Submitted to *Decision Analysis*
manuscript (Please, provide the mansucript number!)

A Bayesian Decision-Theoretic Dose Finding Trial

Peter Müller
M.D. Anderson Cancer Center, Houston, USA, pmueller@mdanderson.org, odin.mdacc.tmc.edu/∼pm/

Don A. Berry
M.D. Anderson Cancer Center, Houston, USA, dberry@mdanderson.org

Andrew P. Grieve
Department of Public Health Sciences, King's College, London, UK, andy.p.grieve@pfizer.com

Michael Krams
Pfizer Global Research & Development, Groton, CT, USA, michael.krams@pfizer.com

We describe the use of a successful combination of Bayesian inference and decision theory in a clinical trial design. The trial involves three important decisions, adaptive dose allocation, optimal stopping of the trial, and the optimal terminal decision upon stopping. For all three decisions we use a formal Bayesian decision-theoretic approach.

The application demonstrates how Bayesian posterior inference and decision-theoretic approaches combine to provide a coherent solution in a complex application. The main challenges are the need for a flexible probability model for the unknown dose-response curve, a delayed response, the sequential nature of the stopping decision, and the complex considerations involved in the terminal decision.

The main methodological features of the proposed solution are the use of decision theory to achieve optimal learning about the unknown dose-response curve, an innovative grid-based approximation method to implement backward induction for the sequential stopping decision, and a utility function for the terminal decision that is based on a posterior predictive description of a future confirmatory trial.

1. Introduction

The planning of a clinical trial provides an ideal setup for the innovative combination of decision-theoretic methods and statistical inference in often complex probability models. At the beginning of any clinical trial the investigator should ask several important questions. Why am I designing this clinical trial; what do I hope to accomplish? What are the possible rewards? What are the possible negatives? What do I know that will help me in setting up the design? What responses are likely to be reported? The first questions are all about actions and utilities, about making decisions and evaluating the consequences. The last two questions are about unknown quantities and probabilities. The relevant unknown quantities include the future data that will eventually be observed, but are unknown at the time of planning the trial. Other quantities are unknown parameters that are never observed. Uncertainties about these quantities are best described by defining appropriate probability models. Probability models that are defined on observable data as

well as parameters are known as Bayesian models. Augmenting the Bayesian model for data and parameters with a formal description of the desired decision and consequences leads to a Bayesian decision problem. Marrying probability with utility is the essence of statistical inference, the yīn yáng of statistics.

The basic ingredients of a Bayesian decision-theoretic setup are an action space $\mathcal{A}$ of possible decisions $d \in \mathcal{A}$, a probability model $p(\theta, y)$ for all relevant random variables, including parameters θ and future data y, and a utility function $u(d, \theta, y)$ that quantifies relative preferences for hypothetical future outcomes y and assumed parameter values θ under alternative decisions d. The probability model is conveniently factored into a prior probability model $p(\theta)$ and a sampling model $p(y \mid \theta)$. The probability model could be indexed by the decision. If this is the case we write $p_d(\theta, y)$.

It can be argued (Robert 2001) that a rational decision maker should choose an action in $\mathcal{A}$ to maximize the expectation of u. The expectation is with respect to p, conditioning on all data observed at the time of decision making, and marginalizing over all parameters and all future data. In the absence of observed data, the optimal decision is formally described as

$$d^* = \arg\max_d \int u(d, \theta, y)\, p(\theta, y)\, \mathrm{d}\theta\, \mathrm{d}y.$$

The integral is known as expected utility, $U(d) = \int u(d, \theta, y)\, p(\theta, y)\, \mathrm{d}\theta\, \mathrm{d}y$. This would be the relevant setup, for example, when a (fixed) sample size decision is carried out before a trial is initiated. If some data is known at the time of decision making, for example, results from a related historical study, the integrating measure changes to $p(\theta, y \mid y^o)$, conditioning on the available information. When we wish to highlight the dependence on the historical data we write $U(d, y^o)$. The optimization for d^* is still carried out with respect to d only, of course, $d^* = \arg\max_d U(d, y^o)$. The decision maker can not change the historical data y^o.

Often decisions are made in stages, with additional data observed between the decisions. This complication often arises in clinical trials, for example when dose allocation decisions are made after each cohort, or interim analyses are planned to allow early stopping of the trial when the experimental therapy is clearly superior or clearly inferior to placebo. A clinical trial with one interim look might consist of the following sequence of decisions and data. First we decide the initial sample size d_1. After observing the responses y_1 for the first d_1 patients, we make a decision about the stage two sample size, d_2. Finally we observe the responses y_2 of the d_2 patients in the second stage. At the time of the second decision d_2, the responses y_1, and the earlier decision d_1,

are known and conditioned upon, $d_2^*(d_1, y_1) = \arg\max \int u(d, \theta, y)\, p(\theta, y_2 \mid y_1)\, \mathrm{d}\theta\, \mathrm{d}y_2$. The optimal solution for d_2 is substituted when evaluating the utility for the first stage decision

$$d_1^* = \arg\max_{d_1} \int u(d_1, d_2^*(d_1, y_1), \theta, y)\, p(\theta, y)\, \mathrm{d}\theta\, \mathrm{d}y = \arg\max_{d_1} \int_{y_1} \max_{d_2} \left\{ \int_{\theta, y_2} u(d_1, d_2, \theta, y)\, p(\theta, y_2 \mid y_1) \mathrm{d}\theta\, \mathrm{d}y_2 \right\} p(y_1)\, \mathrm{d}y_1. \quad (1)$$

Problems with an alternating sequence of decisions and observations as in (1) are known as sequential decision problems. In general, the solution of sequential decision problems requires computationally intensive backward induction. See, for example, DeGroot (1970) for a description. One important feature of the solution (1) is that strict adherence to the Bayesian decision-theoretic framework requires that all decisions be made with respect to the same utility function. Different decisions only require different choices of conditioning and marginalizations.

The need for formal decision-theoretic approaches for clinical trial planning was highlighted as early as Anscombe (1963). More recent general discussions appear in Berry (1993, 2004 and 2006), and Spiegelhalter et al. (2004). In practice, however, unabashedly Bayesian decision-theoretic approaches in clinical trial design are rare. Many Bayesian clinical trial designs use an approach that combines posterior inference for the probability model with reasonable, but ad-hoc rules for the desired decisions. A typical example is Thall et al. (1995) and Yin et al. (2006) who proceed by evaluating posterior probabilities of clinically meaningful events. When these probabilities cross predefined boundaries certain decisions are indicated. The boundaries are set at clinically meaningful values. The design is validated by evaluating frequentist properties. If necessary, boundaries are adjusted to achieve desired frequentist properties, while keeping the value of the boundary within a clinically meaningful range. Such decision rules are referred to as partial Bayes (Berry and Stangl, 1996) or proper Bayes (Spiegelhalter et al., 2004). They stop short of a decision theoretic framework with utility functions and expected utility maximization.

