Adaptive Designs
IMS Lecture Notes - Monograph Series (1995) Volume 25

SEQUENTIAL ELIMINATION PROCEDURES IN CLNICAL TRIALS OF THREE BERNOULLI RESPONSE TREATMENTS

BY CHRISTOPHER R. PALMER

The University of Cambridge

Abstract

Data dependent allocation methods could be advantageously employed in some clinical trials, though in practice, such techniques are rarely used, in part due to their inherent complexity. We consider a practical, decision-theoretic approach for three dichotomous response treatments, using equal allocation until irrevocably dropped from contention. The objective is to maximize the total expected number of successes. We study three elimination procedures: from three treatments to one, both with and without an intermediate pairwise stage; and two serial, pairwise comparisons. The dynamic equations involved are computable for any patient horizon and asymptotic behavior is only marginally worse than if all participants receive superior treatment (in marked contrast to fixed sample size trials). For practical application, prior dependence is effectively removed and smaller horizons are discussed. The procedure with an intermediate stage is preferred, not so much for numerical reasons (by evaluation of rewards and regrets), but for qualitative ethical and practical considerations.

Received October 1992; revised April 1993.

AMS 1991 subject classifications. Primary 62F07, 62L05, 62L10, 62L12.

Key words and phrases. Clinical trials, dynamic programming, ethics, optional stopping.

1. Introduction. The design problem for clinical trials model is notoriously complex, since it requires a delicate balancing of ethical, practical and theoretical considerations. In general, adaptive designs applied to medical trials have the noble aim of minimizing patient suffering while expediting decisions about new treatments. Sometimes, though, ethicists disfavor shifting treatment assignment probabilities within a trial away from equal allocation (by the "if-you-think-one's-better-then-why-randomize?" argument), while practitioners may object if a strategy is so complicated that it is virtually impractical. On the theoretical side, one needs to distinguish in the planning stage of a trial whether its purpose, using terminology from Schwartz, Flamant and Lellouch (1980), is *explanatory* or *pragmatic*. That is, is the trial's objective to estimate treatment efficacies, or merely to identify a potentially beneficial new therapy? The latter is the more appropriate for early phase clinical trials, since it clearly requires fewer patients and seeks only a provisional (as opposed to a confirmatory) answer.

This paper considers three related sequential elimination procedures that are firmly rooted in the pragmatic camp. To circumvent ethical dilemmas, and to keep the designs simple to implement, equal allocation of remaining treatments is used in restricted 3 or 2-armed bandit settings, with immediate Bernoulli responses. This means that the designs discussed are not strictly 'adaptive' in one sense, although one can think of them as adapting (downwards) on the number of treatments remaining according to the results of accruing data. Our approach is decision-theoretic, as per Anscombe (1963), Canner (1970), Colton (1963, 1965) and Sylvester (1988), and involves the exact solution of various dynamic programs. In each case, the equations are computable for arbitrarily large patient 'horizons', N, the total number to be treated during and beyond the comparative stages of each design, a feature not common to all bandit problems. This is a non-trivial consideration and has proven to be a stumbling block in other settings. Armitage (1985), for instance, comments on the two-armed bandit problem for a finite horizon, that "the computation involved is prohibitive except for trivially small horizons". In part, the computability follows from our special choice of a prior distribution which is a generalization of Feldman's (1962) two-point prior in the two treatment setting and for which Berry (1978) demonstrates some sound properties. However, to counteract the problem of specifying a particular prior, a conservative stance is adopted that combines worst-case scenarios across all priors considered. In a related problem, Bather and Simons (1985) similarly

select least favorable parametric values in their Bayes risk minimax approach.

This work extends that of Simons (1986), who discusses the two treatment problem in depth. Each additional treatment adds a new dimension of complexity to the mathematical theory under the approach described herein. Furthermore, in the application of early phase trials, it is very rare to compare more than three treatments at a time, and since this work is primarily aimed at such applications. We do not deal with an arbitrary number of treatments, but instead limit ourselves to precisely three.

