Journal of Pharmacokinetics and Biopharmaceutics, Vol. 27, No. 2, 1999

PHARMACOMETRICS

Construction of an Optimal Destructive Sampling Design for Noncompartmental *AUC* Estimation

Francois Vandenhende,[1] Muriel Comblain,[2] Marie-Helene Delsemme,[1] Walthere Dewe,[1] Wherly P. Hoffman,[3] and Bruno Boulanger[1,4]

Received September 3, 1998—Final April 26, 1999

*Based on toxicokinetic studies of a destructive sampling design, this work was aimed at selecting the number of time points, their locations, and the number of replicates per time point in order to obtain the most accurate and precise noncompartmental estimate of the area under the concentration-time curve (*AUC*). From a prior population pharmacokinetic model, the design is selected to minimize the scaled mean squared error of* AUC. *Designs are found for various sample sizes, number of time points, and a distribution of animals across time points from being very unbalanced to balanced. Their efficiencies are compared both theoretically and based on simulations. An algorithm has been implemented for this purpose using the symbolic resolution and numerical minimization capabilities of Mathematica™ and an example of its use is provided. This method provides efficient tools for constructing, validating, and comparing optimal sampling designs for destructive sampled toxicokinetic studies.*

KEY WORDS: optimal design; nonparametric estimation; mean squared error; toxicokinetics; nonlinear mixed model.

INTRODUCTION

In the course of the preclinical toxicology evaluation and according to the ICH guideline (1), toxicokinetics studies are designed to assess the systemic exposure of animals to a drug under investigation. After single and multiple drug administrations, blood samples are collected at several times postdose and the level of exposure achieved after dosing is quantified by the area under the concentration-time curve (*AUC*).

[1]Statistical and Mathematical Sciences, Eli Lilly & Company, 11 rue Granbonpré, B-1348 Mont-Saint-Guibert, Belgium.
[2]Department of Biostatistics, Limburgs Universitair Centrum, Universitaire Campus, B-3590 Diepenbeek, Belgium.
[3]Statistical and Mathematical Sciences, Eli Lilly & Company, 2001 West Main, P.O. Box 708, Greenfield, Indiana 46140.
[4]To whom correspondence should be addressed.

0090-466X/99/0400-0191$16.00/0

In small animals, only a limited volume of blood can be collected from each subject without impairing its health and without interfering with the toxicological evaluation. A classical approach to avoid this risk is to reduce the sampling frequency of each animal and to collect blood from different animals at each time point, as in a destructive setting. The quadrature (2) method for destructive sampling designs is a nonparametric technique that computes the estimate of the true mean $AUC(t_0, t_k)$ as a linear combination of the mean concentrations recorded at successive time points. Due to the mutual independence between all concentrations induced by the destructive sampling procedures, the variance of the estimated $A\hat{U}C(t_0, t_k)$ is also written in terms of the variance of concentrations at sampling time points. The variance of $A\hat{U}C(t_0, t_k)$ can therefore be minimized by selecting appropriate time points and by distributing optimally the total number of samples among these time points.

Among all first-order approximation techniques, the linear trapezoidal rule (3–5), produces a piecewise linear approximation to the true nonlinear pharmacokinetic (PK) profile of a compound. It may therefore produce biased estimates. A method for reducing this bias is to space the time points so that the sums of areas formed by linear segments below and above the true profile tend to be equal.

The problem of sparse sampling in pharmacokinetic and toxicokinetic studies has already been the topic of numerous research articles. Optimal designs have been suggested for the parametric estimation of PK models and the derivation of $A\hat{U}C(t_0, t_k)$ estimates from these models (6–10). Concerning the noncompartmental linear interpolation methods, the location of time points that minimize the mean squared error (*MSE*) (11) has been discussed by Katz and D'Argenio (12) for the case of one sample per time point. The optimal allocation of several animals to these time points has been suggested by Piegorsch and Bailer (13). More recently, Wei (14) introduced an iterative *MSE* minimization algorithm for the selection of time locations and the allocation of animals for the trapezoidal method when considering uncorrelated kinetic model parameters. Although theoretically appealing, the efficiency of their minimization procedure is questionable, because the problems of time location and animal allocation are addressed sequentially and in view of the illustrated clasping of time points in their example.

In previous articles (12–14) optimal designs were developed based upon simple priors given as fixed-effect models with an additive covariance structure. Following the recent interest in fitting nonlinear mixed models to sparsely sampled toxicokinetic data (7), the optimal design theory is hereafter extended to this class of prior models with a Gaussian but otherwise nonrestrictive covariance structure. A strategy is suggested for selecting optimal

destructive designs for the estimation of $AUC(t_0, t_k)$ by any linear interpolation method. The minimization is performed on the scaled *MSE*, using a method due to Brent (15). The efficiency of optimal designs of various sizes is then compared, both on a theoretical basis and by simulation.

This method is first presented for the trapezoidal (3) technique and then extended to any nonparametric $AUC(t_0, t_k)$ estimation method which is linear in the concentrations (2,4,16). An example of this generalization is provided for an integration using the Simpson's rule (17).

For other estimation methods which are not linear in the concentrations, i.e., the log-trapezoidal rule (4,16), an approximation to the method is suggested by estimating the variance of $A\hat{U}C(t_0, t_k)$ using a first-order Taylor series expansion of mean concentrations.

OPTIMAL DESIGNS FOR THE TRAPEZOIDAL RULE

Method

Applying the Bailer method (3), when a total of N animals are to be sampled at $(k+1)$ time points from t_0 to t_k, and at each of the time points (t_j), n_j animals are sampled, so that $\sum_{j=0}^{k} n_j = N$, the mean $A\hat{U}C(t_0, t_k)$ is computed as a linear interpolation of average concentrations $(\bar{c}_j)$ across time points (t_j), for $j = 0, \ldots, k$,

$$A\hat{U}C(t_0, t_k) = \sum_{j=0}^{k} \delta_j \bar{c}_j$$

where

$$\delta_j = \begin{cases} \frac{1}{2}(t_1 - t_0) & \text{for } j = 0 \\ \frac{1}{2}(t_{j+1} - t_{j-1}) & \text{for } j = 1, \ldots, k-1 \\ \frac{1}{2}(t_k - t_{k-1}) & \text{for } j = k \end{cases} \tag{1}$$

As samples are independently collected from different animals, the variance of mean $A\hat{U}C(t_0, t_k)$ is written in terms of the variance of sampled concentrations (σ_j^2) divided by the number of samples (n_j) collected at that time point

$$\sigma^2[A\hat{U}C(t_0, t_k)] = \sum_{j=0}^{k} \frac{\delta_j^2 \sigma_j^2}{n_j} \tag{2}$$

As seen in Piegorsch and Bailer (13), Eq. (2) can be minimized with respect to the number of animals (n_j) per sampling time, using the Lagrange multipliers method. The relative number of animals to be assigned at each

time point t_j is proportional to the standard deviation of the average concentration at time t_j

$$\hat{n}_j = N\frac{|\delta_j|\sigma_j}{\sum_{l=0}^{k}|\delta_l|\sigma_l} \tag{3}$$

From Eq. (3), it appears that balanced designs, which allocate the same number of animals (N/k) to each time point, will be optimal when times are selected so that intervals δ_j are inversely proportional to the standard deviation of concentrations σ_j. The existence of such designs is not guaranteed, as it depends solely on the variability profile of concentrations. Designs that are close to being balanced can however be created and their efficiencies will be compared to those of unbalanced designs using simulations.

