1 Introduction

2 The System

Need to introduce some notation for the data, and dose levels and (generic) notation for the prob model. I will assume we have defined:

2.1 Notation

Let y_i denote the final week 13 response for the i-th patient, measured as difference to baseline SSS (scandinavian stroke score) at the time of admission. Let y_{ij}, $j = 1, \ldots, M-1$, denote early responses in weeks 1 through $M - 1 = 12$. We will use y without index to generically denote a week 13 response, and $\boldsymbol{y}_i = (y_{ij},\ j = 0, \ldots, M)$ and $\boldsymbol{y}(i) = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_i)$ to denote all data for the i-th patient and for the first i patients, respectively. We will use D to denote all data collected in the study. Let x_i denote a vector of covariates for the i-th patient. In the current implementation the only covariate is the baseline SSS score at the time of admission. In general, x_i could include any relevant patient-specific information, like time between stroke and start of treatment, age, etc.

We shall use z to generically refer to a treatment dose, z_i to refer to the dose assigned to patient i, and $\{Z_j,\ j = 0, \ldots, J\}$ to refer to the set of allowable doses, including placebo $Z_0 = 0$. We will use Y_{jk}, $j = 0, \ldots, J$, to denote the response of the k-th patient who is assigned to dose Z_j. Note the convention of using upper case Z and Y to refer to doses and responses indexed by dose level, and lower case z and y to refer to doses and responses indexed by patient number i.

In Sections 5 and 6 we will define a sampling model for y_i and y_{ij}. Until then the discussion does not rely on a specific probability model. The approaches to define dose assignment (Section 3) and stopping time (Section 4) are valid with any underlying probability model. We will only need generic notation to refer to the dose/response curve for y_i and the longitudinal data model for y_{ij}. We shall use $f(z, \theta)$ to denote the dose/response curve $E(y|\theta)$ as a function of dose z, parametrized by an unknown parameter vector θ, i.e., $E(y_i|\theta) = f(z_i, \theta)$.

Given the mean curve $f(z, \theta)$ we assume normal errors:

$$y_i = f(z_i, \theta) + \epsilon_i \text{ with } \epsilon_i \sim N(0, \sigma^2). \tag{1}$$

Let $df(z, \theta) = f(z, \theta) - f(0, \theta)$ denote the advantage over placebo. For a given parameter vector θ let $\hat{z}_\theta$ denote the dose with maximum expected response, i.e., $f(\hat{z}_\theta, \theta) = \max_j\{f(Z_j, \theta)\}$. Denote with $z95_\theta$ the ED95 of the unknown dose response curve, defined as the minimum dose Z_j with mean improvement greater or equal than the maximum possible improvement over placebo:

$$z95_\theta = \min\{Z_j : \ df(Z_j, \theta) \geq 0.95 \ df(\hat{z}_\theta, \theta)\}$$

We shall use $f(z)$ for the posterior expected dose/resonse curve, $f(z) = E[f(z, \theta) \mid D]$, and $z95$ for the ED95 of the posterior mean curve $f(z)$. Also, we shall use $\overline{z95}$ to denote the posterior mean of $z95_\theta$, i.e., $E(z95_\theta \mid D)$. Of course, due to the non-linear nature of the ED95 $\overline{z95} \neq z95$. For technical reasons, in the program we sometimes consider conditional posterior means of $z95_\theta$, keeping some aspects of the model fixed. We will indicate when this is the case.

2.2 Dose-Finding Study

The first phase of the trial is a dose-finding study. Details of the dose allocation are discussed in Section 3. We will use D_{1t} to denote the data up to week t of the dose-finding phase, and n_{1t} to denote the number of patients enrolled by week t. Similarly D_1 denotes the data from the complete dose-finding phase, and n_1 denotes the total number of patients enrolled in this phase.

2.3 Confirmatory Study

The confirmatory study randomizes patients uniformly to placebo ($z = 0$), recommended treatment dose z^* and a third dose z^{**} chosen as alternative to z^*. As recommended treatment dose we use the ED95 dose $z^* = \overline{z95}$. As alternative dose z^{**} we use the ED50 $z^{**} = \overline{z50}$. We will use D_2 to denote the data from the confirmatory phase, and n_2 to denote

the number of patients enrolled in this phase at each of the three doses (i.e., the total sample size is $3n_2$).