The following challenges currently prevent Bayesian decision-theoretic approaches from being implemented more widely. But all of them can be overcome in principle and can provide unique opportunities. Prior elicitation for complex probability models, like the model described in this article, poses a difficult challenge. Rather than seeing this challenge as a problem, we argue that it provides an opportunity for early team-building and discussion of critical issues pertinent to the drug development program. Another major challenge is the choice of the utility function. We argue that careful consideration of the appropriate utility function provides a great opportunity to involve health economic aspects early on and move the perspective from thinking about an isolated project, to considering a longer term program, portfolio, or even society at large. A third important

challenge involves computational difficulties in implementing expected utility maximization. Carlin et al. (1998) and Brockwell and Kadane (2003) propose simuluation-based implementations of sequential designs. Christen et al. (2004) and Berry and Kadane (1997) justify randomization in a Bayesian decision-theoretic setup by appropriate considerations of the utility function. Stallard (2003) and Rossell et al. (2005) discuss decision-theoretic Bayesian designs that consider multiple studies simultaneously, i.e., designs for a drug development process involving multiple studies.

In this article, we describe a Bayesian decision-theoretic approach to an adaptive design for a dose-finding trial. Central to the decision-theoretic approach is a utility function that formalizes the relative benefits of alternative dose assignments and stopping decisions. For practical reasons we choose to deal separately with the decisions related to dose assignment and to sequential stopping. We first consider the dose allocation problem. We propose a utility function related to learning about the unknown dose-response curve. Specifically, we use the posterior variance of some key summary statistic of the dose-response curve. The suggested dose for the next patient in the trial is the dose resulting in maximum expected utility. After each patient cohort we consider sequential stopping. We decide whether or not the trial should continue to accrue patients. For this decision we use a utility function that is built on a stylized description of a possibly following confirmatory trial. The utility function includes a large reward for a statistically significant outcome at the end of the future confirmatory trial, and a sampling cost for the future trial. The problem of choosing an appropriate reward for an eventual significant outcome illustrates the challenge of utility elicitation. For the study discussed in this article we fixed the reward based on elicitation with marketing experts in the company.

The discussion in this article is motivated by a phase II dose-response finding study as conducted in the context of a drug development program of a neuroprotectant for acute stroke. Background of the motivating clinical trial, implementation details and simulation studies to evaluate the performance of the proposed approach in that study are discussed in Berry et al. (2001). In this article we focus on the underlying decision-theoretic model.

In Section 2, we describe the motivating application. Section 3 introduces the underlying probability model. In Section 4, we develop an approach for optimal dose allocation. In Section 5, we solve the optimal stopping problem. Section 6 concludes with a final discussion.

2. Data

The motivating case study is a phase II dose-response finding study as conducted in the context of a drug development program of a neuroprotectant for acute stroke. Clinical aspects of the

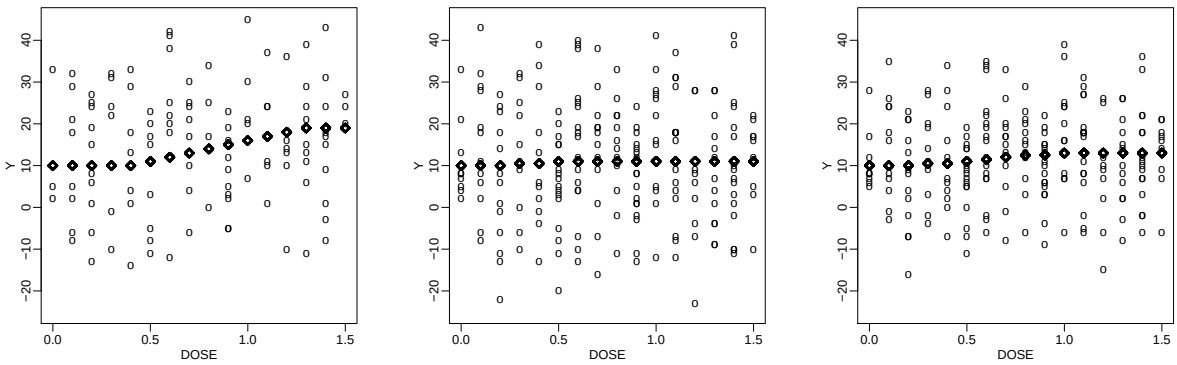


Figure 1 **Typical dose-response curves (diamonds) and simulated patient responses (circles). Note the large signal to noise ratio.**

study are described in Krams et al. (2003) and Grieve and Krams (2005). Zhang et al. (2003) discuss preclinical background. The main interest in the study is on the unknown dose-response curve $f(z) = E(y|z)$ which gives the mean response y at a given dose z. The recorded response is the change of Scandinavian Stroke Scale (SSS; Scandinavian Stroke Study Group, 1985) score relative to baseline. Higher scores of y are desirable. Thus an effective drug will show an increasing dose-response curve $f(z)$. Figure 1 shows hypothetical dose-response curves $f(z)$ together with responses y of simulated patients in a simulation experiment. The levels of measurement variances are plausible for the observed response variable.

Inference is complicated by the fact that final measurements of the desired clinical effect are delayed for up to $M = 13$ weeks after treatment. But early responses are recorded in weeks 1 through $M - 1$. We facilitate the use of early responses by augmenting the probability model for the dose-response curve with a longitudinal data model for repeated measurements in weeks 0 through M for each patient. The longitudinal data model will allow us to impute missing final week M measurements as latent data. We will denote with y_{ij} the measurement on patient i in week j, $j = 0, \ldots, M$. We will use y_i without a second index to denote the final outcome y_{iM}. An important covariate is y_{i0}, the baseline measurement on the clinical outcome of interest at the time of admitting the patient into the study. We will use x_i to generically denote the covariate vector. In the current implementation the only covariate considered is baseline score, i.e., $x_i = y_{i0}$. Other covariates that might be useful in a stroke trial include time delay between the onset of stroke and the administration of therapy, stroke severity, and indicators for ischemic versus hemorrhagic stroke and cortical versus subcortical stroke.

