Adaptive biased-coin designs for skewing the allocation proportion in clinical trials with normal responses

A. C. Atkinson[1,*,†] and A. Biswas[2]

[1]*Department of Statistics, London School of Economics, London WC2A 2AE, U.K.*
[2]*Applied Statistics Unit, Indian Statistical Institute, Barrackpore Trunk Road, Kolkata 700 108, India*

SUMMARY

Response adaptive designs are used in phase III clinical trial for skewing the allocation pattern towards the better treatments. We use optimum design theory to derive adaptive designs when the responses are normally distributed. The performance of the designs is studied with respect to the loss and the proportion of allocation to different treatments. The adaptive design does not affect inference. Copyright

KEY WORDS: 50:50 allocation; D_A-optimal design; ethics; inference; limiting proportion of allocation; minimization; randomization, regularization

1. INTRODUCTION

Patients arrive sequentially for a phase III clinical trial and are to be allocated one of t treatments. The use of statistical methods in pharmaceutical development is described by Senn [1]. Adaptive designs are used to achieve a larger allocation proportion to the eventually best treatment by using earlier responses to determine the next allocation. The extensive literature for binary responses includes Zelen [2], Wei and Durham [3], Durham, Flournoy and Li [4], Rosenberger *et al.* [5] and Rosenberger and Lachin [6, Chapters 10–12]. Applications of adaptive designs are described by Bartlett *et al.* [7], Tamura *et al.* [8] and by Rout *et al.* [9]. There is however little work on the case of main interest in this paper, that of continuous responses. An exception is Bandyopadhyay and Biswas [10].

Virtually all the available adaptive designs for skewing the allocation pattern use conceptual urns to decide which treatment to allocate next. There is no use of the powerful general methods provided by the theory of optimum experimental design, an advantage of which is that information about prognostic factors can be incorporated in a natural and efficient manner.

We build on the work in which optimum design theory has been applied to biased-coin designs. In such designs the next allocation is made on the basis of the previous allocations

and prognostic factors or covariates and on the prognostic factors with which the current patient presents. Although these designs are sometimes called covariate adaptive, they are not adaptive in the sense used here, since the responses from previous patients are not available for allocating the next treatment. To avoid confusion, the class of designs considered in this paper is sometimes called response adaptive.

The purpose of covariate adaptive designs is to provide some randomness against potential biases as well as to ensure that the trial, of unknown length, will be approximately balanced whenever it is stopped. Biased-coin designs were introduced by Efron [11]. Atkinson [12] used optimum design theory to extend Efron's work to balance over prognostic factors. The properties of these and other designs have been studied by Smith [13, 14] and by Burman [15].

The importance of randomness in clinical trials is discussed in the books of Matthews [16] and Rosenberger and Lachin [6], with a sharp comment by Senn [17]. Because biased-coin designs include some randomness, they are not fully balanced over treatments and prognostic factors. The estimates of the treatment effects therefore have variances slightly larger than those from a completely balanced design. Use of the linear model to which the optimum design theory is applied means that this increased variance can be expressed as a 'loss', the number of patients on whom information is not available compared with the balanced design. Atkinson [18] gives plots of loss against number of patients for a variety of allocation rules. The distribution of loss is studied by Atkinson [19].

These biased-coin designs give equal allocation of all treatments. In this paper we show how optimum design theory can be used to provide biased-coin designs with unequal allocation. We calculate the losses for these designs. In adaptive designs we use earlier responses to bias the allocation in favour of the better treatment, whilst both obtaining efficient estimates of parameters and maintaining some randomness in allocation.

Our objective is to quantify the trade-off between efficiency of estimation of treatment differences, partially randomized allocation and reduction of the number of patients receiving inferior treatments. The model and parameter estimates are presented in Section 2. In Section 3 we derive efficient skewed biased-coin designs. The properties, including losses, of these designs are found in Section 4 by simulation and related to earlier theoretical results. Section 5 describes our adaptive design and shows some properties of the designs produced by four allocation rules.

A difficulty with adaptive designs is that they can sometimes be highly unbalanced at the beginning of the trial. An example is the ECMO trial [7, 20] which used an urn design. Accordingly, in Section 5.2, we introduce a regularization of our design that avoids extremes of unbalance. The advantages of our chosen regularization are demonstrated in Section 6. The penultimate section considers inference from our design. Unlike the ECMO trial, we fail to find any effect of the adaptive nature of the design on inference about treatment effects. The paper concludes with brief comments on extensions of our method, particularly to more than two treatments.

