A Bayesian Randomized Clinical Trial: A Decision Theoretic Sequential Design

J. Andrés Christen [1], **Peter Müller** [2], **Kyle Wathen** [2] and **Judith Wolf** [3]

[1] CIMAT, A.P. 402, Guanajuato, GTO, 36000, MEXICO,

[2] Department of Biostatistics, [3] Department of Gynecologic Oncology,

University of Texas M.D. Anderson Cancer Center, Houston, TX

77030, U.S.A.

Summary

We propose a Bayesian decision theoretic framework for randomization in clinical trials. The proposed approach reconciles the apparent contradiction between the practice of clinical trial design and implications of a traditional decision theoretic setup. Although it is universal practice to use randomization in clinical trial design, a standard decision theoretic setup does not justify randomized rules. The proposed approach is based on formally acknowledging that the decision maker might be unable or unwilling to specify a unique utility function. Instead we proceed with a set of possible utility functions. We develop a specific algorithm to implement the proposed approach in a phase II clinical trial comparing k competing experimental treatments. We develop a sequential myopic design that includes randomization justified by a set of utility functions. Randomization is introduced over all "non-dominated" treatments, allowing for interim removal of treatments and early stopping. Results are shown for a study to find the optimal biologic dose of pegylated interferon for platinum resistant ovarian cancer.

Keywords: Clinical trials; Backward induction; Utility functions; Robustness.

1 Introduction

We discuss a Bayesian decision theoretic framework to justify randomization in clinical trial protocols. The proposed approach addresses the conflict between the common practice of using randomization (for example, Piantadosi, 1997, chapter 9) versus the implication of a traditional decision theoretic setup that would assign essentially one unique optimal treatment (for example, Berger, 1985, chapter 1.4). Using the proposed setup we develop a specific algorithm for a decision theoretic design to compare k experimental treatments with placebo (or standard of care), assuming categorical response. The design is sequential, allowing to drop inferior treatment arms at any time in the trial, and allowing early stopping when stopping dominates continuation. Upon stopping the proposed trial design recommends a set of non-dominated treatments.

The planning of clinical trials with human subjects is driven by sometimes conflicting ethical and performance considerations (Thall, 2001, Carlin, Kadane and Gelfand, 1998, Spiegelhalter, Freedman and Parmar, 1996). We feel that any reasonable ethical judgment of a clinical trial should consider the consequences of all foreseeable outcomes of the trial. Utilities are numerical values assigned to such consequences. We therefore believe that a utility based decision theoretic approach to clinical trials provides a useful framework for efficient and ethical evaluation of clinical trial designs. Model based Bayesian inference provides a natural context to perform a formal decision theoretical approach. The need for formal decision theoretic approaches for clinical trial planning was highlighted as early as Ascombe (1963). See Berry (1993) for a recent discussion. Bayesian trial designs have been proposed using both, fully decision theoretic (for example, Carlin, *et al.*, 1998) and stylized approaches (for example, Spiegelhalter et al. 1996, or Stallard, Thall and Whitehead, 1999). The basic ingredients of a decision theoretic setup are an action space $\mathcal{A}$, a utility function v, and a probability model $p(\theta)$ for all relevant random variables θ, including parameters

and future data. The probability model could be a prior probability model, or the posterior predictive model conditional on historical data. It can be argued (DeGroot, 1970) that a rational decision maker should choose an action in $\mathcal{A}$ to maximize the expectation of v with respect to $p(\theta)$.

An important aspect of fully decision theoretic Bayesian approaches is the implication about randomization. Is a randomized decision justifiable as maximizing expected utility? We argue, yes, it is. A special case allows a simple justification. If an action a is equal in expected utility to another action a', that is, we are indifferent between a or a', then we may choose either of them and therefore we may justify a random selection among these two. But this justification misses the main motivations that leads researchers to choose randomization. In practice investigators choose randomization to avoid biases due to lurking variables and time trends, or because constraints of the regulatory process and peer review require them to do so. Berry and Kadane (1997) propose a formal justification of randomization by considering the impact of unknown covariates. Another possible formalization of randomization in a decision theoretic setup is the notion of uncertainty and lack of specification for the probability model and utility function.

Several approaches have been proposed to help utility elicitation for consequences of clinical trials. See Spiegelhalter *et al.* (1996), Bryan and Day (1995) or Wolfson, Kadane and Small (1996). However, in practice the utility function, for lack of time, difficulty to combine competing goals, controversy, etc., is not specified as a single function. Rather, we can at best narrow it down to a set of possible utility functions. In the following we denote with $\mathcal{V}$ a set of utility functions, and assume the decision maker is unwilling or unable to further specify a single utility function $u \in \mathcal{V}$. The problem has been studied from the perspective of sensitivity analysis to mis-specifications in the utility function and prior distribution. See Ríos–Insua, Ruggeri and Martin (2000) for a review. French and Rios–Insua (2000) discuss related axiomatic foundational issues. A central notion in the study of

related problems is the definition of sets of non dominated actions. An action $a \in \mathcal{A}$ is *non dominated*, given information I (*a priori* and data), if there is no other action $a' \in \mathcal{A}$ such that $E\{v(a', \theta) \mid I\} \geq E\{v(a, \theta) \mid I\}$, for all $v \in \mathcal{V}$, and $E\{v(a', \theta) \mid I\} > E\{v(a, \theta) \mid I\}$ for at least one v. The expectation is with respect to $p(\theta \mid I)$. Clearly, one should only consider non dominated actions. If the set of alternatives is finite (as in our case), then the set of non dominated actions (e.g. treatments) is non-empty.

In the next section we briefly describe a Phase II clinical trial for finding optimal dose levels. In Section 3 we use the above ideas to design a sequential clinical trial, using backward induction, where non dominated treatments (actions) are randomized to select the treatment for the next patient in the trial. Dominated treatments are abandoned in the course of the trial and early stopping is entertained for conclusive evidence in favor of a set of treatments. The finally selected set of treatments could be placebo (or standard of care) only, i.e., the trial can be stopped early for lack of evidence in favor of any experimental treatments. In Section 4 we apply the proposed approach to design a clinical trial for the protocol introduced in Section 2.