To decide n_2 we use an argument of predictive power. Let df^* denote the advantage of treatment $df(z^*, \theta)$ over placebo at the recommended dose z^*. Let (m, s) denote posterior mean, $m = E(df^* \mid D_1)$, and standard deviation, $s^2 = Var(df^* \mid D_1)$ at the end of the dose-finding study. Details of this estimation depend on the underlying probability model (See Section 5 for a discussion of the probability model in our implementation). Let $\hat{\sigma}^2$ denote the posterior mean of $E(\sigma^2 \mid D_1)$ conditional on all current data. Fixing σ^2 at $\hat{\sigma}^2$ and approximating $p(df^* \mid D_1) \approx N(m, s)$, we can find the minimum sample size n_2 which gives a desired predictive power (90% in our implementation). See the appendix for a definition of predictive power. [1]

3 Adaptive Dose Allocation: The Allocator

3.1 Introduction

During the dose-finding phase of the trial we need to decide a dose for each newly recruited patient. Conventional solutions are to randomly assign one out of a fixed set of allowable doses (*reference?*); up-down designs (*reference?*); and designs which are based on balancing probabilities of adverse otucomes and efficacy (Thall and Russell, 199x).

We consider an alternative, entirely decision theoretic approach. To introduce notation and to clarify the context, we briefly review the general setup of a Bayesian decision problem. Decision making under uncertainty is choosing an action d to maximize expected utility $U(d) = \int u(d, \theta, y) p_d(\theta, y)$. Here, $u(d, \theta, y)$ is the utility function modeling preferences of consequences and $p_d(\theta, y)$ is a probability distribution of parameter θ and observation y, possibly influenced by the chosen action d. Typically $p_d(\theta, y)$ is specified as a prior $p(\theta)$ on the parameters and a sampling model $p_d(y|\theta)$. Note that utility $u(d, \theta, y)$ needs to be specified for a specific realization (θ, y) of the experiment only. Computing expected $U(d)$ we average

[1] Alternatively we could use a conventional power calculation based on the likely treatment effect m, or at a conservative estimate $(m - s)$ for the treatment effect.

over the quantities which are unknown at the time of decision making. Often, some of the data, say y_0, is already known at the time of decision making. Assume $y = (y_0, y_1)$. Then $p(\theta)$ and $p_d(y \mid \theta)$ are replaced by $p(\theta \mid y_0)$ and $p_d(y_1 \mid \theta, y_0)$. See Chaloner and Verdinelli (1995) and Verdinelli (1992) for reviews of Bayesian approaches to decision problems traditionally known as optimal design. Spiegelhalter, Freedman and Parmar (1994), Berry (1993) and Berry and Stangl (1996) discuss general issues related to the use of Bayesian optimal design methods in medical decision problems.

3.2 Dose Allocation as a Decision Problem

Central to the proposed approach is a utility function which expresses the relative preferences over alternative outcomes. The proposed utility function is related to learning about the unknown dose/response curve. Learning is formalized as minimizing the posterior variance for some key parameter $g(\theta)$ of the dose/response curve. In the current implementation we choose as key parameter $g(\theta)$ the mean response $f(z95_\theta, \theta)$ at the ED95. Note that the posterior variance of $g(\theta)$ includes uncertainty in the unknown ED95 dose, as well as the unknown response at that dose.

The proposed approach is myoptic in the sense that when we consider the optimal dose for the next patient we proceed as if he or she was the last patient to be recruited into the trial. Assume we currently have N patients enrolled in the trial. Let $\tilde{y}_k = y_{N+k}$, $\tilde{x}_k = x_{N+k}$ and $\tilde{z}_k = z_{N+k}$ denote the response, covariate and assigned dose for the next K patients, $k = 1, \ldots, K$. Let $\tilde{y}(k) = (\tilde{y}_1, \ldots, \tilde{y}_k)$, $\tilde{z}(k) = (\tilde{z}_1, \ldots, \tilde{z}_k)$, and $\tilde{x}(k) = (\tilde{x}_1, \ldots, \tilde{x}_k)$, denote responses, assigned doses and covariates up to the k-th new patient. Let D_N denote the observed data for the first N patients, and let $\tilde{D}$ denote the still missing final responses for already enrolled patients. We define the utility function for choosing the dose $\tilde{z}_k$ as

$$u_k[\tilde{z}(k), \tilde{y}(k), \tilde{x}(k), \tilde{D}, D_N] = -Var[g(\theta) \mid D_N, \tilde{D}, \tilde{y}(k), \tilde{x}(k), \tilde{z}(k)].$$

Of course we have to decide upon $\tilde{z}_k$ before observing $\tilde{D}$, $\tilde{y}(k)$ and $\tilde{x}(k)$. Thus we choose the dose $\tilde{z}_k$ by maximizing the utility $u_k(\cdot)$ in expectation, averaging with respect to $\tilde{D}$,