3. Probability Model

The choice of the probability model for $f(z)$ is guided by the following considerations. First, we need a model that allows analytic posterior inference to facilitate efficient computation of expected utilities when solving the decision problem. Second, we want a flexible model which includes *a priori* a wide range of dose-response curves. Although an increasing curve with asymptotes is *a priori* likely, the model should allow for possible lack of monotonicity and other irregular features. Lees et al. (2003) provide evidence for non-monotonic response-curves of the experimental drug. Based on these considerations we chose a normal dynamic linear model (NDLM). See, for example, West and Harrison (1997) for a formal definition and discussion of NDLM's. Denote with Z_j, $j = 1, \ldots, J$, the range of allowable doses, and with $\theta_j = f(Z_j)$, $j = 1, \ldots, J$, the vector of mean responses at the allowable doses. The underlying idea is to formalize a model which locally, for z close to Z_j, fits a straight line $y = \theta_j + (z - Z_j)\delta_j$, with level θ_j and slope δ_j. Let Y_{jk}, $k = 1, \ldots, \nu_j$, denote the k-th response observed at dose Z_j, i.e., $Y_j = (Y_{jk},\, k = 1, \ldots, \nu_j)$ is the vector of responses y_i of all patients with assigned dose $z_i = Z_j$. Note the notational convention of using upper case symbols for quantities Z_j and Y_j indexed by doses, and lower case y_i and z_i for quantities indexed by patients. We use an NDLM that assumes $Y_{jk} = \theta_j + \epsilon_{jk}$ with the following prior probability model for $\theta = (\theta_1, \ldots, \theta_J)$. The prior specifies that the coefficients change between doses by extrapolating the line and adding an evolution error, $\theta_{j+1} = \theta_j + \delta_j(Z_j - Z_{j-1}) + e_{j1}$ and $\delta_{j+1} = \delta_j + e_{j2}$. The residual ϵ_{jk} and the evolution error $e_j = (e_{j1}, e_{j2})$ are assumed normal distributed, $\epsilon_{jk} \sim N(0, V\sigma^2)$ and $e_j \sim N(0, W)$. The pairs (θ_j, δ_j) are known as the state parameters of the NDLM. The model includes the hyperparameters σ^2, V, W and (θ_0, δ_0). We fix V and W and assume conjugate hyperpriors for θ_0, δ_0 and σ^2. The main attraction of the NDLM is the availability of a straightforward recursive algorithm to compute the posterior distribution $p(\theta \mid Y_1, \ldots, Y_J)$ and any other desired posterior inference. The algorithm, known as Forward Filtering Backward Sampling (FFBS), is described in Frühwirt-Schnatter (1994) and Carter and Kohn (1994). It can be shown that the posterior distributions $p(\theta_j \mid Y_1, \ldots, y_J)$ and the posterior predictive distributions $p(Y_{jk} \mid Y_{-jk})$ are normal distributions. Here Y_{-jk} denotes the data with Y_{jk} removed. West and Harrison (1997) give recursive equations to exactly compute the moments of these posterior distributions.

A minor shortcoming of the NDLM in the present application is that the prior specification does not naturally allow one to fix arbitrary desired prior moments for $\theta_j = f(Z_j)$. Only $E(\theta_0)$ and $Var(\theta_0)$ are fixed. Prior expectation and variance for θ_j, $j > 0$, are then implied by the evolution equation. To increase the number of prior parameters and allow for essentially arbitrary prior moments $E(\theta_j)$ and $Var(\theta_j)$ we augment the model by introducing dummy observations $\tilde{Y}_j$,

$j = 0, \ldots, J$, with associated observation variance $\tilde{\sigma}_j^2$. When going through the FFBS scheme for posterior inference in the NDLM we proceed then as if $\tilde{Y}_j$ were data sampled from $\tilde{Y}_j \sim N(\theta_j, \tilde{\sigma}_j^2)$. By appropriate choice of $\tilde{Y}_j$ and $\tilde{\sigma}_j^2$ we can achieve any prior moments for θ_j, subject only to technical constraints (for example, the marginal prior variances $Var(\theta_j)$ can not be larger than those implied without the dummy observations).

For many patients with missing final score y_i we have earlier scores available, y_{ij}, $j < M$. From historical data we expect y_{ij}, even for j as early as week 4, to be a good predictor of y_i. We extend the probability model with a longitudinal data model for $(y_{ij}, j = 1, \ldots, M)$ to allow formal imputation. We assume

$$y_{ij}|y_i \sim N(m_j + a_j y_i, s_j^2).$$

Conditional on y_{ij} we assume that y_i and any earlier response $y_{ij'}$, $j' < j$, are conditionally independent. This formalizes the notion that knowing the latest available score we should ignore earlier responses. This implies

$$p(y_{ij}|y_{i,j+1}, \ldots, y_{iM}) = p(y_{ij}|y_{i,j+1}) = N(m'_j + a'_j y_{i,j+1}, s'^2_j), \tag{2}$$

with parameters m', a' and s'^2 which are easily derived from m, a and s^2.

In the following discussion we will use $\theta = (\theta_1, \delta_1, \ldots, \theta_J, \alpha_J, \sigma^2, m_1, a_1, s_1^2, \ldots, m_M, a_M, s_M^2)$ to denote the vector of all unknown model parameters.

We have introduced the NDLM and (2) as a flexible and computationally convenient probability model for the curve $(\theta_1, \ldots, \theta_J)$ and the observed outcomes y_{ij}. For the upcoming discussion, the only important feature of the probability model is the availability of computationally easy and fast algorithms to evaluate the posterior mean curve and related posterior probabilities. The proposed approach remains equally valid for any other probability model. Examples of alternative probability models that might be appropriate for other applications are non-linear regression models with a shifted and scaled logistic mean function, spline models, and models with non-normal sampling distributions.

4. Adaptive Dose Allocation

Based on the assumed probability model, we can now consider the problem of computing an optimal dose for the next patients. Let N denote the number of currently accrued patients, and let K denote the maximum number of patients who are recruited into the trial on one day. Let θ denote the vector of all unknown model parameters. We shall compute the optimal doses to be assigned to the next K patients, $i = N + 1, \ldots, N + K$.

The idea is to choose a dose Z_j which in expectation will let us learn most about the dose-response curve $f(z,\theta)$, where we write $f(z,\theta)$ instead of $f(z)$ to indicate that the dose-response curve is parameterized by θ. Of course a specific implementation requires a formal definition of this vague notion. How should we formalize learning, and with respect to which probability model is the expectation to be taken?