2. Preliminary model.

2.1. Theoretical framework. Consider the simple model having a total N patients to be treated and three competing therapies. Triplets of patients are randomized, one on each therapy, until one expects more successes by allocating the best appearing treatment alone for the remaining patients. True success probabilities of the treatments are, naturally, unknown, but we assume a six-point prior distribution putting equal mass on each of the permutations of a, b, and c, where $a \geq b \geq c$ are known. More explicitly, if $\Pi = (\pi_1, \pi_2, \pi_3)$ denotes the vector of true success probabilities, $P\{\Pi = (a,b,c)\} = \cdots = P\{\Pi = (c,b,a)\} = \frac{1}{6}$. This stance corresponds to the physician having no initial preference of treatments. The question becomes "precisely when is it optimal to switch from three treatments to one?"

After observing m triplets, suppose the treatments have accumulated s_1, s_2, and s_3 successes, respectively, where we specify, without loss of generality, $s_1 \geq s_2 \geq s_3$. Then, the posterior distribution is spread over the six points according to readily calculable functions of the non-negative parameters t, j and k defined as follows:

$t \equiv N - 3m$ can be thought of as the "time remaining";

$j \equiv s_1 - s_2$ measures the advantage of the leading over the median treatment; and

$k \equiv s_2 - s_3$ measures the advantage of the median over the worst treatment.

Conditional probability considerations show that the "state" or "point" (t,j,k) so created is Markovian. In general, if both "score differences", j and k, are positive, seven immediately attainable states are possible arising from the next triplet after the state $(t+3,j,k)$. Their corresponding conditional transition probabilities can be evaluated easily

and take a form most conveniently described in terms of a function

$$D(j,k) \equiv (\lambda\mu)^j \left(\mu^k + \lambda^k\right) + \mu^j \left(\mu^k + 1\right) + \lambda^j \left(\lambda^k + 1\right),$$

where

$$\lambda \equiv \frac{a(1-b)}{b(1-a)} \text{ and } \mu = \frac{a(1-c)}{c(1-a)}.$$

Note that when $a > b > c$, we have $\mu > \lambda > 1$.
For example, in obvious notation, while sampling by triplets, for $k > 0$,

$$P\{(t,j,k-1) \mid (t+3,j,k)\} = a(1-b)(1-c)\frac{D(j,k-1)}{D(j,k)}.$$

The other transition probabilities are expressed similarly.

Define $S(t,j,k)$ to be the expected number of successes achieved by switching from three treatments to one at the optimal time, starting from the state (t,j,k). Analogously, define $R(t,j,k)$to be the expected number of successes achieved by switching immediately, starting from (t,j,k). Thus, we are interested in the "reward" of the trial, $S(N,0,0)$. The quantity $R(t,j,k)$ can be written in terms of the parameters as follows:

$$\begin{aligned} R(t,j,k) &= \frac{t}{3D(j,k)}\left[(2a-b-c)(\lambda\mu)^j\left(\mu^k+\lambda^k\right)\right. \\ &\quad \left. +(2b-a-c)\mu^j\left(\mu^k+1\right)+(2c-a-b)\lambda^j\left(\lambda^k+1\right)\right]. \end{aligned}$$

Then, we can express the recurrence relation, that is, the "dynamic equation" that S satisfies for $t \geq 0$, by

(1)

$$S(t+3,j,k) = \max\{R(t+3,j,k), a+b+c+\Delta S(t,j,k)\},$$

where $S(t,j,k) = R(t,j,k)$ for $t = 0$, 1 and 2, and Δ denotes a certain linear operator that is defined, in general, by