$\tilde{y}(k)$ and $\tilde{x}(k-1)$. The relevant distributions for $\tilde{D}$ and $\tilde{y}(k)$ are the posterior predictive distributions given the current data. For the covariates $\tilde{x}_1, \ldots \tilde{x}_{k-1}$ we use the empiricial distribution $p(x)$ from the Copenhagen Stroke Study data base, assuming independence, i.e., $p[\tilde{x}(k-1)] = \prod_{h=1}^{k-1} p(\tilde{x}_h)$. And $\tilde{x}_k$ is fixed at "typical" covariate values x^o, i.e., we find the optimal dose for an average next patient. For $\tilde{z}_1, \ldots, \tilde{z}_{k-1}$ we substitute the optimal values found by optimizing the expected utilities $U_1(\cdot), \ldots, U_{k-1}(\cdot)$.

$$U_k[\tilde{z}_k, \tilde{z}(k-1), \tilde{x}_k, D_N] = \int u_k[\tilde{z}(k), \tilde{y}(k), \tilde{x}(k), \tilde{D}, D_N] \; \times \\ \times \, p(\tilde{D} \mid D) \; p[\tilde{y}(k) \mid D, \tilde{z}(k)] \; p[\tilde{x}(k-1)] \; d\tilde{D} \; d\tilde{y}(k) \; d\tilde{x}(k-1). \quad (2)$$

Maximizing $U_k(\cdot)$ over $\tilde{z}_k$ we find the optimal action.

3.3 Evaluating Expected Utility

Critical for a successful implementation of the proposed decision theoretic dose allocation is the availability of analytical or efficient numerical integration to evaluate the integrals in (2). Key to our implementation strategy is to rewrite the expected utility integral (2) as an integral with respect to the posterior distribution $p(\theta \mid D)$

$$U_k[\tilde{z}(k), \tilde{x}_k, D_N] = \int \Big\{ u_k[\tilde{z}(k), \tilde{y}(k), \tilde{x}(k), \tilde{D}, D_N] \; \times \\ \times p(\tilde{D} \mid \theta, D) \; p[\tilde{y}(k) \mid \theta, \tilde{x}(k)\tilde{z}(k)] \; p[\tilde{x}(k-1)] \; d\tilde{D} \; d\tilde{y}(k) \; d\tilde{x}(k-1) \Big\} \; p(\theta \mid D) \; d\theta. \quad (3)$$

Most models allow efficient random variate generation from the posterior distribution $p(\theta \mid D)$ using Markov chain Monte Carlo (MCMC) simulation. See, for example, Tierney (1994) or Gilks et al. (1996) for a summary of MCMC methods. Details of implementing the appropriate MCMC scheme depend on the specific probability model $p(D \mid \theta)$. In Section 5 we will discuss posterior simulation in the model underlying our implementation. But for the following discussion we do not need to refer to specific details of the MCMC simulation. We only assume that by appropriate simulation techniques it is possible to generate an (approximate) posterior Monte Carlo sample $\Theta = \{\theta^1, \ldots, \theta^T\}$ with $\theta^t \sim p(\theta \mid D)$. Using the Monte Carlo sample Θ we can evaluate expected utilities by replacing (3) with a corresponding

Plot of expected utilities $\hat{U}_k$ as a function of dose z.

Figure 1: Approximate expected utilities $\hat{U}_k$ plotted against $\tilde{z}_k$. The circles show the Monte Carlo estimates of expected utilities at the allowable doses. The solid curve shows a smooth fit through the Monte Carlo estimates. For comparison the dashed curve shows a smooth fit through Monte Carlo estimates of expected utilities using half the Monte Carlo sample size.

Monte Carlo sample average. For each θ^t we simulate covariates $\tilde{x}_h^t \sim p(x)$, $h = 1, \ldots, k-1$, responses $\tilde{y}_h^t \sim p(y_h|\tilde{z}_h, \tilde{x}_h^t, \theta^t)$ and missing responses $\tilde{D}^t$ of current patients $\tilde{D} \sim p(\tilde{D} \mid \theta^t)$, assuming that these models are all available for efficient random variate generation. For each simulated experiment we then evaluate observed utility $u_k^t = u_k(\tilde{z}(k), \tilde{y}(k)^t, \tilde{x}(k)^t, \tilde{D}^t, D)$, and replace expected utility by a Monte Carlo average

$$\hat{U}_k(\tilde{z}(k), \tilde{x}_k, D) = \frac{1}{M} \sum_{t=1}^{M} u_k^t$$

Evaluating expected utility $\hat{U}_k$ for a grid of possible choices $\tilde{z}_k$ we find the optimal dose as the dose with maximum $\hat{U}_k(\cdot)$. To reduce numerical error in the approximation of U_k by $\hat{U}_k$ we use common random numbers, i.e., whenever possible we use the same Monte Carlo sample Θ and random variates $\tilde{x}_h^t$, $\tilde{y}_h^t$, $\tilde{D}^t$ when we evaluate $\hat{U}_k(\tilde{z}_k)$ for alternative choices of $\tilde{z}_k$.

