Cost-Constrained Optimal Sampling for System Identification in Pharmacokinetics Applications with Population Priors and Nuisance Parameters

CARLOS OSCAR S. SORZANO,[1,2,3] MARIA ANGELES PÉREZ-DE-LA-CRUZ MORENO,[2] JORDI BURGUET-CASTELL,[1] CONSUELO MONTEJO,[3] ANTONIO AGUILAR ROS[3]

[1]Natl. Center of Biotechnology (CSIC), Madrid, Spain
[2]KineStat Pharma, Madrid, Spain
[3]Univ. San Pablo CEU, Madrid, Spain

Received 18 November 2014; revised 30 January 2015; accepted 2 February 2015

Published online 9 March 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.24417

ABSTRACT: Pharmacokinetics (PK) applications can be seen as a special case of nonlinear, causal systems with memory. There are cases in which prior knowledge exists about the distribution of the system parameters in a population. However, for a specific patient in a clinical setting, we need to determine her system parameters so that the therapy can be personalized. This system identification is performed many times by measuring drug concentrations in plasma. The objective of this work is to provide an irregular sampling strategy that minimizes the uncertainty about the system parameters with a fixed amount of samples (cost constrained). We use Monte Carlo simulations to estimate the average Fisher's information matrix associated to the PK problem, and then estimate the sampling points that minimize the maximum uncertainty associated to system parameters (a minimax criterion). The minimization is performed employing a genetic algorithm. We show that such a sampling scheme can be designed in a way that is adapted to a particular patient and that it can accommodate any dosing regimen as well as it allows flexible therapeutic strategies. © 2015 Wiley Periodicals, Inc. and the American Pharmacists Association J Pharm Sci 104:2103–2109, 2015

Keywords: irregular sampling; Fisher information matrix; dynamic systems; pharmacokinetic time sampling; Pharmacokinetics/pharmacodynamics; Computational; ADME; Dynamic simulation; Dose-response; Clinical pharmacokinetics

INTRODUCTION

Pharmacokinetics (PK) is the study of the time evolution of the amount of a certain drug in the body as well as its concentration in different tissues and plasma.[1] This evolution is of crucial importance because for many drugs there is a therapeutic window within which the drug is effective (below a certain concentration, the drug has no effect; and above a certain concentration, the drug may become toxic). Following safety recommendations, the therapeutic window is assumed to be the same for all patients. However, each patient has a different response to a certain dose regimen. In fact, drug concentration in plasma can be seen as the output of a nonlinear, causal system with memory whose input is the dose applied at each time. In general, it is accepted that the system belongs to a parametric family of systems and that the response of a particular patient corresponds to a particular choice of system parameters. Consequently, personalizing the therapeutic regimen to a particular patient allows identifying her system parameters and specific dosing regimen. Thus, the expected drug concentration in plasma is within the therapeutic window. This is normally performed in an intensive care unit for certain pathologies and with drugs whose therapeutic window is relatively tight.[2–9]

In order to determine the patient's parameters, we need to give a first dose (similar to a delta function) and monitor the patient's response (equivalent to her impulse response). This monitoring is performed by extracting blood samples from the patient and analyzing the drug concentration in plasma. For cost reasons and to avoid unnecessary inconveniences to the patient, the number of blood extractions is limited. Additionally, for certain drugs, it would be preferable to be able to administer multiple doses as there are parameters that do not "manifest" their effects at low drug concentration.

The goal of this work is to provide a time sampling basis that, on average over a population, minimizes the maximum uncertainty about any of the system parameters and that can accommodate any dosing regimen. We will presume that the distribution of parameters within the general patient population is known. Then, we will use Monte Carlo simulations to determine which would be the distribution of the Fisher's information matrix for any sampling scheme. Then, the sampling scheme will be optimized using a global optimization algorithm (in our implementation a genetic algorithm) so that the maximum uncertainty of the worse determined parameter is minimized. If there is a parameter we are particularly interested in, we can minimize its uncertainty instead.

A similar approach has already been proposed,[10–23] and it is known as D-optimal or C-optimal sampling. Most of these algorithms do differ on the optimization algorithm employed and the use or not of the *a priori* distribution of model parameters. However, our approach differs in a number of points: first, previous approaches presume knowledge of the closed-form solution of the differential equation system being solved, which is not true for any arbitrary dosing regimen; second, our approach easily incorporates random nuisance parameters that do not need to be estimated; third, our goal function is a

Correspondence to: Carlos Oscar S. Sorzano (Telephone: +34-915854510; Fax: +34-915854506; E-mail: coss@cnb.csic.es)

Journal of Pharmaceutical Sciences, Vol. 104, 2103–2109 (2015)

minimax function that minimizes the maximum variance of any of the parameters, instead of a global measurement of the overall variance. The first two points make an important step forward in the design of the optimal sampling point for highly nonlinear systems. Additionally, our approach can be applied to patient-specific parameters instead of providing sampling rules for a general population. This is also an appealing feature of our method as it can be readily used in clinical practice.

