Closed-Loop Control in Clinical Pharmacology: Paradigms, Benefits, and Challenges

James M. Bailey[†], **Wassim M. Haddad**[‡], **and Tomohisa Hayakawa**[‡]

[†]Department of Anesthesiology, Northeast Georgia Medical Center, Gainsville, GA 30503
[‡]School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0150

I. Introduction

Control engineering is the underpinning for technological advances in fields as diverse as aerospace, chemical, power, manufacturing, electronic, communication, transportation, and network engineering. However, control technology has had less impact on modern medicine. There have been exciting breakthroughs in areas such as robotic surgery, electrophysiological systems (pacemakers and automatic implantable defibrillators), life support (ventilators, artificial hearts), and image-guided therapy and surgery. However, in general, there are steep barriers to the application of modern control theory and technology to medicine. The steepest barriers are the system uncertainties, inherent to biology, that preclude mathematical modeling and hence application of many of the tools of modern control technology. Another steep barrier is communication between control engineers and the medical community. Future advances will depend on collaboration between control theorists and engineers and biomedical researchers.

One of the areas of medicine most suited for applications of control theory is clinical pharmacology, a discipline in which mathematical modeling has had a prominent role. Some of the most important advances in modern medicine have been in the area of pharmacology. The physician in the 21st century has a broad armamentarium of drugs available for the treatment of disease. This is in contrast to previous generations of physicians, who were largely limited to diagnosis, possible surgery, and often only consolation. Yet, while we have an abundance of therapeutic agents, proper dosing of drugs is often imprecise and may be a significant cause of increased costs, morbidity, and mortality.

It is instructive to consider how dose guidelines are derived. Drug development begins with animal experimentation. Promising agents are then taken to human trials, beginning with healthy volunteers and progressing to patients with the disease for which the drug is being developed. Early stages of these trials focus on safety while the final trials usually entail randomized, blinded administration of placebo and different drug doses for the evaluation of efficacy. Efficacy is statistically defined and even when there is a therapeutic effect in the statistical aggregate, there may still be individual patients for whom the drug is either not efficacious or who experience side-effects. If a therapeutic effect is observed, then the drug may be approved by the Food and Drug Administration and, in general, the recommended dose is that found to be efficacious in the "average" patient. And this is the problem. No patient is an "average" patient. There is very substantial variability among patients in the drug concentration at the locus of the effect (the *effect site concentration*) that results from a given dose, as there is also variability among patients in the therapeutic efficacy of any given effect site concentration. Thus, there is large variability among patients in the therapeutic effect of any given dose. In the vast majority of cases, the appropriate dose for a specific patient is found by trial and error. For example, the internist treating a patient with essential hypertension will begin by prescribing the recommended dose and then, in follow-up, will observe the effect of the drug on blood pressure and adjust the dose empirically. This process can be cumbersome, time consuming, and imprecise.

This research was supported in part by AFOSR under Grant F49620-03-1-0178.

II. A Primer on Clinical Pharmacology

It has been apparent for some time that dosing of drugs could be placed on a more rational basis by using *pharmacokinetic* and *pharmacodynamic* modeling. Pharmacokinetics is the study of the concentration of drugs in various tissues as a function of time and dose schedule. Pharmacodynamics is the study of the relationship between drug concentration and effect. By developing techniques relating dose to resultant drug concentration (pharmacokinetics), and concentration to effect (pharmacodynamics), one can generate a model for drug dosing.

Pharmacokinetic models will be familiar to most control engineers and theorists since they are based on dynamical system theory. The disposition of drugs in the body is a complex interplay of numerous transport and metabolic processes, many of which are still poorly understood [1], [2]. However, *compartmental models* may effectively encapsulate these processes [3]. Common pharmacokinetic models assume that, for the purpose of describing drug disposition, the body is comprised of a few homogenous, well-mixed compartments (so that the drug concentration is constant within the compartment), with linear (proportional to drug concentration) transport to other compartments or elimination from the compartment and the body by metabolic processes. The simplest model, the one-compartment model, assumes that the body is just a single compartment and also typically assumes instantaneous mixing when drug is introduced intravenously, with subsequent linear elimination. The model is characterized by two parameters, the volume of distribution (V_{d}) and the elimination rate constant (a_{e}). With this simple model the concentration (C) immediately after a dose of amount of D is equal to D/V_{d} and drug is subsequently eliminated at a rate equal to $a_{\mathrm{e}}C$ (exponential decay). While the behavior of a few drugs may actually

be described by this model, it is too simplistic for most. The assumption of instantaneous mixing, which is clearly unrealistic in the case of drugs that are taken orally, can be remedied by using a two (or more) compartment model in which there is a compartment representing the gastrointestinal tract that receives the dose and from which drug is transferred irreversibly to a second compartment that represents *intravascular blood* (blood within arteries or veins) and organ systems which receive a large amount of blood flow and hence which equilibrate with intravascular blood rapidly.

For drugs that are administered intravenously, a common model is the two-compartment *mammillary* model [3]. This model assumes that there is a central compartment which receives the intravenous dose with instantaneous mixing. Drug is then either transferred to a peripheral compartment or metabolized and eliminated from the body. Drug elimination from the peripheral compartment is ignored since this compartment is identified with tissues such as muscle or fat which are metabolically inert as far as the drug is concerned. (Most drugs are metabolized in the liver or kidney, organs that, along with the heart and brain, equilibrate rapidly with the intravascular blood and are identified with a central compartment that receives the intravenous dose.) Drug in the peripheral compartment transfers back to the central compartment with linear kinetics. The system is then described by the familiar state space model

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0, \quad t \geq 0, \tag{1}$$

where

$$A = \begin{bmatrix} -(a_{21} + a_{11}) & a_{12} \\ a_{21} & -a_{12} \end{bmatrix},$$

$x = [x_1, x_2]^{\mathrm{T}}$ is the state vector representing the masses in the two compartments, a_{12} and a_{21} are the compartment 2 to compartment 1 and the compartment 1 to compartment 2 transfer coefficients, respectively, and a_{11} is the rate at which drug is eliminated out of the system from (the central) compartment 1. The other system parameter is V_1, the volume of the central compartment (for a total of four pharmacokinetic parameters). Note that with the assumption of instantaneous mixing, the *concentration* at $t = 0$ after dose D is D/V_1. The assumption of instantaneous mixing is unrealistic but has little effect on the predictive accuracy of the model as long as we do not try to model drug concentrations immediately after the initial drug dose. The two-compartment mammillary model is generally useful for drugs that are administered intravenously, although some require an extension of the model to include two distinct peripheral compartments along with the central compartment (the three compartment mammillary model). Other extensions or revisions of the basic model are possible. In most cases the assumption of linear transfer is maintained so that the system equation remains the familiar

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0, \quad t \geq 0, \tag{2}$$

where $x \in \mathbb{R}^n$ represents the system compartmental masses or system compartmental concentrations and $A \in \mathbb{R}^{n \times n}$ is a *compartmental matrix* [3] in the case where x represents compartmental masses and a *nonnegative matrix* [3] in the case where x represents compartmental concentrations. Hence, (2) describes a nonnegative, compartmental dynamical system and there is a substantial body of theoretical work which is relevant for analyzing these systems (see [3] and the numerous references therein).

