Journal of Pharmacokinetics and Biopharmaceutics, Vol. 22, No. 6, 1994

Comparison of Some Control Strategies for Three-Compartment PK/PD Models

Chuanpu Hu,[1] William S. Lovejoy,[2] and Steven L. Shafer[1,3]

Received August 26, 1993—Final January 18, 1995

In drug therapy, effective dosage strategies are needed to maintain target drug effects. The relationship between drug dose and drug effect is often described by pharmacokinetic/pharmacodynamic (PK/PD) models where typically the PK model has a multicompartment form and the PD model is the sigmoidal Emax model. The parameters in the PK/PD model are generally unknown in the individual patient, although prior knowledge may be available and can be updated after measurements of drug effect are taken during the therapy. This fact, together with the complexity of the PK/PD model, makes the control problem complex. This paper investigates several control strategies in the framework of a three-compartment PK model plus an effect site with a PD model. Using computer simulations under different assumptions, we show that a MAP (maximum a posteriori) Bayesian type of strategy is effective, nevertheless in high-risk situations a stochastic control strategy hedging against estimation errors provides better performance at computational cost.

KEY WORDS: Bayesian; compartment model; dose regimen design; pharmacokinetics; pharmacodynamics; stochastic control; effect site.

INTRODUCTION

This paper addresses the problem of choosing an appropriate drug infusion regimen to keep the drug effect in a patient at some target level in the context of incomplete information about the patient's pharmacokinetic/pharmacodynamic (PK/PD) characteristics. Observations of drug effect are taken but are assumed to be imprecise due to measurement errors.

Partially funded by Palo Alto Institute for Research and Education Inc., and the Veterans Administration Merit Review Program.

[1]Stanford University School of Medicine, Anesthesiology Service (112A), Veterans Administration Medical Center, 3801 Miranda Avenue, Palo Alto, California 94304.
[2]Graduate School of Business, Stanford University, Stanford, California 94305.
[3]To whom correspondence should be addressed.

0090-466X/94/1200-0525$07.00/0

The pharmacokinetics of many drugs can be described by a poly-exponential disposition function of time having the form: $\sum_{i=1}^{m} c_i e^{-\lambda_i t}$ where λ_i are rate constants, and $\sum_{i=1}^{m} c_i$ equals the concentration at Time 0 following a bolus injection. In practice m is usually no larger than 3 because there are not enough data to estimate large numbers of parameters. The poly-exponential disposition function is mathematically compatible with a biological compartment model of the human body (1) with unknown parameters. Compartment models represent the body as a series of finite-volume compartments with transfer rates connecting compartments to each other and to the outside world. In this paper we restrict our attention to three-compartment models with an effect site (see below). These models have been found useful in describing the behavior of intravenous drugs used in anesthesia practice.

Controlling the plasma drug concentration is not necessarily desirable because the plasma is not the site of drug effect. The pharmacodynamics are often modeled by using an additional compartment, called "the effect site," combined with a nonlinear function relating the effect site concentration and the drug effect (2,3). Thus we have the model shown in Fig. 1, where Compartment 1 represents the site sampled (often the plasma), 2 and 3 are peripheral compartments, and 4 is the effect site compartment. We denote by v_i the volume (liter) of the ith compartment. The volume of the effect site compartment v_4 is assumed to be very small so as to not affect the plasma concentration. This effect site models the time lag between drug dosing and drug effect. As in Shafer and Gregg (4), we arbitrarily define

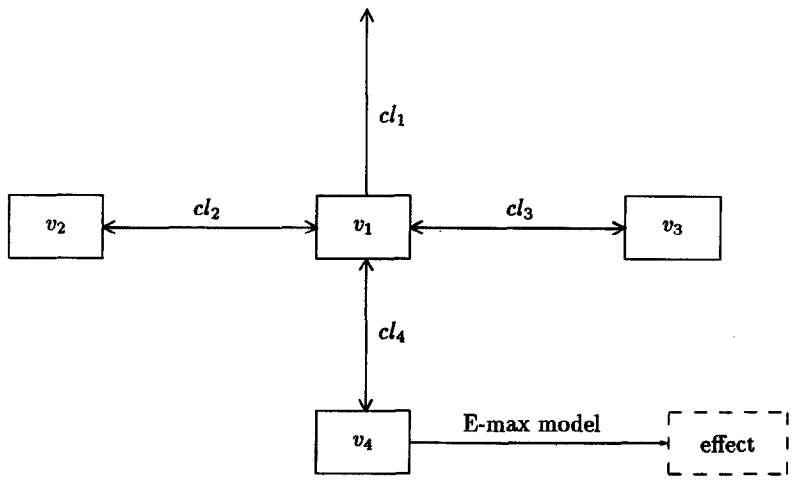

Fig. 1. Three-compartment model with effect site.

$v_4 = v_1/10000$. cl_i is the clearance (liter/time) between the central and the ith compartment, with the exception that cl_1 is the clearance to the outside world. Note that the clearance from Compartment 1 to 2 equals that from Compartment 2 to 1, which follows from the assumption that the concentrations at all compartments are equal at steady state, which can be assumed without loss of generality in relating the drug dose to effect.

Therefore, given the drug infusion rate $d(t)$ (g/time) at time t, the trajectory of the drug concentrations is represented by the system of differential equations:

$$\begin{aligned}
C_1'(t) &= -\frac{1}{v_1}(cl_2 + cl_3 + cl_0)C_1(t) + \frac{1}{v_1} cl_2 C_2(t) + \frac{1}{v_1} cl_3 C_3(t) + \frac{d(t)}{v_1} \\
C_2'(t) &= \frac{1}{v_2} cl_2 C_1(t) - \frac{1}{v_2} cl_2 C_2(t) \\
C_3'(t) &= \frac{1}{v_3} cl_3 C_1(t) - \frac{1}{v_3} cl_3 C_3(t) \\
C_4'(t) &= \frac{1}{v_4} cl_4 C_1(t) - \frac{1}{v_4} cl_4 C_4(t)
\end{aligned} \tag{1}$$

where $C_i(t)$ is the drug concentration (g/L) at the ith compartment at time t. For computational reasons we will use the parameter $k_{41} \equiv cl_4 v_1/10000$ instead of cl_4.

The relationship between the effect site concentration and the drug effect is modeled by the following Emax model

$$E(t) = E_0 + \frac{E_{\max} C_4^{\eta}(t)}{EC_{50}^{\eta} + C_4^{\eta}(t)} \tag{2}$$

This model, also known as Hill equation, represents monotonically increasing sigmoidal curves relating $C_4(t)$ and $E(t)$. EC_{50} is the "apparent" drug concentration in the effect site (this is not a real concentration and thus cannot be measured, cf. ref. 4) that corresponds to the 50% effect and η determines the steepness of the curve. We assume that E_0 and $E_{\max}$ are known, and (without loss of generality) equal to 0 and 1, respectively. Therefore the parameters of interest are: v_1, v_2, v_3, cl_1, cl_2, cl_3, k_{41}, EC_{50}, η. For a more detailed description of this PK/PD model, see Shafer and Gregg (4).

The model parameters are generally not exactly known prior to the therapy, although some information may be available and can be updated during the therapy as new observations become available. Thus the drug effect for a given dose cannot be exactly predicted. We would like to find a

dosing strategy that optimally balances the risk of overdosing vs. underdosing strategy that optimally balances the risk of overdosing vs. underdosing during the whole course of therapy. The mathematical definition of this notion is given in Methods.

Several difficulties arise in controlling $E(t)$ optimally at target. First, even if the PK/PD parameters were perfectly known, the optimal solution still would not be available. Because of the delay between the peak of effect site concentration and that of the central compartment concentration, optimal policies in general, when the effect is below target, result in the effect rising above target and then dropping to the target. How much the effect should go over the target depends on the extent that overdosing and underdosing is penalized, i.e., depend on the cost structure. [Appendix A provides the optimal solution if one restricts the strategies to the class that do not overdose (4)]. Second, even if the optimal solution in the known parameter case were available, with incompletely known parameters, the problem still becomes a Bayesian control problem of such complexity that only approximate solutions are known. In the setting of incompletely known parameters the most popular approximate solution strategy in the medical literature is the MAP Bayesian strategy (5–9), which simply uses the policy corresponding to the posterior mode of the parameter distribution. However, even this simple strategy is not easily available for the second reason above. With a sigmoidal PD model, the information contained in an observation decreases as the predicted effect approaches E_0 or E_{max}. Therefore when the desired target effect is close to E_0 or E_{max}, one might wish to compromise the near term objective to obtain more information for controlling the system in some fashion.