2. MODELS AND ESTIMATES

2.1. Models

Two treatments are to be compared in a clinical trial, with the results for the ith patient adjusted by a vector of covariates x_i. We assume a normal theory linear model

$$y_i = \sum_{j=1}^{q+1} \omega_j g_{ij} + \varepsilon_i = \alpha_1 h_1 + \alpha_2 h_2 + \sum_{j=1}^{q-1} \theta_j z_{ij} + \varepsilon_i \qquad (1)$$

where z_i is a $q-1 \times 1$ vector of extended covariates which may include powers and products of the x_i. The parameters of interest are the treatment effects α; the h_j are indicator variables for treatment allocation and the θ_j are nuisance parameters. Because of the way we have parameterized the treatment effects, the z_{ij} do not include a constant column. With two treatments we are interested in one linear combination of α_1 and α_2. There is therefore a second linear combination which is treated as a nuisance variable. Together with the $q-1$ elements of θ there are therefore in all q nuisance parameters. As we shall see, the properties of the designs depend on q. Although much of our theory is readily extended to any number of treatments, the numerical examples, like this description, are for only two.

For the normal theory linear model to be appropriate, it may be necessary to transform the response, for example with a power transformation. Our results also hold, at least approximately, for generalized linear models in which the variation in response is sufficiently small that the iterative weights are sensibly constant [21].

2.2. *Variances and efficiencies*

With two treatments on an equal footing designs are required for estimating $\alpha_1 - \alpha_2$ with low variance. In the presence of covariates this can be formulated as minimizing the variance of the linear combination $a^{\mathrm{T}}\widehat{\omega}$ with

$$a^{\mathrm{T}} = \{1 \ -1 \ 0 \ \cdots \ 0\} \tag{2}$$

a vector with $q+1$ elements, which from (1) treats the θ_j as nuisance parameters. In the absence of covariates $\mathrm{var}\,(\widehat{\alpha}_1 - \widehat{\alpha}_2)$ is minimized when each treatment is allocated to the same number of patients. In the presence of covariates, biased-coin designs and other sequential procedures based on optimum design theory were used by Atkinson [12, 18] to find sequences of treatment allocations minimizing the variance of $\widehat{\alpha}_1 - \widehat{\alpha}_2$. As the trials continue, the designs become increasingly balanced and the proportion of patients receiving each treatment tends to one half. Then $\mathrm{var}\,(\widehat{\alpha}_1 - \widehat{\alpha}_2) = 4\sigma^2/n$.

Now consider unequal allocation either in the absence of covariates, or when there is balance across covariates. Suppose that of n patients $n_1 = nr$ are treated by treatment 1, when the efficiency (with respect to the optimum equal allocation) is

$$E_n = \left(\frac{4\sigma^2}{n}\right) \Big/ \left\{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n-n_1}\right)\right\} = \frac{4n_1(n-n_1)}{n^2} = 4r(1-r) \tag{3}$$

This efficiency does not change much for central values of r. A value of $r = 0.75$ gives a 3:1 allocation in favour of the better treatment and an efficiency of 75 per cent for estimation of $\alpha_1 - \alpha_2$. Even an allocation as extreme as $r = 0.85$ gives an efficiency of 50 per cent.

3. SKEWED DESIGNS

3.1. *Efficient designs*

Let the design now be such that a known proportion p of the patients should receive treatment 1. For example, treatment 2 might be the control, but the primary focus of interest is in

the new treatment, which should be allocated more often. Designs with the desired skewed allocation can be found by seeking to minimize the variance of the linear combination

$$a^{\mathrm{T}} = \{p \ \ -(1-p)\ 0\ \cdots\ 0\}, \quad 0 \leqslant p \leqslant 1 \tag{4}$$

In the absence of covariates the design is therefore found to minimize $\mathrm{var}\{p\hat{\alpha}_1 - (1-p)\hat{\alpha}_2\}$, a design that indeed allocates a proportion p of the patients to treatment 1. For this design $\mathrm{var}\ a^{\mathrm{T}}\hat{\alpha} = \sigma^2/n$.

3.2. *Sequential skewed biased-coin D_A-optimum design*

The purpose of this sequential skewed allocation scheme is, in the long run, to allocate treatment 1 to a proportion p of the patients. Optimum design theory is used to generate designs which are unlikely to be far from balance if the trial ceases at an arbitrary time point as well as providing some randomness in allocation.

The matrix form of the model (1) for the first n patients is

$$E(Y_n) = G_n\omega = H_n\alpha + Z_n\theta$$

where Y_n is the $n \times 1$ vector of responses for the first n patients, H_n is the $n \times t$ matrix of indicator variables with one non-zero entry per row, and Z_n is the $n \times (q-1)$ extended matrix of prognostic factors, which does not include a constant column. In this notation, the variance of the linear combination $a^{\mathrm{T}}\hat{\alpha}$ in the presence of prognostic factors is

$$\mathrm{var}\{a^{\mathrm{T}}\hat{\omega}\} = \sigma^2 a^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}a \tag{5}$$

a special case of (A1).