It should be readily apparent that pharmacokinetic models, especially mammillary models, are coarse grained oversimplifications. Consider the injection of a drug into a small peripheral vein in the hand. The drug will be transported in the venous stream of flow to the right heart, binding to blood cells or proteins and mixing with other venous streams as various veins coalesce, with large scale mixing in the right atrium and ventricle, from where it will be transported to the lung. In the lung some of the drug may bind to lung tissue. From the lung the drug returns to the left heart from which it is expelled into the aorta for transport to other inert tissues, where drug binding occurs, and to tissues (like the liver and kidney) where the drug is metabolized. Modeling this with a small number of compartments is clearly a coarse-grained approximation. It is almost a cliche to note that the clinical utility of these models depends entirely on the time scale of the application. For example, these simplified models work quite well for estimation of dosing intervals for drugs administered orally. As another example, there has been interest among anesthesiologists, to be discussed in more detail later, in using pharmacokinetic models to produce and maintain therapeutic drug concentrations. Using simplified mammillary models one can achieve median absolute performance errors (the normalized offset of target and measured drug concentrations) of less than 20%, when drug concentrations are sampled on the order of every 15 minutes. This is clinically quite acceptable in the sense that drug concentrations within this range of the target generally achieve the desired effect. But consider the problem of predicting drug concentrations during the induction of anesthesia. Anesthesia is typically initiated by intravenously administering a bolus (impulse function) dose of a hypnotic drug. During the minute needed to induce anesthesia large-scale mammillary models fail to predict drug concentrations. This is, indeed, obvious, since mammillary models assume instantaneous mixing. To predict drug concentrations one needs more elaborate models, such as those which incorporate mixing chambers or *catenary* models [3]. For example, models which utilize catenary structures to approximate transport of drug from the injection site to the central circulation (the heart, brain, etc.) and additional parallel compartments (similar to mammillary models) to account for distribution of drug to peripheral (muscle and fat) tissues have very effectively described the process of induction of anesthesia [4]. Alternatively, we can consider the pharmacokinetics of inhaled anesthetics. Modern technology allows the on-line measurement of the anesthetic concentration at the end of each breath. This measurement rate requires a much more fine-grained pharmacokinetic model than the typical 2 or 3 compartment mammillary model.

While the most commonly used pharmacokinetic models are linear, it is clear that the underlying processes that determine pharmacokinetic behavior are nonlinear. For example, the molecular processes of drug metabolism are typically described by Michaelis-Menten kinetics in which the rate of drug metabolism is given by $V_{\mathrm{m}}C/(K_{\mathrm{m}} + C)$, where V_{m} and K_{m} are constants and C is the drug concentration, while the large scale pharmacokinetic models assume linear drug metabolism or elimination (which is valid only in the limit of decreasing drug concentration). Similarly, transport of drug between various tissues will be proportional to blood flow between the tissues, so the transport of drug would be linear. However, many drugs can alter cardiovascular function and, hence, pharmacokinetic behavior becomes nonlinear. The impact of these nonlinearities is unclear

and again depends on the time scale of the model. It appears that linear models are adequate for coarse-grained prediction of drug concentration (> 10 minutes) as long as the range of observed drug concentrations is not too great. This latter observation follows from the practical fact that drugs are not approved for clinical use if the ratio of toxic concentrations (as in cardiovascular depression) to therapeutic effect is small. The impact of nonlinearity (and also model oversimplification) for application of control theory can only be assessed by actual clinical testing of the control application. And while animal experimentation is very useful, inter-species differences will dictate that human clinical testing will be the final measure of the utility of the models. In addition, as we will see, at this time there is a paucity of human data available from the application of control theory to clinical pharmacology.

Parenthetically, it is important for the control engineer or theorist who wants to approach the pharmacokinetic literature to realize that the conventions of nomenclature are somewhat different than those used in this article. For example, pharmacokineticists denote the transfer coefficient from compartment i to compartment j as k_{ij} rather than a_{ji}. Pharmacokinetists also often parameterize models differently. For example, most pharmacokinetic papers will report the *terminal elimination half-life*, the time required for drug concentration to decrease by 50% if all tissues are equilibrated with the blood concentration. Another commonly reported parameter is the *clearance*, which is the volume of tissue or blood "cleared" of drug per unit time. Many pharmacokinetic investigations will be parameterized in terms of compartment volumes and intercompartmental clearances. These parameters are simply transformations of the basic elements of the system matrix A, along with a scale parameter, which in the case of the two compartment mammillary model is the volume of the central compartment.

The experimental data used for pharmacokinetic modeling is typically collected by administering drug to patients and then drawing blood samples at various times after the initiation of dosing, and determining the concentration of drug as a function of time. Consequently, most pharmacokinetic investigations focus on blood concentrations. One of the goals of the analysis for drugs administered intravenously is to derive an expression for the *unit disposition function*, the blood concentration that results from a single unit bolus dose (impulse function) of drug. In the case of linear kinetics, if the unit disposition function (f_{ud}) is known then the blood concentration that results from any arbitrary dose schedule is easily calculated by the convolution integral

$$C(t) = \int_0^t f_{\mathrm{ud}}(\tau) D(t-\tau) \mathrm{d}\tau, \tag{3}$$

where $D(t)$ is the dose as a function of time [5]. Note that it is seldom technically feasible to actually measure drug concentrations in the tissue thought to be the site of the therapeutic effect, and it is often assumed that effect site concentration and blood concentration are linearly related, if not equal. The vast majority of drugs are distributed to the site of action by blood flow and in general the effect site rapidly equilibrates with blood. If the finite equilibration time between the central intravascular blood volume and the effect site is clinically relevant, then the pharmacokinetic model should be revised to include a distinct effect site compartment.

Pharmacokinetic parameters (the entries of the system matrix A) are estimated by fitting models to the data. The models, of course, are approximations and there are numerous sources of noise in the data, from assay error to human recording error. Thus there is always an offset between the concentration predicted by the model and the observed data, the prediction error. One common method for estimating pharmacokinetic parameters is to use the method of maximum likelihood [6]. In this type of analysis one assumes a specific statistical distribution for the prediction error and then determines the parameter values that would maximize the likelihood of the observed results. For example, suppose we have conducted a study in a single patient in which we have collected blood samples at 10 different points in time after a single bolus intravenous dose of the drug. If we assume that the prediction error has a simple normal or Gaussian distribution, then the likelihood of the observed results will be proportional to

$$\prod_{i=1}^{r} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-PE_i^2/2\sigma^2}, \tag{4}$$

where PE_i is the prediction error of the ith observation and is given by $PE_i = C_{\mathrm{p}i} - C_{\mathrm{m}i}$, where C_{p_i} is the predicted ith drug concentration and $C_{\mathrm{m}i}$ is the measured ith drug concentration, σ^2 is the variance of the assumed Gaussian distribution of prediction errors, and r is the number of observations (measured concentrations). We refer to this as the *intrapatient error model*. Note that the above expression is a function of σ and the pharmacokinetic parameters (the entries of the system matrix A). By maximizing the above expression (or more commonly its logarithm) with respect to the pharmacokinetic parameters and σ one may estimate the structural model parameters (the entries of the system matrix A) and the error model parameters (in this simple case, σ) that maximize the likelihood of the observed results. The reader familiar with statistical estimation theory will realize that the above example reduces to simple least squares estimation. However, using a more sophisticated error model (for example, by assuming that prediction error has a normal distribution with variance proportional to the predicted concentration raised to an unknown power) leads to more complex methods of parameter estimation [6].