2 PEG Intron Protocol

PEG Intron, a pegylated form of alpha interferon, is proposed for treatment of patients with platinum resistant ovarian, fallopian tube or peritoneal cancer. Preliminary laboratory studies suggest that chronic exposure of interferon alpha may be an antiangiogenic agent in ovarian cancer. The main advantages of the PEG modification are a lengthened plasma half-life and increase in area under the time/concentration curve (AUC). Investigators at M.D. Anderson Cancer Center (Tedjarati et al., 2002) are currently conducting a phase II trial that includes adaptive dose allocation to three doses, 1.0, 1.25 and 1.5 μg/kg/week.

One course of the therapy goes over 4 weeks, consisting of a weekly injection of PEG

Intron. Injections are subcutaneous, and after the first injection, the drug is provided to the patients for self–administration. There are no within patient dose modifications. The protocol design decides on the dose for each new patient and this dose remains the same for the 4 week treatment. After 8 weeks a response is measured. Responses are classified as complete remission (CR), partial response (PR), stable disease (SD) and increasing disease (ID). Previous studies demonstrated good tolerability of PEG Intron, with mild to moderate flu-like symptoms in the majority of cases. Besides some haematologic changes no other significant adverse events to PEG Intron administration have been reported for the dose levels considered in this trial.

3 Trial Design

3.1 The Sequential Decision Problem

Suppose we have treatments $t = 1, 2, \ldots, T$ and responses $r = 1, 2, \ldots, R$ and a set $\mathcal{V}$ of possible utility functions, assuming that the investigator can not further narrow down the utility set. For any $v \in \mathcal{V}$, the value $v(t, r)$ denotes the utility of consequence r under treatment t, according to v. We fix a maximum number N of patients to be enrolled in the trial. At every stage $n = 0, 1, \ldots, N-1$ of the trial we have to make a two-step decision. First we decide whether or not to continue the trial for one more period. If we decide to continue we select a dose to assign to the next enrolled patient. If we decide to stop we decide upon the best dose that is recommended for further development. The latter decision includes $t = 0$ if we do not wish to recommend any of the proposed treatments for further development, i.e., recommend no treatment or current standard of care. Figure 1 illustrates the decision process.

– **Place Figure 1 around here.**

The following discussion will necessarily involve extensive notation for utilities and expected utilities with different levels of expectations and maximizations carried out. We already have introduced $v(t, r)$ for the utility corresponding to the response for one patient. We will need notation for the utility of the entire trial. We will use $u(\cdot)$ for the utility of the entire trial, and $U_n(\cdot)$ for the corresponding expected utility with the expectation taken with respect to future responses r. Finally, we will use $U_n^*(\cdot)$ for the expected utility after substituting the optimal treatment decision at time n. Details are defined below when the various utility functions appear first in the argument.

Assume for the moment that a single utility function $v(t, r) \in \mathcal{V}$ is chosen for the utility of consequence r under treatment t. Based on v we now define a utility function u for the entire trial. Let t_i and r_i be the assigned treatment and the recorded response for the ith patient, $i = 1, \ldots, N$. If the trial is stopped at stage n we let t_{n+1} indicate the recommended treatment. If $n < N$, we assume that t_{n+1} will be used for the remaining $N - n$ patients. Let $d_n = 1$, $n = 0, \ldots, N-1$, indicate that the trial is stopped at stage n and $d_n = 0$ otherwise. Together, the pair (d_n, t_{n+1}) constitutes the decision at any stage n. If $d_n = 0$, the trial continues and the next patient is assigned to treatment $t_{n+1} \in \{1, 2, \ldots, T\}$. If $d_n = 1$ the trial stops and $t_{n+1} \in \{0, 1, \ldots, T\}$ is the recommended treatment. Let $t_{1\ldots n} = (t_1, t_2, \ldots, t_n)$ be the sequence of treatments and $r_{1\ldots n} = (r_1, r_2, \ldots, r_n)$ be the sequence of responses for the first n patients in the trial, and $r_{m\ldots n} = (r_m, \ldots, r_n)$. We introduce $i = N + 1$ as a generic future patient and define the utility $u(d_n = 1, t_{n+1}, t_{1\ldots n}, r_{1\ldots N+1}, v)$ for the entire trial as a weighted average of the utilities v for each of the patients in the trial plus the utility of the recommended treatment for future patients. The option of recommending no treatment is

included as $t_{n+1} = 0$.

$$u(d_n = 1, t_{n+1}, t_{1\ldots n}, r_{1\ldots N+1}, v) = \\ = \alpha\ v(t_{n+1}, r_{N+1}) + (1-\alpha)\frac{1}{N}\left\{\sum_{i=1}^{n} v(t_i, r_i) + \sum_{i=n+1}^{N} v(t_{n+1}, r_i)\right\}, \quad (1)$$

where $\alpha \in [0, 1]$ is the relative weight of the benefit for future patients treated with the recommended treatment versus the patients in the trial. A special case of (1) is u equal to the total number of responses, for patients in and beyond the trial (Berry and Stangl, 1996). A default choice for the weight is $\alpha = 1/(N+1)$. The hypothetical future patient, beneficiary of the recommended treatment of the trial, is just like any other patient in the trial. This is what we used in our implementation. The optimal decision t_{n+1} is found by computing the expected utility

$$U_n(d_n = 1, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v) = E\{u(d_n = 1, t_{n+1}, r_{n+1\ldots N+1}, t_{1\ldots n}, r_{1\ldots n}, v) \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n}\} \quad (2)$$

for all possible decisions t_{n+1}. The expectation is with respect to the future outcomes $r_{n+1\ldots N+1}$. The relevant distribution is the posterior predictive distribution of $r_{n+1\ldots N+1}$ under the treatment t_{n+1}, given the current history, $p(r_{n+1\ldots N+1} \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n})$. However, since (1) is a weighted sum over $v(\cdot, r_i)$, the only relevant probabilities are the marginal posterior predictive probabilities $p(r_i \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n})$ for each r_i, $i = n+1, \ldots, N+1$. Note that $p(r_i \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n})$ is identical across $i = n+1, \ldots, N+1$. We generically denote the common distribution by $p(r \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n})$ and evaluate U_n as

$$U_n(d_n = 1, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v) = (1-\alpha)\frac{1}{N}\sum_{i=1}^{n} v(t_i, r_i) + \\ \sum_{r=1}^{R}\left\{\left(\alpha\ + (1-\alpha)\frac{N-n}{N}\right) v(t_{n+1}, r)\right\} p(r \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n}).$$