Finally, the fact that we have a large number of unknown parameters makes the stochastic control problem difficult due to what is called "the curse of dimensionality" (10). Previously, stochastic control methods have been applied to two-compartment models by D'Argenio and Rodman (11) and D'Argenio and Park (12), where the "myopic" approach, which is also known as "open-loop stochastic control," were used. The approach is relatively computationally easy, however it is not appropirate in our case, as discussed in the next section.

In Hu *et al.* (13), a variety of alternative control strategies were introduced and compared in simulated patients in the context of controlling plasma concentration in one-compartment PK models. It was shown that a strategy called "VU" is effective in a variety of problem settings. In this paper we adapt some of those strategies to the much more complex situation of three-compartment PK/PD models. However, direct extention is not possible for two reasons. First, in the one-compartment case, it is easy to calculate optimal policies if the parameters are known; these policies can then be

used to obtain stochastic control strategies to handle the situation where the parameters are unknown. This is not directly possible with the three-compartment PK/PD case, because optimal policies in the case of known parameters are not easily available. Second, stochastic control strategies generally use discrete representations of the distribution of the unknown parameters. This is very difficult in the three-compartment PK/PD case, where the number of the uncertain parameters is large. We present our computational approaches to address these two difficulties in Methods. Then, in Results, we present the numerical comparisons of our candidate control strategies. We shown by computer simulation that due to the presence of the PD model and the high computational complexity in this three-compartment PK/PD environment, a MAP Bayesian type of strategy can be quite efficient. We also show, however, that the VU procedure is still more advantageous in the high-risk case where it is important to hedge against estimation errors. The Discussion summarizes the findings.

METHODS

The Three-Compartment PK/PD Model

As stated, our objective is to control the system described by Gibaldi and Perrier (1) and Hull *et al.* (2), where the PK/PD parameters (v_1, v_2, v_3, cl_1, cl_2, cl_3, k_{41}, EC_{50}, η) for the individual are assumed unknown at the beginning of the dosage regimen. It is assumed however that the parameters come from a population prior with a known joint probability distribution.

Observations of the drug effect can be made at any time with measurement error. Specifically, a reverse form of the multiplicative noise model (14) is assumed

$$\tilde{E}(t) = 1 - [1 - E(t)]\, e^{\varepsilon}$$

where $\tilde{E}(t)$ is the observed effect at time t, and ε is a normal random variable with mean zero and standard deviation σ. This noise model has the feature that the variance of the measured effects decreases with increasing predicted effect, a feature that is empirically observed in measuring muscle relaxation in anesthesia. An example for this is that with 100% paralysis there is no measurement error (provided that the equipment is working).

For any specific set of PK/PD parameters and a given dosage regimen, the differential Eq. (1) can be solved and this solution together with Eq. (2) gives the drug effect as a function of time, as described in Appendix A. Let t_i, $i = 1, 2, \ldots$, be the times when observations are taken and decisions on

dosage rates made. For simplicity we assume that the times are evenly spaced, i.e., $t_{i+1} = t_i + L$ for some constant L.

In our problem scenario, we assume that at each time t_i the decision maker inspects the entire therepeutic history (i.e., historical dosages, prior distribution, and observations) and chooses a constant infusion rate for the time interval (t_i, t_{i+1}). Since the observations are imperfect, the exact PK/PD parameters that characterize the patient, and hence the true drug effect at any point in time, are only imperfectly known. Our objective is to maintain a constant target effect T over some unspecified, long time horizon. A cost function $g[\cdot]$ penalizes the deviation from target. Different forms of the cost function $g[\cdot]$ represent different objectives, i.e., how close the drug effect should be kept at target and how much overshooting (or undershooting) should be penalized. Then our objective is to find a strategy that minimizes the expected total cost

$$E \int_0^\infty e^{-\beta t} g[E(t) - T]\, dt$$

where β is taken to be very small. The presence of the term $e^{-\beta t}$ implies that our interest in optimizing decreases as we look further into the future. This term is necessary to keep the integral finite.

An equivalent formulation of the above is to find the best strategy that sequentially, at each time t_i, takes an observation, updates the prior distribution (which is the posterior distribution at time t_{i-1} after the observation at time t_{i-1} was taken), then chooses an infusion rate to be used from t_i until t_{i+1}.

Unfortunately, no optimal solution is currently available for this problem and we have to look for approximation methods, which are discussed in the next section.

Control Strategies

Here we introduce four suboptimal but potentially well-performing strategies. The first strategy is intuitive and easy to implement. The second strategy is an improved version of the first. The third and fourth strategies use stochastic control techniques, and offer potential further improvement.

Certainty Equivalent Strategies

These types of strategies separate the estimation and the dosage administration process in that they first obtain an estimate of the parameters and then compute the control assuming that the estimate is exact. The term "certainty equivalence" comes from the fact that the control is chosen as if

the parameters were known with certainty and equal to the estimate. Certainty equivalence strategies are conceptually simple and easy to implement.

MAPSG. Map Bayesian is currently the most commonly invoked strategy in the context of incomplete information of model parameters, where MAP stands for "maximum a posteriori." The MAPSG strategy directly uses this idea and is defined as follows. To emphasize the dependence of $E(t)$ on the PK/PD model parameters which we shall call θ, we temporarily write $E(t)$ as $E(\theta, t)$. Recall that for any dosage history $E(\theta, t)$ can be computed for any θ and time. After the jth observation (z_j) is taken, the parameter estimate $\hat{\theta}$ is chosen to maximize

$$-\frac{1}{2}\sum_{i=1}^{j}\left(\frac{\ln z_i - \ln E(\theta, t_i)}{\sigma}\right)^2 + f_p(\theta)$$

where f_p is the log of a multivariate normal density of the prior distribution. Thus the estimate $\hat{\theta}$ maximizes the Bayesian posterior probability for θ (hence MAP). In this study we used an IMSL least squares routine to carry out the maximization in this process. MAPSG then computes the policy described by Shafer and Gregg (4) where $\hat{\theta}$ are assumed to be the true parameters. This policy is presented in Appendix A. The rate used in an interval results in an increasing effect (if the rate is large enough), a moment where the effect peaks, and a decreasing effect (4). Once the effect is on target at or after the peak time, it can be maintained at target by maintaining the plasma concentration at the appropriate level using the method proposed by Schwilden (15). Finding the optimal policy that does not overshoot therefore is equivalent to finding the rate such that its resulting effect peaks at target [followed by a maintenance procedure (15)]. This requires a double search: One needs to search for the correct infusion rate, and for each rate one needs to search for the peak drug effect. More on the search procedure is given in Appendix B. Figure 2 illustrates how MAPSG would perform.

