Gaussian distributions). The methods are based on the separability of the logarithm of $p(\mathbf{D})$ into two elements as shown in Eqs. 3 to 5.

$$\ln p(\mathbf{D}) = \mathcal{L}(q) + \mathrm{KL}(q||p) \tag{3}$$

$$\mathcal{L}(q) = -\int_{\boldsymbol{\phi}} q(\boldsymbol{\phi}|\theta) \ln\left[\frac{q(\boldsymbol{\phi}|\theta)}{L(\mathbf{D}|\boldsymbol{\phi})p(\boldsymbol{\phi})}\right] \mathrm{d}\boldsymbol{\phi} \tag{4}$$

$$\mathrm{KL}(q||p) = \int_{\boldsymbol{\phi}} q(\boldsymbol{\phi}|\theta) \ln\left[\frac{q(\boldsymbol{\phi}|\theta)}{p(\boldsymbol{\phi}|\mathbf{D})}\right] \mathrm{d}\boldsymbol{\phi} \tag{5}$$

Here, $q(\boldsymbol{\phi}|\theta)$ is the parametric probability distribution that approximates the posterior distribution $p(\boldsymbol{\phi}|\mathbf{D})$. θ is the set of parameters characterizing q (e.g., mean and covariance for a Gaussian distribution). If $q(\boldsymbol{\phi}|\theta)$ is free to be any probability function, then the maximum lower bound ($\mathcal{L}$) is obtained when the Kullback-Leibler (KL) divergence is zero. This occurs when $q(\boldsymbol{\phi}|\theta)$ exactly matches the posterior distribution.[15] Then, minimizing the KL divergence is equivalent to maximizing the lower bound ($\mathcal{L}$). Therefore, the set of parameters θ is determined as that which maximizes $\mathcal{L}$. Laínez et al. (submitted) develop a decomposition strategy to deal with the variational inference of models that are described by a set of DAE (Eq. 1). For purposes of the present work, the VB approximation is advantageous as it generates a probability density function for the characterization of the uncertain parameters instead of the collection of samples that MCMC methods provide. This facilitates the selection of the scenarios to be incorporated into the stochastic optimization as will be later discussed.

The Dosage Regimen Individualization Problem

The goal of dosage regimen individualization is to select the dose amount (*Dose*) and/or administration interval (τ), which are represented by the control variables u in the general pharmacometric model, so that the risk of drug exposure (e.g., drug concentration in blood, areas under the curve

concentration-time) departing from the desired therapeutic window is minimized. The therapeutic window is defined as a range of drug exposure $\left[\mathbf{TW}^l, \mathbf{TW}^u\right]$, which is below a threshold defining an acceptable toxic level $\mathbf{TW}^u$ and above a threshold defining a minimum acceptable level of efficacy $\mathbf{TW}^l$. In order to mathematically pose this problem, a binary variable $\boldsymbol{\rho}$ is defined as shown in Eq. 6. This variable takes the value of 0 if the predicted drug exposure $\hat{y}$ is inside the therapeutic window, and 1 otherwise. Then, the dosage regimen optimization is formulated as given by Eq. 7 where the risk is defined as the expected value of $\boldsymbol{\rho}$.

$$\boldsymbol{\rho}(u,\boldsymbol{\phi}) = \begin{cases} 0, & \text{if } \mathbf{TW}^l \le \hat{y}(u,\boldsymbol{\phi}) \le \mathbf{TW}^u \\ 1, & \text{otherwise} \end{cases} \tag{6}$$

$$\min_u \left\{ Risk = \int_{\boldsymbol{\phi}} p(\boldsymbol{\phi}|\mathbf{D})\boldsymbol{\rho}(u,\boldsymbol{\phi})\mathrm{d}\boldsymbol{\phi} \right\} \tag{7}$$

subject to Eqs. 1 and 6

A discretized stochastic optimization

For computational purposes, an approximation to the optimization problem represented by Eq. 7 can be readily constructed, which translates the problem into an MINLP. To do so, a discretization of time t and of the integral involving the probability distribution $p(\boldsymbol{\phi}|\mathbf{D})$ is carried out. The former is accomplished using orthogonal collocation on finite elements and the latter by using a scenario based approach. For this problem, each scenario s represents a different set of pharmacometric parameters.

Orthogonal Collocation Approach. The idea behind this well-known approach is to approximate the profiles of state variables (x) by subdividing their domain into a set of finite elements and using Lagrange polynomials to describe the profile within each element. Then, the state profile is further discretized using a finite number of collocation points within each element. For further details the interested reader is directed to the work of Cuthrell and Biegler.[16] The model described by Eq. 1 can be then discretized as follows

$$\sum_c \boldsymbol{\chi}_{ecs}\dot{\ell}_{ec}\left(t_{ec'}\right) = g\left(\boldsymbol{\chi}_{ec's}, u, \boldsymbol{\phi}_s\right) \qquad \forall e, c', s \tag{8}$$

$$\dot{\ell}_{ec}\left(t_{ec'''}\right) = \frac{\sum_{c' \ne c}\left(\prod_{c'' \ne c'}\left(t_{ec'''} - t_{ec'''}\right)\right)}{\prod_{c' \ne c}\left(t_{ec} - t_{ec'}\right)} \qquad \forall e, c''' \tag{9}$$

$$\sum_c \boldsymbol{\chi}_{e-1,c,s}\ell_{e-1,c}\left(t_{e1}\right) = \boldsymbol{\chi}_{e1s} \qquad \forall e \ge 2, s \tag{10}$$

$$h(\boldsymbol{\chi}_{ecs}, u, \boldsymbol{\phi}_s) = 0 \qquad \forall e, c, s \tag{11}$$

$$\boldsymbol{\chi}_{11s} = x_o \tag{12}$$

$$\hat{y}_{is} = f\left(\sum_c \boldsymbol{\chi}_{ecs}\ell_{ec}(t_i)\right) \qquad \forall i, e \in E_i, s \tag{13}$$

Equations 8 and 9 are for the discretization of the differential equations. Equation 10 imposes the state profile continuity between elements. The algebraic constraints of the model are evaluated only in the collocation points as expressed by Eq. 11, while Eq. 12 fixes the initial conditions at the first collocation point on the first element. Equation 13 computes the predictions for selected times. E_i represents the element in which t_i is located.

Risk Formulation. As defined in Eq. 7, *Risk* is the probability that the drug exposure is outside the therapeutic window. A related metric usually employed by clinicians is confidence which is understood as the probability that the drug exposure lies within the therapeutic window (i.e., 1-*Risk*). One disadvantage of optimizing these metrics is that they require the utilization of discrete variables, thus increasing the complexity of the optimization problem. In addition, the *Risk* obtained using a scenario based approach is sensitive to the reduced number of scenarios considered as a result of the definition of the binary variable $\boldsymbol{\rho}$. However, there have been a number of alternative metrics proposed for quantifying the risk associated with a decision made under uncertainty, among them robustness, conditional value-at-risk (CVaR), and downside risk. All of them can be formulated without the use of discrete variables.