When Eq. (3) is used as weights in Eq. (2), the variance of $A\hat{U}C(t_0, t_k)$ is rewritten as a function of the sample locations (t_j) and the total number of animals (N) only

$$\sigma^2[A\hat{U}C(t_0, t_k)] = \sum_{j=0}^{k}\frac{\delta_j\sigma_j\sum_{l=0}^{k}\delta_l\sigma_l}{N} \tag{4}$$

The variance of $A\hat{U}C(t_0, t_k)$ is inversely proportional to N. When the number of animals doubles, the variance is divided by two. For a specified sampling design, when prior information about the variability of concentrations is available, the total number of animals N can be determined to attain a desired level of precision, that is, a minimum width for the confidence interval on mean $A\hat{U}C(t_0, t_k)$.

The squared bias of $A\hat{U}C(t_0, t_k)$, (11,12) is calculated as the squared difference between the expected value of $A\hat{U}C(t_0, t_k)$ and of the true area under the concentration profile $AUC_{\text{True}}(t_0, t_k)$

$$Bias^2[A\hat{U}C(t_0, t_k)] = \{\text{E}[A\hat{U}C(t_0, t_k)] - AUC_{\text{True}}(t_0, t_k)\}^2 \tag{5}$$

It is a function of the number of time points k and their locations $(t_0, \ldots, t_k)$ but not of the N nor of their assignment to time points $(n_0, \ldots, n_k)$. As stated before, an unbiased design will be selected by equaling the areas formed by the true nonlinear profile above and below the linear segments between time points.

After a first-order oral absorption, the pharmacokinetic profile is usually convex in the absorption phase and then concave through the elimination phase. When considering profiles with both a convex and a concave part, it is possible to find unbiased designs by properly selecting time points that produce equal areas in the convex and concave parts. A well-chosen single intermediate time point is, in this case, sufficient to annul the bias of $A\hat{U}C(t_0, t_k)$. Under such circumstances, there is no correlation between the

number of time points (k) and the bias. The bias is not necessarily changed when the number of time points is increased. However, the number of unbiased designs increases as the number of time points (k) increases.

On the other hand, some profiles can only be either concave like after an iv bolus administration, or convex like after a continued infusion. In these cases, the trapezoidal interpolation always produces positively or negatively biased estimates whatever the sampling scheme is and the bias will be negatively correlated to the number of time points. As the number of time points increases, the bias decreases and tends asymptotically to zero.

Assuming that the concentration profile of the investigated compound can be fairly well characterized by a nonlinear mixed effect model of the time (8), the determination of σ_j^2, $E[A\hat{U}C(t_0, t_k)]$, and $AUC_{\text{True}}(t_0, t_k)$, which is required for evaluating both the bias and variance of $A\hat{U}C(t_0, t_k)$, will be made based on some prior experiences with the compound.

In this framework, the drug concentration at each time point t_j, $j = 0, \ldots, k$, in the ith animal sampled at that time, $i = 1, \ldots, n_j$, is often characterized by a nonlinear function of time, with animal-specific parameters β_i and residual error ε_{ij}.

In the present context of pharmacokinetics (7), parameters β_i are usually selected to characterize the individual rates and extents of absorption, distribution, and elimination of the compound. These parameters are animal-specific and assumed to be multivariate normally distributed with population mean β and block-diagonal covariance matrix Σ. The residual error is also assumed to be normally distributed and independent of the animal effects. Finally, the nonlinear function should be differentiable with respect to its variance components ($\beta_i, \varepsilon_{ij}$) around their expected values. The model is written as

$$c_{ij} = f(t_j, \beta_i, \varepsilon_{ij}) \text{ where } \beta_i \sim MVN(\beta, \Sigma) \text{ and } \varepsilon_{ij} \sim N(0, \sigma^2) \tag{6}$$

The constraints of differentiability of $f()$ and normality of $(\beta_i, \varepsilon_{ij})$ are essentially required in order to derive the variance [Eq. (4)] and bias [Eq. (5)] of $A\hat{U}C(t_0, t_k)$ using a first-order Taylor series expansion of the concentration model [Eq. (6)] around its mean parameters $(\beta, 0)$. This yields the following approximation

$$f(t_j, \beta_i, \varepsilon_{ij}) \approx f(t_j, \beta, 0) + \frac{\partial f(t_j, \beta, 0)}{\partial \beta_i}(\beta_i - \beta) + \frac{\partial f(t_j, \beta, 0)}{\partial \varepsilon_{ji}}(\varepsilon_{ij} - 0) \tag{7}$$

From this approximation, the expected mean concentration becomes the function of mean parameters: $E[f(t_j, \beta_i, \varepsilon_{ij})] = f(t_j, \beta, 0)$ and the expected mean $A\hat{U}C(t_0, t_k)$ is the linear combination of that function across design points t_j: $j = 0, \ldots, k$, as in Eq. (1).

As shown in Wei (14), the variance of concentrations can be approximated using the covariance structure of the model specified in Eq. (6), using the Delta method, provided that the model has nonzero differentials with respect to β_i and ε_{ij} at their expected value (β, 0). The variance of concentrations at time t_j is thus approximated by

$$\sigma_j^2 \approx \left(\frac{\partial f}{\partial \beta_i|_\beta} \frac{\partial f}{\partial \varepsilon_{ij}|_0} \right) \begin{pmatrix} \Sigma & 0 \\ 0 & \sigma^2 \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial \beta_i|_\beta} \\ \dfrac{\partial f}{\partial \varepsilon_{ij}|_0} \end{pmatrix} \tag{8}$$

where $\partial f/\partial x|_a$ is the derivative of f with respect to x at $x = a$.

Substituting σ_j^2 defined in Eq. (8) into Eq. (4), the approximated variance of $A\hat{U}C(t_0, t_k)$ becomes a function of the prior model specified in Eq. (6), including the response function of its mean (β) and covariance structure (Σ), the total number of animals (N), and the number (k) and locations of time points (t_j) only. The optimal distribution of animals across time points (n_j) is not explicitly part of that function, but can be retrieved from it, if the location of time points is known from Eq. (3).