There are two distinct approaches to estimating mean pharmacokinetic parameters for a population of patients [7], [8]. In the first, models are fitted to data from individual patients and the pharmacokinetic parameters for individual patients are then averaged (*two-stage analysis*) to provide a measure of the pharmacokinetic parameters for the population. The other approach to data analysis involves pooling of the data from individual patients. It is called *mixed-effects modeling* because in this situation the prediction error is determined not only by the stochastic noise of the experiment but also by the fact that *different patients have different pharmacokinetic parameters*. The error model, the analogue of the simple Gaussian distribution used in the example above, must account not only for variability between the observed and predicted concentrations within the same patient but also for variability between patients. The analyst must assume a statistical distribution for both intrapatient variability and interpatient variability. Most commonly, it is assumed that pharmacokinetic parameters have a log-normal distribution. This sophisticated method of analysis not only estimates the mean structural pharmacokinetic parameters (the elements of the system matrix A) but also the statistical variability of these elements in the population,

the *interpatient variability*. Since the total variance is the sum of interpatient and intrapatient variability, the latter is also estimated. This is a very powerful method of analysis for two reasons. First, it gives the clinician not only an estimate of the pharmacokinetic parameters but also an estimate of their variance. This is extremely important for the clinician since no matter how desirable the properties of a drug are, on average, if there is extreme variability in these properties it may not be safe for clinical use. And second, mixed-effects modeling may allow a reduction in the amount of data that is gathered from each individual patient. In a two-stage analysis, one must have enough data points from each patient to estimate their pharmacokinetic parameters. For example, if one adopts a two compartment mammillary model, there are 4 pharmacokinetic parameters. It is impossible to estimate these parameters for any one patient with 4 or less data points from that patient. However, with mixed-effects modeling it is possible to use sparse data. This also is an important advantage since pharmacokinetic studies may be expensive and time consuming.

In contrast to pharmacokinetic modeling, pharmacodynamic modeling is more empirical. The molecular mechanism of action of many drugs is reasonably well-understood in that most drugs act by binding to some "receptor" on or within target cells [1]. There is a well-developed theory of multiple equilibrium binding of ligands, such as drug molecules, to receptors on larger macromolecules, such as proteins. So in theory pharmacodynamics, the relationship between drug concentration and effect, should follow from these models of molecular binding. However, the physiological effect is a complex interplay of numerous factors and it is generally not possible to quantitatively relate the effect at the level of the intact organism to the number of receptors bound by the drug at the molecular level. Empirical models are needed. It could be assumed that drug effect is proportional to the drug concentration at the effect site but this is clearly unrealistic since it admits the possibility of limitless drug effect. For example, consider a drug which lowers heart rate. It is unrealistic to assume that the drug effect is proportional to drug concentration since there is no limit on the drug concentration but there is a limit on the effect (the heart rate cannot be slower than zero). The empirical model should incorporate a ceiling effect. One model that has been quite effective for a variety of drugs is the *Hill equation*

$$E = E_{\max} C^{\gamma} / (C^{\gamma} + C_{50}^{\gamma}), \tag{5}$$

where E is the drug effect, $E_{\max}$ is the maximum drug effect, C is the drug concentration, C_{50} is the drug concentration associated with 50% of the maximum effect, and γ is a dimensionless parameter that determines the steepness of the concentration-effect relationship [9]. Note that this model reduces the concentration-effect relationship to three parameters, the maximum effect, a measure of the midpoint of the relationship, and a measure of the steepness. It is interesting that this model was first developed in 1906 to describe a *molecular* interaction, the binding of oxygen to hemoglobin. Since that time it has been applied to a wide variety of phenomenon which are far removed from explanations at the molecular level. There are a number of modification of this basic model that have been employed. One important one is when the drug effect is a binary, yes-or-no, variable. An example of a binary variable is anesthesia, for which the patient is either responsive or not. In this case, the pharmacodynamic model based on the Hill equation becomes

$$P = C^{\gamma} / (C^{\gamma} + C_{50}^{\gamma}), \tag{6}$$

where the effect is now the probability P that the patient will not respond to some noxious stimuli (and $E_{\max}$ equals unity) [10], [11].

In typical pharmacodynamic studies, drug is administered and the effect is measured at various points in time. At each point of observation, a blood sample is taken for the determination of the drug concentration at the time of observation of effect. The parameters of the pharmacodynamic model ($E_{\max}$, C_{50}, γ) may then be estimated by the same methods (maximum likelihood, generalized least squares, etc.) described above. Obviously, if blood drug concentrations and effect site concentrations have not equilibrated, this analysis is invalidated.

It should be noted that pharmacodynamic models are inherently nonlinear, in contrast to pharmacokinetic models, which are usually linear. However, the interplay with pharmacodynamics may lead to nonlinear pharmacokinetics also. For example, some intravenous anesthetics depress *cardiac output*, the volume of blood pumped by the heart per unit of time. Since the basic transport processes that determine pharmacokinetic behavior are fundamentally functions of blood flow, administration of the drug alters its kinetics and since the pharmacodynamic relationship between drug concentration and depression of cardiac output is nonlinear, the pharmacokinetics of the drug are, in reality, also nonlinear.

III. Clinical Pharmacology and Drug Dosing

In addition to safety and efficacy, the Food and Drug Administration requires pharmacokinetic evaluation before approval of any new drug. The pharmacokinetic profile may be useful in developing dose guidelines. However, this application of basic principles is usually quite simplified. The disposition of most drugs is determined by both metabolic processes that eliminate the drug and distribution processes, that is, transfer between various tissue groups. The route of distribution is via the intravascular blood volume whether the drug is administered by mouth, intramuscular injection or intravenously. The complexity of these processes implies that the governing dynamical system model is almost always characterized by a vector differential equation. However, the vast majority of drugs are given for chronic conditions, and when the time scale of treatment greatly exceeds the time scale of the distributive processes, one can ignore them. Furthermore, very few patients would comply with the complex dosing schemes ("take 3 pills in the morning and then 2 1/2 at 3:00 pm and then 2 at 8:00 and 10:00 pm and then one the next morning...") needed to account for distributive processes at the onset of therapy. Thus, the application of pharmacokinetic principles must be simplified. In terms of the system equation (2), we assume that A is a scalar. For example, if we know that a dose of 50 mg of an antihypertensive drug is efficacious in the average patient and we also know that the half life in the average patient is 12 hours then we may propose a dosing schedule that begins with an initial dose of 50 mg with subsequent dosing of 25 mg every 12 hours. Or, as another example, suppose we know that a blood concentration of an intravenous anesthetic of 100 μg/ml reliably produces unconsciousness and that we also know that the clearance (the amount of blood cleared of drug per unit time) is 150 ml/minute. Then an infusion of

100 μg/ml $\times$ 150 ml/min = 15000 μg/min will maintain this blood concentration, although this concentration will not be achieved until distributive processes have equilibrated. In point of fact, many of the dosing guideline recommended by the manufacturers of drugs are based on simple calculations like these. And although it is often not perceived as such by the clinician, initial drug dosing is a form of open-loop control, that is, control without feedback.

There have been attempts to develop more precise open-loop control in the acute care environment, especially in the area of anesthetic pharmacology. With the increased availability in the 1980s of small computers that could be taken into the operating room, several groups of investigators developed computer-controlled pump systems that continually adjusted the drug infusion rate to achieve and maintain the drug concentration desired by the clinician [12]–[15]. These algorithms use the appropriate pharmacokinetic model

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \geq 0, \qquad (7)$$

with *average* pharmacokinetic parameters taken from previous investigations to calculate the needed dose $u(t)$, $t \geq 0$, usually via the unit disposition function and the assumption of linearity. The output, which is continually updated, drives the infusion pump.