Robustness[17] or variability is defined as the standard deviation of the selected performance measure, specifically drug exposure in our case. Note that robustness does not involve the consideration of a target which is needed to account for the therapeutic window in the dosage regimen optimization. CVaR[18] is defined as the expected "loss" in the $(1-\alpha)$ quantile. Note that in addition to the "loss" function, which could be defined as the deviation from the therapeutic window, one must choose the confidence level α in order to formulate this metric. The downside risk[19] is defined as the average positive deviation below an established target. It is noteworthy that as discussed by Barbaro and Bagajewicz,[20] it can be demonstrated that the downside risk is the area under the risk curve, thus minimizing downside risk is a surrogate to finding optimal risk solutions [i.e., solutions for problem (7)].

We will use the downside risk metric in our development as it allows consideration of the therapeutic window as a target, and in contrast to CVaR avoids the choice of a confidence level which may derive in different solutions. Notice that our case requires the evaluation both of positive deviations below as well as those above (upside risk) the therapeutic window. Consequently, two continuous variables must be defined ($\boldsymbol{\delta}^l_{is}$ and $\boldsymbol{\delta}^u_{is}$) which represent the positive deviation from the lower and the upper threshold defined by the therapeutic window, respectively. They are defined via Eqs. 14 and 15.

$$\mathbf{TW}^l - \hat{y}_{is} \le \boldsymbol{\delta}^l{}_{is} \qquad \forall i, s \tag{14}$$

$$\hat{y}_{is} - \boldsymbol{\delta}^u{}_s \le \mathbf{TW}^u \qquad \forall i, s \tag{15}$$

The deviation of drug exposure from the therapeutic window for a specific scenario s is thus the summation of these two variables. We define the risk of therapeutic failure (*TRisk*) as the expected value of this positive deviation as stated in Eq. 16, where ϱ_s is the weight/discrete probability associated with scenario s.

$$TRisk = \sum_s \sum_i \varrho_s\left(\boldsymbol{\delta}^u_{is} + \boldsymbol{\delta}^l_{is}\right) \tag{16}$$

The formulation of the risk metric (downside risk) in Eq. 16 results in deviations equal to zero for any drug exposure value that is within the therapeutic window. If evidence exist that the consideration of a specific target for the drug exposure

is more appropriate, the formulation can be easily modified to account for the deviation above and below such a level. Moreover, a combination of these different alternatives could be introduced into the objective function by assigning an additional weight to each type of deviation.

Scenario Selection. The solution quality of scenario-based approaches depends on appropriately selecting the scenarios to be included in the evaluation of Eq. 7. Note that the discretization of probability distributions transforms the integral included in the objective function of Eq. 7 into a weighted sum over a few significant evaluation nodes (i.e., scenarios). To perform this discretization, quadrature rules are a suitable well-known alternative for the univariate case. However, a drawback of using multivariate quadrature rules is that they generally suffer from the curse of dimensionality.

A typical manner of extending univariate quadrature rules to multiple dimensions is to apply a product rule. However, this causes the number of evaluation nodes to increase exponentially. The Smolyak rule or sparse grids use a different approach to extend univariate rules to multiple dimensions such that the number of evaluation points increases polynomially instead of exponentially. The reader is referred to Heiss and Winschel[21] for details. To show the reduction in computational cost that can be achieved using sparse grids, a comparison of three different multidimensional quadrature rules is presented in Table 1.

We use Smolyak rules to select the scenarios to be included in the stochastic optimization. As previously mentioned, quadrature rules in general approximate integrals using a weighted sum of the integrand function over some prespecified nodes. In our case, the integrand in Eq. 7 incorporates the posterior probability of parameters $p(\boldsymbol{\phi}|\mathbf{D})$. This makes more convenient representing the posterior distribution as a parametric probability density function rather than a sample from it. Thus, to facilitate formulating the dosage regimen optimization the variational approach is suggested for carrying out the Bayesian inference and characterizing the variability of the pharmacometric parameters.

Then, an approximation to the optimization problem represented by Eq. 7 can be mathematically posed as follows

$$\min_{u} \; TRisk$$

subject to Eqs. 8 to 16

Case Studies

Two case studies are presented next to demonstrate the capabilities of the proposed optimization approach. In both of them compartmental PK models are used to describe drug concentrations. They implicitly assume perfect mixing within the compartment/organ of interest.

In these studies, the optimization problems have been implemented using GAMS.[22] The CONOPT3[23] and SBB[24] solvers have been used for solving the NLP's and MINLP's, respectively.

Table 1. Number of Evaluation Nodes for Quadrature Rules with a Level of Accuracy $l = 4$

	Product Rule	Smolyak Rule	
Dimensions (d)	Gaussian	Gaussian	Kronrod–Patterson
1	4	4	7
5	1024	286	151
10	1,048,576	1771	1201

The accuracy level is associated with the degree of the polynomial that can be accurately approximated using the univariate rule.

A Smolyak rule derived from univariate Kronrod-Patterson quadratures with Gaussian weight[25] is used for the selection of the scenarios s to be considered in the stochastic optimization. These rules are defined for standard multivariate Gaussian distributions. Then, the following transformation based on the spectral decomposition of the covariance matrix is required

$$\Sigma = \Gamma \Lambda \Gamma^{\mathrm{T}} \tag{17}$$

$$\boldsymbol{\Lambda}^{1/2} \boldsymbol{\phi}_s^{\mathrm{Smolyak}} = \Gamma\left(\boldsymbol{\phi}_s - \bar{\boldsymbol{\phi}}\right) \qquad \forall s \tag{18}$$

where $\bar{\boldsymbol{\phi}}$ and Σ are the mean vector and covariance matrix for the (approximate) posterior distribution of parameters $p(\boldsymbol{\phi}|\mathbf{D})$. Γ is a square matrix whose columns are the covariance eigenvectors and Λ is a diagonal matrix whose elements are its corresponding eigenvalues. $\boldsymbol{\phi}_s^{\mathrm{Smolyak}}$ are the evaluation nodes corresponding to a standard Gaussian distribution derived from a Smolyak rule.

For comparison purposes, the distributions obtained from an MCMC approach, which provides the most accurate representation of the uncertainty of the parameters, are also used to find the optimal dosage regimen. This is done by using a grid search algorithm on the dose variable. This strategy has been implemented using the R[26] software. The DAE models are solved using the deSolve[27] package and have been coded in C++ and integrated with R using the Rccp[28] and the inline[29] packages.

All problems instances have been executed in an Intel i7 2.66 GHz computer.

Gabapentin

The first case study deals with individualizing the dosage regimen of gabapentin which is a drug prescribed to treat neuropathic pain. Data for 19 (G1-G19) patients have been collected during the clinical studies carried out by Urban et al.[30] The therapeutic window for gabapentin has been established[31,32] in terms of the drug concentration: the window is defined by the lower and upper bounds, 2 and 10 μg/cm^3, respectively.

