Use of a simulated annealing algorithm to fit compartmental models with an application to fractal pharmacokinetics.

Rebeccah E. Marsh [1], Terence A. Riauka [2], Steve A. McQuarrie [2]

[1] Department of Physics, Faculty of Science, and [2] Department of Oncology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta

Received June 16, 2006; Revision received February 10, 2007; Accepted February 14, 2007, Published June 14th 2007

Keywords: individual pharmacokinetics; simulated annealing; optimal design; fractal pharmacokinetics

ABSTRACT

Increasingly, fractals are being incorporated into pharmacokinetic models to describe transport and chemical kinetic processes occurring in confined and heterogeneous spaces. However, fractal compartmental models lead to differential equations with power-law time-dependent kinetic rate coefficients that currently are not accommodated by common commercial software programs. This paper describes a parameter optimization method for fitting individual pharmacokinetic curves based on a simulated annealing (SA) algorithm, which always converged towards the global minimum and was independent of the initial parameter values and parameter bounds. In a comparison using a classical compartmental model, similar fits by the Gauss-Newton and Nelder-Mead simplex algorithms required stringent initial estimates and ranges for the model parameters. The SA algorithm is ideal for fitting a wide variety of pharmacokinetic models to clinical data, especially those for which there is weak prior knowledge of the parameter values, such as the fractal models.

INTRODUCTION

Pharmacokinetics (1), the study of the absorption, distribution, metabolism, and eventual elimination of a drug from the body, is a quantitative tool used in drug development and subsequent therapy. Pharmacokinetic models are mathematical constructs whose parameters can be estimated from experimental data, which typically consist of discrete values of the drug concentration as a function of time.

Pharmacokinetic models can be divided broadly into two classes, compartmental models and non-compartmental models. The latter include moment curve and residence time analysis (2). In compartmental modeling (3), the body is divided into compartments, with a compartment being defined as the number of drug molecules having the same probability of undergoing a set of chemical kinetic processes. The exchange of drug molecules between compartments is described by kinetic rate coefficients, which may be related to physiological parameters such as molecular binding rates and organ volumes.

Because the rate of change of the concentration is as important as its magnitude, most pharmacokinetic models are expressed as a set of differential equations. Modeling is most efficient when these equations can be solved analytically to produce algebraic equations that can be fit to experimental data using linear and nonlinear regression techniques. However, some models, especially those with nonlinear or time-dependent terms, lead to equations that can only be solved numerically(4). In such cases, including the growing set of fractal models, alternate methods must be developed to estimate the model parameters. The objectives of this paper are to experimentally determine the optimal implementation of the simulated annealing algorithm, test its performance against existing algorithms, and assess its applicability to fitting compartmental models. Specific attention is given to the case of fractal compartmental models, in which one or more kinetic rate coefficients are power functions of time.

FRACTAL PHARMACOKINETICS

In the past couple of decades, the concept of fractals (5) has emerged in pharmacokinetics to describe the influence of heterogeneous structures and physiology on drug processes occurring within the body.

Corresponding Author: Dr. Rebeccah E. Marsh, Department of Physics, Room #238 CEB, University of Alberta Edmonton, AB, Canada T6G 2G7, Tel: (604) 532-8169, Email: rmarsh@ualberta.ca

Fractals can describe complex objects that cannot be characterized by one spatial scale. Fractal structures in the body include the bifurcating patterns of the bronchial tree, vascular system, bile-duct system, renal urine collection tubules, and the neuronal network (6). In addition, the architecture, growth, and blood supply of tumors exhibit fractal organization (7,8).

The concept of fractals also extends to processes that do not have a characteristic time scale. Drug processes that have been found to exhibit fractal behaviour include drug release (9), aerosol transport in the lungs (10), transport across membranes (11), diffusion (12), binding and dissociation kinetics (13,14), washout from the heart (15), and tissue trapping (16).

Transport and chemical reactions that occur on or within a fractal medium obey anomalous, fractal behaviour (17). Specifically, the kinetic rate coefficient follows a decreasing power of time (18,19)

$$k = k_0 t^{-\alpha} \quad \textbf{(1)}$$

where α is the fractal exponent and $0 \leq \alpha < 1$. The quantity t^{-a} is considered dimensionless, and both k and k_0 are in units of inverse time (h^{-1}).

Because equation **(1)** has a singularity at $t = 0$ for $h > 0$, Schnell and Turner (20) suggested a modified form based on the Zipf-Mandelbrot distribution, $k = k_0(\tau + t)^{-a}$, where the constant τ is the critical time from which the rate constant is driven by fractal effects. However, if τ is very small, equation **(1)** is a good approximation.

Macheras (21) first suggested the application of fractal kinetics to the study of drugs, and subsequently, fractal equivalents have been derived for the Michaelis-Menton formalism (22,23), the volume of distribution (24), and the clearance and half-life (25).

Fuite et al. (26) developed a fractal compartmental model to fit experimental data for the cardiac drug mibefradil (27). Because mibefradil is dispersed relatively quickly in the plasma but the fractal geometry of the liver slows down its rate of elimination, a classical compartment was used to represent the plasma whereas a compartment with a fractal elimination rate coefficient was used to represent the liver.

A classical compartment can be broadly defined as one for which the probability of a molecule undergoing a kinetic process remains constant in time, whereas a fractal compartment is a compartment for which this probability varies in time. Fuite et al. obtained approximate analytical solutions through perturbation analysis, fit them to experimental data, and found a relationship between the fractal exponent from the compartmental model and the fractal dimension of the liver. Simulations of the model showed that the fractal exponent describing drug elimination plays a significant role in determining the shape of the concentration-time curve (28). Advantages of the fractal compartmental model in addressing clinical questions include its traditional compartmental framework and the relatively simple adjustments that can take into account the effects of heterogeneity.