DPMAPSG. In case of known parameters, the Shafer–Gregg (SG) policy is only optimal with the restriction of not overdosing, therefore one can expect that the policy may be improved, especially in the situations where underdosing is undesirable. The DPMAPSG strategy aims at achieving this improvement by using the following dynamic programming procedure. The dynamic programming procedure has as its cornerstone the "principle of optimality" (10) which can be stated as follows. The optimal infusion rate must minimize the following: the cost of deviation from target over the next interval (duration = L), plus the future cost starting at the end of this interval (at time = now + L), after updating our information with the current dosing decision and effect observation at the end of the interval, and then following

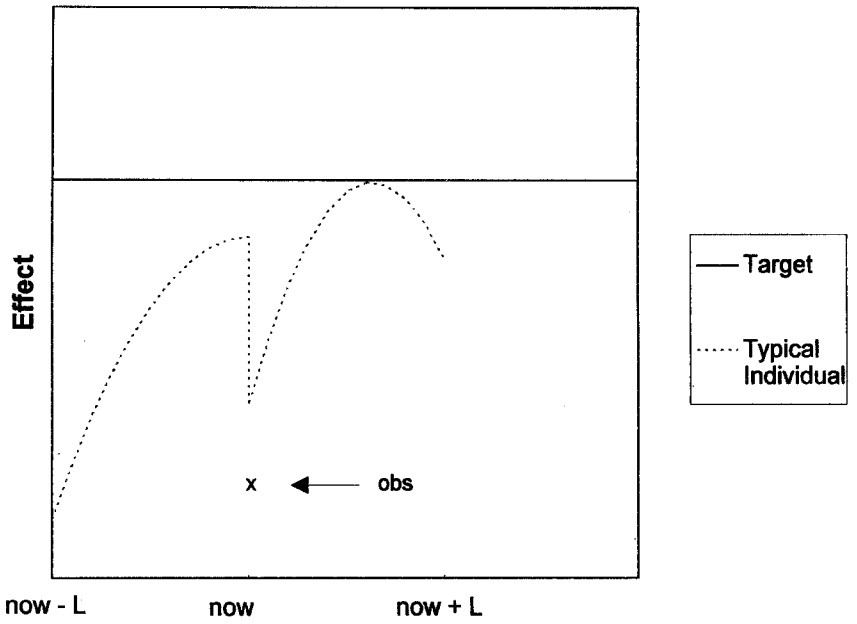

Fig. 2. MAPSG. At time (now $- L$) it used an infusion rate so that the drug effect corresponding to the typical individual would not go over the target. At time (now), a new observation revealed that the typical individual's effect is less than previously predicted. Our belief about the PK/PD model is updated with this observation, and a new infusion rate is chosen according to the posterior distribution of the PK/PD parameters.

an optimal dosing strategy from that point forward. Informally, the optimal cost V satisfies the following recursive form:

$$V|_{\text{now}}^{\infty} = cost|_{\text{now}}^{\text{now}+L} + V|_{\text{now}+L}^{\infty}$$

This principle reduces the task of finding the optimal infusion rate to a simple minimization of a function with two terms. The first term is simple, being the cost integral only over the next time interval (from now until now$+L$). The "future value" function, however, is typically impossible to evaluate for realistically complex problems. Indeed, if we could find this value function the control problem would be essentially solved, i.e., evaluating this value function inherits all of the complexity of the control problem itself. However, if we can find an approximation to this value function, we could use the approximation in the minimization to derive an infusion rate. The better the approximation of the value function is, the better an infusion rate so derived can be expected to perform.

One useful approximation can be generated by a so-called "nominal policy". That is, using the nominal policy results in a trajectory of drug

effect, and the accumulation of the total deviation of this effect trajectory is then used as the approximation of the value function. Using this approximate value function in the dynamic programming recursion can be expected to generate an improved strategy (10). DPMAPSG uses the SG policy as the nominal policy. More precisely, at any particular time t, given the parameter set θ and the past dosage history, the current state of the system can be viewed as the amount of drug in each compartment (including the effect site), which we shall call s. Using an infusion rate d, the effect in the current period (from now until now$+L$), can be fully written as $E(\theta, s, d, t)$, and the resulting "state" of the system at time (now$+L$), written as $s'(s, d)$, can be computed by solving Eq. (1) if the PK/PD parameters θ are known. DPMAPSG first uses the same estimation procedure as MAPSG to obtain an estimate $\hat{\theta}$, then uses the following to obtain an infusion rate:

$$\min_d \left\{ \int_0^L e^{-\beta t} g[E(\hat{\theta}, s, d, t) - T]\, dt + e^{-\beta L} V(\hat{\theta}, s'(s, d)) \right\} \tag{3}$$

where V is the total cost incurred by using the SG policy (4). It is well established in the control literature that this iteration step improves the nominal policy, which in this case is the SG policy (4) [cf. (10)]. Therefore we can expect DPMAPSG to perform better than MAPSG. Figure 3 illustrates how DPMAPSG would perform.

MAPSG and DPMAPSG separate the estimation and the dosage administration process and thus fall into the class of certainty equivalence strategies. These types of strategies have the weakness that the only way the posterior distribution enters into the strategy is via the estimate, thus ignoring the variability when computing the control policy. Consequently, in using a single estimate it behaves as if the patient's characteristics were known with certainty, i.e., as if estimation errors were not possible. To properly account for the uncertainties in the patient's PK/PD parameters, one would want a strategy that considers more than just one of the most likely parameter sets, i.e., a strategy that "hedges" against estimation errors. Below we introduce two strategies having this property. However, any hedging normally makes the strategy more computationally intensive because considering simultaneously a number of possible parameter sets takes more effort than considering only one. Also, in the three-compartment PK/PD case the number of unknown parameters is large and computation is large and already heavy, therefore dense discretizations (11,12,16,17) are not computationally feasible here in the full parameter space.

Stochastic Control Strategies

Stochastic control strategies hedge against estimation errors by using some discretized parameter distribution, instead of a single estimate. A

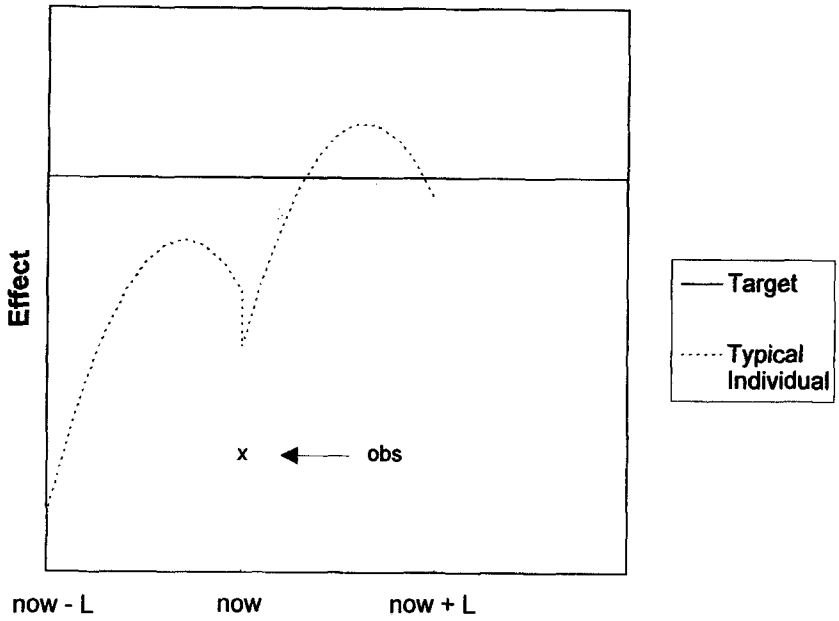

Fig. 3. DPMAPSG. Similar to Fig. 2 (MAPSG), but allows overdosing and gets the drug effect closer to target.

discretization procedure considers a discrete prior probability distribution of the parameters of interest, and the corresponding control strategies use this distribution, rather than using only a single estimate, such as the mode often used in certainty equivalent control strategies. Using a discrete distribution can reduce potential errors caused by invoking specific distributional assumptions such as normalcy. A discrete prior probability distribution assumes that the patient's parameters come from a finite number of points in the parameter space. These points are fixed throughout the control process, therefore the number of points must be relatively large in order to properly represent the original distribution. We have previously reviewed the influence of discretization on one-compartment control strategies (17).