A similar model expansion as in Eq. (7) can be applied to approximate the $AUC_{\text{True}}(t_0, t_k)$ as the integral over the time interval (t_0, t_k) of Eq. (6) of mean parameters

$$AUC_{\text{True}}(t_0, t_k) \approx \int_{t_0}^{t_k} f(s, \beta, 0)\, ds \tag{9}$$

With these simplifications, the bias of $A\hat{U}C(t_0, t_k)$ in Eq. (5) depends only on the response function, mean parameters (β), and on the number (k) and locations of time points (t_j). The bias remains independent of the distribution of animals across time points (n_j).

With priors on the shape and variability of the concentration profile expressed in the form of a nonlinear mixed-effect model, as in Eq. (6), optimal designs of various sizes (k, N) can be selected by combining the variance and squared bias of $A\hat{U}C(t_0, t_k)$ into a single criterion, for minimization. Following ideas from Katz and D'Argenio (12), the aggregated function is chosen to be the scaled mean squared error (MSE). It is defined as the square root of the mean squared error divided by the mean $A\hat{U}C(t_0, t_k)$.

$$\begin{aligned} MSE_{\text{Scaled}}(A\hat{U}C(t_0, t_k)) &= \frac{\sqrt{MSE[A\hat{U}C(t_0, t_k)]}}{A\hat{U}C(t_0, t_k)} \\ &= \frac{\sqrt{Bias^2[A\hat{U}C(t_0, t_k)] + \sigma^2[A\hat{U}C(t_0, t_k)]}}{A\hat{U}C(t_0, t_k)} \end{aligned} \tag{10}$$

A scaled criterion has been preferred to the pure *MSE*, as consistently suggested in previous papers (11–14) to avoid generating designs that produce negatively biased and less precise $A\hat{U}C(t_0, t_k)$. The linear interpolation technique is not unbiased and as the variability is often proportional to concentrations, designs that minimize the pure *MSE* would be chosen at time points of low expected concentrations, producing a systematic negative bias in the estimation of $A\hat{U}C(t_0, t_k)$. Although, the *MSE* of these estimates is optimized, the precision of the estimated $A\hat{U}C(t_0, t_k)$ in terms of coefficient of variation usually remains poor. The scaling of the *MSE* is recommended in our criterion to tackle this issue.

With the availability of symbolic mathematical packages, the scaled *MSE* can easily be expressed as a function of the total number of animals (N) and the number of times (k) and their locations $t_j: j = 1, \ldots, k-1$. When N and k are held fixed, Eq. (10) becomes a response surface of the times: t_j. Optimal time points are selected in the (t_0, t_k) interval to minimize that function. The minimization is made using the Powell's quadratically convergent method (15). An algorithm (presented in the Appendix) has been developed to implement this complete methodology in Mathematica™ and an example of its use is detailed in the following section.

Illustration

We illustrate the optimal design selection method with a pharmacokinetic profile from our current experimental practice within the Lilly Research Laboratories. The profile of interest can be estimated by a one-compartment first-order absorption model defined in Eq. (11) with normally distributed absorption rate constant ($ka = 0.5$), bioavailability ($F = 0.4$), and elimination rate constant ($ke = 0.04$). A quite large coefficient of variation of 50% is observed in all parameters which are well correlated (see Table I) An error of 10% is injected into our model to mimic the precision of the bioassay.

$$C_{ij} = \frac{Dose\, F_i ka_i}{V_i(ka_i - ke_i)}(e^{-ke_i t_j} - e^{-ka_i t_j})(1 + \varepsilon_{ij}) \tag{11}$$

$$[ka_i, ke_i, (F/V)_i] \sim MVN[(ka, ke, F/V), \Sigma]$$

Table I. Mean, Standard Deviation, and Correlation of Pharmacokinetic Parameters

			Correlation matrix		
Pharmacokinetic parameters	$\bar{x}$	SD	ka	F/V	ke
ka	0.5	0.25	1	−0.8	−0.9
ke	0.04	0.02	−0.8	1	0.9
F/V	0.4	0.2	−0.9	0.9	1

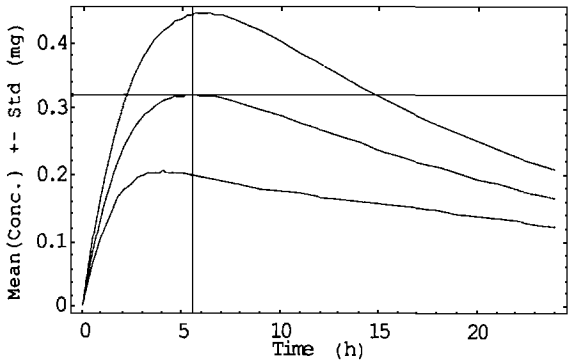

Fig. 1. Mean ± 1 standard deviation of the concentration function. Gridlines locate the expected T_{max} (5.49 hr) and C_{max} (0.321 mg/L).

where

$$\Sigma = \begin{pmatrix} \sigma_{ka}^2 & \sigma_{kake} & \sigma_{kaF/V} \\ \sigma_{kake} & \sigma_{ke}^2 & \sigma_{keF/V} \\ \sigma_{kaF/V} & \sigma_{keF/V} & \sigma_{F/V}^2 \end{pmatrix} \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Following a single administration (*Dose* = 1 mg), as shown in Fig. 1, the expected profile reaches a maximum concentration ($C_{max} = 0.321$ mg/L), 5.49 hr postdose. The $AUC(0, 24\text{ hr})$ is equal to 5.838 mg hr/L.

Optimal designs are found to estimate the $AUC(0, 24\text{ hr})$ by linear interpolation. To do so, the extreme points 0 and 24 hr are fixed in the design and intermediate points are found to minimize the scaled *MSE* [Eq. (10)] of $AUC(0, 24\text{ hr})$. The distribution of animals among time points is derived by Eq. (3), with the additional constraint that, after a single administration, the predose concentration is fixed to 0 and no sample is needed at that time point.

For the two samples postdose problem ($k = 2$), Fig. 2 illustrates the bias, variance, and scaled *MSE* of $\hat{AUC}(0, 24\text{ hr})$ as a function of only one intermediate selected time point (t_1) moving between 0 and 24 hr, when the total number of animals (N) is fixed at 12. The bias is minimized at 4.64 hr and the variance profile is roughly proportional to concentrations. The scaled *MSE* (12.1%) is minimized at 4.3 hr postdose, which provides an optimal design {0, 4.3, 24 hr}.