Figure 1 shows a typical expected utility curve. Of course, the expected curve changes from week to week. Assume at one time it is optimal to allocate to high doses. As more patients are allocated in that part of the dose range it will eventually become more advantageous to allocate new patients in other parts of the dose range to learn about, for example, the response at placebo. While it is difficult to intuitively understand the expected utility curve and the dose assignement at a given time, typical patterns of dose allocations over the course of a clinical trial do seem intuitively meaningful. Figure 2 shows an example of dose assignments in a simulated trial.

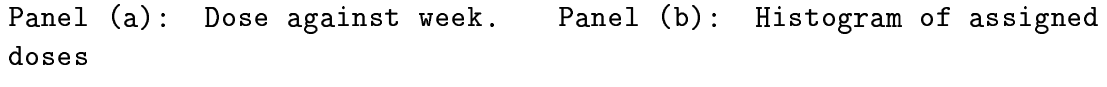

Figure 2: Doses assigned over the course of simulated clinical trial. The left panel plots the assigned doses against week. The right panel shows a histogram of assigned doses with an overlaid plot of the true (diamonds) and estimated (solid line) dose/response curve.

3.4 Additional Randomization: The Recommender

Maximization of $\hat{U}_k(\cdot)$ delivers the optimal dose $\tilde{z}_k^*$ to be assigned to the next patients, $k = 1, \ldots, K$. Assume z_k^* equals Z_{j^*} in the list of allowable doses. Before actually assigning a dose to a new patient we use an additional randomization. First, because of regulatory constraints we need to keep a given minimum percentage p_0 at placebo. Second, because of safety concerns we want to avoid unnecessarily high doses. To achieve these two aims we allocate with probabilties p_0 at placebo and split the remaining probability $(1-p_0)$ uniformly over all doses Z_j within a neighborhood of Z_{j^*}, defined as the set of all doses less than or equal Z_{j^*} with estimated mean response within 10% of the estimated mean response at Z_{j^*}.

4 Optimal Stopping: The Terminator

4.1 Introduction

The dose-finding phase of the trial involves two important decision problems, dose assignment and termination. For the first, we already discussed an adaptive dose allocation scheme in Section 3. In this section we discuss the second problem, i.e., the problem of optimal stopping in the dose-finding trial. At each period t of the trial, say once a week, we have to decide (d_t) whether to terminate the trial and abandon the drug $(d_t = A0)$, continue with the dose-finding phase $(d_t = A1)$, or terminate the dose-finding phase and switch to pivotal mode $(d_t = A2)$.

The following fact significantly complicates the derivation of an optimal decision. When

we consider not to terminate today ($A1$), then to evaluate the expected utility of this action we need to take into account what the optimal termination decision will be next week, and what the expected worth of this decision will be, i.e., we need to solve a sequential decision problem. A standard approach to sequential decision problems is backward induction: Consider the optimal solution under every possible scenario going all the way into the future to the latest possible termination time T (determined by some maximum number of patients in the trial), and record a table of such decisions, starting from T. Thus, when it comes to evaluate today's decision we already have available the optimal actions for tomorrow in the table. See, for example, Berger (1985, chapter 7).

For the *terminator* we implemented a numerical solution of the backward induction problem. The approach is based on a dual strategy of using a reduced action space to constrain the number of scenarios which we need to consider in backward induction; and forward simulation to evaluate expected utility integrals under all relevant scenarios. Central to our approach is a formulation of the problem as a formal decision problem with a probability model describing all relevant uncertainties and a utility function which describes the relative preferences of possible outcomes.

Alternative Bayesian approaches to optimal sequential design in similar medical decision problems are discussed, among other references, in Thall, Simon and Estey (1995) who define stopping criteria based on posterior probabilities of clinically meaningful events. Similarly, Thall and Russell (1998) define a sequential procedure based on monitoring posterior probabilities of certain events. Using ad-hoc rules based on these probabilities they define designs and evaluate their frequentist performance. Vlachos and Gelfand (1998) follow a similar strategy. Whitehead and Brunier (1995) and Whitehead and Williamson (1998) use what is essentially a Bayesian m-step look-ahead procedure to find the optimal dose to assign to the next m patients in a dose-finding study.