Since dense discretizations are not computationally feasible in the three-compartment PK/PD setting, we use a crude hedging around the parameter estimate in the estimation procedure as in MAPSG by singling out some parameters that we think are important and take a three-point discretization in the chosen parameters. More precisely, let $\hat{\theta} = (v_1, v_2, v_3, cl_1, cl_2, cl_3, k_{41}, EC_{50}, \eta) \equiv (\hat{\theta}_1, \ldots, \hat{\theta}_9)$ be the estimated mode of the posterior distribution and say that the jth parameter, with initial variance σ_j, is being discretized. Then the discretized parameter sets are $(\hat{\theta}_1, \ldots, \hat{\theta}_{j-1}, \hat{\theta}_j,$

$\hat{\theta}_{j+1}, \ldots, \hat{\theta}_9)$, $(\hat{\theta}_1, \ldots, \hat{\theta}_{j-1}, \hat{\theta}_j - \sigma_j, \hat{\theta}_{j+1}, \ldots, \hat{\theta}_9)$ and $(\hat{\theta}_1, \ldots, \hat{\theta}_{j-1}, \hat{\theta}_j + \sigma_j, \hat{\theta}_{j+1}, \ldots, \hat{\theta}_9)$. The probability weight associated with the parameter sets in the discretization is assigned as the corresponding posterior probability density function values. If m parameters are discretized, the procedure produces a discrete prior of 3^m points, centered around the Bayesian estimate $\hat{\theta}$. This prior is not as significantly different from the single estimate as that in D'Argenio and Katz (16) and Hu *et al.* (17) because of the crudeness of the discretization. Still, using this crude discretization better hedges against the estimation error than considering only the estimate $\hat{\theta}$. Additional comments on our discretization procedure appear in Appendix D.

The theory of dynamic programming can also be applied in the situation where the parameters are only imperfectly known. In this context, the above stated principle of optimality still holds if we replace the costs by their expectations, taken with respect to the future observation and the prior probability distribution of the parameters (18). Therefore, to generate strategies using this theory, we need to search for approximations of the expected future values. Two such procedures, called VU and VL, were compared in the setting of one-compartment models (13). Below we introduce them in the setting of three-compartment PK/PD case.

VUMAPSG. The VU procedure uses a relatively easily computed approximate value function to derive infusion rates, as described below. The method was originally proposed by Van Hee (19) and extended by Loveyjoy (20). See Hu *et al.* (13,17) for more discussion and application of the strategy in one-compartment models. In the simulations we discretize six parameters, $(v_1, v_2, cl_1, k_{41}, EC_{50}, \eta)$, resulting in a discrete prior of $n = 3^6 = 729$ points. Thus at each period the stragegy considers only a finite number of discrete parameter sets indexed by subscript j, i.e., θ_j for $j = 1$ to n, as the discrete prior. However, unlike in previous applications (11,12,16–17), all points in the discrete prior change from one period to the next. In the one-compartment case, if the parameters were known with certainty then the dosage problem would be easily solvable as a deterministic control problem, so that the corresponding optimal policies can be comptued and the average weighted values are then used as the approximate value function (17). In the three-compartment PK/PD case, if the PK/PD parameters were known with certainty then only the policy given in Appendix A is available. Therefore, we use the (numerically evaluated) cost generated by this policy to approximate the optimal cost in the known parameter case, as follows. Given the dosage history until time t_i and a fixed parameter set θ_j, we can compute the drug concentration levels at compartment $1, \ldots, 4$, denoted as the current "state" s_j. For each infusion rate d, the resulting drug effect $E(\theta_j, s_j, d, t)$ during the interval from t_i to t_{i+1} can be computed (4)

so that the cost accumulated for the interval, denoted by $R_j(s_j, d) \equiv \int_0^L e^{-\beta t} g[E(\theta_j, s_j, d, t) - T]\, dt$, can be evaluated numerically. Similarly, for the same infusion rate we can also compute the resulting state of the system $s_j'(s, d)$ and policy (4) which we denote by P_j, and the resulting drug effect $E(\theta_j, s'(s_j, d), P_j, t)$ from time t_{i+1} to infinity. Thus we can numerically compute the cost accumulated from time t_{i+1} to infinity $V_j(s_j'(s_j, d)) \equiv \int_L^\infty e^{-\beta t} g[E(\theta_j, s'(s_j, d), P_j, t) - T]\, dt$.

Let π_j denote the probability weight associated with the parameter set θ_j. VUMAPSG uses the approximate future cost $\sum_j \pi_j V_j(s_j'(s_j, d))$. The infusion rate d that minimizes the following

$$\sum_j \pi_j R_j(s_j, d) + e^{-\beta L} \sum_j \pi_j V_j(s_j'(s_j, d)) \tag{4}$$

will then be used for the current period. Figure 4 illustrates how VUMAPSG would perform.

In the context of one-compartment models the "myopic" strategy that minimizes only $\sum_j \pi_j R_j(s_j, d)$ may be useful (13), which is similar to the approach in refs. 11 and 12. However, because of the delay between the peak of the plasma concentration and the peak of the effect in our case, the "future" is very much relevant. That is, when the effect $E(t_i)$ is expected to be significantly lower than target, minimizing only the immediate expected cost $\sum_j \pi_j R_j(s_j, d)$ will make the effect exceed the target significantly at some time after t_{i+1}. Therefore the myopic approach is not considered here. VUMAPSG hedges against estimation errors by considering a number of possible parameter values in the discrete prior, rather than using only the mode. With a dense discretization, the original VU procedure performed well in one-compartment models (17). However, here our discretization is very crude. Therefore we may expect VUMAPSG to perform better than DPMAPSG, but only in circumstances where it is important to hedge against estimation errors.

VLMAPSG. The VLMAPSG strategy is obtained similarly to the above strategy, by using the VL procedure to approximate the optimal value function instead of VU. The VL procedure works as follows: For each parameter set θ_j, $j = 1, \ldots, m$, in the discrete prior there corresponds a specific policy P_j (4). The VL procedure chooses actions as if the decision maker has one more chance to gather information, and then must commit to one of the policies P_j, which is implemented from that point forward. (The committed policy P_j is the one that results in the least expected accumulated deviation from target with respect to the posterior distribution after taking an observation in the next period.) The VL procedure differs from the previous three strategies in that it actively gathers information (in the parameters being

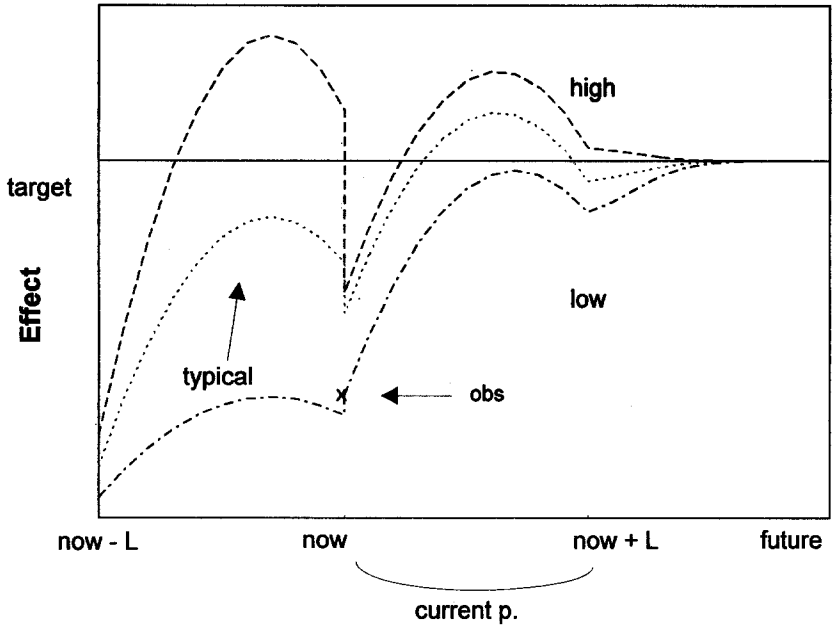

Fig. 4. VUMAPSG. At time (now), a new observation changed the current predicted drug effects (say, for 90% of the population). An infusion rate would result, for each individual in the population, an effect profile in the current period (between now and now+L), and an end effect at (now+L). For each such end effect, if the parameters were going to become known after (now+L), the corresponding optimal policy would generate a drug effect profile. VUMAPSG minimizes the sum (weighted by the posterior probabilities) of the costs, that would be accumulated in the current period (between now and now+L) and the future. The future concentrations all approach target because they are computed with the corresponding optimal strategies.