Maximize $U_n(\cdot)$ in t_{n+1} to find the optimal decision t^*_{n+1}. The expected utility of continuation ($d_n = 0$) is more difficult. To compute the value of continuation we need to know the optimal

decision in the next period $n+1$, and so forth, until the horizon N, when continuation is not possible. Evaluating expected utilities in such sequential decision problems is solved by backward induction (DeGroot 1970, chapter 10). The algorithm starts at time N, when we have no option but to stop the trial and maximize expected utility. The overall utility $U_N^*(t_{1\ldots N}, r_{1\ldots N}, v)$ at time N is equal to the expected utility of stopping and choosing the optimal treatment:

$$U_N^*(t_{1\ldots N}, r_{1\ldots N}, v) = \max_{t_{N+1}=0,1,\ldots,T} U_N(d_N = 1, t_{N+1}, t_{1\ldots N}, r_{1\ldots N}, v).$$

We use U_n^* to denote the value of the optimal decision at time n. At any interim stage $n < N$ the utility of continuing the trial, $d_n = 0$, with treatment t_{n+1} is

$$U_n(d_n = 0, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v) = \sum_{r_{n+1}=1}^{R} U_{n+1}^*(t_{1\ldots n+1}, r_{1\ldots n+1}, v)\, p(r_{n+1} \mid t_{1\ldots n+1}, r_{1\ldots n}). \quad (3)$$

The overall utility is

$$U_n^*(t_{1\ldots n}, r_{1\ldots n}, v) = \max_{d_n=0,1} \{ \max_{t_{n+1}=1,2,\ldots,T} U_n(d_n, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v) \}. \quad (4)$$

The two steps (3) and (4) define the alternating sequence of expectation and maximization characteristic for the backward induction algorithm. The double maximization in (4) is an artifact of our notation, splitting the decision at time n into the pair (d_n, t_{n+1}).

The outlined backward induction algorithm provides, in theory, the optimal design for a given utility function v. In practice, however, backward induction even for moderate N is only computationally feasible with a small number T of treatments and a small number of responses R, say $T \leq 2$ and $R \leq 3$ (see Berry and Stangl, 1996). In general it is not feasible to implement backward induction, let alone the calculation of non dominated sets of treatments discussed below. We therefore propose to use a 2-step look-ahead procedure (Berger, 1985, chapter 7). That is, at any time n assume the trial is finishing after two more patients are enrolled, and perform backward induction under that assumption, i.e., starting with time $n+2$.

3.2 Randomization

We now relax the setup to consider a family of utility functions $v \in \mathcal{V}$. We modify the optimal design described earlier in two directions. First, instead of selecting one unique optimal dose t^*_{n+1} at each time we report the set A^* of all non-dominated doses and allow randomization over A^* to assign a dose to the next patient. Second, we only stop, $d_n = 1$, if stopping dominates continuation under all $v \in \mathcal{V}$.

In summary, we propose the following design with randomized dose allocation

1. For all $v \in \mathcal{V}$, calculate $U_n(d_n, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v)$ using backward induction, i.e., evaluate (3) and (4) for $N, N-1, \ldots, n$.

2. If stopping, $d_n = 1$, dominates, i.e.,

$$\max_{t_{n+1}=0,1,\ldots,T} U_n(d_n = 1, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v) \geq \max_{t_{n+1}=1,2,\ldots,T} U_n(d_n = 0, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v)$$

   for all $v \in \mathcal{V}$, then stop the trial. Otherwise, randomize the next patient among the non-dominated set of treatments, using $U_n(d_n = 0, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v)$ to define the non-dominated set.

3. When the trial stops report the non dominated set A^* of treatments, using $U_n(d_n = 1, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v)$ to define the non-dominated set.

The main features of the proposed design are as follows. Treatments are included or excluded into the randomization based on their performance provided by all available data. No patients are assigned to clearly inferior (dominated) treatments. Early stopping is implemented. Multiple winners are possible. When stopping, the cardinality of the non dominated set may be greater than 1. The first three characteristics have been described as desirable features to be included in a clinical trial design (see Pocock, 1983, chapter 7). The fourth allows for the commonly not considered conclusion that treatments may be considered equivalent.

In general, it might not be trivial to find the non dominated set of treatments at each stage, making even a 2-step look-ahead scheme not feasible. Note that, in principle we would need to calculate a backward–induction for each $v \in \mathcal{V}$. Even for a finite action space, an algorithm for finding non dominated sets in general may be computationally demanding (see Rios-Insua *et al.*, 1997). However, by restricting the utility set $\mathcal{V}$ to have a small cardinality with some fixed structure, calculations may be feasible. A *shuffled* utility set is such that $\mathcal{V} = \{v(t,r) = c_t v_M(t,r) + (1-c_t)v_m(t,r) : c_t = 0,1\}$, for $v_M(t,r) \geq v_m(t,r)$ for all $t = 0,1,2,\ldots,T$, $r = 1,2,\ldots,R$. A shuffled utility set is compatible with the idea of having utilities that vary independently within maximum and minimum values. That is $v(t,r) \in \{m_{tr}, M_{tr}\}$, with no further structure. Thus $\mathcal{V}$ includes 2^{T+1} possible utility functions. Nonetheless, some further structure may be necessarily in particular examples with utilities varying in more systematic ways. For example, $v(t,r) = p\, v_M(t,r) + (1-p)\, v_m(t,r)$ for $p \in [0,1]$. In such problems, the solution assuming a shuffled utility set provides a reasonable approximation to the optimal solution.

We consider a shuffled utility set in the PEG intron case study below. A detailed algorithm for the general k-arm clinical trial design is presented in the Appendix.