(2)

$$\Delta S(t,j,k)$$

$$\equiv \frac{1}{D(j,k)}\left[abc\Psi_1(t,j,k) + (1-a)(1-b)(1-c)\Psi_2(t,j,k)\right],$$

and Ψ_1 and Ψ_2 are functionals given by

$$\Psi_1(t,j,k)$$

$$\equiv \gamma^{-2}\phi(j+1,k) + \gamma^{-1}\{\phi(j+1,k-1) + \phi(j,k+1)\} + \phi(j,k),$$

$$\Psi_2(t,j,k)$$

$$\equiv \gamma^2\phi(j-1,k) + \gamma\{\phi(j,k-1) + \phi(j-1,k+1)\} + \phi(j,k),$$

where $\phi(j,k) \equiv \phi(t,j,k) \equiv S(t,j,k)\,D(j,k)$and $\gamma = a/(1-a)$.

Note that (1) encapsulates the optimal policy, that is, switch to the best appearing treatment as soon as the first component exceeds the second.

2.2. Results. One can assert for given a, b, and c, certain theoretical properties of the optimal "continuation region" C, that is, the set of points (t,j,k) for which it is optimal to continue sampling by triplets. One such intuitive yet non-trivial result is: if (t,j,k) is in C, then so too, for the same a, b, and c, is $(t+1,j,k)$ in C.

With the aid of the computer, one can thoroughly explore C and perform robustness studies in a, b, and c and N. When $j = 0$, it is always optimal to continue sampling by triplets unless t is very small. Table 1 shows which other values of (j,k) continue for various t, for the stated example choice of (a,b,c).

Table 1.
Optimal continuation region for (a,b,c)=(.6, .5, .4)

t	Continue sampling triplets for (j,k) satisfying
30	$j = 1, k = 0$
60	$j = 1, k \leq 2$
90	$j = 1, k \leq 5$ or $j = 2, k = 0$
120	$j = 1, k \leq 13$ or $j = 2, k \leq 1$

The continuation region is not completely, but largely, described by those minimal values of t for which $(t, j, 0)$ is a "stopping point" for a given j. Define, for given (a, b, c) and j a non-negative integer, $\tau_j = \min\{t : (t, j, 0) \in C\}$. Then, the first few values of τ_j for $(a, b, c) = (.6, .5, .4)$ are 6, 24, 69, 159, 312, 567 and 975. The question arises: Are there universal optimal continuation points? That is, is there a maximal C to negate the effect of choosing the prior (a, b, c)? It turns out that no set of values (a, b, c) uniformly minimizes the τ_j's. However, those sets that approximate doing so have $a \simeq 0.5$ and $b = c \simeq 0.47$. Further redefining $\tau_j \equiv \min\{t : (t, j, 0) \in C, \text{ all } (a, b, c)\}$, numerical evidence suggests that the first few values of τ_j are 6, 21, 60, 123, 207, 309, 438 and 585. This means that, regardless of the prior values (a, b, c) assigned, it is always optimal to switch to one treatment when in state (t, j, k) if $t < \tau_j$.

2.3. Conclusions. This preliminary design suffers in that it continues sampling by triplets even in the presence of a clear loser, if the top two treatments are succeeding approximately equally well. For instance, Table 1 reveals that it is optimal to continue when $j = 0$ or 1 and k is large, presumably entirely because there is insufficient information to distinguish between the better two treatments. Therefore, the design only allowing one decision to switch from three treatments to one is clearly unsatisfactory. However, besides introducing the mathematical approach in a straightforward setting, the model suggests that the "$b = c$" case makes the identification of the superior treatment the most difficult. Since this situation also conveniently renders ensuing technicalities manageable, $b = c$ is assumed throughout the rest of the paper. In multiple treatment settings, it is frequently assumed that the inferior treatments are equivalent [see, for example, Rodman (1976) or Zaborskis (1978)].

3. Main model.

3.1. Theoretical framework. We now generalize the preliminary model to allow an additional, intermediate stage during which patients are randomized in pairs to the best two appearing treatments. Otherwise, the objective function and the assumptions remain unchanged, save for now, prior success probabilities of the worst two treatments, b and c, respectively, are equal.