Unfortunately, most fractal models cannot be solved analytically. There are several commercially-available software packages, including WinNonLin® (29) and Boomer® (30), that make use of the Gauss-Newton (31), and Nelder-Mead simplex (32) algorithms to numerically fit differential equations to experimental data. For models based on classical and Michaelis-Menton kinetics, these programs are excellent. However, even with the option of user-defined models, these programs currently do not have the capability to handle power-law time-dependent kinetic rate coefficients. Furthermore, the Gauss-Newton method is a gradient-based method that involves the calculation of derivatives, and the simplex algorithm is sensitive to initial conditions (33). To fit their model for fractal Michaelis-Menten kinetics to experimental data, Kosmidis et al. (23) used the Levenberg-Marquardt (LM) algorithm. However, the LM algorithm is also gradient-based; therefore, although it is appropriate for fitting their one-compartment model that can be solved analytically, it may not be the best method for fitting models that do not have an algebraic solution. Furthermore, the LM algorithm has the disadvantage of converging towards local minima when the initial parameter estimates are poor (34). This paper describes the simulated annealing

algorithm and explores its ability to optimize functions through an efficient exploration of the parameter space.

SIMULATED ANNEALING

Kirkpatrick and colleagues developed simulated annealing to optimize the design of integrated circuits, and they later applied it to the optimization of many-variable functions (35). Simulated annealing (SA) derives its name from an analogy to the cooling of heated metals. As a metal cools, the atoms fluctuate between relatively higher and lower energy levels. If the temperature is dropped slowly enough, the atoms will all reach their ground state. However, if the temperature is dropped too quickly, the system will get trapped in a less-than-optimum configuration. If the energy function of this physical system is instead replaced by an objective function $f(\{x_i\})$, where $\{x_i\}$ is a set of independent variables, then the progression of this function towards the global minimum is analogous to the physical progression towards the ground state.

The SA algorithm also requires a control parameter, T (an effective temperature), a strategy for changing T, and a method for exploring the parameter space. The algorithm, illustrated in Figure 1, begins at a random initial position in parameter space with a corresponding objective function value, f_0. The parameters are then perturbed to generate a trial point with a new objective function, f_1, and the move is either accepted or rejected. All downhill moves, corresponding to $\Delta f = f_i - f_{i-1} < 0$, are accepted, while uphill moves, corresponding to $\Delta f > 0$, are accepted in a probabilistic manner using the Metropolis algorithm (35), where

$$P(\Delta f) = \exp\left[-\frac{\Delta f}{T}\right] \quad \textbf{(2)}$$

A random number is generated over the range $(0,1)$, and if $P(\Delta f) > random(0,1)$, the trial point is accepted. Otherwise, the point is rejected and another trial point is generated. The control parameter is decreased every n steps, and the step lengths can be adjusted every m steps. The process continues in an iterative manner until the program converges on a solution. This convergence can be expressed as a function of the control parameter, the number of iterations, the acceptance rate of new moves, or the absolute or relative change in the objective function.

At the start of the annealing process, the control parameter is relatively high compared to the standard deviation of the objective function, so the probability of accepting an uphill move is great. Hence, the random walk is able to explore a wide area of parameter space without getting trapped in local minima. As the control parameter is decreased, the algorithm is able to focus on the most promising areas.

In addition to manipulating the control parameter, the magnitude of the step length in parameter space can be controlled. It has been found experimentally that the space is best explored when the acceptance rate of new steps is 50% (36). A relatively high rejection rate means that the space is being ineffectively explored, while a relatively low rejection rate means the algorithm is being explored with too-small steps.

The SA algorithm has many advantages over other optimization methods. It is largely independent of the starting values, it can escape local minima through selective uphill moves, and the underlying function need not be continuous. The SA method has been found to be superior to the simplex method, the Adaptive Random Search, and the quasi-Newton algorithm in finding the optimum of a continuous function (33,37). Eftaxias et al. (34) compared the SA and Levenberg-Marquardt algorithms and found that the LM algorithm was only advantageous when few model parameters must be optimized and a good initial estimate was provided. They observed that the SA algorithm was more robust and found more accurate and meaningful fits as the number of model parameters was increased.

The SA algorithm has been applied to population pharmacokinetics (38), optimal design (39), and in the physiologically-based program PKQuest (40). However, in the latter, the method is built into an application for Maple, a commercial mathematical problem-solving program, and not explicitly described. To our knowledge, this algorithm has not yet been applied to solve compartmental models in individual pharmacokinetics.