4 The PEG Intron Trial – Results

We consider the Phase I/II trial described in Section 2. The protocol allows for a maximum of $N = 100$ patients, $T = 3$ dose levels and $R = 3$ responses. The dose levels are 1.0 μg/kg/wk ($t = 1$), 1.25 μg/kg/wk ($t = 2$) and 1.5 μg/kg/wk ($t = 3$). Recall that $t = 0$ indicates no treatment. We code responses as $r = 1$ for CR/PR, $r = 2$ for SD, and $r = 3$ for ID.

For each treatment we assume a multinomial sampling model $p(r_i|t_i = j) = Mn(\pi_{j1}, \pi_{j2}, \pi_{j3})$, with a conjugate Dirichlet prior $(\pi_{j1}, \pi_{j2}, \pi_{j3}) \sim Dir(\alpha_{j1}, \alpha_{j2}, \alpha_{j3})$. The prior model as-

sumes independence across treatments $j = 1, \ldots, T$. Alternatively, a dose-response curve for response probabilities across treatments could be assumed. This would allow borrowing strength across treatments and to enforce monotonicity if appropriate. However, the proposed design approach remains the same, independently of the underlying probability model. All we need to assume of the probability model is that it be possible to compute the posterior predictive probabilities $p(r_{n+1} \mid t_{n+1}, t_{1\ldots n}, r_{1\ldots n})$. For the conjugate multinomial/Dirichlet model these posterior predictive probabilities are evaluated analytically. For more general models the evaluation might require numerical integration techniques.

4.1 The Utility Set

Consideration of the utility set $\mathcal{V}$ is a delicate matter. The strategy we follow here is to establish the utility in terms of Quality Adjusted Life Years (QALY, see Kaplan, 1995), relative to failure response. For treatment t and response r, let a_{tr} be the life expectancy (in weeks, months, etc.), and $p_{tr} \in [0, 1]$ the quality of life, for a hypothetical patient with the eligibility characteristics of the trial. The exact value of a_{tr} and p_{tr} are difficult to fix. Instead we allow them to take on maximum and minimum values, that is $a_{tr} \in \{a_{tr}^m, a_{tr}^M\}$ and $p_{tr} \in \{p_{tr}^m, p_{tr}^M\}$. The ratio $a_{tr}p_{tr}/a_{t3}p_{t3}$ represents the relative QALY with respect to failure response ($r = 3$). We will use it as utility for response r to treatment t, i.e., $v(t, r) = a_{tr}p_{tr}/a_{t3}p_{t3}$.

Since no considerable adverse events are expected for the selected three doses, we set $p_{tr}^m = p_{tr}^M = 1$ and also $a_{tr}^m = m_r$ and $a_{tr}^M = M_r$. That is, utility is solely dependent on relative QALY's of responses. Setting $m_3 = M_3 = 1$ the range of utilities then simplifies to $v(t, r) = a_{tr}$, with $a_{tr} \in \{m_r, M_r\})$, $t \in \{0, 1, 2, 3\}$ and $r \in \{1, 2, 3\}$. For $r = 1, 2$, we define $m_1 = 1.75$, $M_1 = 2$, $m_2 = 1.2$ and $M_2 = 1.5$. This reflects a judgment of a 75 – 100% increase in QALY for CR/PR relative to ID, and 20 – 50% relative increase for a SD

outcome.

4.2 Operating Characteristics

The proposed approach is based on an expected utility argument, leading to an algorithm that decides treatment allocation and stopping to maximize an underlying utility function in expectation under the appropriate posterior predictive probabilities. While we consider this a desirable property, we are also concerned about frequentist properties of the derived decision rules. This leads us to investigate frequentist operating characteristics of the proposed trial design. The use of frequentist properties to validate Bayesian inference is common practice in the context of medical decision making. Thall, Simone and Estey (1995), who use frequentist error rates to calibrate tuning parameters in a Bayesian decision rule, is a typical example. We consider six different scenarios, i.e., assumed values for the multinomial probabilities $(\pi_{j1}, \pi_{j2}, \pi_{j3})$, $t = 0, \ldots, 3$. Under each assumed scenario we generated 5,000 possible realizations of the entire trial. Results are summarized in Tables 1 through 3. The tables report for each scenario the assumed true multinomial cell probabilities, the average number of patients allocated to each treatment, and for each treatment the probability of selecting the treatment in the recommended (non-dominated) set upon termination of the trial. Here, the average and the probability are with respect to repeated simulation under the respective assumed true scenario, i.e., frequentist expectations and probabilities.

– Place Tables 1 through 3 around here.

In all scenarios in Tables 1 through 3 the utility set is as described in Section 4.1, the prior parameters for the experimental treatments are $\alpha_{tr} = 1/3$, $r = 1, \ldots, R$ and $t = 1, \ldots, R$. The prior parameters for no treatment ("standard of care"), $t = 0$, are $(\alpha_{01}, \alpha_{02}, \alpha_{03}) = (5, 5, 90)$. In all six scenarios the assumed true parameters for the simulation are $(\pi_{01}, \pi_{02}, \pi_{03}) = (0.05, 0.05, 0.90)$ for no treatment. Thus the chosen prior parameters

represent a vague prior for the response probabilities under the experimental treatments and a highly informative prior for standard of care. A maximum of $N = 100$ patients is set for all scenarios.

Scenario 1 represents the case where all experimental treatments are ineffective. There is no difference between treatments and standard of care. Patients are randomly allocated between treatments. The average trial length is 87.8 patients. The second scenario represents the case where the standard of care is better (4–7% increase in QALY) than any of the experimental treatments (0.8% to 1% increase in QALY). Patients are randomly allocated between treatments, with an average trial length of 75.2 (with a high probability of stopping early, 95%). As desired, 92% of the cases $t = 0$ is included in the recommended set. Scenarios 3 and 4 represent two cases where one treatment is superior. In both cases, randomization is more likely to include only the superior treatment, and the superior treatment is commonly included in the recommend set (88% and 97%). Scenario 5 contains two non-dominated treatments. These treatments are far more often included in the randomization than the dominated treatments, and with high probability they are included in the recommended set. The trial length is shorter, with a probability of 35% of stopping early. Finally, Scenario 6 is set up with treatment 3 superior to the other treatments (17% to 26% increase in QALY compared to 5% to 8% increase in QALY). We find an average number of patients of only 60.4, and a probability of stopping early of 70%.