Define $M(t, j, k)$ to be the optimal expected number of successes when currently in state (t, j, k). We are interested in $M(N, O, O)$,

the reward prior to sampling, or else the "regret": $Na - M(N,O,O)$, the expected shortfall compared to knowing and applying the superior treatment throughout the trial. The relevant dynamic equation, for $t \geq 0$, takes the form

(3)

$$M(t+3,j,k) = \max\{a+2b+\Delta M(t,j,k),$$
$$(1-\rho_3)(a-b)S(t+3,j) + \frac{1}{2}(t+3)[a+b-(a-b)\rho_3]\Big\},$$

with $M(t,j,k) = t[b+(a-b)\rho_1]$ for $t = 0, 1$, and 2, where the linear operator Δ is defined in (2), which simplifies because $b = c$. Also, $S(t,j)$ itself satisfies the dynamic equation

(4)

$$S(t+2,y) = \max\left\{\frac{1}{2}(t+2)\tanh\alpha|y|,\right.$$
$$\left. u_y S(t,y-1) + [ab+(1-a)(1-b)]S(t,y) + w_y S(t,y+1)\right\}$$

with $S(t,y) = \frac{1}{2}t\tanh\alpha|y|$ for $t = 0$ and 1, where

(5)

$$\begin{aligned} u_y &= \beta\frac{\cosh(y-1)\alpha}{\cosh y\alpha}, \\ w_y &= \beta\frac{\cosh(y+1)\alpha}{\cosh y\alpha}, \\ \alpha &= \frac{1}{2}\ln\lambda, \beta = \sqrt{ab(1-a)(1-b)}. \end{aligned}$$

$S(t,j)$ represents the optimal reward in a two treatment setting, one of which has success probability a, the other b, when t patients remain and the initial score difference is j. Defining ρ_1, ρ_2, and ρ_3 to be posterior probabilities on the permutations of (a,b,b) arising out of the triplets stage, we have

$$\begin{aligned} \rho_1 &= \rho_1(j,k) = \frac{\lambda^{j+k}}{[\lambda^{j+k}+\lambda^k+1]} \\ \rho_2 &= \rho_2(j,k) = \frac{\lambda^k}{[\lambda^{j+k}+\lambda^k+1]} \\ \rho_3 &= \rho_3(j,k) = \frac{1}{[\lambda^{j+k}+\lambda^k+1]} \end{aligned}$$

and we note that (4) and (5) correspond to equations (9) and (10) in Simons (1986), which in effect, considers the case $\rho_1 = \rho_2 = \frac{1}{2}, \rho_3 = 0$. Equation (3) can be programmed efficiently to handle any (a, b) and arbitrarily large horizons, N.

3.2. Results. The continuation region for sampling by triplets in the main model is characterized by an array of minimal numbers of remaining patients, t, for relevant pairs (j, k). For example, if $(a, b) = (.6, .4)$, then $(j, k) = (2, 1)$ is an optimal continuation point provided the number of patients remaining is at least 153. Further, such minimal t-values for other values of (j, k) can be given in a "policy table".

The reward prior to sampling is also computable. As an example, $M(300, 0, 0) = 169.61$, with an overall success rate of 56.54% when two treatments are 40% successful and one treatment is 60% successful. Thus, the regret is just 10.39 patients, or 3.46% of the 300 entered. Note, by contrast, that in a corresponding fixed sample size trial, allocating 100 patients on each treatment, we would expect only 140 successes, for a regret of 40, or 13.33%. Further results will be discussed in the final section.

Empirical evidence suggests that, in common with other adaptive designs, as N increases, the overall success rate approaches that of the single best treatment. Studies of large N, and also the small to moderate sized samples more likely to be encountered in practical applications, all reveal that the design behaves desirably, while robustness studies indicate the choice and role of N are not critical. Other studies show that the parameter λ has about as much influence as the actual values of a and b, with policy tables being quite robust in these parameters. Finally the choice of (a, b) has more bearing on the speed of stopping the testing stage of the trial than on the final decision identifying the superior treatment.