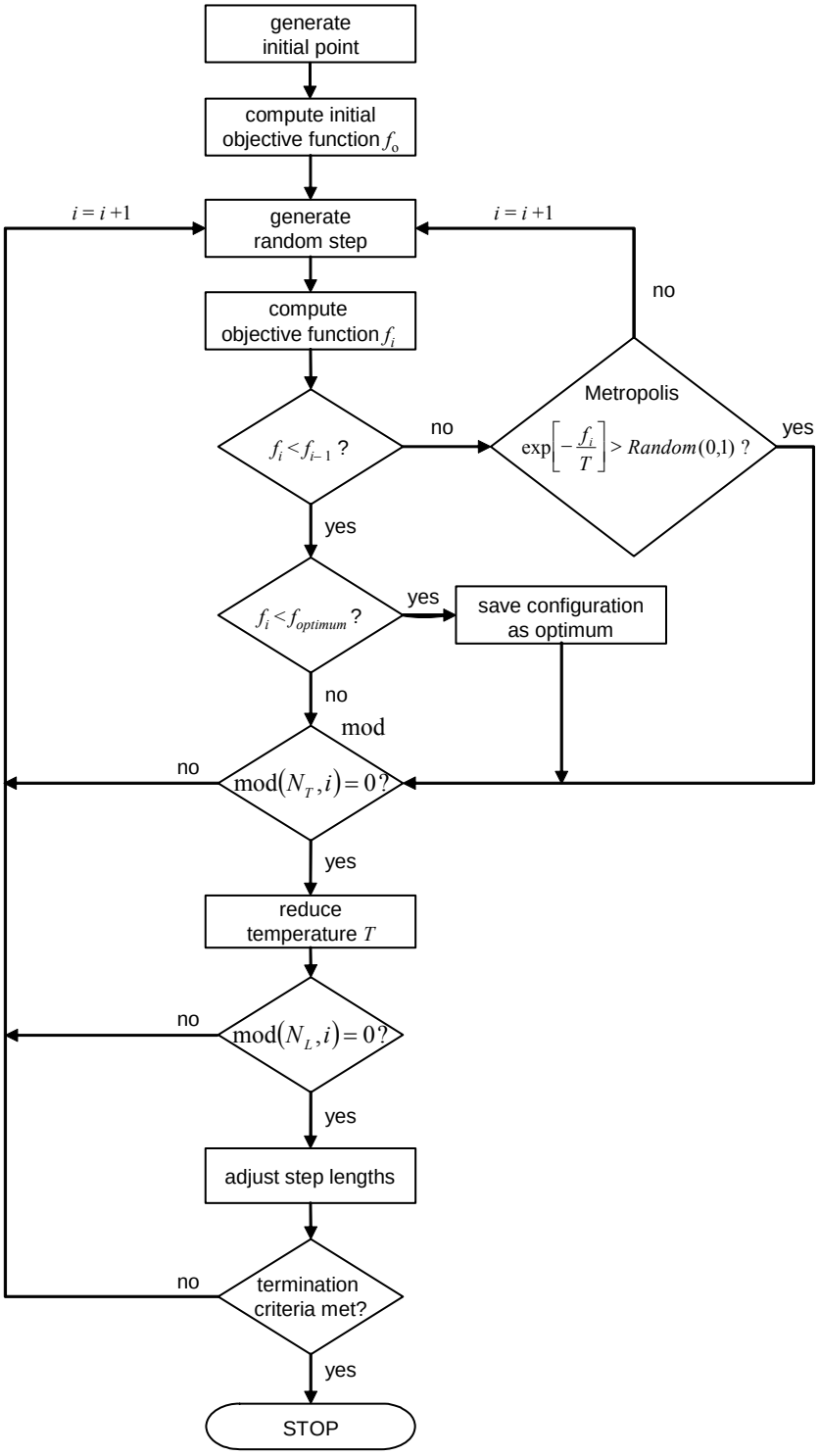

Figure 1. The simulated annealing algorithm. The operator $\mathrm{mod}(x_1, x_2)$ returns the remainder of x_1 divided by x_2.

METHODS

The SA algorithm was implemented in a C++ program, PKPhit. The differential equations were solved numerically using a fourth-order Runge-Kutta routine with 500 steps (41).

Annealing schedule

The program begins by generating a random starting position in parameter space and calculating the corresponding objective function, f_0. The form of the objective function, f, was chosen to be the weighted residual sum of squares (WRSS)

$$WRSS = \sum_{i=1}^{n} w_i \left(C_i - \hat{C}_i\right)^2 \quad \textbf{(3)}$$

where $\hat{C}_i$ denotes the predicted value of C_i based on the chosen model, and w_i is the weighting factor. The weight is commonly chosen to be the inverse of the variance of the observation, and here an iterative reweighting scheme was chosen where the variance was taken to be proportional to the predicted concentration, $\hat{C}_i$, so that (42)

$$w_i = \frac{1}{\hat{C}_i^2} \quad \textbf{(4)}$$

The algorithm progresses as new steps are generated and either accepted or rejected as a function of the control parameter T, which is reduced every m iterations. The magnitude of a step is calculated by multiplying a random number between 0 and 1 by the step length. The acceptance rate (AR) of new trials is checked every n iterations. The step lengths are increased if AR > 60%, decreased if AR < 40%, and left unchanged otherwise. Following Corana et al. (33), the new step length, L'_j, for the j^{th} parameter is calculated from the previous step length L_j according to

$$L_j' = L_j\left(1 + c_j \frac{AR - 0.6}{0.4}\right) \text{ if AR} > 0.6 \quad \textbf{(5)}$$

$$L_j' = \frac{L_j}{1 + c_j \dfrac{0.4 - AR}{0.4}} \text{ if AR} < 0.4 \quad \textbf{(6)}$$

where c_j is the step length adjustor for parameter j.

The program terminates when the convergence criteria are reached. Several different criteria were investigated alone and in combination, including the total number of iterations, the absolute or relative change in WRSS, the acceptance rate, and the current value of the control parameter.

Test models

The robustness of the SA algorithm in fitting pharmacokinetic parameters was tested using three two-compartment models. Figure 2 shows a mamillary model with a central compartment. The case with constant kinetic rate coefficients, which we will refer to as the classical Model 1, is described by the equations

$$\dot{C}_1 = k_{21}C_2 - \left(k_{12} + k_{10}\right)C_1 + \frac{i(t)}{V_d} \quad \textbf{(7)}$$

$$\dot{C}_2 = k_{12}C_1 - k_{21}C_2 \quad \textbf{(8)}$$

where C_1 is the concentration in compartment 1, C_2 is the concentration in compartment 2, k_{ij} is the kinetic rate coefficient for the transfer of drug molecules from compartment i to compartment j, and $i(t)$ is the infusion rate of the drug in mass or moles per hour. The term V_d is the apparent volume of distribution, defined as the volume of fluid into which the dose would have to be dispersed in order to produce the concentration observed in the plasma, and it is expressed in liters. Compartment 1 typically corresponds to the plasma, and compartment 2 could represent a tumor, the brain, a bound state, or simply a mathematical construct. In the case of a bolus dose or a constant infusion, Equations **(7)** and **(8)** have an exact algebraic solution in which the compartment concentrations are typically sums of terms that are exponential in time.

