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