METHODS

Most PK models can be described with a first-order linear or nonlinear differential equation of the form

$$\frac{\mathrm{d}\mathbf{C}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{C}, \boldsymbol{\Theta}, \boldsymbol{\alpha}) + \mathbf{g}(t, \mathbf{X}, \boldsymbol{\Theta}, \boldsymbol{\alpha}) \tag{1}$$

where t is the continuous time variable, $\mathbf{C}(t)$ is a vector of concentrations measured at multiple locations (e.g., blood plasma and urine), $\boldsymbol{\Theta}$ is a vector with the model parameters (those that we are interested in determining by the measurement process), $\boldsymbol{\alpha}$is a vector of nuisance parameters (in which we are not interested but that also affect the concentration levels), and $\mathbf{X}(t)$ is the input driving signal [in our case the dose given to the patient as a function of time; note that this dose is also a vector allowing multiple dosage routes (oral, intravenous, ...)].

The objective of system identification is to find the Θ parameters from a set of (t_n, C_n) measurements. This is carried out by least-squares regression of the model above evaluated at the sampling times, producing the predicted observations $[t_n, C(t_n)]$, and comparing these predictions to the actual measurements (t_n, C_n). Measurements are supposed to be independent and normally distributed with zero mean and a variance $\sigma^2{}_C$. The variance on its turn depends on the concentration being measured.[24] Concisely, it depends on the assay sensitivity, AS, and the coefficient of variation, $\mathrm{CV}_{\mathrm{assay}}$,

$$\sigma_C^2 = (\mathrm{AS} + \mathrm{CV}_{\mathrm{assay}} C)^2 \tag{2}$$

It can be proven[25] that the asymptotic maximum-likelihood estimate of the system parameters is unbiased and distributed as a Gaussian

$$\hat{\boldsymbol{\Theta}}_{\mathrm{MLE}} \sim N\left(\boldsymbol{\Theta}_{\mathrm{true}}, I_T^{-1}\right) \tag{3}$$

where I_T is Fisher's information matrix calculated on the N measurements performed at the time points in the set T. Obviously, N must be larger than the number of Θ parameters, otherwise there would not be any spare degree of freedom to perform the regression, and the fitting would become an interpolation problem highly exposed to measurement errors.

The ij-th element of Fisher's information matrix can be calculated as:

$$\begin{aligned} I_{T,ij} &= \sum_{n=1}^{N} \left(\frac{\partial(C_n - C(t_n))}{\partial\Theta_i}\right)^T \Sigma_{\mathbf{C}_n}^{-1} \frac{\partial(C_n - C(t_n))}{\partial\Theta_j} \\ &= \sum_{n=1}^{N} \left(\frac{\partial C(t_n)}{\partial\Theta_i}\right)^T \Sigma_{C_n}^{-1} \frac{\partial C(t_n)}{\partial\Theta_j} \end{aligned} \tag{4}$$

where Σ_{C_n} is a diagonal matrix whose ii-th entry is the variance associated to the i-th concentration measurement at the n-th time point (Eq. (2). If we have some *a priori* distribution for the system parameters, as is the case in the problem addressed in this article, we should incorporate this information into the Fisher's information matrix. For instance, it can be shown[26] that assuming that the parameters are independent and normally distributed amounts to add in the diagonal terms the inverse of the variance of each one of the prior distributions. In this way, the diagonal terms become

$$I_{T,ii} = \frac{1}{\sigma_{\Theta_i}^2} + \sum_{n=1}^{N} \left(\frac{\partial C(t_n)}{\partial\Theta_i}\right)^T \Sigma_{C_n}^{-1} \frac{\partial C(t_n)}{\partial\Theta_j}. \tag{5}$$

We need to calculate the term $\frac{\partial C(t_n)}{\partial\Theta_i}$. For doing so, let us define the sensitivity with respect to the parameter Θ_i as:

$$\mathbf{s}_{\Theta_i} = \frac{\partial\mathbf{C}}{\partial\Theta_i} \tag{6}$$

Obviously, this sensitivity is a vector that depends on t. In Refs. 27 and 28, a similar derivation was performed for the case of scalar, instead of vector, functions. Let us find a differential equation that the sensitivity must satisfy in order to be able to solve for the sensitivity at any time and, in particular, at the time points t_n. For doing so, we differentiate the previous equation with respect to time

$$\frac{\mathrm{d}\mathbf{s}_{\Theta_i}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathbf{C}}{\partial\Theta_i}\right) \tag{7}$$

Assuming that $C(t)$ is a C^2 function, we can interchange the differentiation order (Clairaut's theorem) to get

$$\begin{aligned} \frac{\mathrm{d}\mathbf{s}_{\Theta_i}}{\mathrm{d}t} &= \frac{\partial}{\partial\Theta_i}\left(\frac{\mathrm{d}\mathbf{C}}{\mathrm{d}t}\right) \\ &= \frac{\partial}{\partial\Theta_i}\left(\mathbf{f} + \mathbf{g}\right) \\ &= \frac{\partial f}{\partial\mathbf{C}}\frac{\partial\mathbf{C}}{\partial\Theta_i} + \frac{\partial\mathbf{f}}{\partial\Theta_i} + \frac{\partial\mathbf{g}}{\partial\Theta_i} \\ &= \frac{\partial\mathbf{f}}{\partial\mathbf{C}}\mathbf{s}_{\Theta_i} + \frac{\partial\mathbf{f}}{\partial\Theta_i} + \frac{\partial\mathbf{g}}{\partial\Theta_i} \end{aligned} \tag{8}$$