The optimal dosage regimen for gabapentin consist of a dose amount (*Dose*) and an administration interval (τ) that confines the drug concentration to the therapeutic window when steady-state is reached. The pharmacokinetic model used for predicting plasma concentrations is a steady-state multidose one compartment model with first-order absorption. The analytical form resulting from an integration of the simple DAE model is as follows

$$\hat{C}(t) = Dose\left(\frac{F}{V}\right)\frac{k_a}{k_a - k_e}\left[\frac{\exp\left(-k_e(t-t_o)\right)}{1-\exp\left(-k_e\tau\right)} - \frac{\exp\left(-k_a(t-t_o)\right)}{1-\exp\left(-k_a\tau\right)}\right] \tag{19}$$

where F [%] is the bioavailability (i.e., the portion of the dose that reaches the blood stream), $V[\mathrm{cm}^3]$ is the volume of distribution, $Dose\,[\mathrm{mg}]$ is the administered dose, $k_e\left[\mathrm{h}^{-1}\right]$ is the elimination rate, $k_a\left[\mathrm{h}^{-1}\right]$ is the absorption rate, $t_o[\mathrm{h}]$ is the lag time, $t[\mathrm{h}]$ is the time after administration, and τ the interval of administration.

Four parameters ($\boldsymbol{\phi}$), $(F/V)[\mathrm{cm}^{-3}]$, k_e, k_a, and t_o are considered to be uncertain in this model. They are characterized using a Gaussian multivariate probability distribution which is the result of applying the VB method to the concentration data obtained for each patient. Note that there is a different

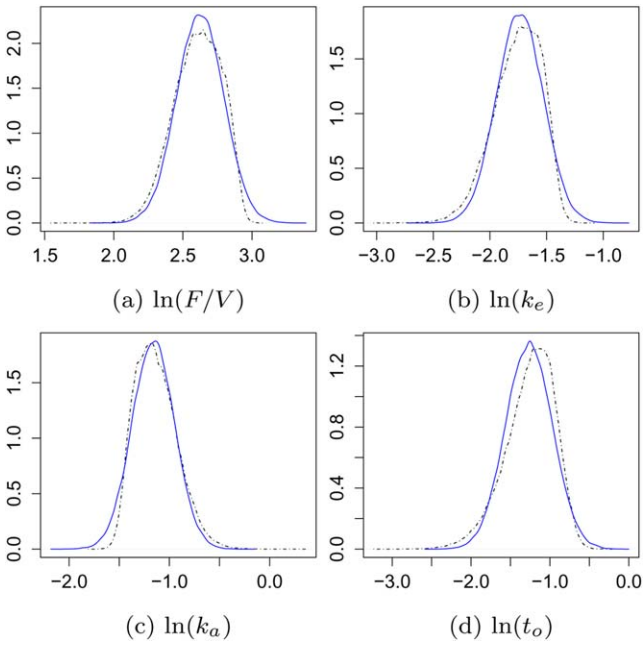


Figure 2. Marginal distributions of the PK parameters for oral administration of gabapentin for patient G11 (continuous line: VB, dashed line: MCMC).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

probability distribution for each patient. The marginal distributions for the PK parameters of patient G11 are shown in Figure 2. Note that the Gaussian distribution obtained using the VB method is a good approximation to the posterior given by the MCMC approach.

The optimization variables for this problem are *Dose* and τ. Note from Eq. 19 that the concentration approaches a constant value with time as τ tends to zero. Consequently, the optimal τ is given by the shortest possible administration interval. We will fix the lower bound for τ at three convenient intervals for the patient to follow, namely 4 h, 8 h, and 12 h, and find their corresponding optimal *Dose*.

Tables 2 and 3 present the optimal dosage regimens that are obtained with scenarios derived from quadrature rules with different levels of accuracy for patients G6 and G11, respectively. An approximate value for the risk is provided as a means of comparison between different optimal dosage regimens. Once the optimal dose is found, the approximate risk is computed from a sample of the parameters drawn from their

Table 2. Individualized Dosage Regimens of Gabapentin for Different Accuracy Levels for Patient G6

Interval (h)	Dose Amount (mg)	Approximate Risk (%)	Accuracy Level	CPU s
4	524	0.1	5	1.0
	516	0.1	6	5.1
	508	0.1	7	18.1
	500	0.1	8	61.2
8	1048	2.9	5	0.9
	1033	2.7	6	5.4
	1016	2.7	7	22.8
	1000	2.6	8	74.9
12	1573	22.6	5	0.9
	1549	20.9	6	7.8
	1524	19.0	7	21.5
	1500	17.3	8	77.7

Table 3. Individualized Dosage Regimens of Gabapentin for Different Accuracy Levels for Patient G11

Interval (h)	Dose Amount (mg)	Approximate Risk (%)	Accuracy Level	CPUs
4	302	≈ 0.0	5	1.1
	301	≈ 0.0	6	8.3
	300	≈ 0.0	7	32.5
	299	≈ 0.0	8	129.2
8	605	≈ 0.0	5	1.0
	602	≈ 0.0	6	8.7
	599	≈ 0.0	7	37.2
	598	≈ 0.0	8	135.5
12	907	≈ 0.0	5	2.5
	902	≈ 0.0	6	8.4
	899	≈ 0.0	7	32.8
	896	≈ 0.0	8	120.4

posterior probability distribution. This post-processing takes an average of 2.7 Central Processing Unit (CPU) seconds.

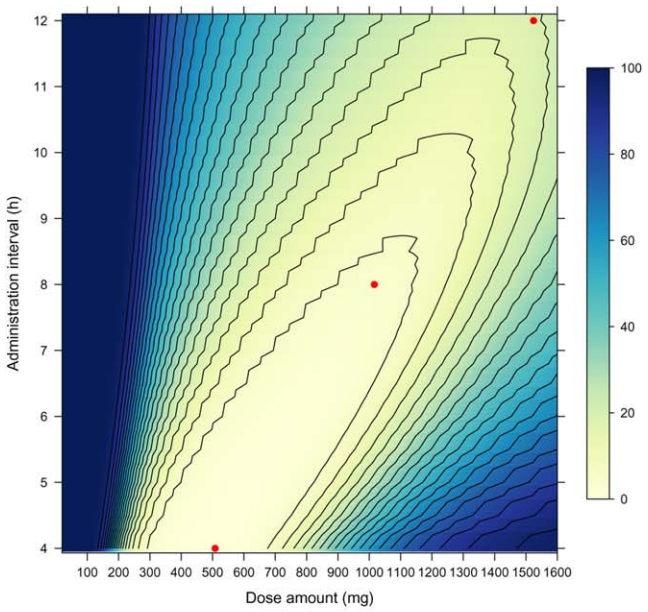
The discretization approximation employed in the proposed approach requires selection of the accuracy level l to use for the numerical integration. Given that a higher level provides a superior accuracy but at the cost of additional computational effort, one should use accuracy levels higher than l only if the difference between the performance metric (i.e., risk) of accuracy levels l and $l - 1$ is sufficient to justify that computational effort. Note that the optimization programs formulated using accuracy level 7 do not generate very different results from those obtained using accuracy level 8. However, the computational time associated with accuracy levels 7 is around 28 CPU seconds, which represent approximately 28% of the time required to find the optimal solution using an accuracy level of 8. These are reasonable times in which to offer support for clinical decision making. In the subsequent discussion of the results for this case study we will utilize those obtained with an accuracy level of 7. The computational time of the optimization programs is given in the fifth column (CPU s) of Tables 2 and 3, and their associated size is presented in Table 4.