This is clearly open-loop control since, as previously emphasized, no one patient is an average patient and there is no mechanism for measuring the concentrations in the individual patient for feedback control. It is technically not feasible to actually measure blood concentrations of intravenous anesthetics in real time. But even with the lack of feedback, numerous studies have demonstrated better control of drug concentrations than the standard empirical dosing used by most clinicians. The clinical relevance of this is unclear. While open-loop control systems have not yet been approved by the Food and Drug Administration for routine clinical use in the United States, several European countries have approved a device for the infusion of the intravenous anesthetic, propofol, and this device is currently in use for clinical delivery of anesthesia.

While initial dosing guidelines may be based on the average patient, the very significant interpatient pharmacokinetic and pharmacodynamic variability observed for most drugs leads to the inevitable conclusion that *precise* drug dosing will require closed-loop control. As noted in the introduction, in one sense most drug dosing is a form of closed-loop control. Patients are quite familiar with this. The physician prescribes a drug, usually given orally, and an initial dose, observes the response, and adjusts the dose. An experienced physician can be quite adept at this process, but, in general, it is certainly not systematic and is usually time consuming. Most individuals who have been treated for a chronic disease know this well.

The process of dose titration can be made somewhat more precise by the use of mixed-effects pharmacokinetic modeling and *post-hoc* Bayesian estimation of individual patient pharmacokinetic parameters [6]–[8]. It will be recalled that mixed-effects modeling provides not only estimates of pharmacokinetic parameters but also their variance within the population. Suppose one has measured one or more drug concentrations in an individual patient. Using Bayesian probability principles, the likelihood of a given value of some pharmacokinetic parameter, Θ, is proportional to $P(C|\Theta)P(\Theta)$. $P(C|\Theta)$ is the probability of the observed concentration(s) as a function of Θ and is simply the intrapatient error model cited earlier (an example is equation (4)). $P(\Theta)$ is the *a priori* probability of a given value of Θ and is given by the assumed distribution for Θ (as noted above, usually log-normal) and the variance of Θ estimated from the mixed-effects analysis. By determining the mode of $P(C|\Theta)P(\Theta)$ with respect to Θ one can derive a maximum likelihood estimate of Θ for the specific patient. By estimating patient-specific parameters one can more accurately calculate the necessary dose to achieve a given drug concentration. This process has been demonstrated to improve the precision of drug dosing [16]. But note that it only improves the precision of achieving a given drug *concentration* which may or may not lead to better control of drug *effect*, given pharmacodynamic variability. Also this process requires measurement of drug concentration, something that cannot usually be done quickly (a typical drug assay takes hours, if not more than a day, to complete).

While the process of titrating drug dose to the desired effect may be acceptable (if often frustrating) for chronic outpatient therapy, in the acute care environment, such as the operating room or the intensive care unit, this process may be dangerously slow or imprecise. It is in this environment that control technology has much to offer modern medicine and for the remainder of this article we will restrict ourselves to drugs used in the acute care setting.

In order to implement closed-loop control in an acute care environment one must have a real-time nearly instantaneously measurable performance or control variable. Early attempts at closed-loop control have of necessity focused on control of variables that are conveniently measured. By their very nature, cardiovascular and central nervous system function are critical in the acute care environment, and so mature technologies have evolved for their measurement. Thus, the primary applications of closed-loop control of drug administration have been to hemodynamic management and control of levels of consciousness. Before discussing our investigations of closed-loop control of anesthesia, we will briefly review closed-loop control of cardiovascular function, as it illustrates many of the general problems inherent in the application of control technology to physiological function.

IV. Closed-Loop Control of Cardiovascular Function

After major surgery, especially cardiac surgery, many patients become profoundly hypertensive [17]. While this syndrome is distinct from the essential hypertension well known to both patients and medical professionals, it does require treatment since elevated blood pressure may cause cardiac dysfunction, leading to pulmonary edema or myocardial ischemia, may be a risk factor for stroke, and may exacerbate bleeding from fragile surgical suture lines. There are a number of potent drugs available for the treatment of post-operative hypertension but titrating these drugs to achieve the desired blood pressure may be difficult. Underdosing leaves the patient hypertensive and overdosing can reduce the blood pressure to levels associated with shock. There has been interest since the late 1970s in developing controllers for the administration of sodium nitroprusside (SNP), a commonly used and potent anti-hypertensive. The problems encountered in this endeavor are enlightening. The initial attempts used simple nonadaptive methods such as proportional-derivative or proportional-integral-derivative controllers that assumed a linear relationship between infusion rate and effect [18], [19]. This was a tenuous assump-

tion. While the drug concentration may be the simple convolution of the infusion rate and a transfer function (equation (3)), the relationship between effect and infusion rate is not likely to be so simple (see equation (5)). Also, one of the significant challenges to the design of a blood pressure controller is the fact that there is a time delay between administration of the drug and the clinical effect. Failure to account for this time delay can lead to significant system oscillations. These early blood pressure controllers included time delays in the system model; however, the delays were assumed to be the same for each patient. While these early controllers were successful in some patients, in general they have not had wide clinical implementation. The barriers to clinical implementation were the nonlinear patient response and significant interpatient differences in drug sensitivity. It was very evident that interpatient variability, and also the fact that an individual patient's sensitivity to the drug varies in time, made adaptive controllers essential. Subsequently, single model and multiple model adaptive controllers were developed [20], [21]. Single model adaptive controllers are based on on-line estimation of system parameters using minimum variance or least squares methodology. These controllers were also not acceptable due to large amplitude transients. Multiple model adaptive controllers represent the system by one of a finite number of models. For each model there is a separate controller. The probabilities that the system is represented by each of the different models are calculated from the relative offsets of the system response and the response predicted by each model. The output of the controller is the probability-weighted sum of the outputs from each model [22], [23]. Multiple model adaptive controllers have proven to be somewhat more satisfactory. Subsequent refinements to blood pressure control have included single model reference adaptive control [24], which appeared promising in simulations, and neural network-based methods [25]. There has also been substantial interest in optimal control since sodium nitroprusside has toxic side effects when the dose is too high [26].

These investigations into control of blood pressure reveal the challenges inherent to biological systems, specifically nonlinearity, interpatient variability (system uncertainty), and time delays. Despite the refinements of closed-loop blood pressure controllers, they are seldom used clinically. While this is due, in part, to the cost of technology acquisition, this is probably not the most important impediment to their clinical use. Blood pressure control is important, but cardiovascular function involves several other important variables and all these variables are interrelated [17]. The intensive care unit clinician (nurse or physician) must not only insure that blood pressure is within appropriate limits but that also cardiac output (the amount of blood pumped by the heart per minute) is acceptable and heart rate is within reasonable limits. Mean arterial blood pressure is proportional to cardiac output, with the proportionality constant denoted the systemic vascular resistance, in analogy to Ohm's law. Cardiac output is equal to the product of heart rate and *stroke volume*, the volume of blood pumped with each beat of the heart. Stroke volume, in turn, is a function of *contractility* (the intrinsic strength of the cardiac contraction), *preload* (the volume of blood in the heart at the beginning of the contraction), and *afterload* (the impedance to ejection by the heart). The intensive care unit clinician must balance all these variables. There are drugs (inotropic agents) that increase contractility, but will also have variable effects on heart rate and afterload. There are also drugs which increase (vasopressors) or decrease (vasodilators) afterload. Finally, stroke volume may be increased by increasing preload and this can be accomplished by giving the patients fluid. However, giving too much fluid may be deleterious since it can lead to impaired pulmonary function as fluid builds up in the lungs. The fact that closed-loop control of blood pressure has not widely adopted by clinicians is not too surprising when one considers the complex interrelationships of hemodynamic variables. However, this also indicates an area where future applications of control theory could be invaluable. The technology is currently available to measure heart rate, blood pressure, cardiac output, and measures of preload continuously and in real time. Adaptive and robust optimal controllers which control the administration of multiple drugs (inotropes, vasopressors, vasodilators) and fluids would be a major advance in critical care medicine. There have been some preliminary investigation of the control of multiple hemodynamic drugs [27], [28] but this is an area of great potential for future research.