The proposed randomization scheme is adaptive. A treatment is dropped when it is found dominated by other treatment, but possibly included in the randomization again at a later time when posterior probabilities change and the treatment reenters the non-dominated set A^*. Figure 2 shows an illustration of how treatments enter and leave the non-dominated set A^*, using a simulation under Scenario 6.

5 Discussion

We presented a Bayesian approach for phase II clinical trials including randomization. The mechanism to formally construct randomization is to stop short of identifying a single utility function and proceed with a set of possible utilities. Treatments are randomly allocated within the non-dominated set. The non-dominated set is computed using myopic sequential design. Early stopping is included as a possible action in the sequential design. In this paper we used a simple multinomial model. More complicated models may be used with no changes in the sequential algorithm.

Here we assume to have one single prior distribution π_0. More generally, π_0 may be considered to vary within some prior set $\mathcal{D}_0$ (see Chaloner, 1996). The same line of argument may be followed as above, with the premise that domination is now considered for all $v \in \mathcal{V}$ and all $\pi_0 \in \mathcal{D}_0$ (see Rios–Insua *et al.*, 2000). Conceptually the design remains basically the same, however with the substantial increase in implementation difficulties.

We have considered that patients are accepted one at a time and, ideally, their response is obtained before the next patient is considered. A more generally applicable strategy would be to accept patients in blocks or batches (see Pocock, 1977). This mitigates problems arising from delayed responses and other logistical complications.

Interested readers can obtain copies of the programs used to implement the examples in this paper. A program written in R, for illustration and comparison purposes is available from *http://www.cimat.mx/∼jac/software.html*. For computing intensive simulations, a C++ program is available. Contact Kyle Wathen *wathen@odin.mdacc.tmc.edu*.

6 Acknowledgments

J.A. Christen is partially funded by a CONACYT-NSF grant. This work was completed while J.A. Christen was visiting MDACC. Peter Müller was partially supported by NIH/NCI

under grant R21 CA97534-01.

References

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

Berry, D.A. (1993). "A case for Bayesianism in clinical trials". *Statistics in Medicine*, **12**, 1377–1393.

Berry, D.A. and Stangl, D.K. (1996). "Bayesian methods in health-related research". in: *Bayesian Biostatistics*, Berry, D.A. and Stangl, D.K. (ed.), Marcel Dekker: NY. 3–66.

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

Bryan, J. and Day, R. (1995). "Incorporating toxicity considerations into the design of two–stage phase II clinical trials". *Biometrics*, **51**, 1372–1383.

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Second Edition. Springer-Verlag: NY.

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

Chaloner (1996). "Elicitation of prior distributions". in: *Bayesian Biostatistics*, Berry, D.A. and Stangl, D.K. (ed.), Marcel Dekker: NY. 141–156.

DeGroot, M.H. (1970). *Optimal statistical decisions*. McGraw-Hill:NY.

Hasselblad, V. and Jarabek, A.M. (1996). "Dose-response analysis of toxic chemicals". in: *Bayesian Biostatistics*, Berry, D.A. and Stangl, D.K. (ed.), Marcel Dekker: NY. 235–259.

Kaplan, R.M. (1995). "Utility assessment for estimating quality-adjusted life years". in: *Valuing health care*, Sloan, F.A. (ed.), Cambridge University Press: NY. 31–60.

Lavori, P.W. and Dawson, R. (2000). "A design for testing clinical strategies: biased adaptive within-subject randomization". *Journal of the Royal Statistical Society – A*, **163**(1), 29–38.

Palmer, J.L. and Müller, P. (1998). "Bayesian optimal design in population models for haematologic data". *Statistics in Medicine*, **17**(14), 1613–1622.

Pocock, S.J. (1977). "Group sequential methods in the design and analysis of clinical trials". *Biometrika*, **64**(2), 191–199.

Pocock, S.J. (1983). *Clinical trials*. John Wiley & Sons: Chichester.

Rios-Insua, D., Martin, J., Proll, L., French, S. and Salhi, A. (1997). "Sensitivity analysis in statistical decision theory: a decision analytic view". *Journal of Statistical Computing and Simulation*, **57**, 197–218.

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

Tedjarati S, Baker CH, Huang S, Wolf JK, Apte S, Killion JJ, Fidler IJ. (2002). "Synergistic Therapy of Human Ovarian implanted orthotopically in nude mice by optimal biological dose of pegylated interferon-a combined with paclitaxel." *Clinical Cancer Research*, **8**, 2413-2422.

Thall, P.F. (2001). "Bayesian clinical trial design in a cancer center". *Chance*, **14**(3), 23–28.

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

Wolfson, L.J., Kadane, J.B. and Small, M.J. (1996). "Expected utility as a policy making tool: an environmental health example". in: *Bayesian Biostatistics*, Berry, D.A. and Stangl, D.K. (ed.), Marcel Dekker: NY. 261–277.

Appendix

Let π_{tr} be the probability of result r under treatment t; $t = 0, 1, 2, \ldots, T$ and $r = 1, 2, \ldots, R$. We assume that $(\pi_{t1}, \pi_{t2}, \ldots, \pi_{tR})$ has a Dirichlet distribution with parameters $(\alpha_{t1}, \alpha_{t2}, \ldots, \alpha_{tR})$. At any given time in the trial, let $S[t, r]$ denote the number of responses r observed under treatment t. The $T \times R$ matrix S is a sufficient statistic for the posterior distribution on the multinomial probabilities π_{tr} given all data observed so far. Substituting S for $(t_{1\ldots n}, r_{1\ldots n})$ we use notation like $U_n(d_n, t_{n+1}, S, v)$ for the expected utilities $U_n(d_n, t_{n+1}, t_{1\ldots n}, r_{1\ldots n}, v)$, etc. Also, we use the notation $S+(t_1, r_1)$ (and $S+(t_1, r_1)+(t_2, r_2)$) to indicate S plus a treatment t_1 with a response r_1 (and a treatment t_2 with a response r_2). We use $p(r \mid t, S)$ to denote the predictive probability of obtaining a response r, using treatment t, given the current summary statistic S. We find

$$p(r \mid t, S) = \frac{\alpha_{tr} + S[t, r]}{\sum_{r'=1}^{R} \alpha_{tr'} + S[t, r']},$$

for $t = 1, 2, \ldots, T$ and $r = 1, 2, \ldots, R$.