In addition, the approximate risk allows comparing optimal doses for different intervals of administration. For instance, patient G6 could take 1524 mg twice a day which has approximately 19.0% risk of drug concentrations departing from the therapeutic window. Such a risk can be reduced by 16.3% if the patient takes 1016 mg every 8 h. Similar therapeutic results could be expected for the different proposed optimal dosage regimens (e.g., 599 mg every 8 h, 899 mg twice a day) for patient G11.

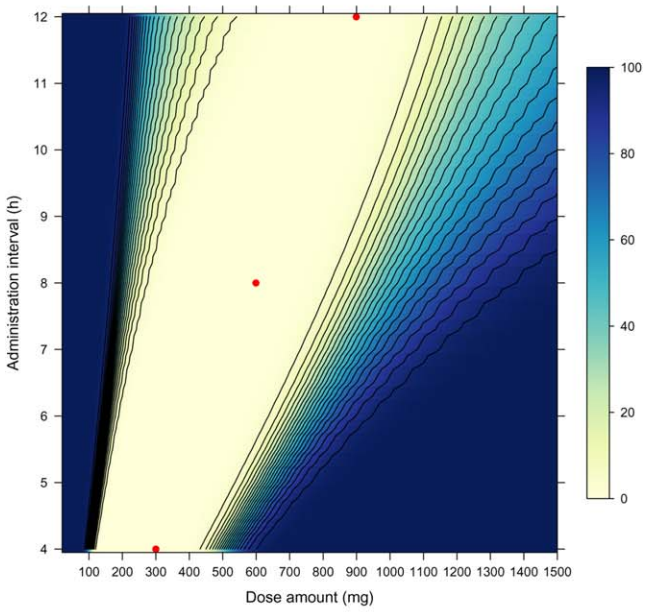
Figure 3 shows the risk level plots obtained using the posterior probability distribution of parameters resulting from the MCMC approach for patients G6 and G11. Contour plots are also drawn in increments of 5%. In this figure, the dosage regimens proposed by the optimization approach are shown as dots. Notice that they are located inside the innermost contour indicating that they are near optimal solutions.

Table 4. Size of Optimization Problems for Different Levels of Accuracy for the Gabapentin Case

Accuracy Level	Scenarios	Number of Equations	Number of Variables
5	401	10,046	10,048
6	993	24,846	24,848
7	2033	50,846	50,848
8	3793	94,846	94,848



(a) Patient G6



(b) Patient G11

Figure 3. Overlaid risk contour and level plots for dosage regimens of patients G6 and G11.

These plots are obtained using the results from the MCMC approach. The individualized dosage regimens proposed by the stochastic optimization are shown as dots. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Tables 5, 6, and 7 present a quantitative assessment of the quality of the solutions proposed by the optimization approach for all 19 patients for the 4 h, 8 h, and 12 h administration interval, respectively. These tables show in the fifth and sixth columns the "true" optimal doses and their corresponding risk both obtained by a grid search for the dose using a step of 5 mg and the probability distribution of parameters resulting from the MCMC approach. Also, the "true" risk (i.e., computed using the posterior distribution from the MCMC) associated with the optimal dose proposed by the stochastic optimization approach and its deviation or gap from the "true" optimal solution are presented in the fourth and last column, respectively. The average gaps are 0.01, 0.01, and 2.94%, while the worst gaps found are 0.11, 0.02, and 11.49% for intervals of administration values of 4, 8, and 12 h, respectively. In general, greater solution gaps are found for longer administration intervals because of the associated increase in concentration variability. Notice that for patient G18, the gap of the stochastic optimization solution for an interval value of 12 h is 11.49%; however, this gap is eliminated for a dosing interval of 8 h. The computational time required to find the "true" optimal value is 126 CPU seconds in average. Using the stochastic optimization approach this burden is reduced by 78%.

Constraints regarding the dose amount are commonly faced when dealing with drugs administered orally. Tablets or capsules are normally manufactured and marketed in only a few different drug quantities, thus restricting the dosing amount to be a multiple of them. This fact requires the stochastic optimization problem to be treated as an MINLP. Suppose that the dose amount can take only values that are multiple of 50 mg. For illustrative purposes, we have solved this form of the problem for patient G6 and shown the results in Table 8. The size of this problem is the same as its NLP counterpart except that one of the variables is integer. The optimal doses obtained using the MCMC—grid search are different from the optimization approach for the 8 h and 12 h interval solutions. The "true" optimal doses are 1050 mg and 1400 mg which results in an associated gap value of 0.02% and 1.66% for an interval of 8 h and 12 h, respectively. The computational time associated with the search grid is 105 CPU seconds on average which is favored by the fact that the dose amount is an integer multiple of 50 mg. Nonetheless, the computational time can be reduced by approximately 70% using the MINLP optimization.

Cyclophosphamide

Cyclophosphamide (CY) is a drug which is used for treating a variety of malignancies as well as an immunosuppressant agent in hematopoietic stem cell transplant therapies.[33] CY undergoes various metabolic transformations, relevantly: CY is activated in the body to Hydroxy-cyclophosphamide (HCY) which is then metabolized to Carboxyethylphosphoramide mustard (CEPM).[5,34] One of the side effects associated with the use of CY as therapy for hematopoietic cell transplant is liver damage. A study carried out by McDonald et al.[34] has found that liver toxicity developed after the therapy is associated with increased exposure to CEPM. The exposure has been quantified as the area under the curve (AUC) of the concentration-time profile over the interval from zero to 48 h. A therapeutic window of 325 ± 25 mol/dm^{3}h has been established for the CEPM AUC, while maintaining HCY AUC over 50 mol/dm^{3}h.

This example is based on the PK data obtained in the clinical trial carried out by Salinger et al.[5] The clinical trial was conducted for 20 (C1-C20) patients. Plasma concentrations of HCY and CEPM were quantified. Blood sampling was done after administration of 45 mg/kg ($Dose_1$) of a CY infusion over approximately 1 h. A second dose ($Dose_2$) was then individualized and administered 24 h after the start of the first CY dose to achieve the target AUCs for CEPM and HCY.