Table II displays optimal 12-animal designs of up to five time points postdose. The number of animals to be sampled at each time point is presented; the scaled *MSE* (%), *bias* (%), and standard error of the mean (*SEM* in %) are also reported. The scaled *MSE* decreases up to four time points

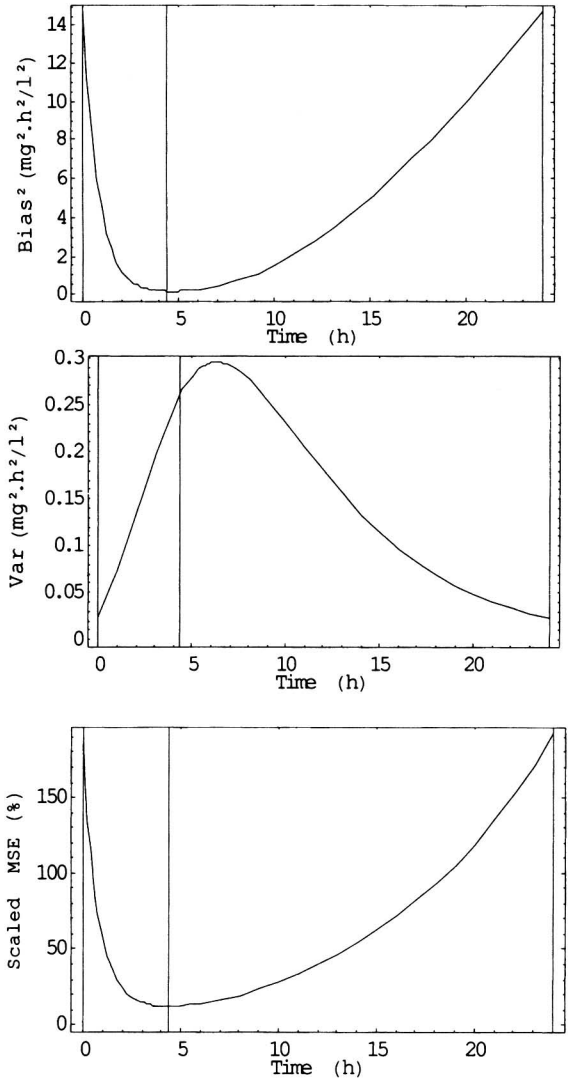

Fig. 2. Squared bias, variance, and scaled mean squared error (*MSE*) of $\hat{AUC}(0, 24\text{ hr})$ versus time for a 12-animal design including one sampling time between 0 and 24 hr. The design which minimizes the scaled *MSE* (0, 4.3, 24 hr) is highlighted by vertical lines in all three graphs.

Table II. Optimal 12-Animal Designs Including Two to Five Samples Postdose

No. of points (k)	Designs: $\begin{Bmatrix} t_1 \dots t_j \dots t_k \\ n_1 \dots n_j \dots n_k \end{Bmatrix}$	Scaled *MSE* (%)	*Bias* (%)	*SEM* (%)
2	4.3 24 9.1 2.9	12.1	7.0	9.4
3	1.9 4.9 24 1.0 8.3 2.7	9.8	2.4	9.5
4	1.2 2.6 4.7 24 0.4 0.9 8.0 2.7	9.6	1.9	9.4
5	1.2 2.7 4.8 22.6 24 0.4 1.0 7.5 2.9 0.2	9.6	2.0	9.4

and then stabilizes around 9.6%. There is no gain in adding a fifth sampling time to the design. The optimal distribution of animals across time points is very unbalanced, with very few animals at early samples and more animals at 4.7 and 24 hr, in order to cover the variability generated over the large time interval. As seen in Fig. 3, no sample is selected between 4.7 and 24 hr because the expected profile is fairly linear in the elimination phase.

In Table III, optimal 4-point designs with the number of animals N varying from 4 up to 120 are displayed. As the number of animals increases, the overall precision of $\hat{AUC}$ is improved. Not surprisingly, the variance of $\hat{AUC}$ is inversely proportional to the number of animals. As the *MSE* is a sum of both squared bias and variance, when the variance is reduced, the contribution of the bias term becomes predominant in the criterion. Therefore, designs optimized for more animals tend to be less biased. As the number of animals increases ($N = 120$), the best design will be chosen to optimize the bias essentially. To do so, time points are placed in regions of higher variability of concentrations (Fig. 4).

Optimal designs can also be found near to a prespecified distribution of animals across time points. Table IV presents such designs with a distribution from very unbalanced to balanced.

The theoretical precision of mean $\hat{AUC}(0, 24\text{ hr})$, as calculated using the Delta method (18), has been validated with a simulation study. The median and interquartiles are reported from 10,000 simulated profiles.

When balancing animals across time points, optimal designs are found to be more equally spaced (Fig. 5). The bias of each design is similar (around 2%), except for the balanced design (3.6%), and the precision is improved when choosing unbalanced designs. With an unrestricted weighing scheme, time points are more flexibly chosen to produce a better fit to the nonlinear pharmacokinetic profile and therefore the bias of such designs will be decreased. With balanced designs, the variance function dictates the location

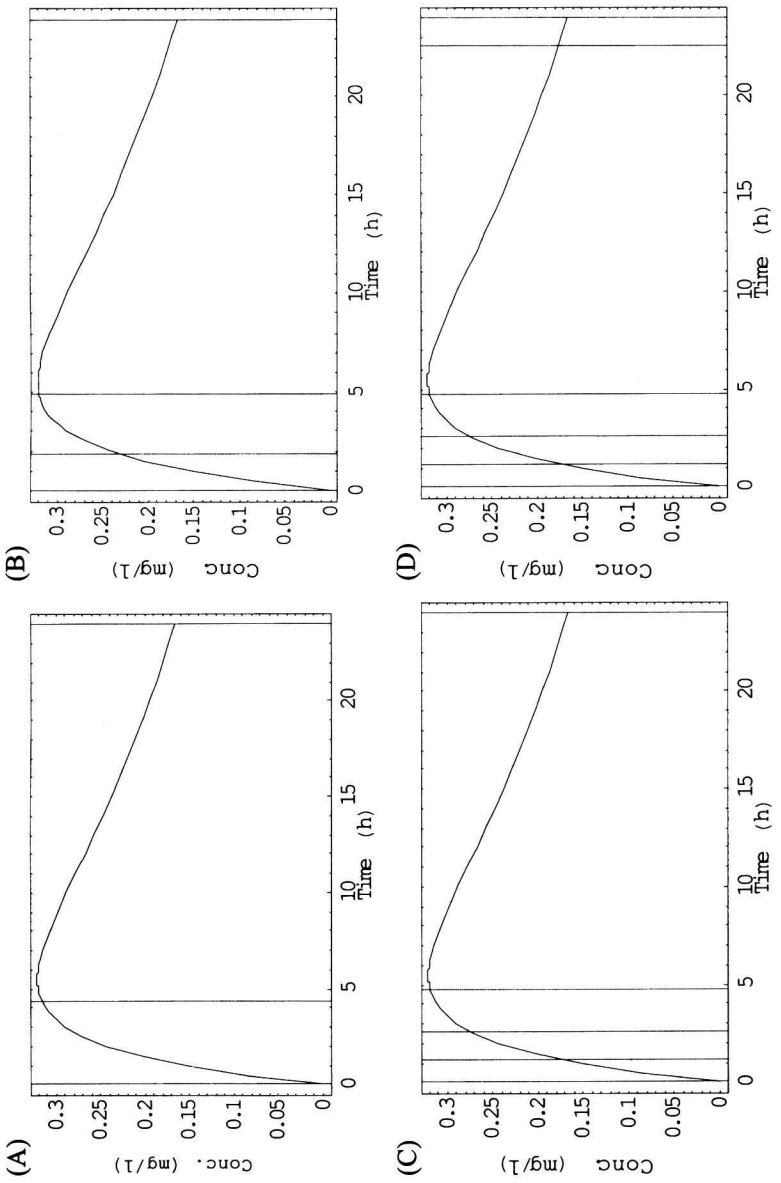

Fig. 3. Expected concentration by time profile and 12-animal optimal designs for the estimation of $AUC(0, 24\,\text{hr})$, with: (A) $k = 2$, (B) $k = 3$, (C) $k = 4$, and (D) $k = 5$ samples postdose. The last point is fixed at 24 hr.