By replacing one or more of the constant kinetic rate coefficients by the fractal equivalent given by equation **(1)**, a fractal compartmental model can be created. For example, by making the substitution $k_{21}=k_{21}t^{-\alpha}$ in equations **(7)** and **(8)**, a fractal trapping model is obtained, with a power-law release of drug molecules from compartment 2 back into compartment 1. This Model 2 is expressed mathematically as

$$\dot{C}_1 = k_{21}t^{-\alpha}C_2 - \left(k_{12} + k_{10}\right)C_1 + \frac{i(t)}{V_d} \quad \textbf{(9)}$$

$$\dot{C}_2 = k_{12}C_1 - k_{21}t^{-\alpha}C_2 \quad \textbf{(10)}$$

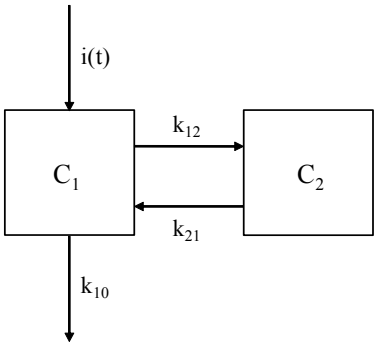

Figure 2. A two-compartment mamillary model.

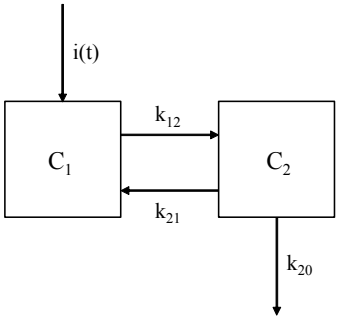

Figure 3. A two-compartment catenary model.

Alternatively, a fractal elimination model can be created from the concatenary configuration shown in Figure 3. Fuite et al. (26) designated compartment 1 as the plasma and compartment 2 as the liver. By making the substitution $k_{20}=k_{20}t^{-\alpha}$, elimination from the liver now occurs at a fractal rate. This case comprises Model 3 and is described by the equations

$$\dot{C}_1 = k_{21}C_2 - k_{12}C_1 + \frac{i(t)}{V_d} \quad \textbf{(11)}$$

$$\dot{C}_2 = k_{12}C_1 - \left(k_{21} + k_{20}t^{-\alpha}\right)C_2 \quad \textbf{(12)}$$

The fractal compartmental models are linear, since the kinetics remain first-order. However, the value of the fractal kinetic rate coefficient, and thus the probability of drug release from the fractal compartment, changes with time. Unlike Model 1, the equations for Models 2 and 3 cannot be solved exactly using analytical methods, except for a few special cases of the parameter values (28).

Simulated data sets

Sets of error-free data were generated for each model using parameters (listed in Table 1) that were chosen to reproduce real clinical situations. For Models 1 and 2, the values correspond to parameters reported for the anticancer agent carboplatin in pediatric patients (43). In addition, an arbitrary value of $\alpha = 0.6$ was included for Model 2. The dose and infusion time were taken to be 500 *mg* and 1.5 *h*, respectively.

Twelve concentration values, C_i, were calculated for times t_i = 0, 0.5, 1, 1.5, 2, 3, 4, 6, 8, 10, 12, and 18 h. Beyond 18 hours, the concentration falls below the quantification limit of 0.0025 $mg\ L^{-1}$ reported for free platinum using atomic absorption spectrometry (44). Model 3 was fit to mibefradil data (27), and the estimated parameters were used to generate 12 points for times t_i = 0, 0.8333, 0.1667, 0.25, 0.3333, 0.5, 1, 1.5, 2, 6, 12, and 24 *h*. A dose of 25 *mg* and an infusion time of 0.1667 *h* was selected for a hypothetical 25-*kg* dog.

To create sets of realistic, noisy data, $\left(t_i, C_i^*\right)$, an independent error value, ε_i, was added to each concentration value:

$$C_i^* = C_i + \varepsilon_i \quad \textbf{(13)}$$

A Gaussian distribution with zero mean was assumed for ε_i, such that

$$\varepsilon_i \sim N\left(0, \sigma_i^2\right) \quad \textbf{(14)}$$

where $N(\mu, \sigma^2)$ is a normally-distributed random variate with mean μ and variance. The variance was chosen to induce a coefficient of variation (CV) of 10% in C_i (45), where

$$CV = \frac{\sigma}{\mu} \times 100\% = \frac{\sigma}{C_i} \times 100\% \quad \textbf{(15)}$$

Table 1. Theoretical coefficients for the simulated two-compartment models

Model	k_{12} (h^{-1})	k_{21} (h^{-1})	k_{10} (h^{-1})	k_{20} (h^{-1})	V_d (L)	α
Classical	1.5	1.1	1.25	-	5.0	-
Fractal trapping	1.5	1.1	1.25	-	5.0	0.6
Fractal elimination	28.0	3.0	-	0.6	10.0	0.8

In order to generate $N\left(0,\sigma_i^2\right)$, we used the polar form of the Box-Muller method to transform two uniformly-distributed random numbers into two $N(0,1)$ variates (41) Further modification provides

$$\varepsilon_i = N\left(0,\sigma_i^2\right) = \sigma_i \times N(0,1) \quad \textbf{(16)}$$

with $\sigma_i = \left(\frac{CV}{100\%} \times C_i\right)$ from equation **(15)**.

Finally, substituting into equation **(13)** yields

$$C_i^* = C_i\left(1 + \left[\frac{CV}{100\%}\right] N(0,1)\right) \quad \textbf{(17)}$$

Figure 4 shows the range in data points for the three models.