Patient $n+1$ arrives with a vector x_{n+1} of covariates and prognostic factors. If treatment 1 is allocated, G_{n+1} is formed from G_n by addition of the row

$$g_{n+1} = \left(1 \ \ 0 \ \ z_{n+1}^{\mathrm{T}}\right)^{\mathrm{T}} \tag{6}$$

where z_{n+1} is a known function of x_{n+1}. If treatment 2 is allocated

$$g_{n+1} = \left(0 \ \ 1 \ \ z_{n+1}^{\mathrm{T}}\right)^{\mathrm{T}} \tag{7}$$

Results in the appendix show that the maximum decrease in the variance of the combination $a^{\mathrm{T}}\hat{\omega}$ is achieved by allocating that treatment for which the variance of prediction $d_{\mathrm{A}}(j, n, x_{n+1})$ (A2) is larger. For skewed allocation designs the suggestion of Atkinson [12] becomes to allocate treatment j with probability

$$\pi_{\mathrm{A}}(j \mid x_{n+1}) = \frac{p_j d_{\mathrm{A}}(j, n, x_{n+1})}{\sum_{s=1}^{t} p_s d_{\mathrm{A}}(s, n, x_{n+1})} \tag{8}$$

The numerical studies mentioned in Section 1 are all for equal allocation. There has been no investigation of the properties of skewed designs with unequal p_j.

3.3. *Efficiencies*

The variance of the estimated treatment combination $a^{\mathrm{T}}\hat{\omega}$ for a general design is given by (5). Since, from Section 3.1, the minimum value is σ^2/n, the generalization of the

efficiency E_n (3) is

$$E_n = 1/\{na^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}a\} \tag{9}$$

The loss L_n is defined by writing the variance (5) as

$$\mathrm{var}\{a^{\mathrm{T}}\widehat{\omega}\} = \sigma^2/(n - L_n)$$

so that

$$L_n = n(1 - E_n) \tag{10}$$

With a random element in treatment allocation, the loss L_n is a random variable, depending upon the particular trial and pattern of covariates. We give values for $E(L_n)$ in Section 4.1. One advantage of loss as a measure of design performance is that it approaches the informative asymptotic value relatively quickly. For the schemes considered in Atkinson [18] (10) shows that the efficiency of all designs asymptotically tends to one.

4. A COMPARISON OF SEQUENTIAL DESIGNS FOR SKEWED ALLOCATION

4.1. Allocation rules

The allocation rules are expressed in terms of probabilities $\pi(j|x_{n+1})$. In some case these depend upon the ordering of the treatments by the variances $d_{\mathrm{A}}(j,n,z_{n+1})$. We use $\pi([j]|x_{n+1})$ to represent the probability of allocating the treatment with the jth largest value of the variance.

D: Deterministic (*sequential design construction*).

$$\pi_{\mathrm{D}}([1]\,|\,x_{n+1}) = 1$$

The treatment with the larger variance of prediction is always selected. Asymptotically, for any reasonable distribution over time of prognostic factors, the design will be balanced over the factors, when allowance is made for the skewing induced by the vector of contrasts a, and there will be no loss. Let $E(L_n) = \mathscr{L}_n$. Then $\mathscr{L}_\infty = 0$.

R: Completely Randomized. For skewed designs with $t = 2$

$$\pi_{\mathrm{R}}(1|x_{n+1}) = p$$

since we know which treatment is to be more highly allocated. For this rule $\mathscr{L}_\infty = q$.

These two rules represent the extremes of rules which aim for skewing and balance over both the short and long term. The losses of the other rules considered here are bounded by these values.

A: D_{A}-Optimality. The probabilities for Atkinson's D_{A}-optimality are given in (8). The desired skewed allocation is obtained by setting $p_1 = p$ and $p_2 = 1 - p$, when $\mathscr{L}_\infty = 1$.

E: Efron's Biased-Coin. In Efron's original biased-coin design [11] with two treatments and no prognostic factors $\pi_E([1]\,|\,x_{n+1}) = \frac{2}{3}$. In the skewed case we need different rules depending on which treatment would be allocated by an unrandomized rule. If $d_A(1,n,x_{n+1}) > d_A(2,n,x_{n+1})$ we find the probabilities of allocating treatment 1 by replacing $d_A(1,n,x_{n+1})$ in (8) by 2/3: if treatment 2 is under-represented we replace $d_A(1,n,x_{n+1})$ with 1/3. Thus

If $d_A(1,n,x_{n+1}) > d_A(2,n,x_{n+1},)$

$$\pi_E(1\,|\,x_{n+1}) = \frac{\frac{2}{3}p}{\frac{2}{3}p + \frac{1}{3}(1-p)} = \frac{2p}{1+p}$$

otherwise

$$\pi_E(1\,|\,x_{n+1}) = \frac{p}{2-p}$$

In the unskewed case, that is $p = 0.5$, we recover the values of $\frac{2}{3}$ and $\frac{1}{3}$. As $p \to 1$, both probabilities tend to one. As for the deterministic rule, $\mathscr{L}_\infty = 0$.