We formalize learning as reducing posterior variance in some key parameters of the dose-response curve. In the current implementation we use as key parameter $Z95_\theta$, the ED95 of the dose-response curve. Here the ED95 is defined as follows. For given parameters θ, let z^o_θ denote the dose with maximum response $f(z,\theta)$, i.e., $f(z^o_\theta,\theta) = \max\{f(Z_j,\theta),\, j=0,\ldots,J\}$. The ED95 is defined as

$$Z95_\theta = \min\{Z_j :\, f(Z_j,\theta) \geq f(0,\theta) + 0.95\,[f(z^o_\theta,\theta) - f(0,\theta)]\}.$$

In words, the ED95 dose $Z95_\theta$ is the minimum dose for which the dose response curve $f(z,\theta)$ achieves at least 95% of the maximum possible improvement.

For a formal description of the optimal dose finding, let $\tilde{y} = (\tilde{y}_1,\ldots,\tilde{y}_K) = (y_{N+1},\ldots,y_{N+K}) =$, $k = 1,\ldots,K$ denote the (still unknown) responses of the future patients with covariates $\tilde{x} = (\tilde{x}_1,\ldots,\tilde{x}_K) = (x_{N+1},\ldots,x_{N+K})$. Let $\tilde{z}_k$ denote the – to be determined – dose for the future patient $N+k$. Let D denote the already observed data on previous patients, including early responses for patients with still missing final response, and let $\tilde{D}$ denote the missing final responses for previous patients. Let $g(\theta)$ denote the key parameter, for example $g(\theta) = Z95_\theta$. We consider posterior variance of $g(\theta)$ to define a utility function for the choice of the optimal dose $\tilde{z}_k$:

$$u_k(D,\tilde{D},\tilde{y}_1,\tilde{x}_1,\ldots,\tilde{y}_k,\tilde{x}_k) = -\int [g(\theta)-\bar{g}]^2\, p(\theta \mid D,\tilde{D},\tilde{y}_1,\ldots,\tilde{y}_k)d\theta, \tag{3}$$

where $\bar{g} = \int g(\theta)\, dp(\theta \mid D,\tilde{D},\tilde{y}_1,\ldots,\tilde{y}_k)$ denotes the posterior mean. Note that u_k does not (yet) depend on the action $\tilde{z}_k$. The dependence on $\tilde{z}_k$ will be introduced when we take the expectation with respect to the response $\tilde{y}_k$. Matching the notation in (3) with the notation for the generic Bayesian decision problem in the introduction, note that $(D,\tilde{D},\tilde{y}_1,\tilde{x}_1,\ldots,\tilde{y}_k,\tilde{x}_k)$ are all observable random quantities, and u_k does in this case not include the parameters θ or the action $\tilde{z}_k$. In other words, the definition of u_k already includes the integration with respect to θ.

Of course we have to choose dosage for the future patients before observing $(\tilde{x},\tilde{y},\tilde{D})$. Thus choice of $\tilde{z}_k$ is based on averaging $u_k(\cdot)$ with respect to the relevant posterior predictive distribution on $(\tilde{D},\tilde{y}_1,\ldots,\tilde{y}_k)$, and with respect to a distribution $p(\tilde{x})$ on $\tilde{x}_1,\ldots,\tilde{x}_{k-1}$, derived from historical data. For $\tilde{x}_k$ we substitute covariates x_0 of a "typical" patient, i.e., we compute the optimal dose allocation for a typical patient. Reporting a patient-specific recommended dose that depends on

the baseline covariate $\tilde{x}_k$ would create excessive logistical difficulties for the administration of this multi-center trial.

$$\begin{aligned} U_k(Z_j) &= \int u_k(D, \tilde{D}, \tilde{y}_1, \ldots, \tilde{y}_k, \tilde{x}_1, \ldots, \tilde{x}_k)\, p(\tilde{D}, \tilde{y}_1, \ldots, \tilde{y}_k \mid \tilde{z}_k = Z_j, \tilde{x}, D)\, d\tilde{D}\, p(\tilde{x})\, d\tilde{x}\, d\tilde{y} \\ &= \int\int\int u_k(D, \tilde{D}, \tilde{y}_1, \tilde{x}_1, \ldots, \tilde{y}_k, \tilde{x}_k)\, p(\tilde{y}_1, \ldots, \tilde{y}_k \mid \theta, \tilde{z}_k = Z_j, \tilde{x}, D)\, p(\tilde{x}_1, \ldots, \tilde{x}_{k-1})\, p(\theta \mid D) \\ &\qquad\qquad d\tilde{x}_1 \ \ldots\ d\tilde{x}_{k-1}\, d\tilde{y}_1 \ \ldots\ d\tilde{y}_k\, d\theta. \quad (4) \end{aligned}$$

Here $U_k(z_j)$ is the expected utility of decision $\tilde{z}_k = Z_j$ for a future patient, i.e., $U_k(z_j)$ expresses how much deciding on dose Z_j is worth to us. See below for the distribution $p(\tilde{x})$ and the assumed doses for $\tilde{z}_1, \ldots, \tilde{z}_{k-1}$. The solution to the optimal dose problem is then formalized as

$$\tilde{z}_k = arg\ \max_{Z_j} U_k(Z_j).$$

In words, recommend the dose which in expectation maximizes the utility defined by expected posterior variance on the key parameters $g(\theta)$. The expectation is with respect to responses y_{N+k} and covariates x_{N+h} of future patients $h = 1, \ldots, k$ (except $\tilde{x}_k$), and with respect to still missing final responses y_i of current patients, $i = 1, \ldots, N$.

For a practical implementation of the proposed optimal design approach it is critical that the posterior variance integral in (3) and the expected utility integral (4) be available for efficient evaluation, analytically or by numerical integration. One of the reasons for choosing the NDLM model was that the posterior variance of the response at the ED95, the choice for $g(\theta)$ in our implementation, can be evaluated analytically. See the appendix for a description of a detailed algorithm based on the commonly used recursive equations to compute posterior moments for the state parameters.

However, the expected utility integral (4) is not analytically tractable. Instead we used independent Monte Carlo simulation. First we generate by computer simulation posterior samples from $p(\theta \mid D)$, using the FFBS algorithm; and $\tilde{x}_h \sim p(x_h)$, $h = 1, \ldots, k-1$, using the empirical distribution from a historical data base of patients treated for the same medical condition. For each patient $h = 1, \ldots, k-1$ we then substitute the optimal dose $\tilde{z}_h$ that we found earlier by maximizing U_h, $h = 1, \ldots, k-1$. Based on these imputed values for θ and $\tilde{x}_h$ and the assigned doses $\tilde{z}_h$, $h = 1, \ldots, k-1$, we can generate the missing data values $\tilde{D}$ and $\tilde{y}_1, \ldots, \tilde{y}_{k-1}$. Given these simulations we evaluate the posterior variance $u_k(D, \tilde{D}, \tilde{y}_1, \tilde{x}_1, \ldots, \tilde{y}_k, \tilde{x}_k)$. Repeating the same simulation many times, say M times, we can replace the integral (4) by a sample average over the evaluated posterior variances.