V. Closed-Loop Control of Anesthesia

There has been long-standing interest in closed-loop control of anesthesia. Adequate anesthesia is comprised of several components; *analgesia*, lack of reflex response, such as increased blood pressure or heart rate, to surgical stimulus, *areflexia*, lack of movement (which simplifies the task of the surgeon), and *hypnosis* or lack of consciousness. In order to implement closed-loop control it is necessary to measure the state and the assessment of consciousness. Attempting to measure and control consciousness has been challenging. However, two technical innovations have facilitated the development of feedback controllers. The first (historically) is the routine clinical implementation of real-time spectroscopic methods for measuring the concentration of inhaled anesthetic agent in exhaled gases from the lung, in particular end-expiratory (routinely called end-tidal) gases. End-tidal anesthetic gas concentration is a reasonable surrogate for arterial blood anesthetic concentration [29]. Since end-tidal anesthetic agent concentrations can be measured in real time with this technology, this has allowed closed-loop control of end-tidal anesthetic concentration. However, anesthetic concentration cannot be equated with anesthetic effect. More recently, real time processed *electroencephalograph* (EEG) measurement has held open the possibility of closed-loop control of anesthetic effect. It has been known for decades that the EEG changes with induction of anesthesia [30]. However, *quantitatively* relating the EEG to anesthetic effect has been challenging. In the last decade, there has been substantial progress in developing processed EEG monitors that provide a measure of the depth of anesthesia and are candidates for performance variables for closed-loop controllers.

Inhaled anesthetic agents have been the mainstay of clinical practice since the first delivery of anesthesia. A fundamental characteristic of every inhaled anesthetic agent is its "MAC" value, for *m*inimum *a*lveolar (*alveoli* are the fundamental units of the lung) *c*oncentration that is associated with a 50% probability of patient movement or no movement in response to surgical stimulus [27]. By maintaining end-tidal concentrations well above MAC, the practitioner is relatively assured of hypnosis. The ready availability of spectroscopic systems for measuring end-tidal anesthetic concentration in real time has led several investigators to develop closed-loop controllers. The earliest of these controllers used proportional-integral-derivative

algorithms [31], [32]. As noted above, these controllers share the weaknesses of assuming that all patients are the same. More recently, adaptive model-based controllers have been developed [33], [34]. These typically rely on least-squares methods to estimate the specific system parameters for the individual patient. In animal studies, the adaptive controllers have performed, not surprisingly, more robustly than the fixed gain controllers. However, they have not been widely adopted clinically. The primary reason is that because of interpatient pharmacodynamic variability, control of anesthetic concentration does not translate into control of anesthetic effect, and most clinicians would value control technology only if it prevented the possible overdoses inherent in maintaining end-tidal concentration in each individual patient well above the MAC value, an average from a population of patients. Closed-loop control of anesthesia requires a monitor of anesthetic effect, specifically consciousness.

The development of a monitor of consciousness has been an elusive challenge for anesthesiologists. The EEG, a global measure of electrical activity in the brain, has been an obvious candidate. In particular, neurophysiologists have observed that the EEG of an anesthetized patient contains slower waves with higher amplitudes. However, the EEG is a complex of multiple time series and multiple spectra and while there are characteristic changes in the EEG with the induction of anesthesia, it has not been clear which, if any, characteristic of the EEG best reflects the anesthetic state. Building on pioneering work by Bickford [35], Schwilden and his colleagues developed and clinically tested a closed-loop model-based adaptive controller for the delivery of intravenous anesthesia using the median frequency of the EEG power spectrum as the control variable [36]. Their model assumed a two compartment pharmacokinetic model for which the concentration of drug $C(t)$ as a function of time (t) after a single bolus dose was given by

$$C(t) = Ae^{-\alpha t} + Be^{-\beta t}, \tag{8}$$

where A, B, α, β are patient-specific pharmacokinetic parameters. It was also assumed that the control variable, median EEG frequency (denoted by E), was related to the drug concentration by the modified Hill equation

$$E = E_0 - E_{\max}[C^\gamma/(C^\gamma + C_{50}^\gamma)], \tag{9}$$

where E_0 is the baseline signal, $E_{\max}$ is the maximum decrease in signal with increasing drug concentration, C_{50} is the drug concentration associated with 50% of the maximum effect, and γ is a parameter describing the steepness of the concentration-effect curve. From the above equation it can be seen that the drug effect is a function of the pharmacokinetic parameters (A, B, α, β) as well as the pharmacodynamic parameters (E_0, $E_{\max}$, C_{50}, and γ). If these parameters are known, calculation of the dose regimen needed to achieve the target EEG signal is straightforward. However, these parameters are not known for individual patients. The algorithm developed by Schwilden and his colleagues assumed that each of the pharmacodynamic parameters (E_0, $E_{\max}$, C_{50}, and γ) and the pharmacokinetic parameters α and β were equal to the mean values reported in prior studies. Then using the mean population values of the pharmacokinetic parameters A and B as starting values, estimates of these parameters were refined by analysis of the difference between the target and observed EEG signal (ΔE). Linearizing ΔE with respect to A and B we find

$$\Delta E = (\partial E/\partial A)\delta A + (\partial E/\partial B)\delta B, \tag{10}$$

where δA, δB represent the updates to the values of A and B in the adaptive control algorithm. In conjunction with minimization of $\delta A^2 + \delta B^2$ this equation was used to solve for δA and δB. It is important to note that this algorithm was only partially adaptive in that the only parameters of the model that were updated were A and B. This algorithm was implemented for the intravenous anesthetic agents methohexital and propofol but did not appear to offer great advantage over standard manual control [36], [37]. This may have been due to the approximations of the algorithm or due to the deficiencies of the median EEG frequency as a measure of the depth of anesthesia.

Since the early work by Schwilden *et al.*, other EEG measures of depth of anesthesia have been developed. Possibly the most notable of these is the *bispectral index* or BIS [38], [39]. The BIS is a single composite EEG measure that appears to be closely related to the level of consciousness and that can track changes in latency of some of the frequency components of the EEG signal. Recently, Struys and colleagues have described a closed-loop controller of the delivery of the intravenous anesthetic propofol using a model-based adaptive algorithm with the BIS as the control variable [40]. The algorithm is similar to that of Schwilden and his colleagues in that it is based on a pharmacokinetic model predicting the drug concentration as a function of infusion rate and time, and a pharmacodynamic model analogous to that used by Schwilden *et al.* [36], [37] relating the BIS signal to concentration. However, in contrast to Schwilden and his colleagues, Struys *et al.* [40] assume that the pharmacokinetic parameters are always correct and that any variability in individual patient response is due to pharmacodynamic variability. More specifically, with induction they calculated a predicted concentration using the pharmacokinetic model and then constructed a BIS-concentration relationship using the observed BIS during induction and the predicted propofol concentration. With each time epoch, the difference between the target BIS signal and the observed BIS signal is used to update the pharmacodynamic parameters relating concentration and BIS signal for the individual patient. Note that this algorithm is only partially adaptive in the sense that there is no adaptive updating of pharmacokinetic parameters. Using this algorithm, Struys *et al.* [40] demonstrated excellent performance as measured by the difference between the target and observed BIS signals. However, as pointed out by Glass and Rampil, the excellent performance of the system may have been because the system was not fully stressed [41]. In their study, Struys *et al.* [40] administered a relatively high fixed dose of the opioid remifentanil, in conjunction with propofol. This dose blunted the patient response to surgical stimuli and meant that the propofol was needed only to produce unconsciousness in patients who were profoundly analgesic. The result was that only small adjustments in propofol concentrations were necessary. Whether the system would have been robust in the absence of deep narcotization is an open question.