However, in order to deflate objections to any sensitivity at all with this approach, it is easy to generate a "global policy table" making the choice of a and b irrelevant. This is analogous to searching for the universal optimal continuation points i the preliminary model. It turns out that, once again, no single pair (a, b) gives rise to a smallest set of entries in a policy table, perhaps not surprisingly, due to the complexity of the problem. Nevertheless, one can in principle form a hybrid table from the set of all tables that have minimal entries in any one position. This pointwise minimization suggests that independent of the values of (a, b), with the score differences $(j, k) = (2, 1)$ for example, one should

always terminate allocating by triplets if fewer than 133 patients remain to be treated. Thus, the practical consequence of not knowing, say, $(a, b) = (.6, .4)$ amounts to about 7 more triplets being sampled in the first stage. Further details and simulation results concerning the practical applications of the main model are given in Palmer (1991).

4. Comparative model.

4.1. Formulation. Earlier we made reference to a comparable fixed sample size trial to contrast relative performances. A more equitable comparison, however, can be made with another simple, sequential trial that samples by at most two treatments at a time, effectively two successive, two-treatment sub-trials, but otherwise maintaining the same assumptions as the main model. Thus, in this "comparative model", there remain three treatments, one of success probability a, the others $b\,(a < b)$, and the goal is to maximize the expected number of successes throughout the horizon, N. But now, one randomly chose treatment is temporarily "shelved" and pairs of patients are randomized between the other treatments, until it is optimal to begin sampling the better one paired with the shelved treatment. Ultimately, just the better of these two is allocated to the remaining patients. To mimic clinical practice, the second phase begins with the treatments on an equal footing. We allow for two extreme options: never using, or switching all assignments, to the shelved treatment. Though one would never trust one's prior beliefs to such an extent to invoke either of these options in practice, they do serve to strengthen the comparative model. It is convenient for computational purposes to parameterize in terms of another Markovian state, (t, f, i), where i represents the score difference between the two treatments first used at a time when t patients remain and the more successful treatment has accumulated f failures. Define $B\,(t, f, i)$ to be the optimal expected number of successes attained from (t, f, i), so the quantity of interest for our comparison is the reward $B\,(N, 0, 0)$ for various horizons, N, and each choice of priors (a, b). Analysis is based on a dynamic equation in $B\,(t, f, i)$, involving a maximum of four terms corresponding to the choices of continuing and three ways of terminating the initial stage of testing a pair of treatments. The details are omitted.

4.2. Results. As with the main model, a computer program can be written to cope with arbitrarily large horizons, yet only requiring the storage of small, two-dimensional matrices. The program used here expresses for a specified horizon, N, and priors (a, b), the optimal strategy

in terms of a "policy vector" and also computes the relevant reward. The policy vector is a sequence of minimal values of t such that it is optimal to continue sampling with the original pair of chosen treatments so long as at least Θ_i patients remain and the current score difference is $i, (i = 0,\ 1,\ 2,\ \ldots)$, where Θ_i is the ith term in the said sequence. By way of illustration, when $(a, b) = (.6, .4)$, the policy vector begins:

$$(2,\ 16,\ 50,\ 118,\ 258,\ \ldots),$$

meaning one should continue with the initial treatment pair if, say, $i = 4$ and $t \geq 258$. Equally, if $i = 4$ and fewer than 258 remain, one should terminate this stage of the trial.