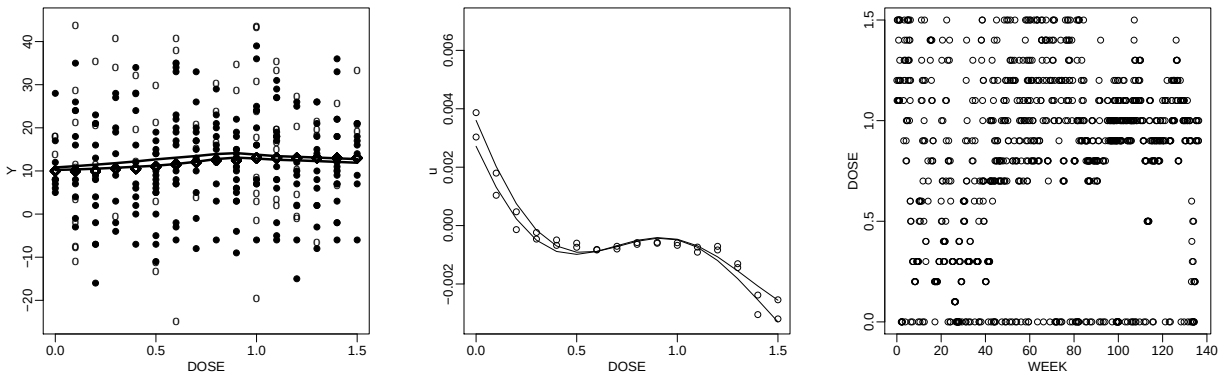
Note that the criterion (4) is myopic in the sense that we only think ahead for the next K patients. Also, the choice of the posterior variance of $Z95_\theta$ as a formal definition of learning about the curve is arbitrary. Possible alternatives are to consider posterior variance on more than one key parameter, or some measure of information gain, like Kulback-Liebler divergence of prior and posterior.

Figure 2a shows the data at a given time t of a simulated realization of the study. The solid bullets indicate observed final responses y_i. The open circles show imputed final responses for recently enrolled patients who have not yet reported the final week 13 response. The diamonds show the assumed simulation truth for the dose-response curve. The smooth curve shows the estimated fit, using the NDLM model. The assumed true dose-response curve rises from $f(0) = 10$ to $f(1.5) = 13$. The small signal to noise ratio is realistic for the reported response. Figure 2b shows the estimated expected utility curve $U(Z_j)$, defined in (4), for the first patient of the next cohort ($k = 1$). The circles show Monte Carlo estimates of $U(Z_j)$ as a function of dose. We carried out two sets of Monte Carlo simulations to evaluate numerical uncertainty. Both are shown in the figure, and are connected by a smooth curve fit through the approximated expected utilities. The optimal dose allocation for the next patient is at placebo, $\tilde{z}_k = 0$. The low utility for high doses reflects the fact that the model has already learned about the flat nature of the dose-response curve in the high dose region. Compare with the estimate shown in panel (a). Figure 2c shows a trace of assigned doses over the course of the entire simulated trial. Note how the system correctly narrows in around $Z_j = 1.0$. The ED95 in the true dose-response curve that was used to simulate the data is at 1.0.

The utility function (3) focuses on the learning about the curve. In other applications, alternative and additional goals might be of interest. For example, one might want to add a term related to the outcomes $\tilde{y}_1, \ldots, \tilde{y}_k$. Technically, any expression that can be evaluated for an assumed set of decisions, future outcomes and hypothetical parameter values could be used. We recommend that the utility function should include a reward for learning about the unknown curve. However, the focus need not be exclusively on this goal. Other terms can be added to the utilty function without changing the proposed approach.

5. Optimal Stopping

At periodic time intervals, say once a week, we consider the decision of stopping the trial. Let D_t denote all data available at time t. This includes the observed final responses for patients that were enrolled more than M weeks before t, and available early responses y_{ij} for recently recruited patients. At each time we make a decision about stopping the trial, $d_t \in \{0, 1, 2\}$, with $d_t = 0$



(a) Data and truth (b) $U_k(Z_j)$ for $k = 1$ (c) z_i vs. i

Figure 2 **Panel (a) shows the data at a given time in a simulated run of the study. The solid bullets show observed final responses. The open circles are imputed final responses for patients recruited within the last 13 weeks only. The diamonds show the assumed simulation truth. The smooth curves shows the estimated dose-response curve using the NDLM and a spline smoother. The two curves are almost indistinguishable. Panel (b) shows the expected utility function $U(Z_j)$ for the dose allocation for the first patient, $k = 1$, in the next cohort. The circles are Monte Carlo estimates of the expected utility integral (4). The line is a smooth curve fit through the Monte Carlo simulations. Two sets of Monte Carlo simulations were used to evaluate numerical errors. Panel (c) shows the assigned doses in a simulated complete trial. The figure shows assigned doses $\tilde{z}$ against weeks in a hypothetical trial with patient responses simulated using an assumed nominal dose-response curve.**

indicating termination with the recommendation to abandon the drug (stopping for futility), $d_t = 1$ indicating continuation, and $d_t = 2$ indicating termination with the recommendation to initiate a confirmatory phase III study that compares the drug at the recommended dose against placebo (stopping for efficacy). The recommended dose is the currently estimated ED95, $E(Z95_\theta \mid D_t)$. We assume that the confirmatory trial is set up as a three-arm trial comparing placebo ($z = 0$), recommended dose $z^\star = E(Z95_\theta \mid D_t)$, and $z^{\star\star} = E(Z50_\theta \mid D_t)$. The third arm is included as an alternative dose. The data D_t to compute the posterior expectation is the data at the time of stopping, i.e., when $d_t = 2$. Let $df^\star = f(z^\star, \theta) - f(0, \theta)$ denote the advantage over placebo at the recommended dose, and let (m_t, s_t) denote the posterior mean and standard deviation of $df^\star$ (again, conditional on D_t). We assume that at the conclusion of the pivotal trial a hypothesis test is carried out to compare $H_0 : df^\star = 0$ versus H_1: $df^\star > 0$, using an approximately normal distributed z-statistic to compare the two treatment arms $z = 0$ and $z = z^\star$. Let D_P denote the observed responses in the future pivotal trial, let n_P denote the number of patients enrolled in the future pivotal trial. We assume that the sample size n_P for the pivotal trial is chosen to achieve 90% Bayesian power. Bayesian power is calculated using a normal approximation to the posterior predictive distribution