s[t]													
1.40	1.50	1.60	1.70	1.80	1.90	2.01	2.11	2.21	2.31	2.41	2.51	2.61	
0	0	0	0	0	0	0	0	0	0	0	0	0	-0.36
0	0	0	0	0	0	0	0	0	0	0	0	0	0.10
0	0	0	0	0	0	0	0	0	0	0	0	0	0.57
0	0	0	0	0	0	0	0	0	0	0	0	0	1.03
1	1	1	1	1	0	0	0	0	0	0	0	0	1.49
1	1	1	1	1	1	1	1	1	1	1	0	0	1.95
1	1	1	1	1	1	1	1	1	1	1	1	1	2.41
1	1	1	1	1	1	1	1	1	1	1	1	1	2.88
1	1	1	1	1	1	1	1	1	1	1	1	1	3.34
1	1	1	1	1	1	1	1	1	1	1	1	1	3.80 m[t]
1	1	1	1	1	1	1	1	1	1	1	1	1	4.26
1	1	1	1	1	1	1	1	1	1	1	1	1	4.72
2	2	1	1	1	1	1	1	1	1	1	1	1	5.18
2	2	2	2	2	2	2	2	2	1	1	1	1	5.65
2	2	2	2	2	2	2	2	2	2	1	1	1	6.11
2	2	2	2	2	2	2	2	2	2	2	1	1	6.57
2	2	2	2	2	2	2	2	2	2	2	2	1	7.03
2	2	2	2	2	2	2	2	2	2	2	2	2	7.49
2	2	2	2	2	2	2	2	2	2	2	2	2	7.96
2	2	2	2	2	2	2	2	2	2	2	2	2	8.42

Figure 3: Terminator. The table shows the optimal decision for each pair (m, s). An entry of 0,1,2 indicates A0, A1, and A2, respectively.

4.2 The Terminator

Before we discuss the underlying decision theoretic framework, we present the form of the final implementation and argue why it is intuitively appealing. Let $z^* = \overline{z95}$ denote the ED95 dose. Let (m_t, s_t) denote the posterior mean and standard deviation of the advantage over placebo at the ED95 dose conditional on the data available at time t, i.e., $m_t = E[df(z^*, \theta) \mid D_{1t}]$ and $s_t^2 = Var[df(z^*, \theta) \mid D_{1t}]$. The proposed stopping rule is a function of (m_t, s_t). We define cutoffs for (m_t, s_t) which partition the space into three subsets, corresponding to decisions $A0, A1$ and $A2$. For example, the decision rule could take the form shown in Figure 3. The overall pattern of the partition is intuitive. For very small, and very large values of m we recommend $A0$ and $A2$, respectively, i.e., the dose-finding phase is terminated and we make a decision for abandoning the drug or for continuation to the confirmatory phase. For intermediate values we recommend to continue the dose-finding

trial.

When we have to decide about termination at time t, we compute (m_t, s_t) and use the cutoffs to look up the recommended decision. Note that the decision depends on the current data only indirectly through (m_t, s_t). The main reason for constraining the decision space are the resulting computational simplifications. The rationale for choosing this particular constraint is that the effect at the finally recommended dose and the uncertainty about that effect are the features of the unknown dose/response curve which are most relevant to the desired decision.

4.3 A Decision Theoretic Stopping Rule

To fix the boundaries in the (m_t, s_t) table we use a decision theoretic argument. At the core of the argument is a utility function which for a given outcome of the experiment and a given decision gives the worth of the observed consequences. Maximizing this utility function in expectation defines the optimal termination decision. As mentioned, the sequential nature of the problem significantly complicates this maximization.

4.3.1 The Utility Function

We start the discussion by stating the utility function used in our implementation. Note that we only need to formalize the utilities assuming that the outcome of the whole experiment is known, i.e., we define utility as a function of possible outcomes. The expected utility function will then be derived from this by considering the appropriate expectations over all unknown random variables. For a given realization the utility includes a sampling cost for the number of patients recruited into the trial, and a payoff for successfully developing and marketing the drug. Of course that payoff is only included if the final results of the experiment are such that the drug can approved by the regulatory authorities.

Utility under $d_t = A2$. If we were to decide $d_t = A2$ then the resulting utility will include the sampling cost (negative utility) for the confirmatory phase of the study plus the payoff if the confirmatory study eventually concludes a significant treatment effect.

Let B denote the event of observing data in the confirmatory phase which in the end lets us reject the null hypothesis of no treatment effect at some fixed significance level α. Let $\bar{y}_0$ and $\bar{y}^*$ denote the sample average in the confirmatory phase of the responses under placebo and treatment (z^*), respectively. B is the event

$$B = \{(\bar{y}^* - \bar{y}_0)/\sqrt{2\,\hat{\sigma}^2/n_2} > q_\alpha\}$$

where q_α is the $(1-\alpha)$ standard normal quantile. Remember from Section 2.3 that n_2 is the sample size in the confirmatory study, and $\hat{\sigma}^2$ is the posterior mean of the measurement variance.