Table III. Optimal Designs with Four Samples Postdose with Varying Numbers of Animals

No. of animals	Design $\begin{Bmatrix} t_1 & t_2 & t_3 & t_4 \\ n_1/N & n_2/N & n_3/N & n_4/N \end{Bmatrix}$				Scaled *MSE* (%)	*Bias* (%)	*SEM* (%)
4	1.1	2.3	4.0	24	16.1	3.8	15.6
	3%	6%	66%	25%			
9	1.2	2.5	4.6	24	11.0	2.3	10.7
	3%	7%	67%	23%			
12	1.2	2.6	4.7	24	9.6	1.9	9.4
	3%	8%	67%	22%			
15	1.2	2.7	4.9	24	8.6	1.6	8.4
	3%	8%	67%	22%			
120	1.4	3.1	5.9	24	3.1	0.4	3.1
	4%	10%	66%	20%			

of time points and the compared optimality of selected designs is affected in terms of both bias and variance.

Results from the simulation are consistent with theoretical predictions and the stability of *SEM*(%), measured by the width of the interquartile range, is not affected by the unbalanced distribution of animals. The small discrepancies between predictions and simulations, as observed in this particular case, could however be much larger for other models. The Taylor series expansion, as used in Eq. (8) might be the principal cause of discrepancies when parameters (ka, ke) involved in the nonlinearity parts of the model defined in Eq. (6) have large variances. In that case, the linear approximation remains locally correct around means, but tends to be less accurate as the allowed variability from the means becomes important. Another cause of the discrepancies could be related to the simulation mechanism, when normality assumptions are made on strictly positive parameters.

EXTENSION TO OTHER *AUC* ESTIMATORS

Linear Methods

Many other $A\hat{U}C(t_0, t_k)$ estimation techniques than the trapezoidal rule are based on the interpolation of observed concentrations (4,5). Among these methods, some, as the trapezoidal rule or the Simpson's rule (17), are built upon a Taylor series expansion of the integrated *AUC* and they remain linear in the concentration. Other methods, as the log-trapezoidal rule (16), are created from other mechanisms and are not linear in the concentrations. The optimal design theory will first be generalized from the trapezoidal to any linear method and then to nonlinear techniques.

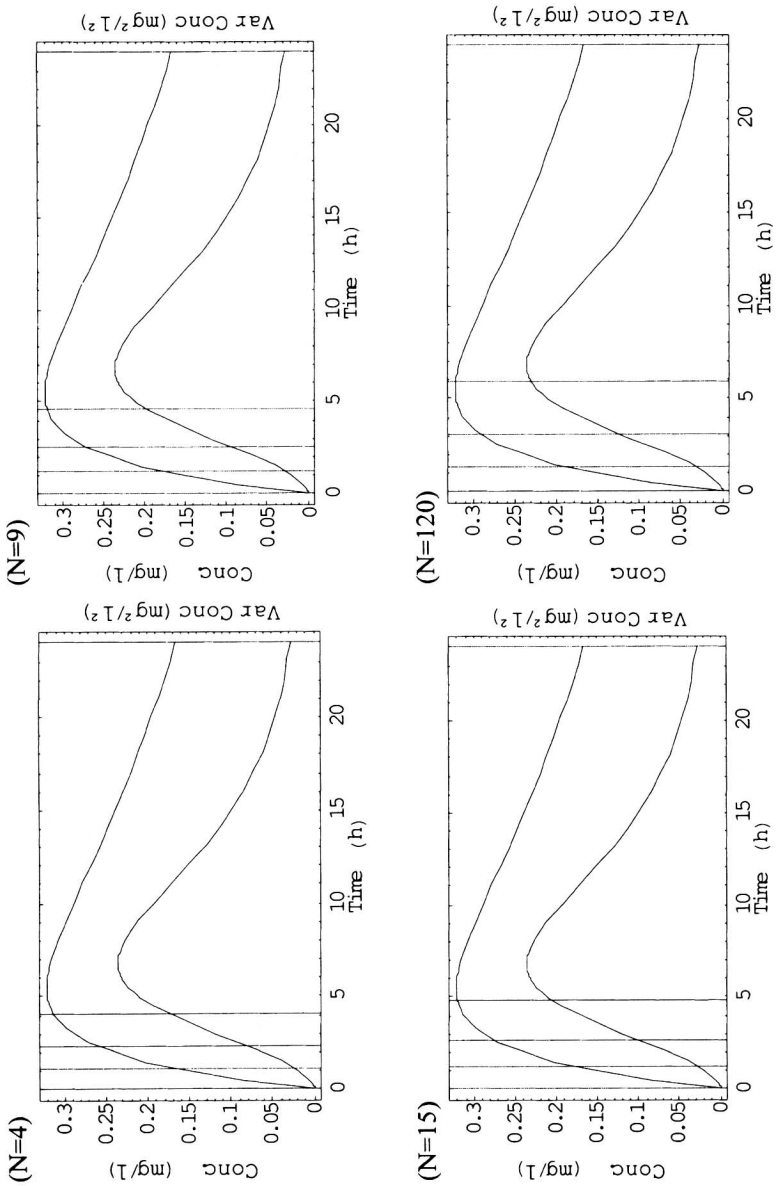

Fig. 4. Expected and variance of concentration by time profile for four-sample postdose optimal designs for the estimation of $AUC(0, 24\,\text{hr})$, including 4, 9, 15, or 120 animals in total (N).