$p(df^\star \mid D_t)$. See, for example, Spiegelhalter et al. (2004, chapter 6.5.3) for a discussion of Bayesian power. Briefly, it is the probability of rejecting the null hypothesis, taking the expectation with respect to (in this case) the joint posterior predictive distribution of $df^\star$ and D_P given D_t. We write $n_P(D_t)$ to highlight the nature of n_P as a function of the data. Alternative sample size calculations could be used. In particular, one could use traditional (frequentist) power instead of Bayesian power. Also, as usual in sample size arguments, the target of 90% is a reasonable but arbitrary choice. Any alternative threshold could be used without changing the approach, if desired.

The evaluation of alternative actions d_t requires us to compare the sampling cost for a possible future pivotal trial, and weight it against the benefit of a significant outcome of the pivotal trial, weighted by the appropriate probability of such an outcome. If we decide to continue, $d_t = 1$, then we need to consider the cost of one more patient cohort and the optimal decision in the next period. A utility function that formalizes these aspects allows us to formally compare the alternative actions. Let $B(D_P)$ denote the event of finding at the conclusion of the pivotal trial a statistically significant advantage over placebo, and let $m_P = E(df^\star \mid D_t, D_P)$ denote the estimated effect at the recommended dose $z^\star$. Let K denote the expected number of accrued patients in the time interval between t and $t+1$. We use the utility function

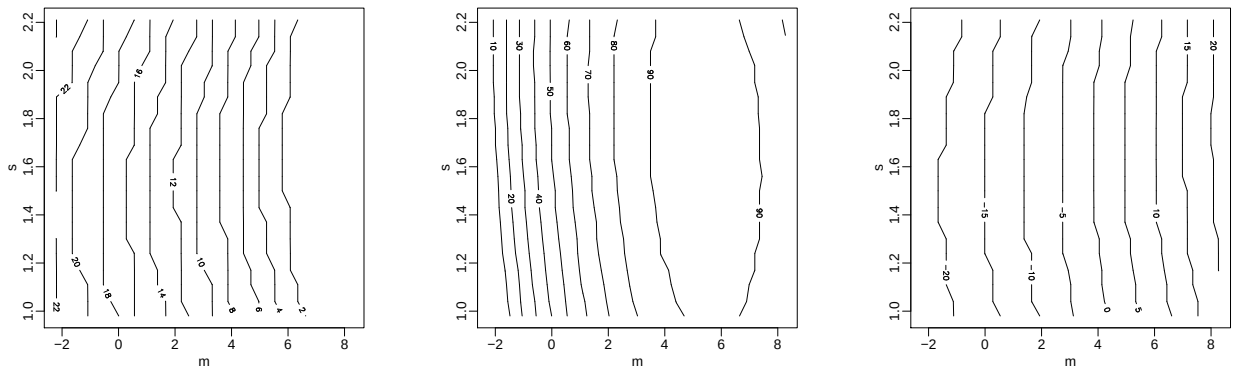
$$u(d_t, D_t, D_P) = \begin{cases} 0 & \text{if } d_t = 0 \text{ (abandon)} \\ -c_1 K + E\{\max_d E[u(d_{t+1} = d, D_{t+1}, D_P)]\} & \text{if } d_t = 1 \text{ (continue)} \\ -n_P(D_t)c_1 + c_2 I\{B(D_P)\}m_P - c_0 & \text{if } d_t = 2 \text{ (phase III)} \end{cases} \tag{5}$$

In words, if we decide to terminate the trial and abandon the experimental therapy, no further loss or payoff occurs. If we decide to stop and move to the pivotal trial, we incur a fixed cost c_0 to set up the study, and the sampling cost (c_1 per patient) for the pivotal trial. A major payoff ($c_2 m_P$) is awarded if the trial concludes with a statistically significant effect of the new therapy. The payoff is proportional to the estimated effect, with c_2 specified as payoff per point advantage over placebo. If we decide to continue, then we pay the sampling cost for the next cohort of patients, and get the optimal reward in the next period. The outer expectation, in the second line of (5), is with respect to all patients who are enrolled between time t and $t+1$. The inner expectation is with respect to all other future outcomes that are required in the evaluation of $u(d_{t+1}, D_{t+1}, D_P)$, including in particular the future pivotal trial data.

The evaluation of the utility of continuation, $d_t = 1$, requires the optimal solution for the stopping decision d_{t+1} in the next period. This characterizes a sequential decision problem, as in (1), and greatly complicates the solution of the decision problem. Maximizing the expected utility function to find the optimal decision is a computationally challenging problem. A full solution involves

the use of backward induction. See, for example, DeGroot (1970). Instead we use a simulation-based approximate solution, based on the method proposed in Brockwell and Kadane (2003) and Müller et al. (1999). First we restrict the decision $d_t(D_t)$ to depend on the data only indirectly through (m_t, s_t). Recall that (m_t, s_t) were defined as the posterior moments of $df^\star$. We consider a bivariate grid on (m_t, s_t). Let $U(d_t = d, D_t) = E[u(d_t = d, D_t, D_P) \mid D_t]$ denote the expected utility for decision $d_t = d$. We use one common approximation $U(d_t = d, D_t) \approx U_G(d, m, s)$ for all data D_t with summaries $(m_t, s_t) \approx (m, s)$ with (m, s) being the closest values on the (m_t, s_t) grid. We use Monte Carlo integration to evaluate the expectation $U_G(d, m, s)$. Carlin et al. (1998) refer to this strategy as forward simulation. There remains the problem that the evaluation of $U_G(d = 1, m, s)$ requires to substitute the optimal decision for the next period. We solve this by using an iterative algorithm. We start out with initial values for all fields $U_G(d, m, s)$, and continue to update the currently imputed values $U_G(d, m, s)$ over the entire grid until the number of changes in one iteration is below a certain threshold.

Figures 3 and 4 show some aspects of the described solution of the sequential decision problem. Figure 3a shows n_P, on a grid over (m_t, s_t). In particular, we use one common value n_P for all D_t with (m_t, s_t) within the same grid cell. Panel (b) of the same figure shows $P(B \mid D_t)$, again on a grid over (m_t, s_t). Panel (c) shows $U(d_t = 2, D_t) = E(u(d_t = 2, D_t, D_P) \mid D_t)$ as a linear combination of the two previous figures. Again, we assume a common value $U(d_t = 2, D_t) = U_G(d_t = 2, m, s)$ for all D_t with summaries (m_t, s_t) in the same grid cell. The optimal rule is shown in Figure 4.