discretized) by perturbing the system in order to obtain a more favorable posterior distribution. This is especially true when one is very uncertain which of θ_j is the true parameter set, in which case a little additional information in favour of one parameter set θ_j will be valuable. Figure 5 depicts how VLMAPSG would perform. We have previously described the VL procedure in detail (13). Jelliffe and Schumitzky (21) noted that an optimal strategy is actively gathering information by perturbing the system in order to get a more favorable posterior distribution after taking the next observation. Therefore one might expect the VL procedure to perform better than the other three strategies. However, in the one-compartment case, VL did not perform well because it sacrificed too much on near term costs in order to gather information, and thus actively gathering information may not always be superior (13). In our three-compartment PK/PD case, we are using a very crude discretization, which should reduce the aggressiveness of VL in

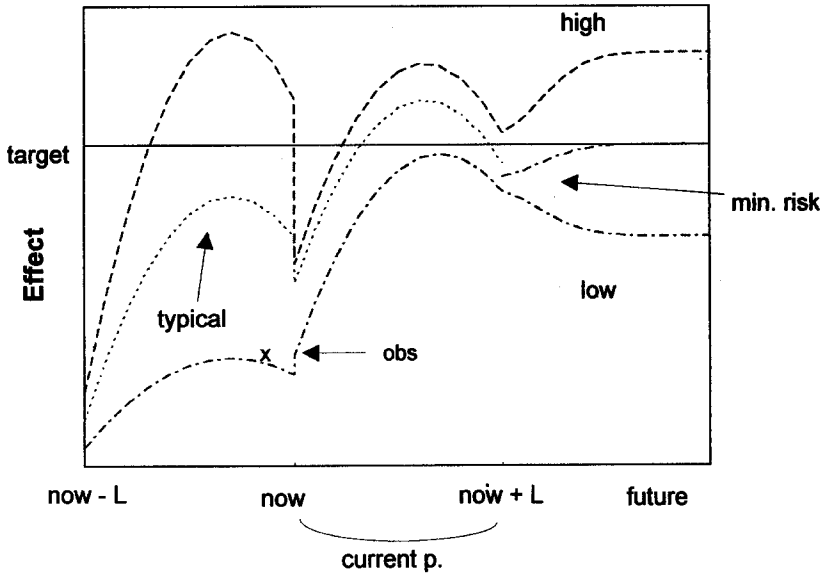

Fig. 5. VLMAPSG. Similar to Fig. 4, but with a different approximation of the future cost. Corresponding to each individual in the population with an end effect at (now + L), there is an optimal policy. One such optimal policy will generate relatively the least risk for the entire population. The expected cost coressponding to this policy is taken as the approximation of the future cost. VLMAPSG minimizes the sum of the expected cost (accumulated in the current period) and this future cost. None of the future concentrations approach the target, except for the individual whose optimal policy has smaller risk than everyone else.

searching for information. This is because in the simple discretization used here, the parameters are relatively less spread out (i.e., represent the whole distribution worse) than in the one-compartment case, therefore different parameter sets do not give rise to significantly different policies. In such case more information would not lead to significant improvement of "future values," so that VL would not seek information as aggressively as in Hu *et al.* (13), where the discretization was dense. We note that the VL procedure seeks information for only those parameters being discretized. More on this appears later. Because of the actively gathering information feature, VL is computationally intensive since additional numerical procedures are necessary to evaluate the future expected value with respect to the possible observation outcome in the next period, i.e., one must consider every possible future observation outcome and plan what to do accordingly. In the simulations we discretized two parameters: (cl_1, EC_{50}) for the VLMAPSG strategy, resulting in total 9 parameter sets at each period. Details of our numerical methods appear in Appendix B. The relative computational complexity of these strategies is discussed in Appendix C.

RESULTS

In the numerical trials we simulated all patients for 30 time periods. At each period, an infusion rate was computed, then an error term simulated to obtain the observed drug effect at the end of the period, and the procedure continues. The target effect T was taken to be 0.9 (i.e., 90% maximal effect). The period length L was chosen to be 5 min.

Maitre *et al.* (14) studied the population kinetics of alfantanil. Here we followed their result by assuming a log-normal prior of the parameters $(v_1, v_2, v_3, cl_1, cl_2, cl_3, k_{41}, EC_{50}, \eta)$ such that the coefficient of variation was 30%. The means of $(v_1, v_2, v_3, cl_1, cl_2, cl_3, k_{41}, EC_{50}, \eta)$ were set to correspond tot he estimates (14) and the means of (k_{41}, EC_{50}, η) were set to be (0.77, 0.479, 4.8). The discount coefficient β was taken as 0.001, corresponding to a discount factor $e^{-\beta L}$ as 0.995. Being so close to 1, the discount factor played little role in determining the strategies. The sample size was 500 for MAPSG, DPMAPSG, and 200 for VLMAPSG, VUMAPSG. The samples are smaller for strategies VUMAPSG and VLMAPSG, because they use more information on the parameters and consequently are more difficult to compute.

Quadratic Cost

In this setting deviation from target was penalized quadratically, and the cost function was $g[E(t)-T]=[E(t)-T]^2$, so that the cost for each period was

$$\int_0^5 [E(t)-T]^2\,dt$$

We computed the costs accumulated by the strategies, as well as their sample standard deviations, in situations of low and high observation error variances ($\sigma^2=0.05$ and 0.5, respectively) as shown in Figs. 6 and 7.

Figure 6 shows that, in the low observation error variance case, DPMAPSG performed better than MAPSG in minimizing average deviation from target. While we are optimzing an objective that depended only on the expectation of the discounted costs, one may also be interested in the risk involved in applying a strategy to a specific patient. Over the range of patients generated from these priors, DPMAPSG generated significantly lower standard deviation results than MAPSG, suggesting that it is also much more robust. This improvement of DPMAPSG relative to MAPSG is due to the dynamic programming iteration operation. Comparing Figs. 6 and 7 shows that changing measurement errors had no siginficant impacts on these results.

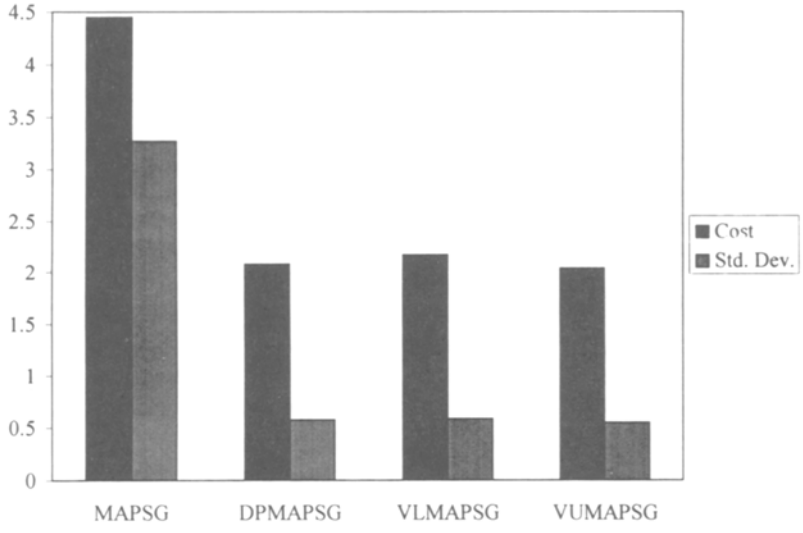

Fig. 6. Quadratic cost, obs. error var. = 0.05.