Table IV. Optimal 12-Animal Designs with Fixed Distribution of Animals

Design $\begin{Bmatrix} t_1 & t_2 & t_3 & t_4 \\ n_1 & n_2 & n_3 & n_4 \end{Bmatrix}$				Theoretical			Simulated *SEM* (%)		
				Scaled *MSE* (%)	*Bias* (%)	*SEM* (%)	Q1	Median	Q3
1.7	3.6	4.8	24	9.7	2.0	9.4	7.6	9.7	12.9
1	1	7	3						
1.7	4.5	5.1	24	10.0	1.8	9.5	8.0	10.3	13.6
1	2	6	3						
2.5	6.3	11.8	24	10.3	1.7	10.0	8.6	10.9	14.4
2	3	5	2						
3.0	9.7	12.2	24	11.1	3.6	9.9	8.5	10.9	14.6
3	3	3	3						

When considering linear interpolation methods, the estimated mean $A\hat{U}C(t_0, t_k)$ can again be calculated as a linear combination of mean concentrations ($\bar{c}_j$) across time points t_j: $j = 0, \ldots, k$, with weights (w_j) defined as a function of time points, according to the assumed relationship between the times and concentrations (linear, quadratic, hyperbolic...)

$$A\hat{U}C(t_0, t_k) = \sum_{j=0}^{k} w_j \bar{c}_j \tag{12}$$

For instance, the generalized Simpson's rule (17) is an estimation method in which the concentration profile is approximated by a piecewise quadratic function. For three successive time points: (t_{j-1}, t_j, t_{j+1}), the $A\hat{U}C(t_{j-1}, t_{j+1})$ is estimated over two time intervals θ_j and θ_{j+1} as

$$A\hat{U}C(t_{j-1}, t_{j+1}) = \left(\frac{\theta_j}{2} - \frac{\theta_{j+1}}{6}\right)\bar{c}_{j-1} + \left(\frac{2\theta_j}{3} - \frac{2\theta_{j+1}}{3}\right)\bar{c}_j + \left(\frac{\theta_{j+1}}{2} - \frac{\theta_j}{6}\right)\bar{c}_{j+1} \tag{13}$$

where $\theta_j = (t_j - t_{j-1})$. For any even number of time points $(k+1)$, the complete $A\hat{U}C(t_0, t_k)$ is then calculated as the sum of partial $A\hat{U}C(t_{j-1}, t_{j+1})$. In that summation, the weights (w_j) are a linear combination of the time intervals θ_j and θ_{j+1}.

When the number of time points $(k+1)$ is odd, the $A\hat{U}C(t_0, t_k)$ cannot be fully calculated based on the piecewise quadratic curves. Additional trapezoidal estimations are required, for an odd number of time intervals (θ_j). The $A\hat{U}C(t_0, t_k)$ is then estimated by a mixture of piecewise quadratic and linear segments. As this estimator depends on the allocation of sets of time intervals to quadratic or linear interpolation methods, optimal designs can

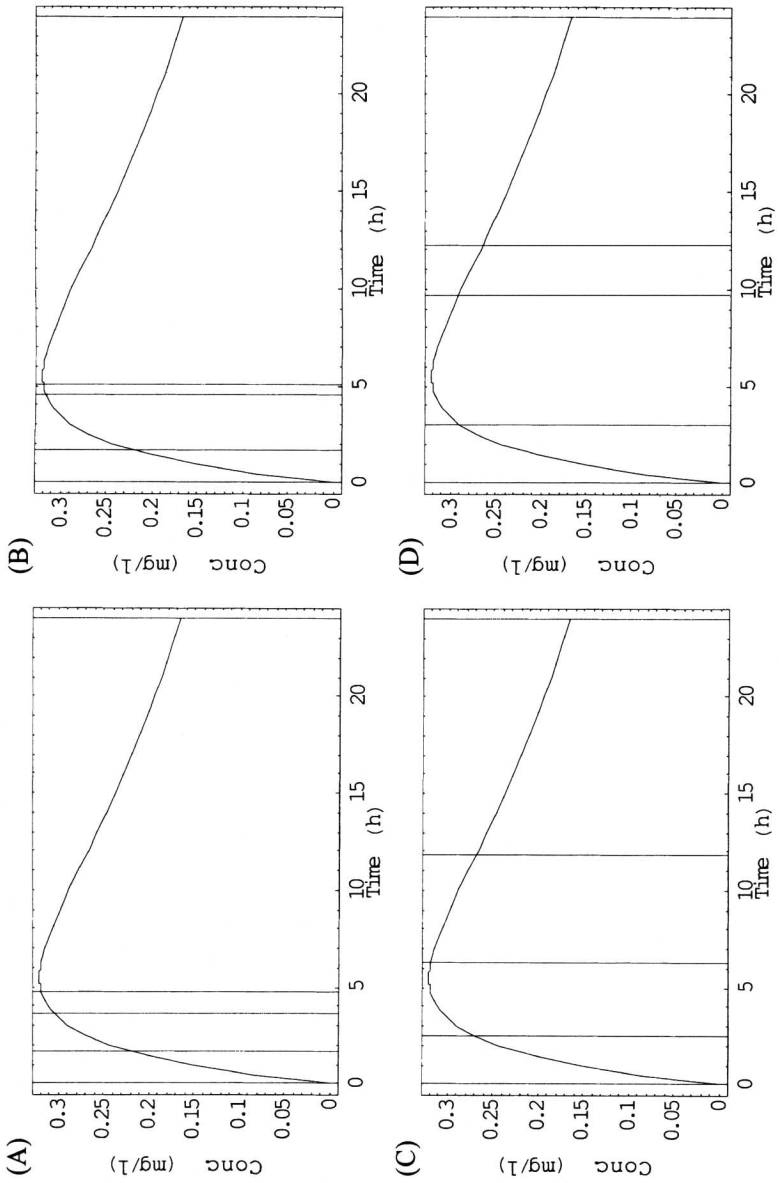

Fig. 5. Expected concentration by time profile with 12 animals and five-sample postdose optimal designs for the estimation of $AUC(0, 24\,hr)$, including very unbalanced to balanced distribution of animals across time points. Number of animals per time point: (A) 1–1–7–3, (B) 1–2–6–3, (C) 2–3–5–2, and (D) 3–3–3–3.

Table V. Optimal Simpson Designs with Four Samples Postdose with Varying Numbers of Animals

No. of animals	Designs: $\begin{matrix} t_1 \\ n_1/N \end{matrix}$	$\begin{matrix} t_2 \\ n_2/N \end{matrix}$	$\begin{matrix} t_3 \\ n_3/N \end{matrix}$	$\begin{matrix} t_4 \\ n_4/N \end{matrix}$	Scaled *MSE* (%)	*Bias* (%)	*SEM* (%)
4	1.6 8.5%	4.0 46.2%	19.0 45.4%	24 0.01%	16.0	3.2	15.6
9	1.7 9.6%	4.6 48.4%	19.2 42.1%	24 0.004%	10.8	2.0	10.6
12	1.7 9.9%	4.9 49.0%	19.2 41.0%	24 0.008%	9.4	1.7	9.3
15	1.7 10.2%	5.0 49.5%	19.3 40.2%	24 < 0.000%	8.5	1.5	8.3
120	2.17 15.37%	6.33 49.33%	19.58 35.29%	24 0.004%	3.0	0.1	3.0

only be selected if the assignment has been initially defined for all intervals. This case of an odd number of time points $(k + 1)$ is not addressed here.

As the Simpson's Eq. (13) is still linear in mean concentrations, the optimization method presented above remains valid provided that the δ_j are replaced by the newly defined weights (w_j) in Eqs. (2), (3), and (4).