(a) $n_P(D_t)$ on (m, s) grid (b) $P(B \mid D_t)$ on (m, s) grid (c) $U_G(2, m, s) \approx U(d_t = 2, D_t)$

Figure 3 **Sample size for a future pivotal study (panel a), posterior predictive probability for a statistically significant outcome at the end of the pivotal study (panel b), and approximate expected utility for $d_t = 2$. All summaries are shown on a grid over (m, s).**

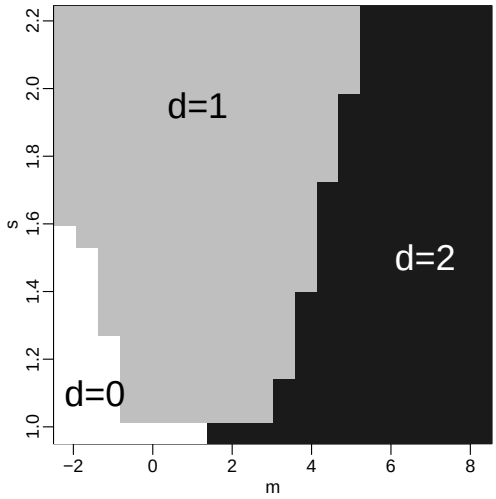


Figure 4 **Optimal rule $d^*(D_t) \approx d^*_G(m_t, s_t)$. Using the grid approximations for $U(d, D_t)$, as shown in Figure 3, the optimal rule is a function of the data D_t only indirectly through (m_t, s_t). Regions in the (m, s) grid with $d^*_G(m, s) = 0$, 1, and 2 are identified in white, gray and black, respectively.**

Similar to (3), the choice of (5) is arbitrary. We believe that the proposed utility is a reasonable stylized description of the drug developer's goals. An alternative choice could, for example, include the treatment success for future patients who will benefit from a successful drug development. Or one could consider minor changes, such as replacing the estimated advantage over placebo $E(df^\star \mid D_t, D_P)$ by the true advantage over placebo, $df^\star$.

6. Discussion

We proposed a Bayesian decision-theoretic approach to a phase II dose-response finding study. For the adaptive dose allocation we use a utility function which formalizes learning about the unknown dose-response curve. We showed how the required computations become practically feasible when using an appropriate probability model for the dose-response curve and, if necessary, for the longitudinal model of delayed responses. For the sequential stopping decision we used a utility function that involves a reward for a statistically significant outcome of a subsequent pivotal trial.

The discussed implementation includes several compromises with a strict decision-theoretic textbook approach. First, we used several utility functions for the different decisions. In a strict implementation of the paradigm of expected utility maximization one would state only one utility function, and all decisions would be made with respect to that utility function. Also, in the discussion of the adaptive dose allocation we ignored the sequential aspect of the problem.

Implementation of the proposed approach involved several, to some level arbitrary, decisions to formalize general concepts. For example, many alternatives exist to the chosen specification for the

utility function (3). Also, the probability model could be modified in several directions without compromising the overall goal of achieving a computationally feasible framework for a decision-theoretic optimal solution to the given design problem. For example, the NDLM could include additional covariates (in the observation equation); we could choose alternative NDLM models which specify different levels of smoothing; the longitudinal data model could include a regression on the baseline score y_{i0}, etc.

We discussed in some detail the two important decisions related to the dose allocation for the next patients and optimal stopping to decide continuation of the trial. Carrying out the trial requires a loop over alternating steps that invoke the two decisions. Practical implementation also needs an elaborate system to make the data available for the continuous updating of posterior and posterior predictive distributions. For multicenter trials this is not a trivial effort. We used an automated fax system in the study underlying the discussion of this paper.

Before the trial is initiated the same system is used in a simulation mode, that is, with simulated patient responses in place of actual data, to evaluate frequentist operating characteristics under selected scenarios. Frequentist operating characteristics are summaries of the system performance under a specific assumed true dose-response curve. Summaries are evaluated as averages over repeated simulations of possible patient outcomes, keeping the assumed fixed truth. Such summaries are traditional criteria to evaluate and critique clinical trial designs.

Bayesian clinical trial design is not a panacea. It only provides a framework for a principled approach to the development of study designs. Actual implementation usually requires extensive consideration and judgement to specify the sampling models, the prior distributions, the utility functions, and finally to find practicable implementations. The rewards for the effort are manifold. A good Bayesian study design will allow early stopping for futility and will indicate when an expansion of the originally planned sample size is called for. The sequential stopping rule presented here is one way of achieving this flexibility. Another important advantage of a model based Bayesian decision-theoretic approach is that there always is a gold standard, namely the Bayes rule that is defined by the expected utility optimization. A practical implementation often calls for approximations and simplifications. But these choices are always guided by the theoretically optimal solution. One of the practically important aspects of Bayesian inference is the ease of hiearchical modeling. Most studies include some notions of hierarchical structure. Examples are different patient subpopulations defined by, e.g., biomarkers, treatment history, family history, or treatment centers; different treatments; different treatment schedules; multiple cycles of treatment;

different phases of drug development, etc. Such hierarchies are easily accomodated in a Bayesian framework, but could pose challenging problems for a traditional design approach.

Acknowledgments

Research was supported by NIH/NCI grant 1 R01 CA075981.

Appendix A: Posterior Inference for $Z95_\theta$

The expected utility integral (4) requires evaluation of $Var(Z95_\theta \mid D)$. Here D generically indicates observed data in the NDLM. We show how posterior inference on $Z95_\theta$ can be derived based on standard FFBS algorithm. Assume first that z^o_θ, the dose with respect to which we define $Z95_\theta$, is fixed. The FFBS provides the joint multivariate normal posterior distribution of the mean dose-response curve $(\theta_0, \ldots, \theta_J) = (f(Z_0, \theta), \ldots, f(Z_J, \theta))$. Note that $Pr(Z_j \geq Z95_\theta \mid D, z^o_\theta) = Pr\{\theta_j \geq f(0,\theta) + 0.95\,[f(z^o_\theta, \theta) - f(0,\theta)] \mid D, z^o_\theta\}$, which is readily available from the FFBS output. Having computed $P_j = Pr(Z_j \geq Z95_\theta \mid D)$ we easily find $Var(Z95_\theta \mid D, z^o_\theta)$. The desired $Var(Z95_\theta \mid D)$ follows from this as $Var(Z95_\theta \mid D) = E\{Var(Z95_\theta \mid D, z^o_\theta) \mid D\} + Var\{E(Z95_\theta \mid D, z^o_\theta) \mid D\}$.