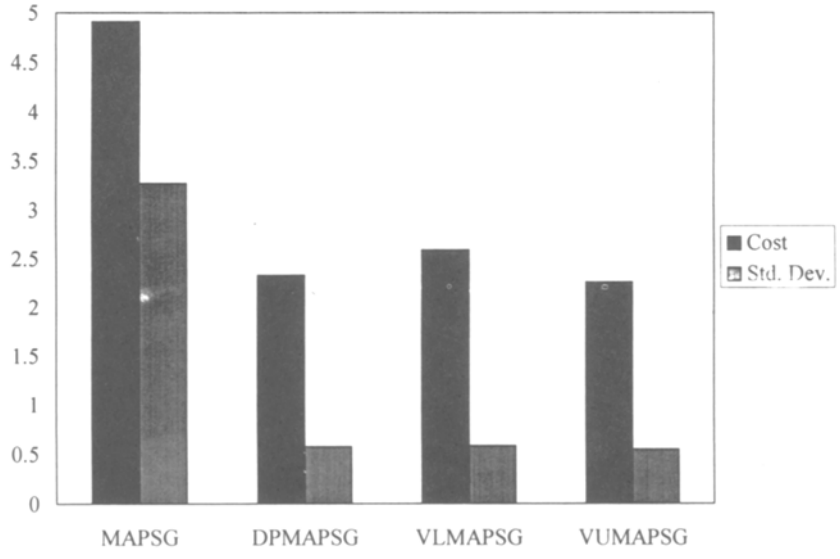

Fig. 7. Quadratic cost, obs. error var. = 0.5

VUMAPSG and VLMAPSG further refine DPMAPSG by a simple discretization procedure. The VU approach did not achieve any significant improvement, contrary to the one-compartment case (17), because the discretization was very crude.

For the same reason, the VL procedure did not perform as badly as in the one-compartment case analyzed in Hu *et al.* (13). As discussed in Methods, in this case of a simple discretization where different parameter sets are relatively close to each other, the VL procedure did not seek information (in cl_1 and EC_{50}) as aggressively. However, our experience suggests that the strategy still puts too much weight on information gathering and performed poorer compared to DPMAPSG and VUMAPSG. Additional simulations could be performed to verify this statement, but this is not important because the performances are close. For an illustration, we provide a sample path for 11 periods with observation error variance 0.05, shown in Fig. 8.

Figure 8 shows that because MAPSG aimed at never overdosing, it first dosed the effect to target, then waited for the effect to drop, and dosed it up again. VLMAPSG did not do well for this particular sample, but achieved good overall behavior. The other strategies, allowing overdosing, maintained the effect on target relatively smoothly.

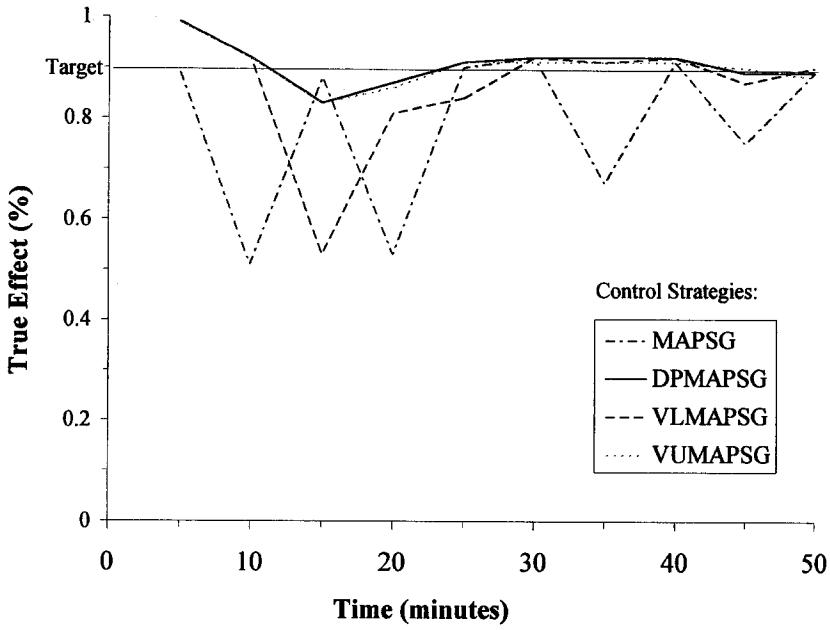

Fig. 8. Sample path, obs. error var. = 0.05.

Skewed Cost

The quadratic cost function penalized deviations from target equally in either direction, i.e., whether those deviations represent overdosing or underdosing. However, in practice one may wish to penalize overdosing (or underdosing) more severely. In this section we explore the effect of using a skewed cost function. Here the deviation from target was penalized by a quadratic function plus a linear underdosing penalty. Precisely, the cost for each period was

$$\int_0^5 [E(t)-T]^2\,dt+5\int [T-E(t)]\,dt$$

where the second integral was taken over the range of t such that $E(t)<T$. The result for low and high observation error variance ($\sigma^2=0.05$ and 0.5 respectively) are shown in Figs. 9 and 10.

Figures 9 and 10 show that in this case the performance of MAPSG significantly deteriorated, because it aimed at never overdosing despite a heavy underdosing penalty.

VUMAPSG appeared to show some advantage relative to DPMAPSG, especially in the high observation variance case as shown in Fig. 10. This suggests that in a risky situation the VU procedure should be more robust. VLMAPSG still seemed to perform slightly worse than VUMAPSG and DPMAPSG due to its relatively aggressive information-seeking feature.

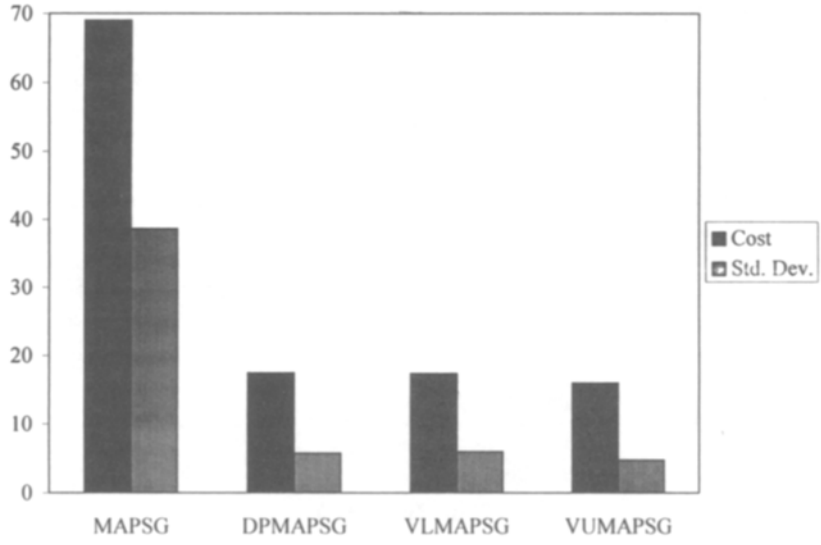

Fig. 9. Skewed cost, obs. error var. = 0.05.

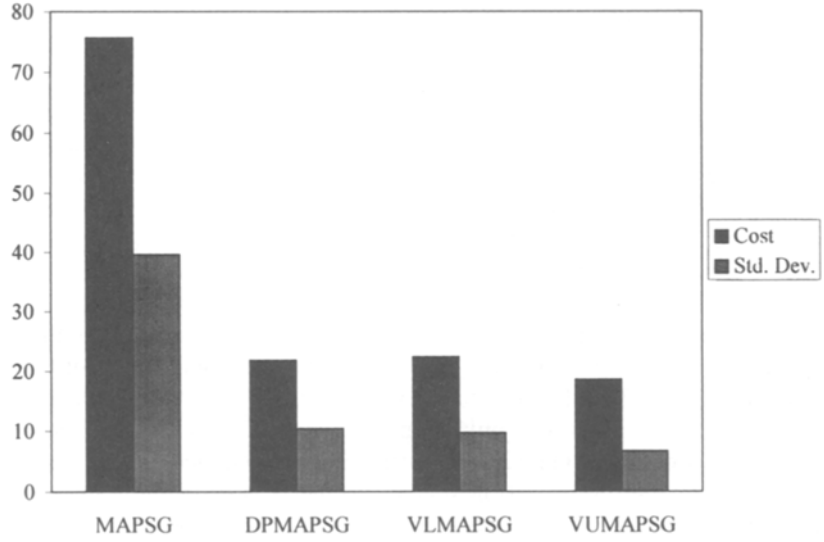

Fig. 10. Skewed cost, obs. error var. = 0.5.