4.2. Numerical results

We compared these four allocation rules for two treatments and two values of p: 0.5 and 0.75. We calculated the balance, that is the value of the proportion r, and following Atkinson [18], the loss for up to 200 trials for $q = 5$ and 10. The results shown are the averages of 10 000 simulations of the 200 patient trials.

The plots of Figure 1 show the average losses $\bar{L}_n$ as functions of patient number, when $q = 5$. The left-hand panel is for $p = 0.5$ and is, as it should be, the same, apart from variability due to the simulation, as the left-hand panel of Figure 1 of Atkinson [18]: the loss for rule R is approximately five throughout and that for D is close to zero for n above 100, while that for A is close to one, which is $q/5$, from $n = 50$. The numbers for $\bar{L}_{200}$ are in Table I.

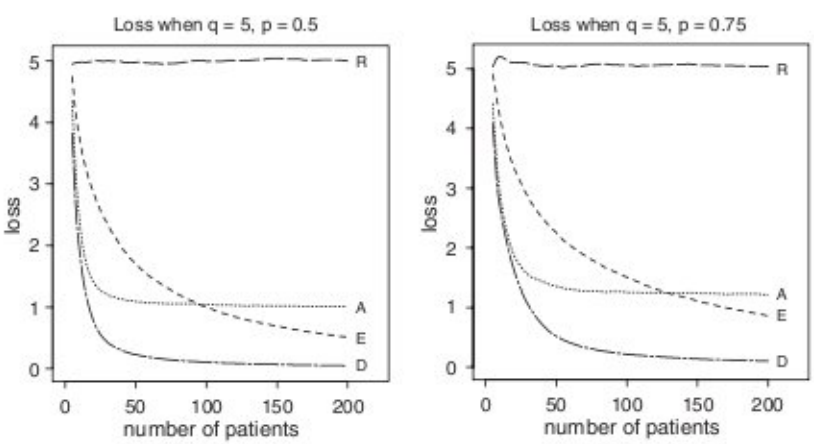

Figure 1. Designs for known skewing proportion. Average losses $\bar{L}_n$ when $q = 5$ for four allocation rules: A, D_A-optimality; D, deterministic; E, Efron's biased coin and R, random. Left-hand panel, unskewed ($p = 0.5$); right-hand panel, $p = 0.75$. Averages of 10 000 simulations.

Table I. Average loss $\bar{L}_{200}$ for unskewed and skewed allocations from 10000 simulations.

	$q=5$		$q=10$	
Rule	$p=0.5$	$p=0.75$	$p=0.5$	$p=0.75$
A	1.013	1.218	2.074	2.540
D	0.051	0.104	0.211	0.451
E	0.513	0.860	1.937	2.930
R	5.001	5.036	9.986	9.994

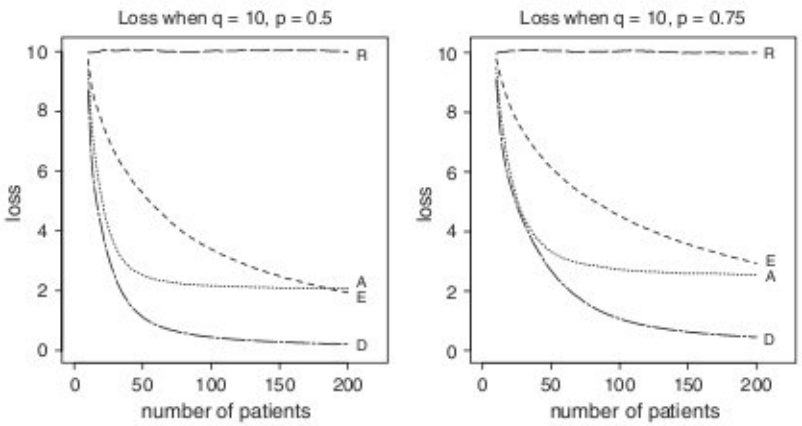

Figure 2. Designs for known skewing proportion. Average losses $\bar{L}_n$ when $q=10$ for four allocation rules: A, D_A-optimality; D, deterministic; E, Efron's biased coin and R, random. Left-hand panel, unskewed ($p=0.5$); right-hand panel, $p=0.75$. Averages of 10000 simulations.

The right-hand panel of the figure shows the average loss for the target proportion $p=0.75$. The losses are similar to those for $p=0.5$, although slightly higher. Table I also gives these values of $\bar{L}_{200}$: the largest increase in going from $p=0.5$ to 0.75 is for rule E, but is only 0.347. As a result the loss for E is lower than that of A over a smaller part of the range.

Similar comments can be made about the plots for $q=10$ in Figure 2. The losses for R and D are now around 10 and zero, for the larger values of n. That for A is asymptotically near to $2=q/5$. Again E is the rule whose performance is most changed when $p=0.75$, rather than 0.5. The numbers are again in Table I.