In contrast to these model-based adaptive controllers, Absalom *et al.* have developed a proportional-integral-derivative controller using the BIS signal as the variable to control the infusion of propofol [42]. The median absolute performance error (the median value of the absolute value of $\Delta E/E_{\text{target}}$) of this system was good (8.0%), although in 3 of 10 patients oscillations of the BIS signal around the set point were observed and anesthesia was deemed clinically inadequate in 1 of the 10 patients. This same system has also been used with an auditory evoked potential (somatosensory

information provided by auditory stimulation generating oscillations within the EEG signal) as the control variable [43]. Intravenous propofol anesthesia has also been delivered by a closed-loop controller that uses both auditory evoked responses and cardiovascular responses as the control variables with a fuzzy-logic algorithm. This system has had only very minimal clinical testing [44]. More recently, Gentilini and his colleagues have described model-based controllers for inhalation anesthetic agents that attempt to control the BIS signal or mean arterial blood pressure, while keeping end-tidal anesthetic concentrations within pre-specified limits [45].

Given the uncertainties of both pharmacokinetic and pharmacodynamic models, and the magnitude of interpatient variability, we have been investigating parameter-independent adaptive controllers that could be implemented using the processed EEG as a performance variable. Specifically, in a recent series of papers [46]–[50] we develop direct adaptive and neural network adaptive control algorithms for nonnegative and compartmental systems. As mentioned above, nonnegative and compartmental models provide a broad framework for biological and physiological systems, including clinical pharmacology, and are well suited for the problem of closed-loop control of drug administration. Specifically, nonnegative and compartmental dynamical systems [3], [51] are composed of homogeneous interconnected subsystems (or compartments) which exchange variable nonnegative quantities of material with conservation laws describing transfer, accumulation, and elimination between the compartments and the environment. It thus follows from physical considerations that the state trajectory of such systems remains in the nonnegative orthant of the state space for nonnegative initial conditions. Using nonnegative and compartmental model structures, a Lyapunov-based direct adaptive control framework is developed in [46], [47], [49] that guarantees partial asymptotic set-point stability of the closed-loop system; that is, asymptotic set-point stability with respect to part of the closed-loop system states associated with the physiological state variables. Furthermore, the remainder of the state associated with the adaptive controller gains is shown to be Lyapunov stable. In addition, the adaptive controllers are constructed *without* requiring knowledge of the system pharmacokinetic and pharmacodynamic parameters while providing a nonnegative control (source) input for robust stabilization with respect to a given set point in the nonnegative orthant.

Neural network adaptive control algorithms have also been recently developed in [48], [50] for addressing closed-loop control of drug administration. Neural networks consist of a weighted interconnection of fundamental elements called *neurons*, which are functions consisting of a summing junction and a nonlinear operation involving an activation function. One of the primary reasons for the large interest in neural networks is their capability to approximate a large class of continuous nonlinear maps from the collective action of very simple, autonomous processing units interconnected in simple ways. In addition, neural networks have attracted attention due to their inherently parallel and highly redundant processing architecture that makes it possible to develop parallel weight update laws. This parallelism makes it possible to effectively update a neural network on line. These properties make neural networks a viable paradigm for adaptive system identification and control in clinical pharmacology. In [48], [50], we present a neural network adaptive control framework that accounts for combined interpatient pharmacokinetic and pharmacodynamic variability. In particular, we develop a neural adaptive output feedback control framework for adaptive set-point regulation of nonlinear uncertain nonnegative and compartmental systems. We emphasize that the formulation in [48], [50] addresses adaptive *output feedback* controllers for nonlinear compartmental systems with *unmodeled dynamics* of *unknown dimension* while guaranteeing ultimate boundedness of the error signals corresponding to the physical system states as well as the neural network weighting gains. Output feedback controllers are crucial in clinical pharmacology since key physiological (state) variables cannot be measured in practice.

VI. Challenges and Opportunities in Pharmacological Control

Even though there has been several control algorithms proposed in recent years for active drug administration as reported in this paper, closed-loop control for clinical pharmacology is still at its infancy. There are numerous challenges and opportunities that lie ahead. In particular, an implicit assumption inherent in all the proposed control frameworks discussed in this paper is that the control law is implemented without any regard to actuator amplitude and rate saturation constraints. Of course, any electromechanical control actuation device is subject to amplitude and/or rate constraints leading to saturation nonlinearities enforcing limitations on control amplitudes and control rates. More importantly, in pharmacological applications, drug infusion rates can vary from patient to patient and it is vital that they do not exceed certain threshold values. As a consequence, actuator nonlinearities and actuator constraints (that is, infusion pump rate constraints) need to be accounted for in drug delivery systems since they can severely degrade closed-loop system performance, and in some cases drive the system to instability. These effects are even more pronounced for adaptive controllers, which continue to adapt when the feedback loop has been severed due to the presence of actuator saturation, causing unstable controller modes to drift, which in turn leads to severe windup effects.

Another important issue not considered by most of the control algorithms discussed in this paper is sensor measurement noise. In particular, EEG signals may have as much as 10% variation due to noise. For example, the BIS signal may be corrupted by *electromyographic noise*; that is, signals emanating form muscle rather that the central nervous system. Even though electromyographic noise can be minimized by muscle paralysis, there are other sources of measurement noise (electrocautery, x-ray, movement) that are stochastic in nature and need to be accounted for within the control design processes.

In many compartmental pharmacokinetic system models, transfers between compartments are assumed to be instantaneous; that is, the model does not account for material in transit. Even though this is a valid assumption for certain biological and physiological systems, it is not true in general; especially in certain pharmacokinetic and pharmacodynamic models. For example, if a bolus of drug is injected into the circulation and we seek its concentration level in the extracellular and intercellular space of some organ, there exists a time lag before it is detected in that organ [3]. In this case, assuming instantaneous mass transfer between compartments will yield erroneous models. Hence, to accurately describe the distribution of pharmacological agents in the human body, it is necessary to include in

any mathematical pharmacokinetic model some information of the past system states. In this case, the state of the system at any given time involves a *piece of trajectories* in the space of continuous functions defined on an interval in the nonnegative orthant. This of course leads to (infinite-dimensional) delay dynamical systems [52]. This is especially relevant to correctly address the time delay inherent in equilibrating the effect site compartment with the central compartment and would have ramifications in the control design processes. For example, for adaptive control, a nonlinear adaptive algorithm for compartmental systems with *unknown* time delay would need to be developed.

Optimal control for drug administration is also often necessary in clinical pharmacology. For therapeutic reasons in the intensive care unit, it may be desirable to regulate (maintain) the amount of a drug in one compartment above a certain minimum threshold (dosage) level, while maintaining the amount below a certain maximum level in another compartment. Furthermore, to minimize drug side effects, it is desirable to minimize the total amount (dosage) of drugs used. Drug administration in clinics and hospitals do not generally satisfy the aforementioned conditions. To enforce the specialized structure of compartmental and nonnegative systems, nonnegative state and control constraints will need to be enforced as part of the controller design. The optimal nonnegative control law will need to be designed to maintain desired drug concentrations in the plasma dictated by therapeutic effects while minimizing drug dosage to reduce side effects.