RESULTS AND DISCUSSION

Optimum algorithm parameters

Perhaps the most important step in evaluating the SA algorithm is the development of an appropriate annealing schedule. To ensure successful optimization, the control parameter T should be lowered slowly enough to find the global minimum but quickly enough to minimize computer run time. For the two-compartment models, it was found that a linear decrease in the control parameter by a factor R was the most efficient, as compared to exponential or power law functions of T, or a function of the number of iterations. Furthermore, the drop in the control parameter was best achieved when the ratio of m, the number of iterations at each value of the control parameter, to T_0, the initial value of the control parameter, was 0.10 (with $m = 1$ and $T_0 = 10$) and $R = 0.999$.

When $\frac{m}{T_0} = 1.0$, the algorithm failed to converge, and when $\frac{m}{T_0} = 0.01$, the algorithm was inefficient and required more than twice as many iterations. An increase in R to 0.9999 also proved inefficient, while a decrease to 0.99 led to a lower accuracy in the final results.

The choice of appropriate convergence criteria was also investigated. WinNonLin® monitors the relative change in the objective function. In this work, a modified method based on results of Goffe et al. (37) was developed for PKPhit. Every N_δ iterations, the average of the previous N_δ accepted values of the objective function was calculated and compared to both the current function value, f_i, and the optimum value, $f_{optimum}$. The relative changes were compared to a predetermined value δ, and the algorithm converged

$$\text{if } \frac{\left|f_{previous} - f_i\right|}{f_{previous}} < \delta \ldots \textbf{(18)}$$

$$\text{and } \frac{\left|f_{previous} - f_{optimum}\right|}{f_{previous}} < \delta \ldots \textbf{(19)}$$

$$\text{where } f_{previous} = \frac{\sum_{j=1}^{N_\delta} f_j}{N_\delta} \quad \textbf{(20)}$$

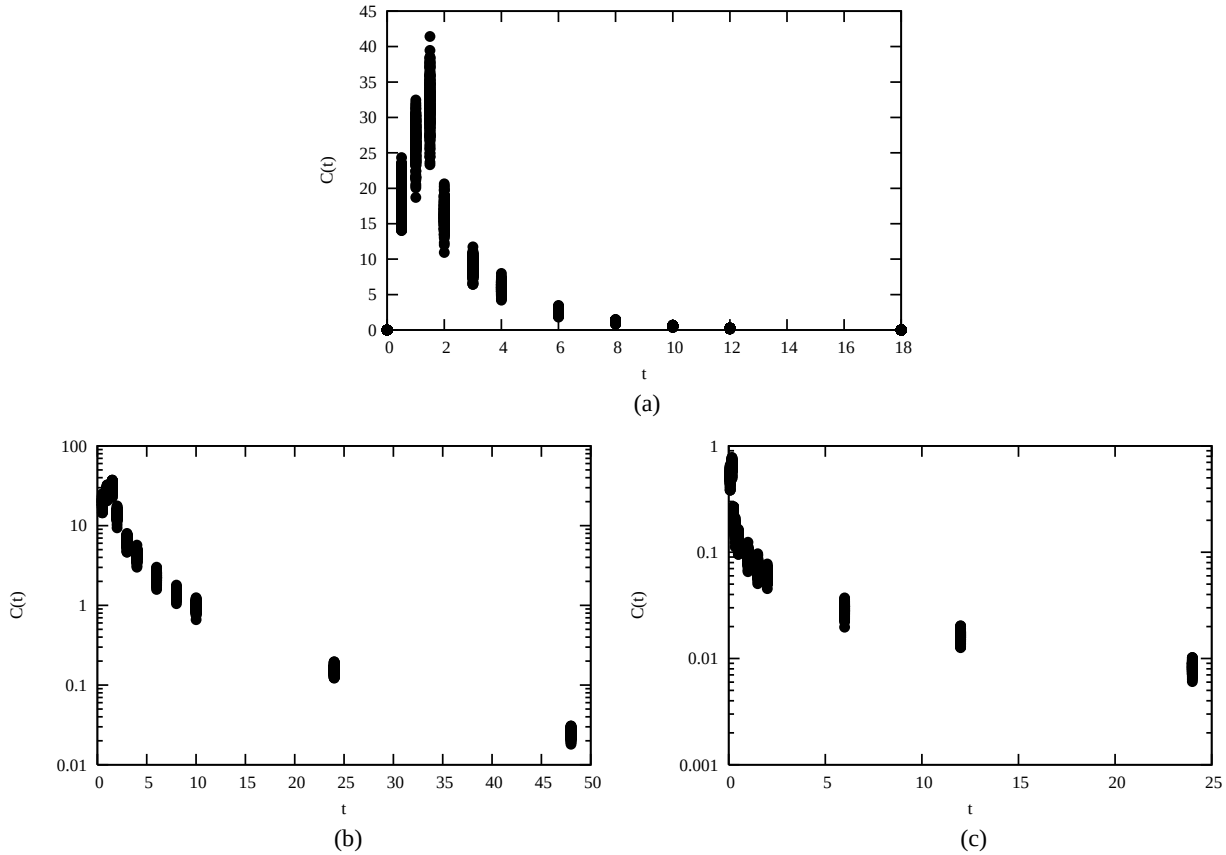

Figure 4. The simulated data sets for (a) Model 1, (b) Model 2, and (c) Model 3.

For flexibility in fitting difficult data sets, an additional condition was imposed such that the total number of iterations did not exceed $N_{\max} = 30{,}000$. Using the optimum parameters discussed above and values of $N_\delta = 10$ and $\delta = 0.0001$, convergence was found to occur after $15{,}000 \pm 5{,}000$ iterations (taking less than a minute on a Pentium 4 computer, 3.2 GHz with 1 G of RAM).