Illustration of the Simpson's Method

Back to the initial example, optimal sampling designs with $(k = 3)$ are constructed, based on the Simpson's interpolation method for a varying number of animals N, from 4 up to 120 (Table V). Similarly to optimal designs for the trapezoidal rule, when the number of animals increases, time points are moved to produce less biased $A\hat{U}C(0, 24\text{ hr})$. The scaled *MSE* of optimal designs for the trapezoidal and Simpson's rules of an equal size (k, N) is similar (refer to Tables III and V). This suggests that, for the same cost, both estimation methods lead to similar accuracy and precision of $A\hat{U}C$.

Since the trapezoidal rule is a first-order approximation to the Simpson's rule, the location of time points can be compared between both methods, in view of the shape of the expected profile. In both cases, the first sampling points are placed to absorb the initial nonlinearity of the curve, up to T_{max}. In this absorption phase, more time points are required for the trapezoidal than for the Simpson's rule. Then, the descending phase is rather linear from T_{max} up to 24 hr, so that, in theory (9) only two time points located at the extremes are required for both methods. In light of this, the last time point selected using the Simpson's method (> 19 hr) is probably

useless, because it is located very close to 24 hr and the weight of the 24 hr point is very low.

Nonlinear Methods

For completeness, we conclude by presenting a theoretical extension of the above methodology, for the case when the $A\hat{U}C(t_0, t_k)$ estimator is not linear in the mean concentrations. When $A\hat{U}C(t_0, t_k)$ is estimated by the log-trapezoidal rule (16), then

$$A\hat{U}C(t_0, t_k) = \sum_{j=0}^{k-1} \frac{\theta_{j+1}(\bar{c}_{j+1} - \bar{c}_j)}{\log[\bar{c}_{j+1}/\bar{c}_j]} \tag{14}$$

If $\bar{c}_j = 0$ or $\bar{c}_{j+1} = 0$ or $\bar{c}_j = \bar{c}_{j+1}$, a linear trapezoidal step is taken instead, by replacing summand_j of Eq. (14) with summand_j of Eq. (1).

Due to the nonlinearity, the variability of $A\hat{U}C(t_0, t_k)$ is no longer computed in terms of the variability of concentrations. It could however be underestimated using another first-order Taylor series expansion of $A\hat{U}C(t_0, t_k)$ around expected concentrations $E(\bar{c}_j)$

$$A\hat{U}C(t_0, t_k) \approx \sum_{j=0}^{k-1} \left\{ \theta_{j+1} \left[\frac{E(\bar{c}_{j+1}) - E(\bar{c}_j)}{\log[E(\bar{c}_{j+1})/E(\bar{c}_j)]} + [\bar{c}_{j+1} - E(\bar{c}_{j+1})]a_{j+1} + [\bar{c}_j - E(\bar{c}_j)]b_{j+1} \right] \right\} \tag{15}$$

where

$$a_{j+1} = \frac{\log[E(\bar{c}_{j+1})/E(\bar{c}_j)] - 1 + E(\bar{c}_j)/E(\bar{c}_{j+1})}{\log[E(\bar{c}_{j+1})/E(\bar{c}_j)]^2}$$

and

$$b_{j+1} = \frac{\log[E(\bar{c}_j)/E(\bar{c}_{j+1})] - 1 + E(\bar{c}_{j+1})/E(\bar{c}_j)}{\log[E(\bar{c}_{j+1})/E(\bar{c}_j)]^2}$$

The approximation of $A\hat{U}C(t_0, t_k)$ is again expressed as a linear combination of mean concentrations, from which the variance of $A\hat{U}C(t_0, t_k)$ can be written in terms of the variance of mean concentrations, as

$$\sigma^2[A\hat{U}C(t_0, t_k)] \approx \sum_{j=0}^{k-1} \left\{ \theta_{j+1}^2 \left[\frac{a_{j+1}^2 \sigma_{j+1}^2}{n_{j+1}} + \frac{b_{j+1}^2 \sigma_j^2}{n_j} \right] \right\} \tag{16}$$

Substituting this variance term into the *MSE* formula should yield approximated optimal designs for the log-trapezoidal method.

DISCUSSION

The selection of the number and location of sampling times and the number of animals per time point is a critical operation when designing a destructive toxicokinetic study. This paper presents a practical strategy to select study designs that optimize the scaled mean squared error of $A\hat{U}C(t_0, t_k)$, estimated by linear quadratures. A scaled function was chosen to avoid systematic selection of negatively biased designs and to optimize the precision of $A\hat{U}C(t_0, t_k)$ estimates. The performance of the method is detailed in an example and illustrated with a simulation study.

The algorithm uses the symbolic resolution capability available in any modern mathematical packages to provide an analytical form of the variance and squared bias functions. Advantages of having such a form for each function are that optimal solutions are exact, computations are very fast, and the algorithm remains flexible and adaptable to any type of prior model.

A vast panel of population pharmacokinetic models can be envisaged in the optimization algorithm. As the variance of concentrations is calculated by applying the Delta method to this nonlinear model, the unique restriction for constructing optimal designs is to start with models having a normal error structure. When no prior variability estimates are available for the model parameters, design can still be constructed, provided the total variability is *a priori* given as a function of concentrations.

The variance of $A\hat{U}C(t_0, t_k)$ is inversely proportional to the total number of animals; therefore, the sample size can be calculated based on power considerations to control the precision of estimates. By allowing a flexible redistribution of animals to time points, designs tend to be more efficient but still as stable as balanced designs.

Further refinements of the method could be envisaged in order to control the robustness of designs to variations in prior model assumptions. In a Bayesian framework (7–10), a solution should be optimizing the expected squared bias of $A\hat{U}C(t_0, t_k)$ over the prior parameters' distributions. Although theoretically appealing, an analytical form of this function remains difficult to manage.

The scope of the method could also be extended to designs with replicated sampling per animal. With these new conditions, the variances of $A\hat{U}C(t_0, t_k)$ in Eq. (2) might be changed to include covariance terms for dependent samplings. An analytical form of the updated variance formula is given by Yeh (7) and comparison of the efficiency of given designs is

possible. However, the algorithm for selecting optimal designs in this framework has not yet been implemented.