Let c_1 denote the sampling cost per patient, and let c_2 denote the payoff for a successful drug. Assuming that this payoff is proportional to the size of the effect we specify c_2 as payoff per point advantage over placebo. Let $m_2 = E(df^* \mid D_1, D_2)$.

$$u(d_t = A2, D_1, D_2) = \begin{cases} -3\ n_2\ c_1 & \text{if not } B \\ -3\ n_2\ c_1 + c_2\ m_2 & \text{if } B \end{cases}$$

Note that the data enters into the definition of $u(\cdot)$ only implicitely, through n_2 which depends on D_1, m_2 which is a statistic of D_2 and B_2 which is an event in D_2. The sampling cost includes only the patients for the confirmatory study, but omits the sampling cost for the first n_1 patients in the dose-finding phase. Including it would add the same term to the utility under all three alternative actions ($A0, A1$ and $A2$) and would thus not change the decision. Of course, decision d_t needs to be decided before observing D_2. The relevant expected utility averages over D_2

$$U(d_2 = A2, D_1) = \int u(d_t = A2, D_1, D_2)\ dp(D_2 \mid D_1). \tag{4}$$

Recall from Section 2.3 that n_2 is a function of (m, s), leaving only $P(B \mid D_1)$ and $E(m_2 \mid D_1, B)$ to be computed. Approximating $p(df^* \mid D_1) \approx N(m, s)$, both can be computed analytically.

Utility under $d_t = A0$. If we were to decide $d_t = A0$ then the trial is over and thus

$$U(d_t = A0) = 0,$$

again omitting the sampling cost for the first n_1 patients.

Utility under $d_t = A1$. Finally, if we were to decide $d_t = A1$, then the utility depends on what we will decide in the next period $t+1$. Let n^+ denote the number of patients enrolled in week $(t+1)$ and let D^+ denote the data collected in week $(t+1)$.

$$u(d_t = A1, D_{1t}, D^+) = -n^+ \cdot c_1 + U^*_{t+1}(D_{1t} \cup D^+),$$

where $U^*_t(D_{1t})$ is the expected utility under the optimal action at time $(t+1)$. By taking the expectation with respect to D^+ we obtain

$$U(d_t = A1, D_{1t}) = \int u(d_t = A1, D_{1t}, D^+)\ dp(D^+ \mid D_{1t}). \tag{5}$$

We are assuming here that n^+ is fixed or known. Extension to random n^+ is straightforward. Evaluation of the integral (5) is discussed below in Section 4.5.

The optimal decision $d^*_t(D_{1t})$ at time t can, in principle, be derived as

$$U^*_t(D_{1t}) = U(d_t = d^*_t, D_{1t}) = \max_{d_t \in \{A0, A1, A2\}} \{U(d_t, D_{1t})\}. \tag{6}$$

There are at least two impediments to a straightforward implementation of (6). First, the definition of $U(d_t = A1, \ldots)$ requires the solution to (6) for period $(t+1)$, i.e., we need backward induction. Second, the definition of $U(d_t = A2, \ldots)$ as well as $U(d_t = A1, \ldots)$ involve typically analtyically intractable integrals. In the following two subsections we outline an implementation strategy based on constrainied backward induction and forward simulation. See Müller, Berry, Grieve, Smith and Krams (1999) for a detailed discussion.

4.3.2 Constrained Backward Induction

To allow a practical solution to the backward induction problem we constrain the decisions d_t to depend on the current data only indirectly through (m_t, s_t), with (m_t, s_t) reported on a finite grid, i.e., we report $d^*(m_t, s_t)$. We will use $U^*(m, s)$ to denote the expected utility of the optimal decision $d^*(m, s)$. Effectively this amounts to constraining our action space. Instead of allowing decisions to depend on the full information set D_{1t} they are only allowed to depend on the current data indirectly through (m_t, s_t). How much we loose by this constraint depends upon how important the abandoned information in the data is, i.e.,

how important for the desired decision is the information beyond what is summarized by (m_t, s_t). Note that $U(d_t = A2, D_{1t})$ (and $U(d_t = A0)$, trivially) depend on D_{1t} only indirectly through (m, s). We will write $U(d = A2, m, s)$ to emphasize this. Only the computation of $U(d_t = A1, \dots)$ depends on D_{1t} beyond the summary statistic (m, s). See Section 4.5 for the definition and evaluation of $U(d_t = A1, m, s)$. We will use $U^*(m, s) = U\left(d_t^*(m, s), m, s\right)$ for the expected utility under the optimal decision.