Because it was found that adjusting the step length actually led to longer run times and in many cases an inability to converge, constant step length values were used. A step length of 0.01 to 0.05 was optimal for the kinetic rate parameters and volume of distribution, while a value of 0.001 was ideal for the fractal exponent. When these values were increased, the algorithm quickly found the correct area in parameter space, but subsequently took over 30,000 iterations to narrow in on the optimum. Smaller step lengths resulted in inefficient exploration of the parameter space and the same excess of iterations. The bounds in parameter space were taken to be 0.01 - 10.0 for the kinetic rate coefficients and volume of distribution, and 0 - 5 for the fractal exponent; however, widening the ranges, even by a factor of 10^4, had no effect on the progression of the algorithm.

In order to investigate the sensitivity of the SA algorithm to the initial parameter values, Model 1 was fit to a data set 50 times, each time with different values of the random seed and initial model parameter values (randomly chosen over the range (0,10]). The mean WRSS was 0.3389 ± 0.0004, and the coefficient of variation was less than 5% for each of the estimated model parameters. Although the algorithm is most efficient when the range of the initial values is limited, it never failed to reach the global optimum, even when the ranges were extended to (0,50] and the initial function value was as high as 10^{100}. Consequently, the performance of the SA algorithm is independent of the starting point.

Finally, performing occasional restarts of the program using the current optimum parameter values as the new initial values was not found to be beneficial; the algorithm explored the surrounding space but still converged on the minimum at the same value of the control parameter.

Comparison with other algorithms

To evaluate the performance of the SA algorithm, the WRSS and model parameter values found by PKPhit for Model 1 were compared to those determined by the Gauss-Newton (Levenberg and Hartley) and Nelder-Mead simplex algorithms as implemented by WinNonLin® Version 4.1. Table 2 lists the results for 50 data sets, the number found to achieve meaningful results since two independent groups of 50 were statistically equivalent (with WRSS values of 0.27 ± 0.10 and 0.28 ± 0.10). The confidence intervals should include the theoretical values used to generate the noise-free data, and this is indeed the case for all of the parameters.

Because the same data and model is involved, the WRSS can be used to compare the goodness of fit of the three algorithms. The WRSS values were the same for all three algorithms, and thus PKPhit performs as well as the commercially-available implementations of the Gauss-Newton and simplex algorithms. However, in order to get the similar degree of accuracy and precision in the Gauss-Newton and simplex results, the initial parameter values had to be within 40% of the actual values, and the lower and upper bounds had to be within an order of magnitude for the Gauss-Newton algorithm and more than twice as narrow again for the simplex algorithm. Doubling the range of the parameter bounds for the simplex algorithm resulted in a decrease in the accuracy of the parameter estimates by an average of 30% and an increase in their standard deviation by an average of 140%. In the case of the Gauss-Newton algorithm, simply increasing the initial estimates of the kinetic rate coefficients from 1.0 h^{-1} to 5.0 h^{-1} resulted in an inability of the algorithm to converge to a solution for any of the 100 data sets.

In the extreme case where no initial estimates were provided to WinNonLin®, the program was unable to find solutions for 26 of the data sets when using the Gauss-Newton algorithm and 24 of the sets when using the simplex algorithm. Furthermore, although the mean objective functions did not rise significantly, the standard deviations in the parameter estimates increased by 15-50% in the Gauss-Newton case and by 50-350% in the simplex case. Therefore, although the three algorithms are capable of achieving equivalent model fits, the Gauss-Newton and simplex algorithms require strict initial conditions, whereas the SA algorithm is able to explore the parameter space and focus on the most promising area without prior knowledge of its location.

Applicability to fractal models

Fifty noisy data sets were generated for Model 2, the fractal trapping model, and Model 3, the fractal elimination model. Based on the performance of PKPhit demonstrated in the previous section, the parameters for the fractal models were estimated using PKPhit, and the results are listed in Tables 3 and 4, respectively. For both the models, the parameter estimates agree with the predicted values, and the confidence intervals are reasonable. These values are similarly independent of the initial values and bounds of the parameters. In these cases, the WRSS values are strictly the metric for minimization and as such cannot be compared across Tables 2-4.

On a final note, while the WRSS is a good metric for function optimization, we recommend using the Akaike Information Criterion (AIC) and/or the Schwarz-Bayesian Criterion (SBC) to guide selection between different models (46). These metrics adjust for the number of model parameters and the number of data points.

CONCLUSION

This paper introduced the first detailed application of the simulated annealing optimization routine to fit individual pharmacokinetic data. The robustness of the SA algorithm in fitting both classical and fractal compartmental models has been demonstrated. Although a technique was used to modify the step length, a constant step length was found to lead to a more stable solution.

Convergence of the SA algorithm was most efficient with a linear decrease in the control parameter by a factor R=0.999 every m=0.10 T_0 iterations and best estimated using the relative decrease in the objective function.

Table 2. Estimated parameters for the classical two-compartment model

Parameter	Theoretical Value		Estimated Value		
			PKPhit	Gauss-Newton[a]	Simplex[b]
k_{12} (h^{-1})	1.5	mean	2.26	2.23	2.48
		SD	0.86	1.06	0.68
		range	0.83-4.1	0.68-5.5	0.93-3.0
k_{21} (h^{-1})	1.1	mean	1.44	1.43	1.47
		SD	0.22	0.18	0.17
		range	0.95-2.1	1.1-1.9	1.0-1.9
k_{10} (h^{-1})	1.25	mean	1.44	1.42	1.51
		SD	0.30	0.35	0.27
		range	0.88-2.3	0.87-2.7	0.92-2.0
V_d (L)	5	mean	4.53	4.64	4.29
		SD	0.93	1.0	0.88
		range	2.7-7.3	2.3-7.6	3.1-7.0
WRSS	-	mean	0.28	0.28	0.28
		SD	0.10	0.11	0.09
		range	0.11-0.50	0.11-0.54	0.12-0.51

WRSS = weighted residual sum of squares; SD = standard deviation;
[a]Initial parameter values: k_{12}= k_{21}= k_{10}= $1.0\ h^{-1}$ and V_d= $7\ L$; parameter bounds: $0.5\text{-}\ 10\ h^{-1}$ and $0.5\text{-}\ 10\text{L}$.
[b]Initial parameter values: k_{12}= k_{21}= k_{10}= $1.0\ h^{-1}$ and V_d= $7\ L$; parameter bounds: $0.5\text{-}\ 3.0\ h^{-1}$ and $2\text{-}\ 7L$.