Note that the term $\frac{\partial\mathbf{f}}{\partial\mathbf{C}}$ is a full matrix, not a vector. This is an ordinary differential equation with the initial value $\mathbf{s}_{\Theta_i}(t_0) = 0$.[27] We may use this equation to determine the vectors $\frac{\partial\mathbf{C}(t_n)}{\partial\Theta_i}$ needed by Fisher's Information matrix above. Note that these vectors depend on our estimate of the system parameters, $\hat{\boldsymbol{\Theta}}$, and the nuisance parameters, $\boldsymbol{\alpha}$, as well as the time sampling points t_n $(n = 1, 2, \ldots, N)$. As these two sets of parameters are random vectors, the sensitivity vectors are also random with a distribution that, in principle, may not be assumed to follow any known distribution (e.g., Gaussian).

As shown in Eq. (3), the uncertainty on the system parameters estimate depend on Fisher's information matrix, which in its turn is also random (as it is calculated using random vectors). So we propose to minimize this uncertainty by choosing a set of N time points, T^* that minimizes the maximum expected

coefficient of variation of the system parameters

$$T^* = \arg\min_T \max_k E\{\mathrm{CV}_k\} = \arg\min_T \max_k E\left\{\frac{\sqrt{\left(I_T^{-1}\right)_{kk}}}{\Theta_k}\right\} \quad (9)$$

where $(A)_{kk}$ represents the kk-th element of the matrix A. If we are particularly interested in minimizing the uncertainty associated to a particular parameter Θ_k, then we could minimize

$$T^* = \arg\min_T E\left\{\frac{\sqrt{\left(I_T^{-1}\right)_{kk}}}{\Theta_k}\right\} \quad (10)$$

In the absence of any *a priori* preference, in the rest of the article, we will stick to the first goal function instead of the second.

We propose to estimate this expected value through a Monte Carlo process by which we estimate the distribution of these random variables. For doing so, for each time set T, we simply need to randomly sample the distributions of the vectors $\boldsymbol{\Theta}$ and $\boldsymbol{\alpha}$, estimate I_T and calculate the coefficient of variation for each system parameter. After repeating this process many times (in our example below, 100 times), we can estimate the mean of the different coefficients of variation and its maximum expected value. Then, we can use a global optimization algorithm to choose the best time set T. In our example below, we use a genetic algorithm[29] as implemented in Matlab Global Optimization Toolbox, but any other global optimizer may be employed. This optimizer constructs a population of candidate samples. Each one consists of a vector with possible sampling times and its length is constrained by the total cost of the measurement (it is in this sense that we refer to our algorithm as cost constrained). Then, the algorithm evaluates the fitness (the value of the goal function in Eq. (9) for each one of the candidates. The more fit they are, the more chances they have of influencing the new generation (offspring) of candidates. This new population of candidates is, again, evaluated and the algorithm is iterated until convergence. We have implemented all the source code needed for this algorithm, except for the genetic algorithm that comes under the name of ga with the standard distribution of Matlab Global Optimization Toolbox, as a set of Matlab routines that are available from the authors upon request.

RESULTS

In order to show the validity of our methodology, we illustrate it with the design of sampling time points for a patient needing phenytoin. Phenytoin is a drug with antiepileptic activity.[30,31] Its therapeutic window is relatively narrow: a concentration of free drug in plasma below 1 mg/L is ineffective (although it takes several days to reach this level because the maximum intake per day is limited to 15 mg/kg) and above 2 mg/L is toxic.[32] Additionally, it has a nonlinear PK in the therapeutic range. For this reason, it is very important to measure the patient system parameters so that the therapy can be carefully adapted. We must distinguish between free drug in plasma and total drug in plasma. The reason is that a fraction of the total amount of drug is bound to plasma proteins, whereas another fraction is freely dissolved in plasma. It is the fraction of free drug that has a therapeutic effect. Additionally, the measurement assays for free drug are much more accurate than those for the total drug (assay sensitivity for the free fraction $\mathrm{AS}_{\mathrm{free}} = 0.1$ mg/L; assay sensitivity for the total drug concentration $\mathrm{AS}_{\mathrm{total}} = 1$ mg/L, see Eq. (2); $\mathrm{CV}_{\mathrm{assay}}$ in both cases).[24]