Algorithm

At any given time we compute the 2-step myopic sequential design. The following steps parallel the structure shown in the decision tree Figure 1. To clarify structure we first

discuss a straightforward implementation of backward induction for a 2-step myopic design. Many simplifications are possible in the actual implementation. Details are shown in the following subsection. To simplify notation we will use subindices 0,1,2 and 3 instead of n, $n+1$, $n+2$ and $n+3$, respectively. Recall that $U_n^*(\cdot)$ denotes expected utility achieved by the optimal decision (d_n^*, t_{n+1}^*) at time n.

1. Consider continuing, $d_0 = 0$, with $t_1 = 1, \ldots, T$

   1.1. For $r_1 = 1, \ldots, R$

      1.1.1. Consider continuation, $d_1 = 0$:

         (a) Consider continuation with $t_2 = 1, \ldots, T$

$$U_1(d_1 = 0, t_2, S + (t_1, r_1), v) = \sum_{r_2} p(r_2 \mid t_2, S + (t_1, r_1)) \times \max_{t_3} U_2(d_2 = 1, t_3, S + (t_1, r_1) + (t_2, r_2))$$

         (Decision $d_2 = 1$ is forced by the myopic nature of the 2-step look-ahead procedure.)

         (b) Consider continuation with optimal $t_2 = t_2^*(0)$:

$$U_1(d_1 = 0, t_2^*(0), S + (t_1, r_1), v) = \max_{t_2} U_1(d_1 = 0, t_2, S + (t_1, r_1), v).$$

      1.1.2. Consider stopping, $d_1 = 1$. Find the optimal $t_2 = t_2^*(1)$ as

$$U_1(d_1 = 1, t_2^*(1), S + (t_1, r_1)) = \max_{t_2} U_1(d_1 = 1, t_2, S + (t_1, r_1), v)$$

      with $t_2 \in \{0, 1, \ldots, T\}$.

      1.1.3. Optimal utility at $n+1$:

$$U_1^*(S + (t_1, r_1)) = \max_{d_1} U_1^*(d_1, t_2^*(d_1), S + (t_1, r_1)).$$

1.2. Expected utility at n when continuing with t_1

$$U_0(d_0 = 0, t_1, S) = \sum_r U_1^*(S + (t_1, r_1))\, p(r_1 \mid t_1, S).$$

2. Expected utility at n when continuing. Find the optimal treatment t_1^*:

$$U_0(d = 0, t_1^*(0), S) = \max_{t_1} U_0(d_1 = 0, t_1, S).$$

3. Expected utility at n under stopping:

$$U_0(d = 1, t_1^*(1), S) = \max_{t_1} U_0(d_0 = 1, t_1, S).$$

4. Optimal utility at n:

$$U_0^*(S) = \max_d U_0(d, t_1^*(d), S).$$

Implementation Notes

Actual implementation of the algorithm simplifies significantly by noting that many intermediate results need not be saved. In particular, $U_1(d_1 = 0, t_2, S + (t_1, r_1), v)$ is only required to find the optimal $t_2^*(0)$ in $U_1(d_1 = 0, t_2, S + (t_1, r_1), v)$. It can be evaluated as needed, and need not be saved. Evaluation can be implemented as a function. And $U_1(d_1 = 0, t_2^*(0), S + (t_1, r_1), v)$, in turn, is only required to evaluate $U_0(d_0 = 0, t_1, S)$. It can be evaluated when needed.

Evaluating $U_n(d_n = 1, \ldots)$:

The (expected) utility of stopping the trial, and deciding in favor of t, given S, is

$$\begin{aligned} U_n(d_n = 1, t_{n+1} = t, S, v) = \\ = \alpha\ V(t, S, v) + (1 - \alpha)\frac{1}{N}\left\{(N - n)V(t, S, v) + \sum_{t'=1}^{T}\sum_{r=1}^{R} S[t', r]v(t', r)\right\}, \end{aligned}$$

where $V(t, S, v) = \sum_{r=1}^{R} v(t, r)p(r \mid t, S)$, is the current expected utility of treatment t.

Evaluating $U_n(d_n = 0, \ldots)$:

At any time n, $n = 0, \ldots, N-2$, evaluate $U_0(d_0 = 0, S, t_1, v)$ for $t_1 \in \{1, \ldots, T\}$, using the nested steps 1-3, below,

1. *Expected utility under $d_0 = 0$:* Implement step 1.2. in the algorithm by evaluating for $t_1 \in \{1, \ldots, T\}$:

$$U_0(d_0 = 0, t_1, S, v) = \sum_{r_1=1}^{R} p(r_1 \mid t_1, S)\, U_1^*(S + (t_1, r_1), v).$$

   For each r_1, use step 2, below, to evaluate $U_1^*(S + (t_1, r_1), v)$.

2. *Optimal expected utility at time $n+1$:* For each r_1 evaluate $U_1^*(S + (t_1, r_1))$:

$$U_1^*(S + (t_1, r_1), v) = \max \left\{ \max_{t_2=1,2,\ldots,T} U_1(d_1 = 0, t_2, S + (t_1, r_1), v), \right.$$
$$\left. \max_{t_2=0,1,\ldots,T} U_1(d_1 = 1, t_2, S + (t_1, r_1), v) \right\}.$$

   This implements steps 1.1.3 and 1.1.2 in the algorithm.

   For each t_2, use step 3, below, to evaluate $U_1(d_1 = 0, t_2, S + (t_1, r_1), v)$.