Starting with some initial guess for $d^*(m_t, s_t)$ and $U^*(m, s)$ we update the tables using (4) through (6). We continue updating until the table remains unchanged over a complete cycle of updates.

4.3.3 Forward Simulation

There still remains the problem of evaluating the expected utility integral in (5). We use forward simulation (Carlin, Kadane and Gelfand, 1998). The idea is to simulate many, say T, trials all the way into the future until the maximum number $\bar{n}_1$ of possible patients in the dose-finding phase.

Figure 4 illustrates how these forward simulations are used to compute posterior integrals. Assume we need a posterior integral conditional on data D_{1t}. We first compute the corresponding summary statistic (m_t, s_t), and then approximate the desired posterior integral by a sample average over all simulated trials whose trajectories pass through (m_t, s_t). If (m_t, s_t) was in fact a sufficient statistic for the unknown parameter vector then the sample average would provide a (simulation) consistent and unbiased estimate of the desired expectation. Carlin et al. (1998) use forward simulation in such a setup. In general, some approximation is involved.

Let D_{1t} and n_{1t} denote the data and number of enrolled patients at the time of making the decision d_t. Let $\tilde{D}$ denote the future data, including the remaining $\bar{n}_1 - n_{1t}$ patients, as well as still missing final measurements on current patients who have been recruited during the last 13 weeks. The T trials for the forward simulation are generated by simulating from the posterior predictive distribution $p(\tilde{D}|D_{1t})$. Denote the simulated values as $\tilde{D}^{(j)}$. For

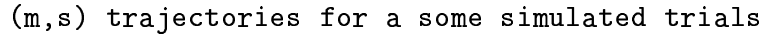

Figure 4: Forward simulation. The figure shows trajectories in the (m_t, s_t) space for some simulated trials. To compute a posterior integral for ....

each week of each simulation we record the moments $(m_t^{(j)}, s_t^{(j)})$.

Assume in the constrained backward induction we have current estimates $d^*(m, s)$ and $U^*(m, s)$. Assume we are considering to update a specific cell (m, s) and want to find the optimal decision. Expected utility $U(d = A2, m, s)$ is easily computed using (4), and $U(d = A0, m, s) = 0$ is fixed. To compute integral (6) for $U(d = A1, \dots)$ we consider the the subset A of simulations which pass through cell (m, s) at some time, i.e., for all $j \in A$ we find $(m_{t_j}, s_{t_j}) = (m, s)$ in some week t_j. Approximate (6) by the sample average

$$\hat{U}(d = A1, m, s) = -n^+ \; c_1 + \frac{1}{T'} \sum_{j \in A} U^*(m_{t_j+1}^{(j)}, s_{t_j+1}^{(j)}),$$

plugging in the current estimates for $U^*(m, s)$. Note that the approximation for $U(d = A1, \dots)$ is the same for all D_{1t} with the same summary statistic (m, s). Thus we write $U(d = A1, m, s)$. Compare with $U(d = A2, m, s)$ and $U(d = A0, m, s)$ to find the optimal decision $d^*(m, s)$ and it's value $U^*(m, s)$. Repeating the same process for all cells (m, s) updates the currently imputed values for $d^*(m, s)$ and $U^*(m, s)$. We repeat updating until nothing changes in one cycle of updates.

4.4 From the Decision Theoretic Stopping Rule to the Terminator

The process described in the last two sections is very comutation intensive. It is impractical to repeat that same computation each week, each time with slightly different data D_{1t}. Instead we build one static decision table like in Figure 3 by combining estimated utilities $\hat{U}(d_t = a, m, s)$, $a \in \{A0, A1, A2\}$, computed under a set of typical dose/response curves. In the current implmentation we use four dose/response curves. For each cell in the (m, s)

```
four figures showing the four d/r curves
four figures showing the ‘‘dynamic’’ term’s
```

Figure 5: The four "typical" dose/response curves used to construct the terminator. The four panels in the first row show the true curves, the four panels in the bottom row show the derived decision rules $d^*(m, s)$. Combining these four decision rules results in the rule shown in Figure 3

table we compute the average expected utility approximation under the four curves. For $A2$ we use equal weights. For $A1$ we use weights proportional to the number of simulated trajectories in each experiment which did pass through the specific (m, s) cell.