The system dynamics are defined by a constant rate absorption of the drug in the intestine and an enzymatically mediated degradation.[24] The following first-order differential equation represents this process

$$V_{\mathrm{d}}\frac{\mathrm{d}C(t)}{\mathrm{d}t} = -\frac{V_{\max}C(t)}{K_{\mathrm{m}} + C(t)} + K_0\left[u(t) - u\left(t - \frac{Dsb}{K_0}\right)\right] \quad (11)$$

where $C(t)$ is the total concentration of drug in plasma, V_{d} is the apparent distribution volume (which is normally larger than the volume of plasma because of the binding effect), $V_{\max}$ is the maximum degradation rate, K_{m} is the drug concentration at which half of the maximum degradation rate is attained, K_0 represents the constant rate absorption of the drug, $u(t)$ is the Heaviside step function, D is the administered dose (in milligram), s is the tablet salt factor (drugs are many times given in a salt form because of its better dissolution and storage properties), and b is the bioavailability (not all the administered drug is capable of crossing the intestine and hepatic first pass barriers to reach the blood stream). Note that K_0 refers to the amount of drug effectively reaching the blood stream, and $\frac{Dsb}{K_0}$ is the time to exhaust the tablet content.

As the assay sensitivity for the free drug is much more accurate than that for the total amount of drug (free and bound to plasma proteins), the measurements are aimed to the free drug. The relationship between the concentration of free drug and total concentration of drug is given by[24]

$$C_{\mathrm{free}}(t) = F(t)C(t) \quad (12)$$

where

$$F(t) = \frac{1}{1 + f\left[\mathrm{Cl}_{\mathrm{creatinine}}(t)\right]C_{\mathrm{albumin}}(t)} \quad (13)$$

where $\mathrm{Cl}_{\mathrm{creatinine}}(t)$ is the clearance of creatinine over time and $\mathrm{C}_{\mathrm{albumin}}(t)$ is the serum concentration in albumin over time. The clearance of creatinine is, on its turn, estimated to be[33]

$$\mathrm{Cl}_{\mathrm{creatinine}}(t) = (0.85)^{\mathrm{female}}\frac{(140 - \mathrm{age})\mathrm{LBW}}{72C_{\mathrm{creatinine}}(t)} \quad (14)$$

where female is a variable that takes the value 1 (if the patient is female) or 0 (if it is male), age is patient's age in years, LBW is the lean body weight (this the total body weight minus the fat weight, because phenytoin does not dissolve in fatty tissues; the lean body weight can be measured using some scales that estimate it by using a current; otherwise, it is between 0.9 and 0.7 the total body weight for a nonobese person), and $\mathrm{Cl}_{\mathrm{creatinine}}(t)$ is the serum concentration in creatinine over time. The factor

$f[\mathrm{Cl}_{\mathrm{creatinine}}(t)]$ in Eq. (13) can be calculated as:

$$f(x) = \begin{cases} 10^{-4} & 0 \le x \le 10 \\ 1.5 \cdot 10^{-4} & 10 < x \le 24 \\ 1.6 \cdot 10^{-4} & 24 < x \le 80 \\ 1.9 \cdot 10^{-4} & 80 < x \end{cases} \tag{15}$$

Summarizing, the model on which we will apply our methodology will be

$$\frac{\mathrm{d}C_{\mathrm{free}}(t)}{\mathrm{d}t} = -\frac{1}{V_{\mathrm{d}}}\frac{FV_{\max}C_{\mathrm{free}}(t)}{FK_{\mathrm{m}} + C_{\mathrm{free}}(t)} + \frac{FK_0}{V_{\mathrm{d}}}\left[u(t) - u\left(t - \frac{Dsb}{K_0}\right)\right] \tag{16}$$

where F is a nuisance parameter (the fraction of free drug) calculated using two nuisance parameters (the concentrations of albumin and creatinine). Note that, in principle, it is difficult to predict, with a closed-form formula, the influence of the variability of the nuisance parameters into the variability of the optimal sampling times. However, we may investigate this issue by running the proposed algorithm multiple times, each time reducing the variability of the nuisance parameter, and observing how the optimal sampling times are affected by this.

The whole model has a relatively large number of parameters. Some of them can be accurately measured in a not too invasive way (female, age, and LBW). Some others such as K_0, s, and b are assumed to be fixed (with values $K_0 = 0.833$ mg/min, $s = 0.92$, and $b = 0.84$). Finally, the measurement of parameters such as $\mathrm{Cl}_{\mathrm{albumin}}(t)$ and $\mathrm{Cl}_{\mathrm{creatinine}}(t)$ would increase the cost of the blood tests required to determine the free drug concentration. They will be treated in this example as nuisance parameters for which an *a priori* distribution will be assumed. This leaves V_{d}, $V_{\max}$, and K_{m} as the only patient parameters that need to be measured. Consequently, we need to perform four blood tests in order to find these three parameters by weighted least-squares regression (the weights are given by the concentration dependent variance of each measurement).

The distribution of the nuisance parameters can be found in the medical literature. For example, the albumin concentration is expected to be between 34 and 54 g/L,[34,35] whereas the creatinine concentration is expected to be between 8.8 and 11.0 mg/L for women and between 10.0 and 12.9 mg/L for men.[36] In the following, we will assume that the albumin and creatinine serum concentrations of a given patient do not change over time.