Interestingly, large horizon analyses suggest that the first stage of the trial behaves as though only those two treatments were present and that one has success probability a, the other b. Simons (1986) reports, in the two-treatment case when $(a, b) = (.75, .25)$, a policy vector beginning (2, 23, 190, 1652, ...) which agrees exactly with results from our comparative model. Intuitively, this is most likely because one expects an "a vs. b" trial to terminate sooner than a corresponding "b vs. b" trial between two treatments. So in our situation, in spite of the one-third probability of commencing with "b vs. b", it tends to consider itself first sampling "a vs. b". More importantly, in terms of the reward, the comparative model maintains the property seen earlier of minimizing the regret asymptotically. For example, $B\,(10{,}000,\ 0,\ 0) = 5{,}976.69$, or 59.77% when $(a, b) = (.6, .4)$. Furthermore, this model displays similar patterns of robustness in its parameters to the main model and is equally amenable to the construction of a minimax-type policy vector that is independent of the selection of a and b.

5. Discussion. Table 2 compares rewards of the main and comparative models, with a moderate horizon and for illustrative purposes only, a fixed sample size design in which one third of the patients are allocated equally to each treatment. Note that each case has a maximal 180 expected number of successes.

There are at least two observations one can make from these results. First, for given a and b, it is clear that the numeric differences between the sequential models are negligible. Perhaps the similarity is due in part at least to our "$b = c$" restriction tending to reduce the contrast in their expected performances. In these and all other examples studied, the main model just eclipses the comparative model, but

never significantly. Therefore, in view of other above-mentioned similarities, one must consider qualitative differences, not just quantitative ones, in choosing between the models. Ethical and practical issues, such as the demerits of shelving a potentially beneficial treatment and recalling that one model incorporates two unrealistic options, lead one to conclude that while both perform well from theoretical standpoints, the main model is preferable. That is, if conducting a trial among three treatments, it is better to begin sampling in triplets with all the treatments than with a pair alone.

Table 2.
Comparison of rewards and percentage regrets, $N = 300$

Model	(a, b) :	$(.6, .55)$	$(.6, .5)$	$(.6, .4)$
Main	Reward	172.46	169.39	169.61
	Regret	2.51%	3.54%	3.46%
Comparative	Reward	172.42	169.24	169.22
	Regret	2.53%	3.59%	3.59%
Non-sequential	Reward	170	160	140
	Regret	3.33%	6.67%	13.33%

Secondly, it is worth noting that the percentage regrets in Table 2 are satisfactorily small, especially in contrast with the non-sequential model. Of course, in doing so we are comparing apples and oranges, but nonetheless, note that the gap widens as the difference between a and b increases, and furthermore, is magnified as the horizon expands. In a fixed sample size model, the percentage regrets remain as unchanging proportions, whereas in either sequential model, the regret tends to zero. Thus, depending on the purpose of the trial and the nature of the treatments being tested, fixed sample size methodology can perform relatively poorly.

A number of general issues are raised by this paper. These include

(i) the choice and rôle of N;

(ii) the concept of choosing a prior and subsequently minimizing its impact via a conservative, minimax-type approach;

(iii) the programability problems arising to cope with arbitrary N;

(iv) the aim to keep theoretical designs both ethical and practical;

(v) the aim, in early phase applications, to select the most promising treatment, rather than rank or estimate treatment performances;

(vi) the use of expected reward as the chief criterion rather than use the probability of correct selection, the expected number of inferior treatment allocations or the p-values in a hypothesis testing approach.

Suffice it to say, I believe (ii)-(vi) are good ideas, although under (v) suggests the ranking techniques of Bechhofer, Kiefer and Sobel (1968) are helpful, while under (vi) any sensible stopping rule should perform satisfactorily against a variety of measures in agreement with Bather (1985). Regarding (i), a number of authors have postulated conceptual values for N in the finite horizon approach. Simons (1989), for instance, assumes a random (geometric) horizon to avoid pre-assigning a particular N, but this necessitates an additional parameter be included in the model. I propose that one views the decision arising from the comparative stages as provisional, as if the recommended treatment is "under guarantee" until N patients have been treated since the trial's inception. Confirmation (or otherwise) of the treatment's superiority has to await more thorough, later phase trials. In more practical terms, choosing a fixed N between 200 and 400 gives rise to suitably long lasting early phase trials for most applications, but I would caution against any hard and fast rules.