A fundamental constraint for nonnegative linear system stabilization with a nonnegative control signal arises in set-point regulation. In particular, it can be shown that the existence of an equilibrium point in the interior of the nonnegative orthant of the state space is assured only if the nonnegative dynamical system has a system matrix that does not possess eigenvalues in the open right-half plane [46]. This condition implies that the largest eigenvalue of the system lies on the imaginary axis. However, by the Perron-Frobenius theorem this eigenvalue is real and therefore equal to zero. Hence, the system matrix is semistable. In light of this constraint, it can be shown using Brockett's necessary condition for asymptotic stabilizability that there does not exist a *continuous nonnegative* stabilizing feedback for set-point regulation in the nonnegative orthant for a nonnegative system. However, that is not to say that asymptotic feedback set-point regulation using *discontinuous* nonnegative feedback is not possible. Of course, in the case where the system matrix is asymptotically stable, continuous nonnegative feedback for set-point regulation in the nonnegative orthant can be used to improve system performance. In light of the above, it may be desirable to develop hybrid (discontinuous) adaptive controllers for positive set-point regulation of semistable compartmental systems. Hybrid adaptive control is virtually nonexistent in the literature. Furthermore, the problem of active control of sedation using an intermittent clinician assessment with an ordinal sedation scoring system as a performance variable necessitates hybrid control architectures to account for abstract decision making units (nurses or physicians) performing logical checks that identify system mode operation and specify the lower-level continuous-time subcontroller to be activated.

VII. Conclusion

There is no doubt that control-system technology has a great deal to offer pharmacology in general, and anesthesia and critical care unit medicine in particular. Critical care patients, whether undergoing surgery or recovering in intensive care units, require drug administration to regulate key physiological variables (e.g., blood pressure, cardiac output, heart rate, degree of consciousness, etc.) within desired levels. The rate of infusion of each administered drug is *critical*, requiring constant monitoring and frequent adjustments. Open-loop control by clinical personnel can be very tedious, imprecise, time consuming, and sometimes of poor quality. Alternatively, closed-loop control can achieve desirable system performance in the face of the highly uncertain and hostile environment of surgery and the intensive care unit. Since robust and adaptive controllers can achieve system performance without excessive reliance on system models, active robust and adaptive closed-loop control has the potential for improving the quality of medical care as well as curtailing the increasing costs of health care.

It is clear that closed-loop control for clinical pharmacology would significantly advance our understanding of the wide effects of pharmacological agents and anesthetics, as well as advance the state-of-the-art in drug delivery systems. While our focus in this paper has been to survey the recent developments of active control methods to deliver sedation to critically ill patients in an acute care environment and outline some of the future challenges of active sedation control, these control methods will have implications for other uses of closed-loop control of drug delivery. There are numerous potential applications such as control of glucose, heart rate, blood pressure, etc., that may be improved as a result of active drug dosing control. Payoffs would arise from improvements in medical care, health care, reliability of drug dosing equipment, as well as reduced cost for health care.

References

[1] A. G. Gilman, J. G. Hardman, and L. E. Limbird, *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 10th ed. New York, NY: McGraw-Hill, 1996.

[2] P. G. Welling, *Pharmacokinetics: Processes, Mathematics, and Applications*, 2nd ed. Washington DC: American Chemical Society, 1997.

[3] J. A. Jacquez, *Compartmental Analysis in Biology and Medicine*. Ann Arbor, MI: University of Michigan Press, 1985.

[4] T. K. Henthorn, T. C. Krejcie, C. U. Niemann, C. Enders-Klein, C. A. Shanks, and M. J. Avram, "Ketamine distribution described by a recirculatory pharmacokinetic model is not stereo-selective," *Anesthesiology*, vol. 91, pp. 1733–1743, 1999.

[5] H. Schwilden, "A general method for calculating the dosage scheme in linear pharmacokinetics," *Eur. J. Clin. Pharmacol.*, vol. 20, no. 5, pp. 379–386, 1981.

[6] M. Davidian and D. M. Giltinan, *Nonlinear Models for Repeated Measurement Data*. Boca Raton, FL: Chapman and Hall/CRC, 1995.

[7] L. B. Sheiner and S. L. Beal, "Evaluation of methods for estimating population pharmacokinetic parameters II. Biexponential model and experimental pharmacokinetic data," *J. Pharmacokinetics Biopharm.*, vol. 9, no. 5, pp. 635–651, 1981.

[8] L. B. Sheiner, "The population approach to pharmacokinetic data analysis: Rationale and standard data analysis methods," *Drug Metabolism Reviews*, vol. 15, no. 1-2, pp. 153–171, 1984.

[9] A. V. Hill, "The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves," *J. Physiol.*, vol. 40, pp. iv–vii, 1910.