The *a priori* distribution of the kinetic parameters is also known.[24] For instance, the distribution volume can be calculated as

$$V_{\mathrm{d}} = \mathrm{BW}v_{\mathrm{d}} \tag{17}$$

where BW is patient's body weight and v_{d} is the normalized distribution volume between 0.3 and 1.4 L/kg, with a mean of 0.8 L/kg and a SD of 0.16 L/kg. Similarly, the maximum degradation rate can be calculated as

$$V_{\max} = \mathrm{BW}v_{\max} \tag{18}$$

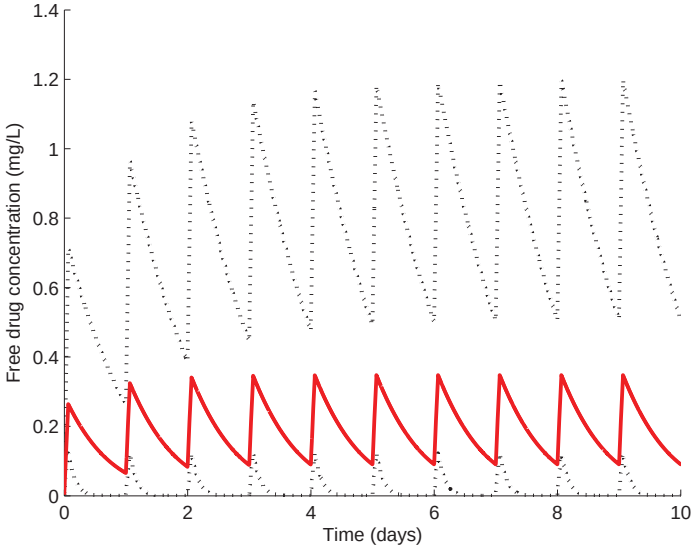

Figure 1. Solid red line: average response of a 40-year-old male patient of 80 kg and 20% of fat weight to 100 mg of phenytoin daily. Dashed lines: minimum and maximum responses according to the distribution of nuisance and system parameters.

where $v_{\max}$ is between 2.48 and 19.84 μg/(kg min), with a mean of 5.46 μg/(kg min) and a SD of 1.63 μg/(kg min). Finally, the concentration at half degradation rate K_{m} is between 2 and 9 mg/L, with a mean of 5.89 mg/L and a SD of 2.95 mg/L.[24] This *a priori* distribution is used, along with the distribution of the nuisance parameters, to calculate the expectation in Eq. (9).

We will exemplify our methodology with a 40-year-old male patient of 70 kg and 20% of fat weight. Typical responses to a tablet of 100 mg of phenytoin are depicted in Figure 1. For the sake of the example, let us say that our plan is to give a patient during 10 days a dose of 100 mg of phenytoin. We have chosen this dosage because in the worse case it does not go above the toxic concentration after 10 days of treatment and it does not exceed the maximal daily dose of 15 mg/kg. However, this dose has to be adjusted to each patient taking into account his gender, age, weight, and body fat.

At this point, we pose the differential equations of the sensitivity functions

$$\begin{aligned} \frac{\mathrm{d}s_{V_{\max}}(t)}{\mathrm{d}t} &= \frac{\partial f}{\partial C_{\mathrm{free}}}(t)s_{V_{\max}}(t) - \frac{1}{V_{\mathrm{d}}}\frac{FC_{\mathrm{free}}(t)}{FK_{\mathrm{m}} + C_{\mathrm{free}}(t)} \\ \frac{\mathrm{d}s_{V_{\mathrm{d}}}(t)}{\mathrm{d}t} &= \frac{\partial f}{\partial C_{\mathrm{free}}}(t)s_{V_{\mathrm{d}}}(t) + \frac{1}{V_{\mathrm{d}}^2}\frac{FV_{\max}C_{\mathrm{free}}(t)}{FK_{\mathrm{m}} + C_{\mathrm{free}}(t)} \\ &\quad - \frac{FK_0}{V_{\mathrm{d}}^2}\left[u(t) - u\left(t - \frac{Dsb}{K_0}\right)\right] \\ \frac{\mathrm{d}s_{K_{\mathrm{m}}}(t)}{\mathrm{d}t} &= \frac{\partial f}{\partial C_{\mathrm{free}}}(t)s_{K_{\mathrm{m}}}(t) + \frac{1}{V_{\mathrm{d}}}\frac{F^2V_{\max}C_{\mathrm{free}}(t)}{\left[FK_{\mathrm{m}} + C_{\mathrm{free}}(t)\right]^2} \end{aligned} \tag{19}$$

where

$$\frac{\partial f}{\partial C_{\mathrm{free}}}(t) = -\frac{F^2K_{\mathrm{m}}V_{\max}}{V_{\mathrm{d}}\left[FK_{\mathrm{m}} + C_{\mathrm{free}}(t)\right]^2}. \tag{20}$$