The field is ripe for further research, especially that directed towards helping this sort of methodology become more favorably viewed and adopted by practitioners of clinical trials. Perhaps the greatest limitation to its current usefulness is the matter of assuming immediate responses. Of course, for adaptive assignment method to be of value, the response time needs to be fairly rapid compared with the patient accrual rate, but there are at least a couple of ways of getting away from instantaneous responses. The first is to allocate, not by triplets of patients, one each per treatment, but by patient blocks of size ℓ, say, per treatment. This could be further refined allowing variable block sizes over time. See Jones, Lewis and Hartley (manuscript in this volume) for a step in this direction, though Pocock's (1977) group sequential approach would raise similar issues. The second is to abandon precisely equal (or matched) allocation in favor of, say, randomizing by

flipping a coin and keeping track of successes and failures on each arm of the trial. At present, we have to assume that patients are given their respective treatments in conveniently arriving triplets and that no missing observations upset our data collection in order to work with out score differences j and k. Some theoretical progress has been made in each of these areas, but clearly more needs to be accomplished to facilitate widespread application.

Acknowledgments. I would like to thank Professor Gordon Simons, University of North Carolina at Chapel Hill, for his guidance during my Ph.D. Research from which this is extracted and acknowledge financial support from the National Science Foundation.

References

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

ARMITAGE, P. (1985). The search for optimality in clinical trials. *International Statistical Review* **53** 15-24.

BATHER, J.A. (1985). On the allocation of treatments in sequential medical trials. *International Statistical Review* **53** 1-13.

BATHER, J.A. and SIMONS, G.D. (1985). The minimax risk for two-stage procedures in clinical trials. *Journal of the Royal Statistical Society B* **47** 466-475.

BECHHOFER, R.E., KIEFER, J. and SOBEL, M. (1968). *Sequential Identification and Ranking Procedures.* Chicago: Chicago.

BERRY, D.A. (1978). Modified two-armed bandit strategies for certain clinical trials. *Journal of the American Statistical Association* **73** 339-345.

CANNER, P.L. (1970). Selecting one of two treatments when the responses are dichotomous. *Journal of the American Statistical Association* **65** 293-306.

COLTON, T. (1963). A model for selecting one of two medical treatments. *Journal of the American Statistical Association* **58** 388-400.

COLTON, T. (1965). A two stage model for selecting one of two treatments. *Biometrics* **21** 169-180.

FELDMAN, D. (1962). Contributions to the "two-armed bandit" problem. *The Annals of Mathematical Statistics* **33** 847-856.

PALMER, C.R. (1991). A comparative phase II clinical trials procedure for choosing the best of three treatments. *Statistics in Medicine* **10** 1327-1340.

POCOCK, S.J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64** 191-199.

RODMAN, L. (1978). On the many armed bandit problem. *The Annals of Probability* **6** 491-498.

SCHWARTZ, D., FLAMANT, R. and LELLOUCH, J. (1980). *Clinical Trials.* London: Academic Press.

SIMONS, G.D. (1986). Bayes rules for a clinical trials model with dichotomous responses. *The Annals of Statistics* **14** 954-70.

SIMONS, G.D. (1989). A random horizon model for sequential clinical trials. *Sequential Analysis* **8** 27-49.

SYLVESTER, R.J. (1988). A Bayesian approach to the design of phase II clinical trials. *Biometrics* **44** 823-36.

ZABORSKIS, A.A. (1976). Sequential Bayesian plan for choosing the best method of medical treatment. *Automation Remote Control* **37** 1750-57.

THE UNIVERSITY OF CAMBRIDGE AND
THE INSTITUTE OF PUBLIC HEALTH
ROBINSON WAY
CAMBRIDGE CB2 2SR
UNITED KINGDOM