5 The Dose/Response Curve

5.1 The Probability Model

The choice of the probability model for $f(z)$ is guided by the following considerations. First, we need a model which 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/reponse curves. Although an increasing curve with asymptotes is *a priori* likely, the model should allow for possible lack of monotonicity and other irregular features. 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. Before we describe details of the model, we outline some important features. Denote with Z_j, $j = 1, \ldots, J$, the range of allowable doses, and with $\theta_j = f(Z_j, \theta)$, $j = 1, \ldots, J$, 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 $\theta_j + (z - Z_j)\delta_j$, with level θ_j and slope δ_j. This is illustrated in Figure 6. When moving from dose Z_{j-1} to Z_j the parameters $\alpha_j = (\theta_j, \delta_j)$ change by adding a (small) so-called evolution noise e_j

Figure 6: The NDLM (xyz) fits a smooth curve to the data by defining for each dose j, $j = 1, \ldots, J$, a local straight line, parametried by α_j. The plot shows the local lines fit at doses $z_j = xyz$ and $z_{j'} = xyz$. The triangle marks the point (z_j, θ_{t_j}), and a solid line segment indicates the locally fit line. Between doses level and slope of the line change by adding an evolution noise e_j.

and adjusting the level $\theta_j = \theta_{j-1} + \delta_{j-1}$. 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 $z_i = Z_j$. The resulting model is

$$Y_{jk} = \theta_j + \epsilon_{jk}, \qquad\qquad j = 1, \ldots, n,\ k = 1, \ldots, \nu_j$$

and

$$\begin{pmatrix} \theta_j \\ \delta_j \end{pmatrix} = \begin{pmatrix} \theta_{j-1} + \delta_{j-1} \\ \delta_{j-1} \end{pmatrix} + e_j, \tag{7}$$

with independent errors $\epsilon_j \sim N(0, V\sigma^2)$ and $e_j \sim N_2(w_j, W_j\sigma^2)$. Here $N_2(m, S)$ denotes a bivariate normal distribution with moments m and S. The first equation describes the distribution of Y_{jk} conditional on the state parameters $\alpha_j = (\theta_j, \delta_j)$ and is referred to as the "observation equation"; the second equation formalizes the change of α_j between doses and is referred to as the "evolution equation". For a given specification of $\{V, W_j,\ j = 1, \ldots, n\}$ and a prior $p(\alpha_0) = N(m_0, C_0)$, $p(\sigma^{-2}) = Gamma(n_0/2, S_0/2)$ with given moments m_0, C_0 and S_0, and degrees of freedom n_0, there exists a straightforward recursive algorithm to compute posterior distributions $p(\alpha_j|Y_1, \ldots, Y_j)$ and any other desired posterior inference. It can be shown that $p(\alpha_j|Y_1, \ldots, Y_j)$ is bivariate normal $N(m_j, C_j)$ with some moments m_j, C_j. For later reference we note also that the predictive distributions $p(Y_j|Y_1, \ldots, Y_{j-1})$ are normal distributions with moments f_j, Q_j, and the posterior distributions $p(\alpha_j|Y_1, \ldots, y_n)$ are

bivariate normal $N(m_j^*, C_j^*)$. West and Harrison (1997) give the recursive equations to compute $m_j, C_j, f_j, Q_j, m_j^*, C_j^*$ and other posterior moments. An algorithm, known as Forward Filtering Backward Sampling (FFBS) allows efficient random variate generation from the full posterior distribution. It is described in Frühwirt-Schnatter (1994) and Carter and Kohn (1994).

In many applications, specification of the evolution variances W_j is not easy. As an alternative to specifying W_j *a priori*, West and Harrison (1997) propose to define W_j as a scalar multiple of the posterior variance C_j, i.e., $W_j = (1-r)/r\ C_j$ for some scalar $r \in (0,1)$. This corresponds to thinking of the evolution noise e_j as discounting some of the current information on α_j as represented by the posterior variance C_j. The scalar factor r is known as "discount factor". A large discount factor implies a small variance W_j and thus small change e_j between times, i.e., strong smoothing. On the other hand, a small discount factor implies large e_j and thus allows for a big change between doses Z_{j-1} and Z_j.

5.2 Prior Specification for the NDLM

A minor shortcoming of the NDLM in the present application is that the prior specification with the hyperparameters m_0, C_0 and r does not naturally allow to fix arbitrary desired prior moments for $\theta_j = f(Z_j)$. Only $E(\theta_0)$ and $Var(\theta_0)$ are fixed as m_0 and C_0. Prior expectation and variance for $\theta_j,\ j > 0$, 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).

6 The Longitudinal Data Model

References

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, New York: Springer-Verlag.

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

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.

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.

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.

Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (eds.) (1996). *Markov Chain Monte Carlo in Practice*, Chapman and Hall.

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*.

Tierney, L., (1994), "Markov chains for exploring posterior distributions," *Annals of Statistics*, 22, 1701–1762, (With discussion).

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.

West, M., and Harrison, P. J. (1989), *Bayesian Forecasting and Dynamic Models*, 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 logisic regression models for dose-finding studies, *Journal of Biopharmaceutical Statistics*, 8, 445-467.