[10] W. Lu and J. M. Bailey, "The reliability of pharmacodynamic analysis by logistic regression: A computer simulation study," *Anesthesiology*, vol. 92, pp. 985–992, 2000.
[11] W. Lu, J. G. Ramsay, and J. M. Bailey, "Reliability of pharmacodynamic analysis by logistic regression: Mixed-effects modeling," *Anesthesiology*, to appear.
[12] J. M. Alvis, J. G. Reves, A. V. Govier, P. G. Menkhaus, C. E. Henling, J. A. Spain, and E. Bradley, "Computer-assisted continuous infusions of fentanyl during cardiac anesthesia: Comparison with a manual method," *Anesthesiology*, vol. 63, pp. 41–49, 1985.
[13] S. L. Shafer, J. R. Varvel, N. Aziz, and J. C. Scott, "Pharmacokinetics of fentanyl administered by computer-controlled infusion pump," *Anesthesiology*, vol. 73, pp. 1092–1102, 1990.
[14] J. M. Bailey and S. L. Shafer, "A simple analytical solution to the three compartment pharmacokinetic model suitable for computer controlled infusion pumps," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 6, pp. 522–525, 1991.
[15] S. L. Shafer and K. Gregg, "Algorithms to rapidly achieve and maintain stable drug concentrations at the site of drug effect with a computer-controlled infusion," *J. Pharmacokinetics Biopharm.*, vol. 20, no. 2, pp. 147–169, 1992.
[16] P. O. Maitre and D. R. Stanski, "Bayesian forecasting improves the prediction of intraoperative plasma concentrations of alfentanil," *Anesthesiology*, vol. 69, no. 5, pp. 652–659, 1988.
[17] J. H. Levy, L. G. Michelsen, J. S. Shanewise, J. M. Bailey, and J. G. Ramsay, "Postoperative cardiovascular management," in *Cardiac Anesthesia*, 4th ed., J. Kaplan, Ed. Philadelphia, PA: WB Saunders, 1999.
[18] J. Slate, "Model-based design of a controller for infusion of sodium nitroprusside during postsurgical hypertension," Ph.D. dissertation, University of Wisconsin-Madison, 1980.
[19] L. C. Sheppard, "Computer control of the infusion of vasoactive drugs," *Ann. Biomed. Eng.*, vol. 8, pp. 431–444, 1980.
[20] J. M. Arnsparger, B. C. McInnis, J. R. Glover, and N. A. Normann, "Adaptive control of blood pressure," *IEEE Trans. Biomed. Eng.*, vol. 30, pp. 168–176, 1983.
[21] W. G. He, H. Kaufman, and R. J. Roy, "Multiple model adaptive control procedure for blood pressure control," *IEEE Trans. Biomed. Eng.*, vol. 33, no. 1, pp. 10–19, 1986.
[22] R. R. Rao, C. C. Palerm, B. Aufderheide, and B. W. Bequette, "Automated regulation of hemodynamic variables," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, pp. 24–38, 2001.
[23] R. R. Rao, B. Aufderheide, and B. W. Bequette, "Experimental studies on multiple-model predictive control for automated regulation of hemodynamic variables," *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 277–288, 2003.
[24] G. A. Pajunen, M. Steinmetz, and R. Shankar, "Model reference adaptive control with constraints for postoperative blood pressure management," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 7, pp. 679–687, 1990.
[25] C. T. Chen, W. L. Lin, and T. S. Kuo, "Adaptive control of arterial blood pressure with a learning controller based on multilayer neural networks," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 7, pp. 601–609, 1997.
[26] K. Behbehani and R. R. Cross, "A controller for regulation of mean arterial blood pressure using optimum nitroprusside infusion rate," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 6, pp. 513–521, 1991.
[27] C. Yu, R. J. Roy, H. Kaufman, and B. Bequette, "Multiple-model adaptive predictive control of mean arterial pressure and cardiac output," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 8, pp. 765–778, 1992.
[28] C. M. Held and R. J. Roy, "Multiple drug hemodynamic control by means of a supervisory-fuzzy rule-based adaptive control system: Validation on a model," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 4, pp. 371–385, 1995.
[29] E. I. Eger, *Anesthetic Uptake and Action*. Baltimore, MD: Williams and Wilkins, 1997.
[30] I. J. Rampil, "A primer for EEG signal processing in anesthesia," *Anesthesiology*, vol. 89, no. 4, pp. 980–1002, 1998.
[31] J. A. Ross, R. T. Wloch, D. C. White, and D. W. Hawes, "Servo-controlled closed-circuit anaesthesia. A method for the automatic control anaesthesia produced by a volatile agent in oxygen," *Brit. J. Anaesth.*, vol. 55, no. 11, pp. 229–230, 1982.
[32] R. G. Ritchie, E. A. Ernst, B. L. Pate, J. D. Pearson, and L. C. Shepherd, "Closed loop control of an anesthesia delivery system: Development and animal testing," *IEEE Trans. Biomed. Eng.*, vol. 34, pp. 437–443, 1987.
[33] R. Vishnoi and R. J. Roy, "Adaptive control of closed-circuit anesthesia," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 1, pp. 39–47, 1991.
[34] G. I. Jee and R. J. Roy, "Adaptive control of multiplexed closed-circuit anesthesia," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 10, pp. 1071–1080, 1992.
[35] R. G. Bickford, "Automatic electroencephalographic control of anesthesia (servo-anesthesia)," *Electroencephalographic Clin. Neurophysiol.*, vol. 3, pp. 83–86, 1951.
[36] H. Schwilden, J. Schuttler, and H. Stoeckel, "Closed-loop feedback control of methohexital anesthesia by quantitative EEG analysis in humans," *Anesthesiology*, vol. 67, pp. 341–347, 1987.
[37] H. Schwilden, H. Stoeckel, and J. Schuttler, "Closed-loop feedback control of propofol anesthesia by quantitative EEG analysis in humans," *Brit. J. Anaesth.*, vol. 62, pp. 290–296, 1989.
[38] P. S. Sebel, E. Lang, I. J. Rampil, P. White, R. C. M. Jopling, N. T. Smith, P. S. Glass, and P. Manberg, "A multicenter study of bispectral electroencephalogram analysis for monitoring anesthetic effect," *Anesth. Analg.*, vol. 84, no. 4, pp. 891–899, 1997.
[39] P. S. Glass, M. Bloom, L. Kearse, C. Rosow, P. Sebel, and P. Manberg, "Bispectral analysis measures sedation and memory effects of propofol, midazolam, isoflurane, and alfentanil in normal volunteers," *Anesthesiology*, vol. 86, pp. 836–847, 1997.
[40] M. Struys, T. De Smet, L. Versichelen, S. Van de Vilde, R. Van den Broecke, and E. Mortier, "Comparison of closed-loop controlled administration of propofol using BIS as the controlled variable versus "standard practice" controlled administration," *Anesthesiology*, vol. 95, pp. 6–17, 2001.
[41] P. S. A. Glass and I. J. Rampil, "Automated anesthesia: Fact or fantasy?" *Anesthesiology*, vol. 95, pp. 1–2, 2001.
[42] A. R. Absalom, N. Sutcliffe, and G. N. Kenny, "Closed-loop control of anesthesia using bispectral index: Performance assessment in patients undergoing major orthopedic surgery under combined general and regional anesthesia," *Anesthesiology*, vol. 96, no. 1, pp. 67–73, 2002.
[43] G. N. Kenny and H. Mantzardis, "Closed-loop control of propofol anaesthesia," *Brit. J. Anaesth.*, vol. 83, no. 2, pp. 223–228, 1999.
[44] D. A. Linkens, M. F. Abbod, and J. E. Peacock, "Clinical implementation of advanced control in anaesthesia," *Trans. Inst. Meas. Control*, vol. 22, pp. 303–330, 2000.
[45] A. Gentilini, M. Rossoni-Gerosa, C. W. Frei, R. Wymann, M. Morari, A. M. Zbinden, and T. W. Schnider, "Modeling and closed-loop control of hypnosis by means of bispectral index (BIS) with isoflurane," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 874–889, 2001.
[46] W. M. Haddad, T. Hayakawa, and J. M. Bailey, "Adaptive control for nonnegative and compartmental dynamical systems with applications to general anesthesia," *Int. J. Adapt. Control Signal Process.*, vol. 17, pp. 209–235, 2003.
[47] ——, "Nonlinear adaptive control for intensive care unit sedation and operating room hypnosis," in *Proc. Amer. Contr. Conf.*, Denver, CO, June 2003, pp. 1808–1813.
[48] T. Hayakawa, W. M. Haddad, J. M. Bailey, and N. Hovakimyan, "Passivity-based neural network adaptive output feedback control for nonlinear nonnegative dynamical systems," in *Proc. IEEE Conf. Dec. Contr.*, Maui, HI, December 2003. Also in *IEEE Trans. Neural Networks*, submitted.
[49] W. M. Haddad, T. Hayakawa, and J. M. Bailey, "Adaptive control for nonlinear compartmental dynamical systems with applications to clinical pharmacology," *Automatica*, to appear.
[50] T. Hayakawa, W. M. Haddad, N. Hovakimyan, and J. M. Bailey, "Neural network adaptive dynamic output feedback control for nonlinear nonnegative systems using tapped delay memory units," in *Proc. Amer. Contr. Conf.*, Boston, MA, July 2004.
[51] W. M. Haddad, V. Chellaboina, and E. August, "Stability and dissipativity theory for nonnegative dynamical systems: A thermodynamic framework for biological and physiological systems," in *Proc. IEEE Conf. Dec. Contr.*, Orlando, FL, December 2001, pp. 442–458.
[52] W. M. Haddad and V. Chellaboina, "Stability and dissipativity theory for nonnegative and compartmental dynamical systems with time delay," in *Advances in Time-Delay Systems*, S.-I. Niculescu and K. Gu, Eds. Springer, pp. 423-437, 2004.