The PK model is described by the following system of nonlinear differential equations

Table 5. Comparison of Dose Amounts Proposed by the Stochastic Optimization and the MCMC Approaches Corresponding to an Administration Interval of 4 h for the Gabapentin Case Study (Accuracy Level = 7)

	NLP			MCMC - Grid Search		
Patient	Dose Amount (mg)	Approximate Risk (%)	Risk (%)	Dose Amount (mg)	Risk (%)	Gap
G1	449	≈ 0.00	0.03	420	0.01	0.02
G2	457	≈ 0.00	≈ 0.00	460	≈ 0.00	0.00
G3	373	≈ 0.00	≈ 0.00	350	≈ 0.00	0.00
G4	323	≈ 0.00	≈ 0.00	300	≈ 0.00	0.00
G5	424	≈ 0.00	≈ 0.00	410	≈ 0.00	0.00
G6	508	0.10	0.27	480	0.23	0.04
G7	421	≈ 0.00	≈ 0.00	400	≈ 0.00	0.00
G8	486	≈ 0.00	≈ 0.00	450	≈ 0.00	0.00
G9	594	≈ 0.00	≈ 0.00	600	≈ 0.00	0.00
G10	481	≈ 0.00	≈ 0.00	480	≈ 0.00	0.00
G11	300	≈ 0.00	0.01	265	≈ 0.00	0.01
G12	404	≈ 0.00	0.12	310	0.01	0.11
G13	314	≈ 0.00	≈ 0.00	305	≈ 0.00	0.00
G14	294	≈ 0.00	0.02	245	≈ 0.00	0.02
G15	372	≈ 0.00	0.01	315	≈ 0.00	0.01
G16	282	≈ 0.00	≈ 0.00	270	≈ 0.00	0.00
G17	352	≈ 0.00	≈ 0.00	340	≈ 0.00	0.00
G18	537	≈ 0.00	≈ 0.00	500	≈ 0.00	0.00
G19	430	≈ 0.00	≈ 0.00	415	≈ 0.00	0.00
Average gap	0.01					

$$
\begin{aligned}
&\frac{d\hat{A}_1(t)}{dt} = k_o - \frac{CL_{non}}{V_{cy}}\hat{A}_1(t) - \frac{CL_{ind}}{V_{cy}}\hat{A}_1(t)\hat{A}_2(t)\\
&\frac{d\hat{A}_2(t)}{dt} = k_{enz}\left(\frac{\frac{E_{max}}{V_{cy}}\hat{A}_1(t)}{1+\frac{\hat{A}_1(t)}{V_{cy}}}\right) - k_{enz}\hat{A}_2(t)\\
&\frac{d\hat{A}_3(t)}{dt} = \frac{CL_{ind}}{V_{cy}}\hat{A}_1(t)\hat{A}_2(t) - k_{hcy}\hat{A}_3(t) - k_{34}\hat{A}_3(t)\\
&\frac{d\hat{A}_4(t)}{dt} = k_{34}\hat{A}_3(t) - k_{cepm}\hat{A}_4(t)\\
&k_0 = \begin{cases} \text{Dose}_1/\text{h}, & \text{if } 0\text{h} \le t \le 1\text{h} \\ \text{Dose}_2/\text{h}, & \text{if } 24\text{h} \le t \le 25\text{h} \\ 0, & \text{otherwise} \end{cases}\\
&\hat{A}_1(0)=0; \hat{A}_2(0)=1; \hat{A}_3(0)=0; \hat{A}_4(0)=0
\end{aligned}
\tag{20}
$$

where $\hat{A}_1(t)$ [μmol/kg], $\hat{A}_3(t)$ [μmol/kg], and $\hat{A}_4(t)$ [μmol/kg] are amounts of CY, HCY, and CEPM, respectively, and $\hat{A}_2(t)$ [dimensionless] is the amount of an autoinducible enzyme. k_o is the infusion rate. CL_{non} [dm^3/(h-kg)] and CL_{ind} [dm^3/(h-kg)] are noninducible and inducible clearances for CY, respectively; whereas, k_{enz} [h^{-1}], k_{hcy} [h^{-1}], and k_{cepm} [h^{-1}] are the elimination rate constants of enzyme, HCY and CEPM, respectively. k_{34} [h^{-1}] is the conversion rate from HCY to CEPM. The previous six parameters are considered to be uncertain. Again, they are characterized for each patient using a Gaussian multivariate probability distribution obtained as the result of applying the VB method to the respective PK data. Figure 4 shows the marginal distributions for these parameters corresponding to patient C11. One can notice from this Figure that the normality assumption for the posterior probability distribution does not hold very well in

Table 6. Comparison of Dose Amounts Proposed by the Stochastic Optimization and the MCMC Approaches Corresponding to an Administration Interval of 8 h for the Gabapentin Case Study (Accuracy Level = 7)

	NLP			MCMC - Grid Search		
Patient	Dose Amount (mg)	Approximate Risk (%)	Risk (%)	Dose Amount (mg)	Risk (%)	Gap
G1	898	0.36	0.54	895	0.54	0.00
G2	914	≈ 0.00	0.02	915	0.02	0.00
G3	745	≈ 0.00	≈ 0.00	740	≈ 0.00	0.00
G4	646	≈ 0.00	≈ 0.00	620	≈ 0.00	0.00
G5	848	≈ 0.00	0.01	800	≈ 0.00	0.01
G6	1016	2.41	4.00	1030	3.99	0.01
G7	843	≈ 0.00	≈ 0.00	875	≈ 0.00	0.00
G8	971	≈ 0.00	0.04	930	0.02	0.01
G9	1188	≈ 0.00	≈ 0.00	1135	≈ 0.00	0.00
G10	962	≈ 0.00	≈ 0.00	940	≈ 0.00	0.00
G11	599	≈ 0.00	0.01	515	≈ 0.00	0.01
G12	808	0.65	0.87	790	0.86	0.02
G13	628	≈ 0.00	≈ 0.00	625	≈ 0.00	0.00
G14	589	≈ 0.00	0.04	550	0.03	0.01
G15	743	≈ 0.00	0.01	605	≈ 0.00	0.01
G16	563	≈ 0.00	≈ 0.00	550	≈ 0.00	0.00
G17	704	≈ 0.00	≈ 0.00	695	≈ 0.00	0.00
G18	1074	0.89	1.32	1075	1.32	0.00
G19	861	≈ 0.00	≈ 0.00	835	≈ 0.00	0.00
Average gap	0.01					

Table 7. Comparison of Dose Amounts Proposed by the Stochastic Optimization and the MCMC Approaches Corresponding to an Administration Interval of 12 h for the Gabapentin Case Study (Accuracy Level = 7)

	NLP			MCMC - Grid Search		
Patient	Dose Amount (mg)	Approximate Risk (%)	Risk (%)	Dose Amount (mg)	Risk (%)	Gap
G1	1347	21.93	18.20	1225	13.51	4.69
G2	1371	33.84	24.15	1180	14.18	9.97
G3	1118	26.80	18.57	1025	9.39	9.18
G4	969	≈ 0.00	0.02	945	0.02	0.00
G5	1272	0.26	0.56	1270	0.56	0.00
G6	1524	19.06	19.18	1395	16.75	2.43
G7	1264	2.69	2.80	1230	2.21	0.59
G8	1457	5.03	5.47	1405	4.87	0.60
G9	1781	12.51	9.70	1575	0.64	9.07
G10	1443	19.61	16.45	1345	10.49	5.96
G11	899	0.01	0.07	880	0.06	0.01
G12	1212	15.74	14.08	1130	12.31	1.77
G13	942	0.10	0.28	940	0.28	0.00
G14	883	0.20	0.71	870	0.69	0.02
G15	1115	0.02	0.12	1125	0.11	0.00
G16	845	0.40	0.49	860	0.45	0.04
G17	1056	≈ 0.00	≈ 0.00	1065	≈ 0.00	0.00
G18	1611	40.05	32.84	1340	21.35	11.49
G19	1291	0.78	1.31	1275	1.27	0.04
Average gap	2.94					

this case. However, it will exemplify the existing trade-off between accuracy and computational expense.