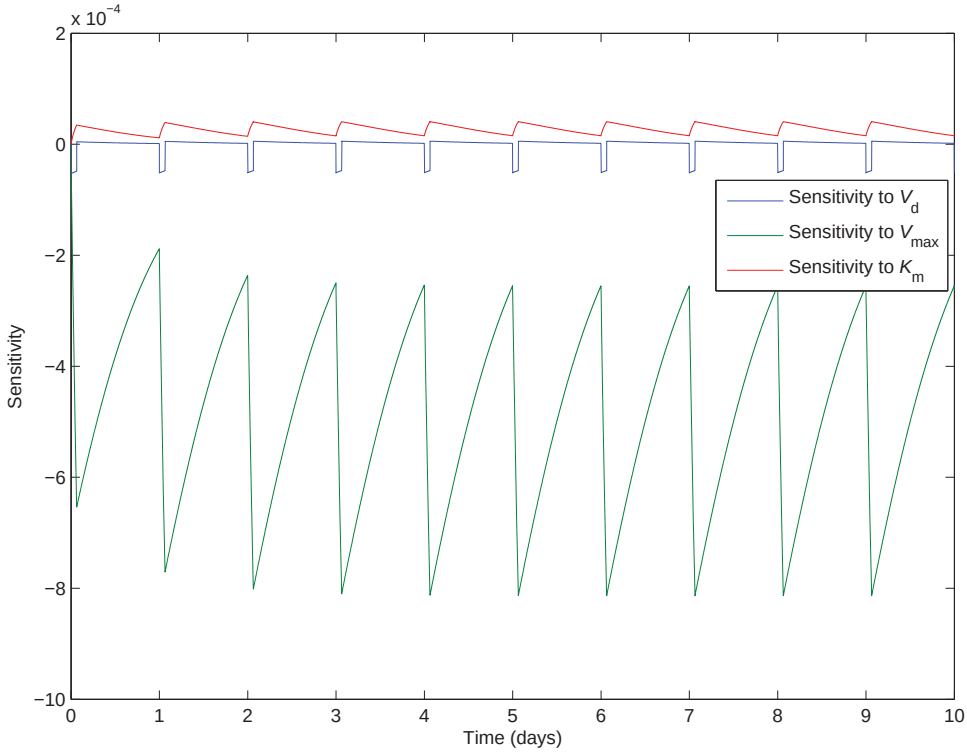

Figure 2. Sensitivity values to each one of the parameters to be determined for an averagely responding person.

Figure 2 shows the sensitivity values for the three parameters V_{max}, V_d, and K_m and the average response of the patient in Figure 1.

We now follow the methodology developed in this article and minimize the maximum coefficient of variation of any of the three system parameters to be determined V_d, V_{max}, and K_m. The methodology suggests to take samples after 7880, 9307, 9439, and 13,894 min after starting the treatment, or what is the same 5.5, 6.5, 6.6, and 9.6 days. The volume of distribution, V_d, can be determined with an average coefficient of variation of 0.2%, K_m with an average coefficient of variation of 7.1%, and V_{max} with an average coefficient of variation of 1233.3%. The reason why V_{max} is so badly determined is that the administered dose (100 mg, daily) causes a plasma concentration of free drug after 10 days of treatment that is far below the concentration needed to induce the maximum degradation.

For comparison purposes, we can easily extend our methodology to the D-optimal design[10–12] and implemented in the free R-package PFIM[37] by calculating

$$T^* = \arg\min_T E\{-|I_T|\} \tag{21}$$

where Fisher's information matrix is calculated as suggested in this article through differential equations for the sensitivities. The D-optimal design for the case above provides the sampling times 649, 6034, 7281, and 11,584 min after starting the treatment. The determinant of the Fisher's information matrix is 14.51 (whereas for our solution, it is 13.29; as expected, the determinant is larger in the D-optimal design as compared with our design, as D-optimal design looks for the sampling times that maximizes this determinant). However, the coefficients of variation for the three parameters are 0.2%, 8.2%, and 1406%, respectively (compared with our coefficients of variation 0.2%, 7.1%, and 1233.3%).

The fact that the concentration is so low that the reaction is far from maximum degradation suggests a two-stage approach to the system identification: in the first stage, of a week of duration, we determine V_d and K_m with a low dose (as we have already performed); in the second stage, we increase the dose, in order to faster reach the therapeutic window, and we take extra samples to better determine V_{max}. We illustrate this second phase in this example. In the second phase, of another 10 days, we increase the dose to 300 mg of phenytoin in the morning and 200 mg after 12 h (the daily maximum for a patient of 70 kg is 1.05 g) so that the therapeutic window can be reached (see Fig. 3). During this second phase, our method suggests to take one sample on day 14.5 (minute 20,987) so that the coefficient of variation of V_{max} drops to 24.5%; the coefficient of variation of V_d drops to about 1.6%, and the one of K_m to about 5.8%. Obviously, these optimal sampling minutes can be modified to fit the clinical needs.

Interestingly, V_{max} is dominating the design because it is the parameter on which samples bring least information about. This is true for our design as well as for D-optimal or any other design based on Fisher's information matrix. However, as shown in this example, we can follow a two-phase design. During the first phase, we use a low dose to have a first estimate of the kinetic parameters while being sure of not reaching the toxic range; in the second phase, we increase the dose to reach more quickly the therapeutic window and increase the information about the most limiting parameter.