DISCUSSION

In this complex three-compartment PK/PD model where little computational experience has been available, we have described several drug infusion strategies and examined their performances in simulations. The immediately available strategy MAPSG, obtained by combining the MAP Bayesian approach with the SG policy (4), did not perform well compared to the other strategies. However, a one-step dynamic programming refinement of the strategy DPMAPSG improved the strategy notably, and was relatively computationally efficient.

Both MAPSG and DPMAPSG separate estimation and control, and it has been established that such strategies share the weakness that they ignore estimation errors. Different stochastic control strategies have been shown to perform better than those that separate estimation and control (12,16,17). D'Argenio and Rodman (11) showed similar results for a two-compartment model, and D'Argenio and Park (12) showed this for a two-compartment PK/PD model. To hedge against estimation errors one needs to take into consideration the whole posterior distribution of the parameters, rather than only the mode. One approach to achieve this without assuming additional properties, such as Linear-Quadratic-Gaussian (21), is to use discrete priors (11,12,16,17). However, to properly represent the posterior distribution, a relatively large number of points are needed in discrete priors. As the number

of unknown parameters increases, the number of discrete points needed explodes, a fact which is known as the "curse of dimensionality" in the control literature. In our three-compartment PK/PD model we have nine parameters: v_1, v_2, v_3, cl_1, cl_2, cl_3, k_{41}, EC_{50} and η, which resulted in heavy computations so that previous discretization procedures are not computationally feasible. Therefore in this paper we introduced a crude discretization scheme that moves with the parameter estimate, rather than fixed schemes as in previous work.

This moving discretization procedure was combined with the control procedures VU and VL that we introduced (13). The VU procedure performed well in one-compartment models (17), but its variant in this three-compartment setting, VUMAPSG, did not achieve a significant improvement over DPMAPSG. The most important reason is that in this context we have a large number of (unknown) parameters which prohibits fine discretization procedures (11,12,17). In our experiments, increasing the number of points in the discretization scheme VUMAPSG did show improvement. This suggests that if we were able to overcome the curse of dimensionality, stochastic control procedures would indeed still be more preferable. Even with the crude discretization, in the high underdosing penalty case where hedging against estimation errors becomes more important, VUMAPSG still shows an advantage.

In our context the observations are taken at no cost, therefore actively gathering information means perturbing the system to get a more favorable posterior distribution after taking the next observation. As noted, active information-gathering strategies may not perform well if the balance between gathering information and exploiting the information to treat the patient is not well maintained (13). We established that in a one-compartment model, VL too aggressively seeks information (13). Its variant in this three-compartment PK/PD model, VLMAPSG, turns out to be less agressive in seeking information. One reason for this is that different parameter sets in this crude discretization did not result in significantly different policies, as mentioned earlier. Another reason might be that in the presence of multiple important parameters (i.e., the parameters being discretized that also have significant impacts on dosage regimen; in this case cl_1 and EC_{50}), the preferred direction (say dosing low) for information search in one parameter may be different than information search in another parameter (say dosing high). Thus as a compromising result VLMAPSG ended up not aggressively searching for information in any direction. Nevertheless, VLMAPSG still appeared to sacrifice too much on near term cost in order to gather information, and hence did not perform as well as VUMAPSG and DPMAPSG. Future research is necessary to establish how to optimally conduct the search for information in a multiparameter context.

As seen above, three-compartment PK/PD models such as considered in this paper are difficult to deal with, and the curse of dimensionality compromises the performance of stochastic control strategies relative to simpler and more naive certainty equivalent strategies. The advantages of stochastic control strategies were shown in simple one-compartment models (16,17,21) as well as in two-compartment settings (11,12). Future research is necessary to find for models with in-between difficulty, such as two-compartment PK/PD models, what type of stochastic control strategies and the associated discretization procedures will perform well. An alternative approach to deal with models with a large number of unknown parameters is to fix some parameters of less importance throughout the process (23), therefore reducing the computational burden.

GLOSSARY SYMBOLS

β	discount coefficient
ε	(normal) observation error
η	slope parameter in Emax model
λ_i	the ith rate constant in disposition function
π_i	probability weight associated with parameter set i
σ	standard deviation of (normal) observation error
θ	pharmacokinetic/pharmacodynamic parameters
c_i	the ith coefficient in disposition function
$C_i(t)$	drug concentration at compartment i at time t
cl_i	clearance of compartment i
d	(constant) drug infusion rate given in one period
$d(t)$	drug infusion rate given at time t
E_0	baseline drug effect
E_{max}	maximum drug effect
EC_{50}	effect site concentration corresponding to 50% drug effect
$E(t)$	drug effect at time t
$g[\cdot]$	cost function penalizing deviation of drug effect from target
k_{ij}	linear transfer rate from compartment i to j
L	period length
P_j	policy associated with parameter set j aimed at never overdosing
s	state of system (amounts of drug in each compartment)
s_j	state of system associated with parameter set j
t	time
T	target drug effect
t_i	time the ith observation is taken

V	optimal cost
V_j	cost value associated with parameter set j
v_i	volume of compartment i
z_i	the ith observed effect

APPENDIX A

Suboptimal Control Policies and Resulting Drug Effect for PK Models with Known Parameters and Effect Site

Shafer and Gregg (4) established the optimal policies with the restriction of never overdosing starting with no drugs in the system, given both in the case where bolus injection is allowed and the case where bolus is not allowed. As stated in the Introduction, optimal policies in general, even with known parameters, are difficult to obtain. Therefore this section is restricted to computing the policy that never overdoses. Given the target concentration C_T at the effect site and the drug amounts $A_1(0), \ldots, A_4(0)$, the policies can be computed in different cases. Here we follow the notation in Shafer and Gregg (4).

Consider the system that describes the amount of drugs in the system:

$$\begin{aligned} A_1'(t) &= A_2(t)k_{21} + A_3(t)k_{31} + A_4(t)k_{41} \\ &\quad - A_1(t)(k_{10} + k_{12} + k_{13} + k_{14}) + I(t)\,dt \\ A_2'(t) &= A_1(t)k_{12} - A_2(t)k_{21} \\ A_3'(t) &= A_1(t)k_{13} - A_3(t)k_{31} \\ A_4'(t) &= A_1(t)k_{14} - A_4(t)k_{41} \end{aligned}$$

where 4 refers to the effect site. We need to derive $A_4(t)$ as a function of the infusion $I(t)$.

Let $\lambda_1, \ldots, \lambda_4$ be the (positive) eigenvalues of

$$\begin{bmatrix} -k_{12}-k_{13}-k_{14}-k_{10} & k_{21} & k_{31} & k_{41} \\ k_{12} & -k_{21} & 0 & 0 \\ k_{13} & 0 & -k_{31} & 0 \\ k_{14} & 0 & 0 & -k_{41} \end{bmatrix}$$

Define

$$p_2(s) \equiv k_{14}[s^2 + (k_{21} + k_{31})s + k_{21}k_{31}]$$

$$p_3(s) \equiv A_4(0)s^3 + [A_1(0)k_{14} + A_4(0)(k_{10} + k_{12} + k_{13} + k_{14} + k_{21} + k_{31})]s^2$$
$$+ \{[A_1(0)(k_{21} + k_{31}) + A_2(0)k_{21} + A_3(0)k_{31}]k_{14}$$
$$+ A_4(0)(k_{10}k_{21} + k_{13}k_{21} + k_{14}k_{21} + k_{10}k_{31} + k_{21}k_{31})\}s$$
$$+ \{[A_1(0) + A_2(0) + A_3(0)]k_{14} + A_4(0)(k_{10} + k_{14})\}k_{21}k_{31}$$

and for $i = 2, 3$, $l = 1, \ldots, 4$,

$$c_{il} \equiv \frac{p_i(-\lambda_i)}{\prod_{j \neq l}(\lambda_j - \lambda_l)}$$

Then

$$A_4(t) = \sum_{i=1}^{4} c_{3i}\, e^{-\lambda_i t} + I(t) * \left(\sum_{i=1}^{4} c_{2i}\, e^{-\lambda_i t} \right)$$

where $*$ denotes convolution.