The additional parameters, V_{cy} [dm^3/kg], E_{max} [dimensionless], and EC_{50} [μmol/dm^3] are assumed constant at 0.694 dm^3/kg, 5, and 0.6 μmol/dm^3, respectively. E_{max} and EC_{50} are enzyme synthesis parameters. V_{cy} [dm^3/kg] is the apparent volume of distribution of CY. Plasma concentrations are predicted by dividing the amounts by their respective volumes of distribution. The HCY and CEPM volumes (V_{hcy}, V_{cepm}) are assumed equal to 0.013 dm^3/kg. Finally, $Dose_2$ is the dose infused after 24 h and is the optimization variable.

We use the orthogonal collocation approach for this optimization model. Ten elements including four collocation points corresponding to the roots of a Legendre polynomial are utilized. Note that the AUC's can be easily formulated using the collocation points which are related to a Gauss-Legendre quadrature rule as shown in Eq. 21. Here, w_c is the weight associated with collocation point c.

$$\mathbf{AUC}_s = \sum_e \left(t_{e+1,1} - t_{e,1}\right) \sum_c w_c \boldsymbol{\chi}_{ecs} \qquad \forall s \tag{21}$$

Table 9 presents the optimal second dose for different accuracy levels for patients C5, C11, and C19. The problem sizes and computational times are also shown in this Table. Note that the solutions computed with accuracy levels 5 and 6 are very similar. However, the computational cost associated with an accuracy level of 5 represents in average just 15% of the cost of accuracy level 6. Therefore, the accuracy level of 5 is selected for further analysis.

Figure 5 depicts the risk vs. $Dose_2$ plots for patients C5, C11, and C19 which are generated using the posterior distribution of parameters resulting from the MCMC approach. The second dose amounts proposed by the stochastic optimization approach are shown as dots. Notice that they are located very close to the "true" minimum risk dose. As in the study carried out for gabapentin, the optimization solutions for all patients are listed in Table 10. The "true" risk (i.e., computed using the posterior distributions derived from the MCMC approach) associated with the doses proposed by the stochastic optimization has been determined and listed in the third column of Table 10. Again, the "true" optimal doses and their corresponding risk are obtained using a grid search with a step of 5 mg/kg for the $Dose_2$ and the probability distribution of parameters resulting from the MCMC approach. They are presented in the fourth and fifth column, respectively.

Finally, the last column of Table 10 shows the optimal gap between the doses proposed by the stochastic optimization and the MCMC approach. The worst case is that of patient C14 who could see the risk of departing from the therapeutic window reduced by 2.31% by having a second infusion of CY equal to 20.0 mg/kg instead of 10.7 mg/kg. On average, the optimal risk gap corresponding to the proposed stochastic optimization is 0.35%. Also note that some improvements can be gained in some patients because of using a continuous variable for $Dose_2$ in the stochastic optimization instead of a discrete step (5 mg/kg) as the MCMC approach does. For instance, the second dose recommended by the stochastic optimization for patient C5 is 23.3 mg/kg which results in a reduction in risk of 6.76% compared to the solution proposed by the MCMC—search grid approach. It is noteworthy that for this case study the average computational burden associated with the MCMC—search grid is 11107.3 CPU seconds, while the one associated with the stochastic optimization is 148.8 CPU seconds. This represents a reduction of two orders of magnitude or 98.7% in terms of computational time.

Table 8. Individualized Dosage Regimens of Gabapentin for Patient G6 Considering Dose Amount as a Multiple of 50 mg

Interval (h)	Dose Amount (mg)	Approximate Risk (%)	Risk (%)	CPU s
4	500	0.09	0.2	27.6
8	1000	2.40	4.06	30.2
12	1500	16.40	18.43	29.1

Concluding Remarks

This work proposes a stochastic optimization approach to the individualization of dosage regimens. This framework can accommodate in the optimization model the distributions

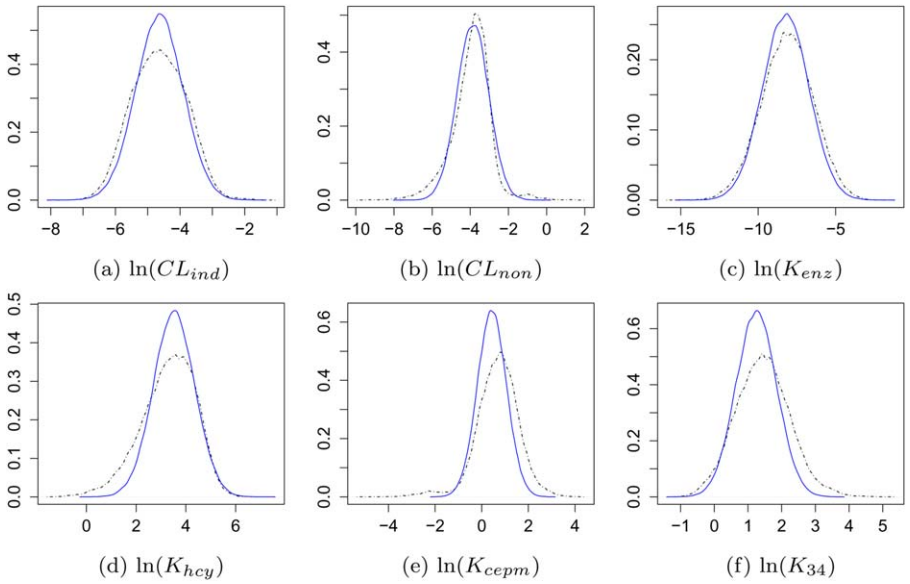


Figure 4. Marginal distributions for the parameters of the CY pharmacokinetic model for patient C11 (continuous line: VB, dashed line: MCMC).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

of pharmacometric parameters that are obtained using Bayesian inference. This is accomplished by selecting a set of scenarios using an efficient multivariate quadrature rule, the Smolyak rule. The case studies presented herein model the uncertainty using a Gaussian distribution and a corresponding Smolyak rule with Gaussian weights. Nevertheless, other more general unweighted Smolyak rules can be used for other type of probability density functions. Moreover, the proposed framework is capable of dealing with general pharmacometric models comprised of a set of differential-algebraic equations. As objective function, the side risk has been chosen. This metric can serve as a surrogate for drug exposure risk and offers the advantage that it does not require integer variables for its formulation.

Solutions obtained from a grid search algorithm which uses posterior distributions of parameters generated by a MCMC approach are used as benchmark. Numerical results from two case studies, the individualization of dosage regimen for gabapentin and CY, have shown that low risk optimal solutions can be obtained by the proposed framework. Indeed, the optimal gap from the "true" solutions (i.e., the one obtained by the search grid algorithm) is on average just 1.0% and 0.35% for the gabapentin and CY case, respectively.