The main change in going from $q=5$ to 10, apart from a doubling in the values of the losses, is that the curves of loss against n descend more slowly to their asymptotic values, which are zero for both D and E. The effect of moving from the unskewed $p=0.5$ to the skewed value of 0.75 is slightly to increase the average loss. This arises because, with a 3:1 ratio of allocation, the skewed designs are on average slightly less well balanced than those for $p=0.5$. However, the observed proportion $r=n_1/n$ is close to 0.75 throughout the simulations. The plots of r against n are not shown here, as they are uninformative. If the values of r were not close to 0.75, calculations similar to those in (3) combined with (10) show that the loss would be larger than those found here.

These numerical results show that the extension of optimum design theory to skewed allocations does not greatly increase the loss due to imbalance. The adaptive use of these designs does however raise several new points, as well as leading to appreciable increases in loss.

5. ADAPTIVE DESIGNS

5.1. Link function based adaptive design

Bandyopadhyay and Biswas [10] introduced an adaptive design that assigns one of two treatments according to the outcome of previous assignments. We assume that large values of the response y are desired and let $\widehat{\Delta} = \widehat{\alpha}_1 - \widehat{\alpha}_2$. They take the adaptive probability of allocating treatment one as

$$\pi_{\mathrm{BB}}(1 \mid x_{n+1}) = J(\hat{\Delta})$$

where $J(\cdot)$ is a suitably chosen distribution function of a random variable, symmetric about zero. If $\widehat{\Delta}$ is positive, that is $\hat{\alpha}_1 > \hat{\alpha}_2$, the probability of allocating the first treatment is > 0.5. Since there is no interaction between the treatments and the covariates, this rule does not depend on x_{n+1}. The most natural choice of $J(x)$ to a statistician is $\Phi(x/T)$, the distribution function of a $\mathrm{N}(0, T^2)$ random variable. Bandyopadhyay and Biswas [10] show that the limiting proportion of allocations to treatment 1 is $\Phi\{(\alpha_1 - \alpha_2)/T\}$.

To combine the adaptive nature of this design with the balance and randomness of the biased-coin designs of Section 4.1 we estimate the skewing proportion p by

$$\hat{p} = \Phi\{(\widehat{\alpha}_1 - \widehat{\alpha}_2)/T\}$$

That is, from the results of n trials we calculate the estimates of the treatment effects and use $\hat{p}$ in the calculation of $d_{\mathrm{A}}(j, n, z_{n+1})$. We then apply one of the allocation rules of Section 4.1. As it becomes clearer that treatment 1 is superior to treatment 2, the allocation proportion is intended to converge to $p = J(\alpha_1 - \alpha_2)$. As we see, the speed of convergence depends on the allocation rule.

5.2. Regularization

A potential problem with adaptive designs is that the observational error may lead to poor parameter estimates. As a result, the design may be highly unbalanced and the inferior treatment may be allocated to too many patients. The ECMO trial, Begg [20] and discussants, is an example of a highly unbalanced adaptive trial for which inference is contentious.

The problem is particularly severe at the beginning of the trial. We chose to allocate five of the first 10 patients to treatment one and the other five to treatment two. It is probably safe to say that this simple rule would have avoided the difficulties in the analysis of the ECMO trial. Thereafter we regularized the design to ensure that each treatment continued to be allocated throughout the trial. A simple rule would be to insist that each treatment was allocated to at least a proportion $r_{\min}$ of the patients. We chose a slightly more complicated rule to ensure that each treatment continued to be allocated, albeit with a decreasing frequency, throughout the trial. Following the initial equal allocation, if the number allocated to either treatment was below $\sqrt{n}$, that treatment was allocated when n was an integer squared. For our 200 trial

design, with 5 allocated initially to each treatment, the first regularization could occur when $n=36$ and the last when $n=189$. An advantage of such a scheme is that, if one treatment is appreciably better than the other, the allocation proportion can fall below $r_{\min}$. We compare regularized and unregularized designs in Section 6.

5.3. Numerical results

The properties of the design depend on the parameters in the function J. We take $\alpha_1 - \alpha_2 = 0.6745$ which, with $T=1$, gives a value of 0.75 for p, so that we can make comparisons with the results of Section 4.2. In practice, T would be chosen to give the desired skewing proportion p for the expected difference in treatment effects. The properties of the design will also depend on the error of measurement σ^2. Larger values of σ^2 will obscure the true treatment difference and lead to designs which converge more slowly to the desired value of p.