References

Anscombe, F.J. (1963). "Sequential medical trials". *Journal of the American Statistical Association*, 58, 365–383.

Berger, J.O. (1985). *Statistical decision theory and Bayesian analysis*, New York: Springer-Verlag.

Berry, D.A. (1993) A case for Bayesianism in clinical trials (with discussion), *Statistics in Medicine*, 12, 1377-1404.

Berry, D.A. (2004). "Bayesian statistics and the efficiency and ethics of clinical trials," *Statistical Science*, 19, 175-187.

Berry, D.A. (2006). "Bayesian clinical trials," *Nature Reviews Drug Discovery*, 5, 27-36.

Berry, S. M. and Kadane, J. B. (1997). "Optimal Bayesian randomization". *Journal of the Royal Statistical Society–B.* 59(4), 813–819.

Berry, D. A., Müller, P., Grieve, A. P., Smith, M., Parke, T., Blazek, R., Mitchard, N. and Krams, M. (2001). Adaptive Bayesian designs for dose-ranging drug trials. In Gatsonis, C., Kass, R. E., Carlin, B., Carriquiry, A., Gelman, A., Verdinelli, I. and West, M., editors, *Case Studies in Bayesian Statistics, Volume V*, Lecture Notes in Statistics. Springer-Verlag, New York.

Berry, D.A. and Stangl, D.K. (1996) Bayesian methods in health related research. In *Bayesian Biostatistics*, (D.A. Berry and D.K. Stangl, eds.), New-York: Dekker, 3-66.

Brockwell, A. E. and Kadane, J. B. (2003), "A gridding method for Bayesian sequential decision problems," *Journal of Computational and Graphical Statistics*, 12, 566–584.

Carlin, B.P., Kadane, J.B. and Gelfand, A. E. (1998). "Approaches for optimal sequential decision analysis in clinical trials". *Biometrics*, **54**, 964–975.

Carter, C. K., and Kohn, R. (1994), "On Gibbs sampling for state space models," *Biometrika*, 81, 541-553.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review, *Statistical Science*, 10, 273-304.

Christen, J.A., Müller, P., Wathen, K., and Wolf, J. (2004). "A Bayesian randomized clinical trial: a decision-theoretic sequential design", *Canadian Journal of Statistics*, 32(4), 387–402.

DeGroot, M. (1970). *Optimal statistical decisions*. New York: McGraw Hill.

Frühwirth-Schnatter, S. (1994), "Data augmentation and dynamic linear models," *Journal of Time Series Analysis*, 15, 183-202.

Grieve, A.P. and Krams, M. (2005). "ASTIN: a Bayesian adaptive dose-response trial in acute stroke," *Clinical Trials* 2, 340-351

Haines, L.M., Perevozskaya, I., and Rosenberger, W.F. (2003). "Bayesian optimal designs for phase I clinical trials," *Biometrics*, 59, 591–600.

Krams, M., Lees, K.R., Hacke, W., Grieve, A.P., Orgogozo, J.M., and Ford, G.A. (2003). "ASTIN: An adaptive dose-response study of UK-279,276 in acute ischemic stroke," *Stroke* 34,2543-2548.

Lees, K.R and Diener, H.-C. and Asplund, K. and Krams, M. (2003). "UK 279,276, a neutrophil inhibitory glycoprotein, in acute stroke: tolerability and pharmacokinetics," *Stroke*, 34, 1704–1709.

Müller, P., Berry, D., Grieve, A., Smith, M., and Krams, M. (1999). "Simulation-based sequential Bayesian design."

Robert, C. (2001) *The Bayesian Choice*, New York: Springer-Verlag.

Rossell, D., Müller, P. and Rosner, G.L. (2005). "Screening designs for drug development," Technical Report, M.D. Anderson Cancer Center.

Scandinavian Stroke Study Group (1985). Multicenter trial of hemodilution in ischemic stroke. Bbackground and study protocol. *Stroke*, 16, 885-890.

Spiegelhalter, D. J, Abrams, K. R., and Myles, J. P. (2004). *Bayesian approaches to clinical trails and health care evaluation*, John Wiley and Sons: Chichester, UK.

Spiegelhalter, D. J., Freedman, L. S., and Parmar, M. K. B. (1996). "Bayesian approaches to randomize trials". in: *Bayesian Biostatistics*, D. A. Berry and D. K. Stangl, (ed.), Marcel Dekker: NY. 67–108.

Stallard, N., Thall, P. F., and Whitehead, J. (1999). "decision-theoretic designs for phase II clinical trials with multiple outcomes". *Biometrics*, 55, 971–977.

Stallard, N. (2003). "decision theoretic designs for phase II clinical trials allowing for competing studies," *Biometrics*, 59, 402–409.

Thall, P.F., Simon, R.M., and Estey, E.H. (1995). Bayesian sequential monitoring designs for single-arm clinical trials with multiple outcomes, *Statistics in Medicine*, 14, 357-379.

Thall, P.F., and Russell, K.E. (199x). A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. to appear in *Biometrics*.

Verdinelli, I. (1992). Advances in Bayesian experimental design. In *Bayesian Statistics IV*, (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds), pp. 467–48, Oxford: Oxford University Press.

Wakefield, J. (1994). An expected loss approach to the design of dosage regimes via sapling-based methods. *The Statistician*, 43, 13–29.

West, M., and Harrison, P. J. (1997), *Bayesian Forecasting and Dynamic Models*, 2nd edition, New York: Springer-Verlag.

Whitehead, J. and Brunier, H. (1995). Bayesian decision procedures for dose determining experiments, *Statistics in Medicine*, 14, 885-893.

Whitehead, J., and Williamson, D. (1998). Bayesian decision procedures based on logistic regression models for dose-finding studies, *Journal of Biopharmaceutical Statistics*, 8, 445-467.

Yin, G., Li, Y., and Ji, Y. (2006). "Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratios," to appear.

Zhang, L., Zhang, Z.G., Zhang, R.L., Lu M., Krams, M., and Chopp, M. (2003). "Effects of a selective CD11b/CD18 antagonist and recombinant human tissue plasminogen activator treatment alone and in combination in a rat embolic model of stroke," *Stroke* 34, 1790-1795.