Table 3. Estimated parameters for the two-compartment model with fractal trapping

	k_{12} (h^{-1})	k_{21} (h^{-1})	k_{10} (h^{-1})	V_d (L)	α	WRSS
theoretical	1.5	1.1	1.25	5	0.6	-
mean	1.75	1.11	1.40	4.76	0.58	0.067
SD	0.75	0.32	0.28	0.94	0.073	0.043
range	0.58-4.01	0.50-1.89	0.96-2.54	1.46-6.99	0.39-0.72	0.0077-0.20

WRSS = weighted residual sum of squares; SD = standard deviation

Table 4. Estimated parameters for the two-compartment model with fractal elimination

	k_{12} (h^{-1})	k_{21} (h^{-1})	k_{20} (h^{-1})	V_d (L)	α	WRSS
theoretical	28.0	3.0	0.6	10.0	0.8	-
mean	24.3	2.68	0.57	10.9	0.77	0.074
SD	2.0	0.58	0.059	1.19	0.056	0.047
range	18.3-29.6	1.42-4.43	0.43-0.71	7.8-15.1	0.63-0.88	0.0068-0.23

WRSS = weighted residual sum of squares; SD = standard deviation

PKPhit always converged towards the global minimum, irrespective of the initial values and bounds of the model parameters, whereas similar fits by the Gauss-Newton and simplex algorithms required estimates of the parameters to within 40% of the actual values as well as narrow parameter ranges. This limitation is significant not only in the case of a new drug, but also for established drugs due to wide ranges in interindividual variability. For example, when Sonnichsen et al. (47) fit a two-compartment model to data for the anticancer agent paclitaxel in 30 pediatric patients, the ranges of values found for three of the parameters were 2.9-47.4 $\mu mol\ h^{-1}$, 6.0- 142.7 $\mu mol\ h^{-1}$, and 0.052- 1.04 $\mu mol\ h^{-1}$.

Due to its versatility and independence on prior knowledge of the parameter values, the SA algorithm is particularly applicable to fitting fractal models that are not solvable using analytical techniques. Eftaxias et al. (34) found that an SA solution may be further improved by around 5% by subsequent application of the LM algorithm, and this possibility could be investigated in a future study.

ACKNOWLEDGEMENTS

Rebeccah Marsh would like to thank Pharsight for the copy of WinNonLin® Version 4.1 for the purpose of academic research. She also received support from Mathematics of Information Technology and Complex Systems (MITACS) and Natural Sciences and Engineering Research Council (NSERC). The authors have no conflicts of interest relevant to the content of this study.

REFERENCES

[1] Gibaldi, M., Perrier, D. Pharmacokinetics, 2nd ed, M. Dekker, New York, NY, USA, 1982.

[2] Weiss, M. The relevance of residence time theory to pharmacokinetics. Eur. J. Clin. Pharmacol., 43:571-579, 1992.

[3] Jacquez, J.A. Compartmental analysis in biology and medicine. BioMedware, Ann Arbor, MI, USA, 1996.

[4] Lànsky, P.A stochastic model for circulatory transport in pharmacokinetics. Math. Biosci., 132:141-167, 1996.

[5] Mandelbrot, B.B. The Fractal Geometry of Nature, W.H. Freeman, San Francisco, CA, USA, 1982.

[6] Bassingthwaighte, J.B., Liebovitch, L., West, B.J. Fractal Physiology, Oxford University Press, Oxford, U.K., 1994.

[7] Iwata, K., Kawasaki, K., Shigesada, N.A dynamical model for the growth and size distribution of multiple metastatic tumors. J. Theor. Biol., 203:177-186, 2000.

[8] Sabo, E., Boltenko, A., Sova, Y., Stein, A., Kleinhaus, S., Resnick, M.B. Microscopic analysis and significance of vascular architectural complexity in renal cell carcinoma. Clin. Cancer Res., 7:533-537, 2001.

[9] Kosmidis, K., Argyrakis, P., Macheras, P. Fractal kinetics in drug release from finite fractal matrices. J. Chem. Phys., 19: 6373-6377, 2003.

[10] Henry, F.S., Butler, J.P., Tsuda, A. Kinematically irreversible acinar flow: a departure from classical dispersive aerosol transport theories. J. Appl. Physiol., 92:835-845, 2002.

[11] Ogihara, T., Tamai, I., Tsuji, A. Application of fractal kinetics for carrier-mediated transport of drugs across intestinal epithelial membrane. Pharm. Res., 4:620-625, 1998.

[12] Fatin-Rouge, N., Starchev, K., Buffle, J. Size effects on diffusion processes within agarose gels. Biophys. J., 86:2710-2719, 2004.

[13] Tiihonen, J., Kuikka, J., Rasanen, P., Lepola, U., Koponen, H., Liuska, A., Lehmusvaara, A., Vainio, P., Kononen, M., Bergstrom, K., Yu, M., Kinnunen, I., Akerman, K., Karhu, J. Cerebral benzodiazepine receptor binding and distribution in generalized anxiety disorder: a fractal analysis. Mol. Psychiatry, 2:463-471, 1997.

[14] Sadana, A. A kinetic study of analyte-receptor binding and dissociation, and dissociation alone, for biosensor applications: a fractal analysis. Anal. Biochem., 291:34-47, 2001.

[15] Bassingthwaighte, J.B., Beard, D.A. Fractal ^{15}O-labeled water washout from the heart. Circ. Res., 77:1212-1221, 1995.