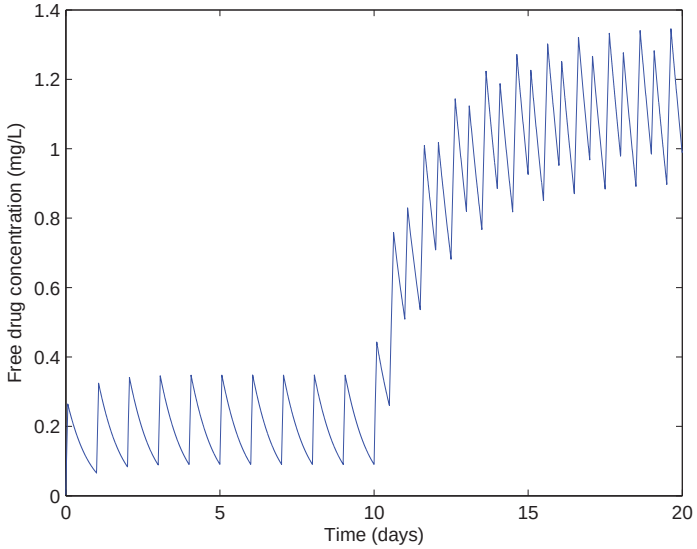

Figure 3. Free drug concentration after 10 days with a daily dose of 100 and 300 mg (morning) and 200 mg (evening) after that period.

CONCLUSIONS

In this article, we have presented a methodology to adapt the sampling time points to any patient taking into account all the *a priori* information available (*a priori* distribution of system and nuisance parameters). The methodology is rather flexible and can deal with any nonlinear PK model, as long as it can be expressed in the form of a differential equation, and any dosing regimen. In this way, we avoid the need of a close-form solution of the model. Notwithstanding, the selection of the time points is performed based on a solid theory using Fisher's information matrix to maximize the information carried by the measurements on the system parameters. Additionally, we can handle nuisance parameters (parameters that affect the concentration but that cannot be measured) through their distribution in a population. Our theory is valid for multiple simultaneous measurements (plasma, urine, ...), although in our example, we have only used plasma concentration. It is our hope that in the future the use of algorithms such as the one presented in this article will help medicine toward a more personalized therapy.

ACKNOWLEDGMENT

Carlos Oscar Sorzano is recipient of a Ramón y Cajal fellowship from the Spanish Ministry of Economy and Competitiveness.