We look at the average properties of 10 000 regularized 200 patient designs. Figure 3 shows the average values of $r=n_1/n$ for each design. The left-hand panel of the figure shows the proportions for the four allocation rules when $\sigma=2$ and $q=5$. Because of our initial design, $r=0.5$ when $n=10$. The plot shows that the average proportion rises slowly to 0.75: rule A reaches this value when $n=166$, the first design to do so. On the other hand, rule R, which is that of Bandyopadhyay and Biswas [10], has only reached a value of 0.728 when $n=200$. It is surprising that A converges faster here than D, which is close to it in performance, or E. In the right-hand panel we consider the proportion for rule A for five values of σ: 0, 0.5, 1, 1.5 and 2. As σ decreases, the proportion approaches 0.75 more quickly with n. For larger n and σ there is some overshoot in the proportion. As we shall see when we look at the properties of individual trials in the next section, this overshoot is caused by some trials over-estimating p and so allocating too many trials to the first treatment.

The corresponding plots of average loss are in Figure 4. These are not what we might expect from Figure 1: for all four rules in the left-hand panel the average losses increases steadily with n. At $n=200$, the loss for rule D is 14.4, whereas for R, A and E the values

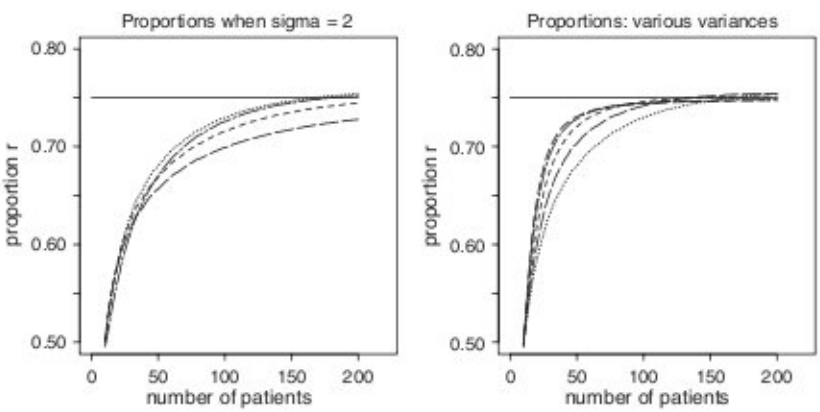

Figure 3. Regularized adaptive designs: average proportion of patients receiving the better treatment when $q=5$. Left-hand panel, reading upwards: rules R, E, D and A with $\sigma=2$. Right-hand panel, rule A for five values of σ; the lowest curve is for $\sigma=2$. Averages of 10 000 simulations.

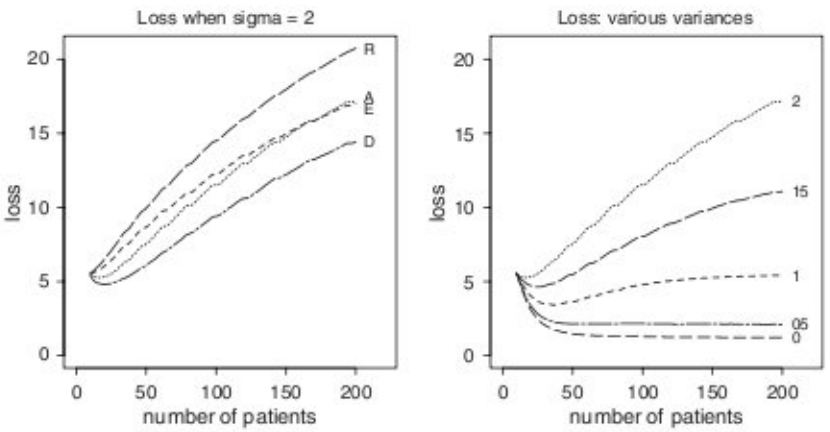

Figure 4. Regularized adaptive designs: average losses $\bar{L}_n$ when $q=5$. Left-hand panel: rules D, E, A and R with $\sigma=2$. Right-hand panel, rule A for five values of σ; the lowest curve is for $\sigma=0$. Averages of 10 000 simulations.

are 20.8, 17.2 and 17.0. Although the difference between the randomized rule R and the other three is roughly what we would expect from the earlier simulations, the values of the average losses are much higher than we have seen before.

The right-hand panel of Figure 4 repeats the plots of average losses for rule A, but with different values of the standard deviation σ of the observations. We have already seen the curves for $\sigma=2$. The losses for $\sigma=1.5$ and 1 are, respectively, 11.08 and 5.42. That for $\sigma=0.5$ is 2.10 and that for $\sigma=0$ is 1.21, as low as the average loss in Table I where the value of p was also known without error.

Figure 4 clearly shows the effect of not knowing p on the efficiency of the designs, which is calculated with $p=0.75$ in (9). As n becomes larger, the value of α becomes better known and the average loss should slowly decrease. For $\sigma=0.5$, the loss initially decreases and then increases slightly, with a local maximum around $n=140$. For $\sigma=1$ the average loss is still gradually increasing at $n=200$, with $\bar{L}_{200}=5.42$. After 1000 trials it is 5.02. The decline in loss comes later for larger values of σ; a poor design with a target value of p appreciably greater than 0.75 gives little chance of sampling the second population and so of correcting the misleading value of p.