APPENDIX B

Numerical Procedures

To compute the SG policy, Shafer and Gregg (4) gave the complicated closed form solution, and presented an efficient algorithm to find the correct infusion rate lasting 10 sec. However, the efficiency of this algorithm is questionable if one needs to search for a rate lasting a longer time, say 5 min. In this work we used a direct search with 10 evaluations in both the rate and time dimension. We tested our results with the number of evaluations increased to 15 without seeing any significant difference in the results, which is an evidence that the search finds good solutions.

Gaussian Quadratures with 5 points are used to compute the R_j and V_j that appear in VUMAPSG and VLMAPSG, as well as the expected future values with respect to future observations in VLMAPSG. We also tried increasing the number of quadrature points to 8 but the results remained the same.

A Fibonacci search with 10 functional evaluations (see ref. 22 for more details on this search procedure) was used to carry out the process of minimizing 4 in computing VUMAPSG and the similar process for VLMAPSG. The result was then compared with rate zero to reduce the chance of getting a local but not global minimum. This caution is necessary (17), because the future value as a function of the current dosage is nonconvex even in the

simplest case, so numerical optimization methods that look only for first-order conditions may not be sufficient to find the desired minimum.

In the simulations we assumed a maximum infusion rate of 3000 μg/min. The Fibonacci search gave a local minimum to within an interval of length 34 μg/min.

APPENDIX C

Computational Complexity

The strategies are computationally intensive for several reasons. First, even in the known parameter case finding a reasonable policy is not easy; the SG policy (4) requires some search work. Second, a large number of unknown parameters makes estimation difficult. It is well known that any optimization package (such as IMSL) does not always stop at the best point. Third, if stochastic control strategies are used, then numerical intergration is required to evaluation the near-term costs as well as the future values such as in Expression (4) in Methods, because they are no longer analytically available (17,21).

The actual computing time the strategies take depends on the particular machine code and optimization package, as well as the number of numercial quadrature points and the discretization levels used. Still, for comparative purposes we provide the CPU time based on our actual experience on a VAX 4000. Table I presents the time needed to simulate the strategies (with the previously mentioned number of quadrature points, discretization level, etc.) for 30 periods.

Note that Table I may quite reflect the relative computational burden, since the time that the optimization package takes is an additive factor for the total computing time.

APPENDIX D

Discretization Level

To explore the effect of the discretization level, we tried varying the number of dimensions of the parameter space being discretized for VUMAPSG and VLMAPSG. VUMAPSG showed slight but consistent improvement as we discretized more parameters, as may be expected. However,

Table I. Computational Complexity

	Strategy			
	MAPSG	DPMAPSG	VUMAPSG	VLMAPSG
CPU(sec)	5	6	54	86

VLMAPSG may perform either better or worse as more parameters are discretized. We tried discretizing only cl_1 as well as cl_1, EC_{50}, and η, in both cases VLMAPSG performed worse. The fact that only discretizing cl_1 resulted in a worse performance of VLMAPSG suggests that search information in the cl direction may hinder that in the EC_{50} direction. To illustrate this, call it Case 1 when cl_1 is discretized and Case 2 when cl and EC_{50} are discretized. Now let us say that dosing low will generate more information about cl_1, then in Case 1 the VL procedure will tend to underdose in order to gather information about cl_1. [This is the case in a one-compartment model (13).] But if dosing high will generate more information about EC_{50}, then in Case 2 VL would dose more than in Case 1, and if this does not result in overdosing, then in Case 2 VL will not perturb the system as much as in Case 1. For similar reasons, the fact that discretizing cl_1, EC_{50}, and η resulted in a worse performance of VLMAPSG suggests that adding η in the discretization increased the aggressiveness in seeking information.

REFERENCES

1. M. Gibaldi and D. Perrier. *Pharmacokinetics*, Marcel Dekker, New York, 1982.
2. C. J. Hull, H. B., Van Beem, K. McLeod, A. Sibbald, and M. J. Watson. A pharmacodynamic model for pancuronium. *Br. J. Anaesth.* **50**:1113–1123 (1978).
3. L. B. Sheiner, D. R. Stanski, S. Vozeh, R. D. Miller, and J. Ham. Simultaneous modelling of pharmacokinetics and pharmacodynamics: Application to d-tubocurarine. *Clin. Pharmacol. Ther.* **25**:358–371 (1979).
4. S. L. Shafer and K. M. Gregg. Algorithms to rapidly achieve and maintain stable drug concentrations at the site of drug effect with a computer-controlled infusion pump. *J. Pharmacokin. Biopharm.* **20**:147–169 (1992).
5. P. O. Maitre and D. R. Stanski. Bayesian forecasting improves the prediction of intraoperative plasma concentration of alfentanil. *Anesthesiology* **69**:652–659 (1988).
6. L. B. Sheiner, S. Beal, B. Rosenberg, and V. V. Marathe. Forecasting individual pharmacokinetics. *Clin. Pharmacol. Ther.* **26**:294–305 (1979).
7. S. Vozeh, M. Berger, M. Wenk, R. Ritz, and F. Follath. Rapid prediction of individual dosage requirements for lignocaine. *Clin. Pharmacokin.* **9**:353–363 (1984).
8. S. Vozeh, R. Hillman, M. Wandell, T. Ludden and L. Sheiner. Computer-assisted drug assay interpretation based on Bayesian estimation of individual pharmacokinetics: Application to lidocaine. *Ther. Drug Monit.* **7**:66–73 (1985).
9. S. Vozeh and C. Steiner. Estimates of the population pharmacokinetic parameters and performance of Bayesian feedback: A sensitivity analysis. *J. Pharmacokin. Biopharm.* **15**:511–528 (1987).
10. D. P. Heyman and M. J. Sobel. *Stochastic Models in Operations Research*, McGraw-Hill, New York, 1984.
11. D. Z. D'Argenio and J. H. Rodman. Targeting the systemic exposure of teniposide in the population and the individual using a stochastic therapeutic objective. *J. Pharmacokin. Biopharm.* **21**:223–251 (1993).
12. D. Z. D'Argenio and K. Park. Stochastic control of pharmacodynamic processes with application to terbutaline. *Proceedings of the IFAC Symposium*, Galveston, TX, 1994, pp. 246–247.
13. C. Hu, W. S. Lovejoy, and S. L. Shafer. Comparison of some suboptimal control policies in medical drug therapy. *Operations Res.*

14. P. O. Maitre, S. Vozeh, J. Heykants, D. A. Thomson, and D. R. Stanski. Population pharmacokinetics of alfentanil: The average dose-plasma concentration relationship and interindividual variability in patients. *Anesthesiology* **66**:3–12 (1987).
15. H. Schwilden. A general method for calculating the dosage scheme in linear pharmacokinetics. *Eur. J. Clin. Pharmacol.* **20**:379–383 (1981).
16. D. Z. D'Argenio and D. Katz. Application of stochastic control methods to the problem of individualising intravenous thephylline therapy. *Biomed. Meas. Inform. Contr.* **2**(3) (1988).
17. C. Hu, W. S. Lovejoy, and S. L. Shafer. An efficient control strategy for dosage regimens. *J. Pharmocokin. Biopharm.* **22**:73–94 (1994).
18. D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case,* Academic Press, New York, 1978.
19. K. M. Van Hee, *Bayesian control of Markov chains*, Mathematical Centre Tract 95, Amsterdam, 1978.
20. W. S. Lovejoy. Suboptimal policies, with bounds, for parameter adaptive decision processes. *Operations Res.* **41**:583–599 (1993).
21. R. W. Jelliffe and A. Schumitzky. Modeling, adaptive control, and optimal drug therapy. *Med. Prog. Technol.* **16**:95–100 (1990).
22. P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*, Academic Press, London, 1981.
23. S. E. Kern and D. R. Westenskow. Development of a closed loop system for neuromuscular blocking agents given in intensive care. *Proceedings of the IFAC Symposium*, Galveston, TX, 1994, pp. 172–173.