REFERENCES

1. Gabrielsson J, Weiner D. 2007. Pharmacokinetic and pharmacodynamic data analysis: Concepts and applications. 4th ed. Swedish Pharmaceutical Press.
2. Drusano GL, Forrest A, Plaisance KI, Wade JC. 1989. A prospective evaluation of optimal sampling theory in the determination of the steady-state pharmacokinetics of piperacillin in febrile neutropenic cancer patients. Clin Pharmacol Ther 45(6):635–641.
3. Forrest A, Ballow CH, Nix DE, Birmingham MC, Schentag JJ. 1993. Development of a population pharmacokinetic model and optimal sampling strategies for intravenous ciprofloxacin. Antimicrob Agents Chemother 37(5):1065–1072.
4. Baille P, Bruno R, Schellens JH, Webster LK, Millward M, Verweij J, Montay G. 1997. Optimal sampling strategies for Bayesian estimation of docetaxel (taxotere) clearance. Clin Cancer Res 3(9):1535–1538.
5. Bruno R, Hille D, Riva A, Vivier N, ten Bokkel Huinnink WW, van Oosterom AT, Kaye SB, Verweij J, Fossella FV, Valero V, Rigas JR, Seidman AD, Chevallier B, Fumoleau P, Burris HA, Ravdin PM, Sheiner LB. 1998. Population pharmacokinetics/pharmacodynamics of docetaxel in phase ii studies in patients with cancer. J Clin Oncol 16(1):187–196.
6. Schoemaker NE, Mathôt RA, Schöffski P, Rosing H, Schellens JH, Beijnen JH. 2002. Development of an optimal pharmacokinetic sampling schedule for rubitecan administered orally in a daily times five schedule. Cancer Chemother Pharmacol 50(6):514–517.
7. Bullock JM, Smith PF, Booker BM, Loughner J, Capozzi D, McCarthy PL Jr, Shaw LM. 2006. Development of a pharmacokinetic and Bayesian optimal sampling model for individualization of oral busulfan in hematopoietic stem cell transplantation. Ther Drug Monit 28(1):62–66.
8. Delattre IK, Musuamba FT, Nyberg J, Taccone FS, Laterre PF, Verbeeck RK, Jacobs F, Wallemacq PE. 2010. Population pharmacokinetic modeling and optimal sampling strategy for bayesian estimation of amikacin exposure in critically ill septic patients. Ther Drug Monit 32(6):749–756.
9. Hope WW. 2012. Population pharmacokinetics of voriconazole in adults. Antimicrob Agents Chemother 56(1):526–531.
10. D'Argenio DZ. 1981. Optimal sampling times for pharmacokinetic experiments. J Pharmacokinet Biopharm 9(6):739–756.
11. D'Argenio DZ, Khakmahd K. 1983. Adaptive control of theophylline therapy: Importance of blood sampling times. J Pharmacokinet Biopharm 11(5):547–559.
12. D'Argenio DZ, Katz D. 1983. Sampling strategies for noncompartmental estimation of mean residence time. J Pharmacokinet Biopharm 11(4):435–446.
13. Tod M, Rocchisani JM. 1996. Implementation of OSPOP, an algorithm for the estimation of optimal sampling times in pharmacokinetics by the ED, EID and API criteria. Comput Methods Programs Biomed 50(1):13–22.
14. Tod M, Mentré F, Merlé Y, Mallet A. 1998. Robust optimal design for the estimation of hyperparameters in population pharmacokinetics. J Pharmacokinet Biopharm 26(6):689–716.
15. Stroud JR, Müller P, Rosner GL. 2001. Optimal sampling times in population pharmacokinetic studies. J Royal Stat Soc 50(3):345–359.
16. Duffull SB, Retout S, Mentré F. 2002. The use of simulated annealing for finding optimal population designs. Comput Methods Programs Biomed 69(1):25–35.
17. Tam VH, Preston SL, Drusano GL. 2003. Optimal sampling schedule design for populations of patients. Antimicrob Agents Chemother 47(9):2888–2891.
18. Foracchia M, Hooker A, Vicini P, Ruggeri A. 2004. POPED, a software for optimal experiment design in populationm. Comput Methods Programs Biomed 74(1):29–46.
19. Dodds MG, Hooker AC, Vicini P. 2005. Robust population pharmacokinetic experiment design. J Pharmacokinet Pharmacodyn 32(1):33–64.
20. Ogungbenro K, Graham G, Gueorguieva I, Aarons L. 2005. The use of a modified fedorov exchange algorithm to optimise sampling times for population pharmacokinetic experiments. Comput Methods Programs Biomed 80(2):115–125.
21. Roy A, Ette EI. 2005. A pragmatic approach to the design of population pharmacokinetic studies. AAPS J 7(2):E408–E420.
22. Foo LK, Duffull S. 2010. Methods of robust design of nonlinear models with an application to pharmacokinetics. J Biopharm Stat 20(4):886–902.
23. Nyberg J, Ueckert S, Strömberg EA, Hennig S, Karlsson MO, Hooker AC. 2012. Poped: An extended, parallelized, nonlinear mixed effects models optimal design tool. Comput Methods Programs Biomed 108(2):789–805.

24. Lacarelle B, Pisano P, Gauthier T, Villard PH, Guder F, Catalin J, Durand A. 1994. Abbott PKS system: A new version for applied pharmacokinetics including Bayesian estimation. Int J Biomed Comput 36(1–2):127–130.
25. Komorowski M, Costa MJ, Rand DA, Stumpf MPH. 2011. Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. Proc Natl Acad Sci USA 108:8645–8650.
26. Sanches J, Sousa I, Figueiredo P. 2010. Bayesian fisher information criterion for sampling optimization in ASL-MRI." In Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), pp 880–883.
27. Banks HT, Dediu S, Ernstberger SL. 2007. Sensitivity functions and their uses in inverse problems. J Inverse Ill Pose P 15(7):683–708.
28. Thomaseth K, Cobelli C. 1999. Generalized sensitivity functions in physiological system identification. Ann Biomed Eng 27(5):607–616.
29. Deb K. 2001. Multi-objective optimization using evolutionary algorithms. John Wiley & Sons.
30. Taylor W, Diers-Caviness MH. 1986. A textbook for the clinical application of therapeutic drug monitoring. Abbott Diagnostic Division.
31. Levy RH, Mattson RH, Meldrum BS, Perucca E, Eds. 2002. Antiepileptic drugs. Lippincott Williams and Wilkins.
32. Burton M, Shaw LM, Schentag JJ, Evans WE. 2006. Applied pharmacokinetics and pharmacodynamics. Lippincott Williams and Wilkins.
33. Cockroft DW, Gault MH. 1976. Prediction of creatinine clearance from serum creatinine. Nephron 16:31–41.
34. Pratt DS. 2010. Liver chemistry and function tests. In Sleisenger and Fordtran's gastrointestinal and liver disease. Saunders Elsevier.
35. Berk PD, Korenblat KM. 2011. Approach to the patient with jaundice or abnormal liver test results. In Cecil medicine. Saunders Elsevier.
36. Jones CA, McQuillan GM, Kusek JW, Eberhardt MS, Herman WH, Coresh J, Salive M, Jones CP, Agodoa LY. 1998. Serum creatinine levels in the us population: Third national health and nutrition examination survey. Am J Kidney Dis 32:992–999.
37. Bazzoli Caroline, Retout Sylvie, MentrÃl' France. 2010. Design evaluation and optimisation in multiple response nonlinear mixed effect models: PFIM 3.0. Comput Methods Programs Biomed 98(1):55–65.