With regard to the computing savings, the proposed optimization framework can reduce very significantly the required computational time compared to the grid-search algorithm when applied to complex models while producing closely similar results. For the gabapentin case for which the PK model is algebraic, this reduction is around 60%. However, we should point out that even for the search grid approach the computational time is very reasonable (around 100 CPU seconds). For the more complex CY case, the computational time can be reduced from 11,107 to 149 CPU seconds using the proposed optimization framework. Notice that despite the fact that the normality assumption we employ for the VB approach does not hold well for the CY case, as shown in Figure 4, a good compromise between accuracy and computational cost can be achieved. Specifically, the computational cost can be reduced up to 99% if the clinician is willing to accept an optimal gap up to 2.3% for the risk of departing from the therapeutic window. Such a reduction can help to partially alleviate deployment issues preventing the implementation of dosage regimen individualization in clinical practice.

We have assumed that the therapeutic window is fixed and equal for every patient. However, variability in PD has been also recognized in the literature.[2] A systematic approach to these additional uncertainties imposes the need for estimating both the PK and the PD parameters in an integrated fashion. This integration increases the number of uncertain parameters to be estimated and thus the

Table 9. Optimal Second Doses of CY Obtained Using Different Accuracy Levels for Patients C5, C11, and C19

Patient	$Dose_2$ (mg/kg)	Accuracy Level	Scenarios	Number of Equations	Number of Variables	CPU s
C5	22.8	3	64	33,601	31,042	2.4
	23.1	4	207	108,676	100,397	16.2
	23.3	5	520	273,001	252,202	107.3
	24.0	6	1193	626,326	578,607	916.0
C11	40.7	3	73	38,326	35,407	1.6
	41.9	4	245	128,626	118,827	10.6
	43.6	5	615	322,876	298,277	59.6
	44.1	6	1402	736,051	679,972	249.1
C19	109.5	3	73	38,326	35,407	2.9
	111.5	4	244	128,101	118,342	31.0
	112.3	5	611	320,776	296,337	160.0
	112.3	6	1389	729,226	673,667	1211.2

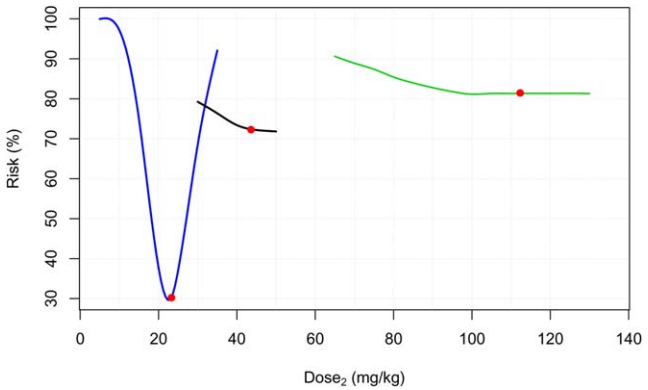


Figure 5. Risk vs. second dose amount curves for, from left, C5, C11, and C19.

These curves were obtained using the MCMC results. The individualized second dose amounts proposed by the stochastic optimization are shown as dots. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

computational burden associated with it. Part of our ongoing research includes the development of computational decomposition alternatives that exploit the structure of PK/PD parameter correlation in integrated models. We believe that the stochastic optimization approach advocated in this work will allow us to handle higher dimensionality parameter spaces whereas with the MCMC approach that becomes computationally prohibitive as suggested by our second case study.

The optimization framework proposed in this work can be easily extended for incorporation into an adaptive stochastic model predictive control scheme for other drugs administered during long periods. Such a control scheme can be suitable for monitoring the dynamics of parameters and adapting them to potential changes as the patient ages and/or disease evolves. The sequential update of the parameter distributions as data is collected can be carried out using the VB framework which is computationally efficient. Then, the posterior distribution derived from this update can be translated into the scenarios to be considered in the stochastic control algorithm.

Since scenario-based stochastic optimization has been extensively used in other domains, the approach proposed in this work which couples (1) Bayesian inference for the uncertainty characterization and (2) sparse grids for scenarios selection may have general relevance beyond the pharmacometrics domain.

Table 10. Comparison of the Individualized Doses of CY Resulting From the Stochastic Optimization and the MCMC Approach

	NLP		MCMC - grid search		
Patient	Dose$_2$ (mg/kg)	Risk(%)	Dose$_2$ (mg/kg)	Risk (%)	Gap
C1	55.1	68.54	60.0	67.80	0.74
C2	75.8	79.00	75.0	78.90	0.10
C3	74.9	78.48	80.0	78.20	0.28
C4	5.9	72.05	5.0	70.70	1.35
C5	23.3	30.24	25.0	37.00	−6.76
C6	3.2	81.20	5.0	82.10	−0.90
C7	58.3	83.33	50.0	83.00	0.33
C8	121.7	81.49	130.0	80.39	1.10
C9	39.0	90.65	30.0	89.70	0.95
C10	60.0	61.02	65.0	60.40	0.62
C11	43.6	72.28	50.0	71.80	0.48
C12	56.5	59.30	60.0	58.20	1.10
C13	11.6	67.93	15.0	66.60	1.33
C14	10.7	72.91	20.0	70.60	2.31
C15	171.9	85.46	165.0	86.20	−0.74
C16	81.3	73.80	90.0	72.10	1.70
C17	166.3	82.19	165.0	82.30	−0.11
C18	95.3	78.58	105.0	77.40	1.18
C19	112.3	81.46	100.0	81.10	0.36
C20	106.8	78.15	120.0	76.60	1.55
Average gap	0.35				

Acknowledgments

The authors gratefully acknowledge support from the United States National Science Foundation (Grant NSF-CBET-0941302), and the collaboration with colleagues Gary Blau and Seza Orçun in this area of research. They also would like to thank University of California, San Francisco for providing the data used in the gabapentin case study.

Notation

c, c', c'', c''' = collocation points
e = finite elements
E_i = element that includes the time point associated with i
i = points (time) of interest
s = scenarios

Parameters

$\mathbf{D}$ = vector/matrix of experimental data
ℓ_{ec} = component of the Lagrange polynomial for element e associated with collocation point c
t_{ec} = time point associated with element e and collocation point c
$\mathbf{TW}^l$ = lower limit for the therapeutic window (i.e., efficacy level)
$\mathbf{TW}^u$ = upper limit for the therapeutic window (i.e., toxicity level)
w_c = weight of the Gauss-Legendre quadrature rule associated with collocation point c
x_o = set of initial conditions
θ = vector of parameters defining the approximate probability distribution q
ϱ_s = weight/discrete probability associated with scenario s
$\boldsymbol{\phi}_s$ = vector of parameters for scenario s
$\bar{\boldsymbol{\phi}}$ = vector of means of a Gaussian distribution of parameters
$\boldsymbol{\phi}_s^{\text{Smolyak}}$ = vector of evaluation nodes from a Smolyak rule for scenario s
Γ = eigenvectors of the variance-covariance matrix
Λ = diagonal matrix of eigenvalues of the variance-covariance matrix
Σ = variance-covariance matrix

Binary Variables

ρ = 0 if the predicted drug exposure is inside the therapeutic window, 1 otherwise

Continuous Variables

$Risk$ = Risk of drug exposure departing from the therapeutic window
$TRisk$ = A surrogate to risk of drug exposure departing from the therapeutic window
u = vector of control/decision variables
x = vector of state variables
$\hat{y}$ = vector of predicted values
δ_{is}^l = positive deviation from the lower threshold of the therapeutic window