APPENDIX

Sample of Mathematica TM Code

Prior Model and Design

```
P: Vector of mean parameters
V: One block of the covariance matrix for P
D: Vector of design points (t0, . . . , tk)
t: Time
Y[t, P]: Model for concentrations
M: Interpolation method (Trapeze or Simpson)
W: Vector of a given allocation of animals to time
   points
```

AUC Integral

$$\texttt{TrueAUC[min\_, max\_, P\_]} := \texttt{Evaluate}\left[\int_{\texttt{min}}^{\texttt{max}} \texttt{Y[t, P]}\,\delta t\right]$$

Trapezoidal Rule

```
Trapeze[D_] := Flatten[{1/2 * (D[[2]] - First[D]),
    Table [1/2 (D[[i+1]] - D[[i - 1]]),
          {i, 2, Length[D] - 1}],
          1/2 (Last[D] - D[[Length[D] - 1]])}]
```

Simpson's Rule

```
Deltaimoins1[D_, i_]
:= (-D[[i+1]]/6 - D[[i - 1]]/2
  +4/6 * D[[i]])

Deltai[D_, i_] := 2/3 (D[[i+1]] - D[[i - 1]])

Deltaiplus1[D_, i_]
:= (D[[i+1]]/2+D[[i - 1]]/6
  - 4/6 * D[[i]])

Simpson[D_
:= Flatten[{Deltaimoins1[D,2], Deltai[D,2],
    Table[{{Deltaiplus1[D, i]+Deltaimoins1[D, i+2]},
         {Deltai[D, i+2]}}, {i, 2, Length[D] - 2, 2}],
         Deltaiplus1[D, Length[D] - 1]}]
```

Bias

```
VectY[D_, P_] := Table[Y[D[[i]], P], {i, 1, Length[D]}]

AUC[D_, P_, M_] := M[D] . VectY[D, P]

Bias[D_, P_, M_] := (AUC[D, P, M]
                     -TrueAUC[First[D],
                     Last[D], P])^2
```

Variance (Optimal Allocation of Animals to Time Points)

```
VectDY[t_, P_] := Evaluate[((∂#1Y[t, P]&)/@{1, eps})]

VarY[t_, P_, V_] := VectDY[t, P] . V . VectDY[t, P]

VectWAUC[D_, P_, V_, M_]
:= Table[VarY[D[[i]], P, V] ^0.5,
       {i, 1, Length[D]}] * Abs[M[D]]

SumWAUC[D_, P_, V_, M_]
:= VectWAUC[D, P, V, M] . Table[{1}, {i, 1,
Length[D]}]

W[N_, D_, P_, V_, M]
:= (N * VectWAUC[D, P, V, M]/SumWAUC[D, P, V, M]
  [[1]]) ^-1/. ComplexInfinity→0

VectVarY[D_, P_, V_]
:= Table[VarY[D[[i]], P, V],
  {i, 1, Length[D]}]

VarAUC[N_, D_, P_, V_, M_]
:= (VectVarY[D, P, V] * W[N, D, P, V, M]). M[D]^2
```

MSE (Optimal Allocation of Animals to Times)

```
CV[N_, D_, P_, V_, M_]
:= (VarAUC[N, D, P, V, M]
  +Bias[D, P, M]) ^0.5/AUC[D, P, M]

MSE[N_, D_, P_, V_, M]
:= VarAUC[N, D, P, V, M]
  +Bias[D, P, M]
```

Variance and MSE (Given Allocation of Animals to Times)

```
VectWVarY[D_, W_, P_, V_]
:= Table[VarY[D[[i]], P, V], {i, 1, Length[D]}]/W
```

```
/.{ComplexInfinity→0, Indeterminate→0}

DiscreteVarAUC[D_, W_, P_, V_, M_]
:= (VectWVarY[D, W, P, V]) . M[D]^2

CVAUCDiscrete[D_, W_, P_, V_, M_]
:= (DiscreteVarAUC[D, W, P, V, M]
  +Bias[D, P, M])  ^0.5/AUC[D, P, M]
```

ACKNOWLEDGMENTS

We thank Jacqueline Carleer, Jean-Marie De Suray, Pierre Gilbert, Glen Rush, and Kristel Vanbelle from the toxicology and drug-metabolism department of Lilly MSG for their interest and support in this research. We also thank two anonymous referees for helpful comments on the manuscript.

REFERENCES

1. ICH. Guideline on the assessment of systemic exposure in toxicity studies. *Fed. Register* **60**:11263–11268 (1995).
2. A. J. Bailer and W. W. Piegorsch. Estimating integrals using quadrature methods with an application in pharmacokinetics. *Biometrics* **46**:1201–1211 (1990).
3. A. J. Bailer. Testing for the equality of area under the curves when using destructive measurement techniques. *J. Pharmacokin. Biopharm.* **16**:303–309 (1988).
4. K. C. Yeh and K. C. Kwan. A comparison of numerical integrating algorithms by trapezoidal, Lagrange and spline approximation. *J. Pharmacokin. Biopharm.* **6**:79–98 (1978).
5. R. D. Purves. Optimum numerical integration methods for estimation of area-under-the curve (*AUC*) and area-under-the moment-curve (*AUMC*). *J. Pharmacokin. Biopharm.* **20**:211–226 (1992).
6. D. Z. D'Argenio. Optimal sampling times for pharmacokinetic experiments. *J. Pharmacokin. Biopharm.* **9**:739–756 (1981).
7. F. Mentré, P. Burtin, Y. Merlé, J. Van Bree, A. Mallet, and J. L. Steimer. Sparse-sampling optimal designs in pharmacokinetics and toxicokinetics. *Drug Inform. J.* **29**:997–1019 (1995).
8. F. Mentré, A. Mallet, and D. Baccar. Optimal design in random-effects regression models. *Biometrika* **84**:429–442 (1997).
9. A. C. Atkinson and A. N. Donev. *Optimum Experimental Design*, Clarendon Press, Oxford, UK, 1992.
10. D. A. Beatty and W. W. Piegorsch. Optimal statistical design for toxicokinetic studies. *Statist. Meth. Med. Res.* **6**:359–376 (1997).
11. A. J. Bailer and W. W. Piegorsch. MSE considerations when using quadrature rules. *ASA Proceedings of the Biopharmaceutical Section*, pp. 177–182 (1990).
12. D. Katz and D. Z. D'Argenio. Experimental design for estimating intervals by numerical quadrature, with applications to pharmacokinetic studies. *Biometrics* **39**:621–628 (1983).
13. W. W. Piegorsch and A. J. Bailer. Optimal design allocations for estimating area under curves for studies employing destructive sampling. *J. Pharmacokin. Biopharm.* **17**:493–507 (1989).
14. G. C. G. Wei. Experimental design for estimating area under the curve by trapezoidal approximation in destructive sampling. *Drug Inform. J.* **31**:1237–1242 (1997).

15. R. P. Brent. *Algorithms for Minimization Without Derivatives*, Prentice-Hall, Englewood Cliffs. NJ, 1973.
16. J. R. Nedelman and E. Gibiansky. The variance of a better *AUC* estimator for sparse, destructive sampling in toxicokinetics. *J. Pharm. Sci.* **85**:884–886 (1996).
17. V. Brun. A generalization of the formula of Simpson for non-equidistant ordinates. *Nord. Mat. Tidskr.* **1**:10–15 (1953).
18. M. M. Bishop, S. E. Fienberg, and P. W. Holland. *Discrete Multivariate Analysis*, MIT Press, Cambridge, MA, 1975.
19. K. C. Yeh. Estimation and significance tests for area under the curve derived from incomplete blood sampling. *ASA Proceedings of the Biopharmaceutical Section*, pp. 74–81 (1990).