3. *Expected utility under $d_1 = 0$:* Implement step 1.1.1. by evaluating for each $t_2 \in \{1, \ldots, T\}$:

$$U_1(d_1 = 0, t_2, S + (t_1, r_1), v) = \sum_{r_2=1}^{R} p(r_2 \mid t_2, S + (t_1, r_1))$$
$$\max_{t_3 \in \{0,1,\ldots,T\}} U_2(d_2 = 1, t_3, S + (t_1, r_1) + (t_2, r_2), v).$$

At stage $n = N-1$, evaluate $U_0(d_0 = 0, S, t_1, v)$ for $t_1 \in \{1, \ldots, T\}$, using step 1, above, substituting $U_1^*(S + (t_1, r_1), v) = \max_{t_2=0,1,\ldots,T} U_1(d_1 = 1, t_2, S + (t_1, r_1), v)$. This is simply step 2 without the option of further continuation.

Optimal decision:

For any stage $n < N$, if $\max_{t_1=0,1,\ldots,T} U_0(d_0 = 1, t_1, S, v) \geq \max_{t_1=1,2,\ldots,T} U_0(d_0 = 0, t_1, S, v)$, then stop the trial and recommend the non-dominated set of treatments, using $U(d = 1, t_1 = t, S, v)$ to define the non-dominated set. Otherwise, randomize a new patient within the non-dominated set of treatments according to $U_0(d_0 = 0, t_1 = t, S, v)$ (not including $t = 0$), increase n by one, add the patient's response to S, and go back to 1. If $n = N$, stop the trial and recommend the non-dominated set of treatments according to $U_0(d_0 = 1, t, S, v)$. An efficient algorithm to calculate non-dominated sets may be found in Rios-Insua *et al.* (1997).

Figures and Tables

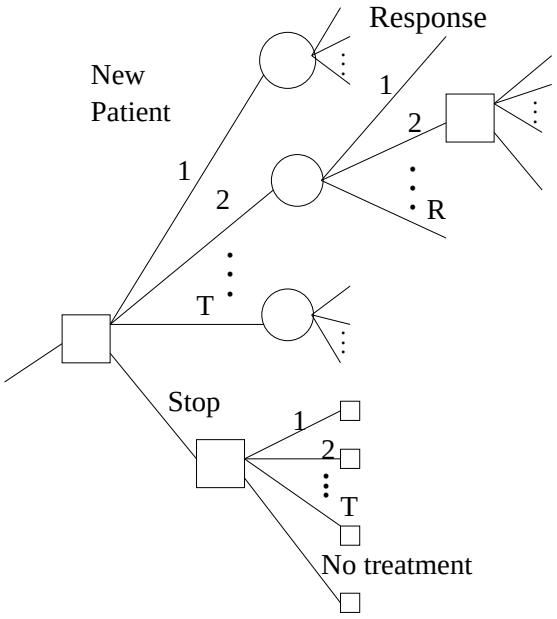


Figure 1: Decision tree at stage n.

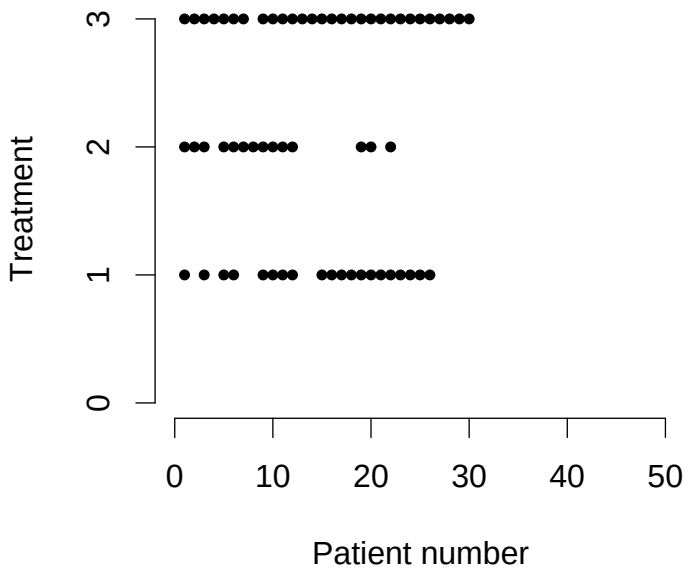
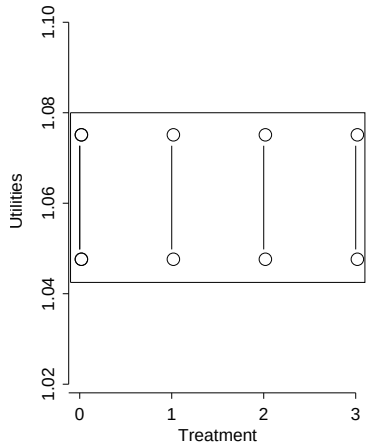


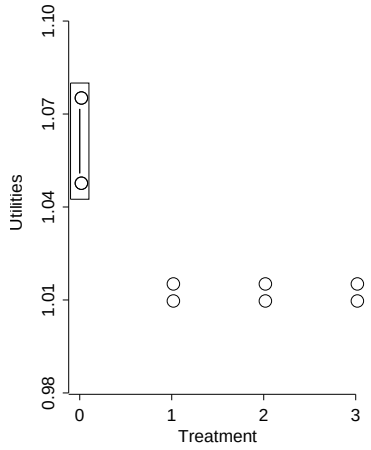
Figure 2: Treatments taken into randomization (dots) as a function of time. Time is measured by the number of enrolled patients. At each time all treatments in the non-dominated set A^* are marked with a bullet. Simulation was done under Scenario 6, with treatment 3 dominating $t = 0, 1, 2$ in the true simulation model (see Table 1). Note how treatments may leave and reenter again the non-dominated set. At the end of this simulated trial, the superior treatment $t = 3$ is the only one considered for allocation, and the trial stops early, well before the boundary of $N = 100$ patients.