[16] Weiss, M. The anomalous pharmacokinetics of amiodarone explained by nonexponential tissue trapping. J. Pharmacokinet. Biopharm., 27:383-396, 1999.

[17] ben-Avraham, D., Havlin, S. Diffusion and Reactions in Fractals and Disordered Systems, Cambridge University Press, Cambridge, U.K., 2000.

[18] Kopelman, R. Rate processes on fractals: theory, simulations, and experiments. J. Stat. Phys., 42:185-200 1986.

[19] Kopelman, R. Fractal reaction kinetics. Science, 241:1620-1626, 1988.

[20] Schnell, S. Turner, T.E. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. Prog. Biophys. Mol. Biol., 85:235-260, 2004.

[21] Macheras, P. Carrier-mediated transport can obey fractal kinetics. Pharm. Res., 12:541-554, 1995.
[22] Heidel, J., Maloney, J. An analysis of a fractal Michaelis-Menten curve. J. Australian Math. Soc. Ser. B. Applied Math., 41:410-422, 2000.
[23] Kosmidis, K., Karalis, V., Argyrakis, P., Macheras, P. Michaelis-Menten kinetics under spatially constrained conditions: application to mibefradil pharmacokinetics. Biophys. J., 87:1498-1506, 2004.
[24] Karalis, V., Claret, L., Iliadis, A., Macheras, P. Fractal volume of drug distribution: it scales proportionally to body mass. Pharm. Res., 18:1056-1060, 2001.
[25] Karalis, V., Macheras, P. Drug disposition viewed in terms of the fractal volume of distribution. Pharm. Res., 19:696-703, 2002.
[26] Fuite, J., Marsh, R., Tuszynski, J., Fractal pharmacokinetics of the drug mibefradil in the liver. Phys. Rev. E, 66:021904, 2002.
[27] Skerjanec, A. Tawfik, S. Tam, Y.K. Mechanisms of nonlinear pharmacokinetics of mibefradil in chronically instrumented dogs. J. Pharmacol. Exp. Ther., 278:817-825, 1996.
[28] Chelminiak, P., Marsh, R.E., Tuszynski, J.A., Dixon, J.M., Vos, K.J. Asymptotic time-dependence in the fractal pharmacokinetics of a two-compartment model. Phys. Rev. E., 72:031903, 2005.
[29] Available from URL: http://www.pharsight.com/winnonlin/
[30] Available from URL: http://www.boomer.org
[31] Hartley, H.O. The modified Gauss-Newton method for the fitting of nonlinear regression functions by least squares. Technometrics, 3: 269-280, 1961.
[32] Nelder, J.A., Mead, R. A simplex method for function minimization. Computer Journal, 7:308-313, 1965.
[33] Corana, A., Marchesi, M., Martini, C., Ridella, S. Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm, ACM Trans. Math. Software, 13:262-280, 1987.
[34] Eftaxias, A. Font, J. Fortuny, A. Fabregat, A. Stüber, F. Nonlinear kinetic parameter estimation using simulated annealing. Comp. Chem. Eng., 26:1725-1733, 2002.
[35] Kirkpatrick, S., Gelatt, C., Vecchi, M., Optimization by simulated annealing. Science, 220:45-54, 1983.
[36] Metropolis, A.W., Rosenbluth, M.N., Rosenbluth, A.H. Equation of state calculations by fast computing machines. J. Chem. Phys., 21:1087-1092, 1953.
[37] Goffe, W.L., Ferrier, G.D., Rogers, J. Global optimization of statistical functions with simulated annealing, J. Econometrics, 60:65-99, 1994.
[38] Duffell, S.B., Retout, S., Mentre, F. The use of simulated annealing for finding optimal population designs. Comput. Methods Programs Biomed., 69:26-35, 2002.
[39] Ji, S., Zeng, Y., Wu, P., Lee, E.J. Sampling schedule design towards optimal drug monitoring for individualizing therapy. Comput. Methods Programs Biomed., 80:57-63, 2005.
[40] Levitt, D.G. PKQuest: a general physiologically based pharmacokinetic model. Introduction and application to propanolol. BMC Clin. Pharmacol., 2:1-21, 2002.
[41] Press, W.H. Numerical recipes in C: the art of scientific computing. Cambridge University Press, New York, NY, USA, 1992.
[42] Gabrielsson, J., Weiner, D. Pharmacokinetic/ Pharmacodynamic Data Analysis: Concepts and Applications, 2nd ed., Swedish Pharmaceutical Press, Stockholm, Sweden, 1997.
[43] Peng, B., Boddy, A.V., Cole, M., Pearson, A.D., Chatelut, E., Rubie, H., Newell, D.R. Comparison of methods for the estimation of carboplatin pharmacokinetics in paediatric cancer patients. Eur. J. Cancer, 31:1804-1810, 1995.
[44] Guichard, S., Arnould, S., Hennebelle, I., Bugat, R., Canal, P. Combination of oxaliplatin and irinotecan on human colon cancer cell lines: activity in vitro and in vivo. Anticancer Drugs, 12: 741-751, 2001.
[45] Hunt, C.A., Givens, G.H., Guzy, S. Bootstrapping for pharmacokinetic models: visualization of predictive and parameter uncertainty. Pharm. Res., 15:690-697, 1998.
[46] Ludden, T.M., Beal, S.L., Sheiner, L.B. Comparison of the Akaike Information Criterion, the Schwarz criterion and the F test as guides to model selection, J. Pharmacokinet. Biopharm., 22:431-445, 1994.
[47] Sonnichsen, D.S., Hurwitz, C.A., Pratt, C.B., Shuster, J.J., Relling, M.V. Saturable pharmaco-kinetics and paclitaxel pharmacodynamics in children with solid tumors. J. Clin. Oncol., 12:532-538, 1994.