δ_{is}^{u} = positive deviation from the upper threshold of the therapeutic window
χ_{ecs} = vector of state values corresponding to finite element e and collocation point c for scenario s

Functions

f = vector function relating the state variables x to the predicted observations
g = vector of differential equations
h = vector of algebraic equations
KL = Kullback-Leibler divergence
$L(\mathbf{D}|\boldsymbol{\phi})$ = likelihood of the data $\mathbf{D}$ given a specific set of parameters
$\mathcal{L}$ = Lower bound for the logarithm of the data marginal distribution
$p(\boldsymbol{\phi})$ = prior distribution for the pharmacometric parameters
$p(\boldsymbol{\phi}|\mathbf{D})$ = posterior distribution provided the pharmacometric data $\mathbf{D}$
q = the approximate probability distribution

Literature Cited

1. Bogle IDL, Allen R, Sumner T. The role of Computer Aided Process Engineering in physiology and clinical medicine. *Comput Chem Eng*. 2010;34:763–769.
2. Neely M, Jelliffe R. Practical, Individualized dosing: 21st century therapeutics and the clinical pharmacometrician. *J Clin Pharmacol*. 2010;50:842–847.
3. Jelliffe R, Schumitzky A, Guilder MV, Liu M, Hu L, Maire P, Gomis P, Barbaut X, Tahani B. Individualizing drug dosage regimens: roles of population pharmacokinetic and dynamic models, Bayesian fitting, and adaptive control. *Ther Drug Monit*. 1993;15:380–393.
4. Mehvar R. Pharmacokinetic-based design and modification of dosage regimens. *Am J Pharm Edu*. 1998;62:189–195.
5. Salinger DH, McCune JS, Ren AG, Shen DD, Slattery JT, Phillips B, McDonald GB, Vicini P. Real-time dose adjustment of cyclophosphamide in a preparative regimen for hematopoietic cell transplant: a Bayesian pharmacokinetic approach. *Clin Cancer Res*. 2006;12(16):4888–4898.
6. McCune JS, Batchelder A, Guthrie KA, Witherspoon1 R, Appelbaum FR, Phillips B, Vicini P, Salinger DH, McDonald GB. Personalized dosing of cyclophosphamide in the total body irradiation-cyclophosphamide conditioning regimen: a phase II trial in patients with hematologic malignancy. *Clin Pharmacol Ther*. 2009;85(6):615–622.
7. Laínez JM, Blau G, Mockus L, Orçun S, Reklaitis GV. Pharmacokinetic based design of individualized dosage regimens using a Bayesian approach. *Ind Eng Chem Res*. 2011;50(9):5114–5130.
8. Bayard D, Milman M, Schumitzky A. Design of dosage regimens: a multiple model stochastic control approach. *Int J Bio-Med Comput*. 1994;36:103–115.
9. Ji S, Zeng Y, Wu P, Lee E. Weighted target interval stochastic control methods with global optimization and their applications in individualizing therapy. *J Pharmacokinet Pharmacodyn*. 2007;34:433–449.
10. Florian JA Jr, Eiseman JL, Parker RS. Nonlinear model predictive control for dosing daily anticancer agents using a novel saturating-rate cell-cycle model. *Comput Biol Med*. 2008;38(3):339–347.
11. Noble SL, Sherer E, Hannemann RE, Ramkrishna D, Vik T, Rundell AE. Using adaptive model predictive control to customize maintenance therapy chemotherapeutic dosing for childhood acute lymphoblastic leukemia. *J Theor Biol*. 2010;264(3):990–1002.
12. Sopasakis P, Sarimveis H. An integer programming approach for optimal drug dose computation. *Comput Methods Programs Biomed*. 2012;108(3):1022–1035.
13. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys*. 1953;21(6):1087–1092.
14. Blau G, Lasinski M, Orcun S, Hsu SH, Caruthers J, Delgass N, Venkatasubramanian V. High fidelity mathematical model building with experimental data: a Bayesian approach. *Comput Chem Eng*. 2008;32(4-5):971–989.
15. Bishop CM. Pattern Recognition and Machine Learning. New York, NY: Springer Science, 2006.
16. Cuthrell JE, Biegler LT. On the optimization of differential-algebraic process systems. *AIChE Journal*. 1987;33(8):1257–1270.
17. Mulvey JM, Vanderbei RJ, Zenios SA. Robust optimization of large-scale systems. *Operation Res*. 1995;43:264–281.
18. Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. *J Risk*. 2000;2:21–41.
19. Eppen GD, Martin RK, Schrage L. A scenario approach to capacity planning. *Operation Res*. 1989;37:517–527.
20. Barbaro AF, Bagajewicz M. Managing financial risk in planning under uncertainty. *AIChE J*. 2004;50:963–989.
21. Heiss F, Winschel V. Likelihood approximation by numerical integration on sparse grids. *J Econometrics*. 2008;144(1):62–80.
22. Rosenthal RE. GAMS—A User's Guide. Washington, DC: GAMS Development Corporation, 2011.
23. Drud A. CONOPT. ARKI Consulting & Development A/S, Bagsvaerd, Denmark, 2011.
24. ARKI Consulting & Development. SBB. ARKI Consulting & Development A/S, Bagsvaerd, Denmark, 2011.
25. Genz A, Keister BD. Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight. *J Comput Appl Math*. 1996;71:299–309.
26. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2008.
27. Soetaert K, Petzoldt T, Setzer RW. Solving differential equations in R. *R J*. 2010;2:5–15.
28. Eddelbuettel D, François R. Rcpp: Seamless R and C++ Integration. *J Stat Software*. 2011;40:1–18.
29. Sklyar O. Inline: Inline C, C++, Fortran function calls from R. Available at: http://cran.r-project.org/web/packages/inline/index.html. 2011, Last accessed September, 2012.
30. Urban T, Brown C, Castro R, Shah N, Mercer R, Huang Y, Brett C, Burchard E, Giacomini K. Effects of genetic variation in the novel organic cation transporter, OCTN1, on the renal clearance of gabapentin. *Clin Pharmacol Ther*. 2008;83(3):416–421.
31. ARUP-Laboratories. ARUP's Laboratory Test Directory—Gabapentin. Available at: http://www.aruplab.com/guides/ug/tests/0090057.jsp, 2010, Last accessed September, 2012.
32. García LL, Shihabi ZK, Oles K. Determination of gabapentin in serum by capillary electrophoresis. *J Chromatogr B*. 1995;669:157–162.
33. Slattery JT, Kalhorn TF, McDonald GB, Lambert K, Buckner CD, Bensinger WI, Anasetti C, Appelbaum FR. Conditioning regimen-dependent disposition of cyclophosphamide and hydroxycyclophosphamide in human marrow transplantation patients. *J Clin Oncol*. 1996;14:1484–1494.
34. McDonald GB, Slattery JT, Bouvier ME, Ren S, Batchelder AL, Kalhorn TF, Schoch HG, Anasetti C, Gooley T. Cyclophosphamide metabolism, liver toxicity, and mortality following hematopoietic stem cell transplantation. *Blood*. 2003;101(5):2043–2048.

Manuscript received Sept. 28, 2012, and revision received Feb. 28, 2013.