Scenario 1



Trt	True $(\pi_{t1}, \pi_{t2}, \pi_{t3})$ (CPR, SD, ID)	$\bar{n}_t$ (St.dev.)	% RS
0	(5, 5, 90)	0 (0)	9
1	(5, 5, 90)	29.2 (14.6)	54
2	(5, 5, 90)	29.2 (14.4)	54
3	(5, 5, 90)	29.3 (14.5)	54
$\bar{n}$ (SD)		87.8 (28.8)	
% Stopped Early		31	

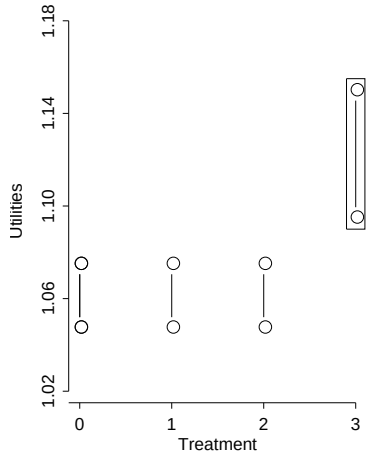
Scenario 2



Trt	True $(\pi_{t1}, \pi_{t2}, \pi_{t3})$ (CPR, SD, ID)	$\bar{n}_t$ (St.dev.)	% RS
0	(5, 5, 90)	0 (0)	92
1	(1, 1, 98)	24.1 (8.7)	11
2	(1, 1, 98)	25.1 (8.8)	11
3	(1, 1, 98)	25.0 (8.7)	11
$\bar{n}$ (SD)		75.2 (15.0)	
% Stopped Early		95	

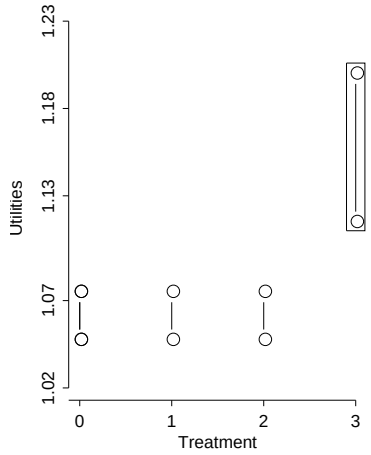
Table 1: Operating characteristics. Treatment 0 is no treatment (standard of care) and treatments 1 - 3 are experimental treatments. In the tables, the second column reports the simulation truth $(\pi_{t1}, \pi_{t2}, \pi_{t3})$, i.e., the assumed probabilities for outcomes CR/PR (CPR in the table), SD and ID (in percent). The column $\bar{n}_t$ reports the average number of patients assigned to each treatment, and % RS is the percentage of simulations when a treatment is contained in the recommend set (RS). For each scenario the graph to the left of the table shows the expected utilities under the (true) parameters assumed in the simulation. The true non-dominated set is indicated by the box around the utility ranges. Note that in all cases no patients are assigned to treatment 0. This is an arbitrary design choice. If desired randomization to placebo could be included in the protocol.

Scenario 3



Trt	True $(\pi_{t1}, \pi_{t2}, \pi_{t3})$ (CPR, SD, ID)	$\bar{n}_t$ (St.dev.)	% RS
0	(5, 5, 90)	0 (0)	92
1	(1, 1, 98)	24.1 (8.7)	11
2	(1, 1, 98)	25.1 (8.8)	11
3	(1, 1, 98)	25.0 (8.7)	11
$\bar{n}$ (SD)		75.2 (15.0)	
% Stopped Early		95	

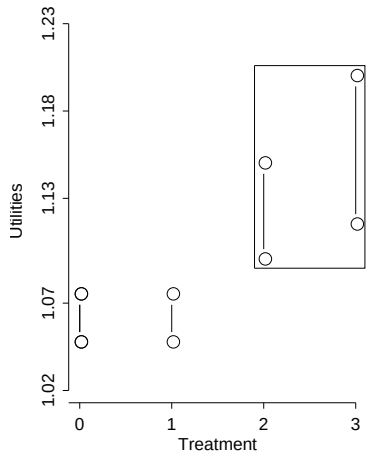
Scenario 4



Trt	True $(\pi_{t1}, \pi_{t2}, \pi_{t3})$ (CPR, SD, ID)	$\bar{n}_t$ (St.dev.)	% RS
0	(5, 5, 90)	0 (0)	1
1	(5, 5, 90)	19.5 (12.5)	19
2	(5, 5, 90)	19.3 (12.2)	10
3	(10, 20, 70)	38.7 (18.7)	97
$\bar{n}$ (SD)		77.4 (29.9)	
% Stopped Early		48	

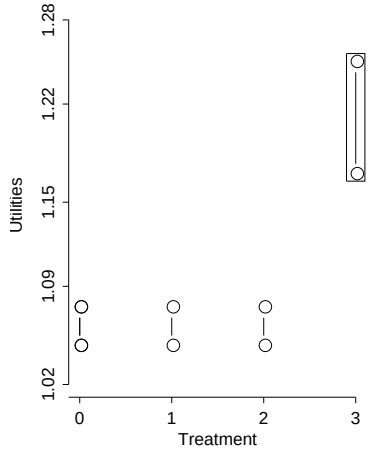
Table 2: Continuation of Table 1, Scenarios 3 and 4.

Scenario 5



Trt	True $(\pi_{t1}, \pi_{t2}, \pi_{t3})$ (CPR, SD, ID)	$\bar{n}_t$ (St.dev.)	% RS
0	(5, 5, 90)	0 (0)	0
1	(5, 5, 90)	16.9 (11.1)	15
2	(10, 10, 80)	29.6 (15.5)	60
3	(10, 20, 70)	37.1 (16.5)	89
$\bar{n}$ (SD)		83.6 (27.6)	
% Stopped Early		35	

Scenario 6



Trt	True $(\pi_{t1}, \pi_{t2}, \pi_{t3})$ (CPR, SD, ID)	$\bar{n}_t$ (St.dev.)	% RS
0	(5, 5, 90)	0 (0)	0
1	(5, 5, 90)	13.2 (9.8)	6
2	(5, 5, 90)	13.4 (10.2)	6
3	(20, 10, 70)	33.7 (22.3)	98
$\bar{n}$ (SD)		60.4 (33.4)	
% Stopped Early		70	

Table 3: Continuation of Table 1, Scenarios 5 and 6.