6. REGULARIZATION AND THE DISTRIBUTION OF LOSS

Although the average properties of the design are of interest, it is important and helpful to look at the individual trials: it is little consolation for a clinician with a poorly balanced trial to be assured that the average properties of trials produced by the randomization scheme are excellent.

Figure 5 shows box plots of the distributions of proportions r for 1000 simulations of a 200-trial design using rule A with $\sigma=2$, both with and without regularization. The results for

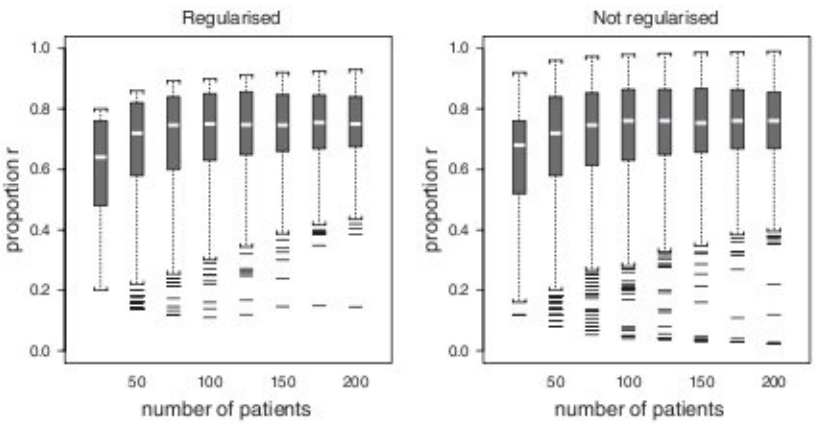

Figure 5. One thousand individual adaptive designs: boxplots of proportion of patients receiving the better treatment. Left-hand panel: regularized; right-hand panel, unregularized. Rule A, $q = 5, \sigma = 2$.

the regularized design are in the left-hand panel. As the trial progresses we have seen that the average value of r converges to around 0.75. However, particularly for small n, there are some designs which have a value of r less than 0.5. Although this number decreases as n increases, even when $n = 200$ there is one trial for which $r = 0.145$. For $n = 200$ the effect of the regularization is such that the maximum proportion is $1 - r_{\min} = (200 - \sqrt{196})/200 = 0.93$ and several trials achieve this value. The results for the unregularized design are in the right-hand panel of the figure. The distribution of r contains more extreme values, both high and low than before. In particular, these values, unlike those for the regularized design, are not bounded away from unity.

The effect of the regularization on average loss is visible in Figure 4, particularly for the curve of loss for $\sigma = 2$ in the right-hand panel. There are occasional inflections in the otherwise steadily increasing curve, which occur when n is the square of an integer. These are caused by the regularization forcing the high values of r in some trials slightly closer to the target value of 0.75. Atkinson [19] demonstrates conditions under which the distribution of loss for unskewed non-adaptive allocations has a χ^2_q distribution. The distribution of loss arising from adaptive designs is more highly skewed and depends on n.

Figure 6 shows the distribution of loss for the adaptive designs of Figure 5. The results for the regularized designs are again in the left-hand panel. The trial evident in the left-hand panel of Figure 5, for which $r = 0.145$ when $n = 200$, gives rise to the large value of 150 for L_{200}. The other large values of loss for smaller n are caused both by trials for which the proportion is too small and for which it is too large. The effect of the lack of regularization is clear in the right-hand panel of Figure 6, plotted on the same scale as the losses for the regularized designs. The greater variability in the distribution of r in the right-hand panel of Figure 5 results in a highly skewed distribution of loss: the tail of this distribution becomes increasingly heavy as n increases. By comparison with the regularized designs, many of the unregularized trials are unsatisfactorily far from balance. Even for the regularized designs, Figure 4 shows that average loss is still increasing when $n = 200$. The left-hand panel of

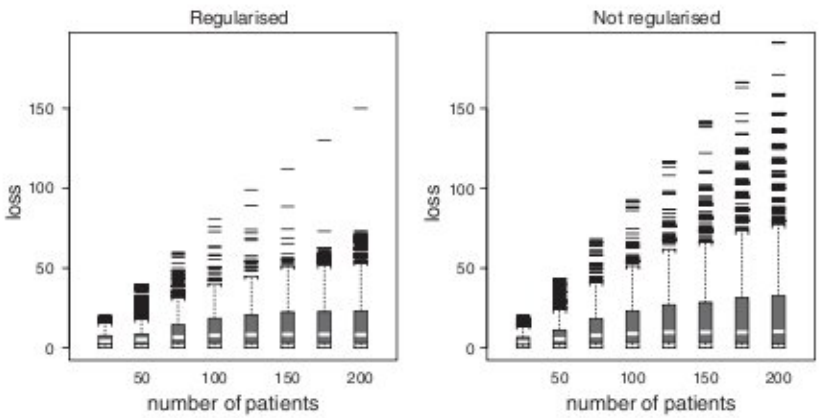

Figure 6. One thousand individual adaptive designs: box plots of loss L_n. Left-hand panel: regularized; right-hand panel, unregularized. Rule A, $q=5, \sigma=2$.

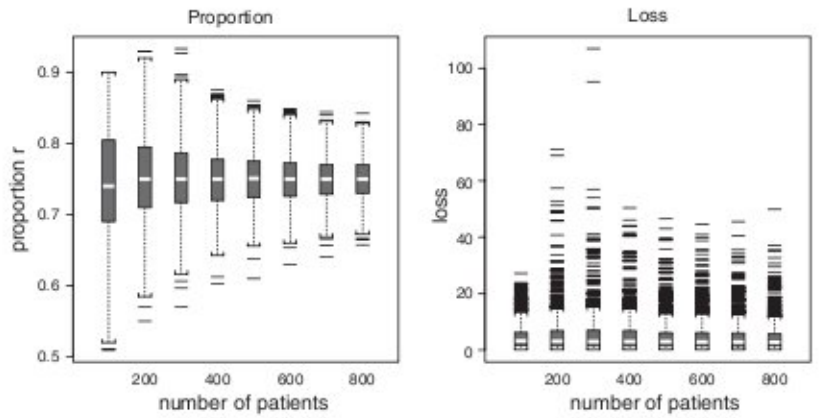

Figure 7. One thousand individual regularized adaptive designs for n up to 800. Left-hand panel: box plots of proportion of patients receiving the better treatment. Right-hand panel, loss L_n. Rule A, $q=5, \sigma=1$.

Figure 6 gives the detail and indicates how the large losses from highly skewed trials continue to grow slightly with n.

The plot of average loss in the right-hand panel of Figure 4 shows that, for sufficiently small σ, the average loss does decrease with increasing n by the time $n=200$. In Figure 7 we repeat the box plots for a regularized design using rule A, but with $\sigma=1$, extending the simulation up to $n=800$. The left-hand panel shows that, for this smaller error standard deviation, there are no trials with $r<0.5$, so that all trials favour the better treatment. For $n>300$ there are no values of r so large that they have been imposed by the regularization,

which has therefore achieved the task of focusing designs around the correct value of r. As the number of trials increases, the distribution of r becomes more compact.

The improved distribution of r compared with those of the simulations of Figures 5 and 6 are reflected in the distribution of loss in the right-hand panel of Figure 7. The median of the distribution is slowly decreasing in the last half of the trial and, arguably, the upper tail of the distribution is becoming slightly lighter. However, the increasing concentration of the distribution of r is being offset by the increase in loss with n for a fixed value of r. Such long tails are a general feature of adaptive methods: frequent small trials are offset by occasional trials in which a large number of patients is needed for satisfactory inference.

7. INFERENCE

The parameter estimates $\widehat{\alpha}$ are derived assuming regression models with independent errors. Since each allocation depends on the earlier responses, the observations are not independent and the likelihood is more complicated. However, the asymptotic normality of the parameter estimates is not affected by the adaptive nature of the design [22] and, for responses modelled by the exponential family, the optimum adaptive designs obtained by sequential use of parameter estimates are indeed optimum [23].

Inference for effects in adaptive designs is much discussed. That for the ECMO trial is the main concern of Begg [20]. For their adaptive trial [8] simulate the adaptive design 500 000 times to find the distribution of the test statistic. We also use simulation to investigate the distribution of a test statistic.

Here the statistic of interest is the t-test for the hypothesis $\Delta = \alpha_1 - \alpha_2 = 0$. It is customary to investigate the null distribution of such test statistics. This is however of limited interest with adaptive designs, where we require the design to be skewing the allocation. Accordingly we investigate the 'pseudo-null' distribution by subtracting off the known value of Δ and look at the statistic

$$t_\Delta = \frac{\widehat{\alpha}_1 - \widehat{\alpha}_2 - (\alpha_1 - \alpha_2)}{\sqrt{s^2(GVG_{1,1} + GVG_{2,2} - 2GVG_{1,2})}}$$

where $GVG_{i,j}$ is element (i,j) of $(G^{\mathrm{T}}G)^{-1}$ and s^2 is the customary mean square estimate of σ^2 on $n - q - 1$ degrees of freedom.

Figure 8 shows box plots of 1000 simulated values of the pseudo-null distribution of t_Δ for regularized designs with rule A for two values of σ as n goes from 25 to 200. Even for the smallest trials, those with 25 patients, the statistics are centred close to zero with symmetrical distributions that are well behaved and seem close to normal. There is also surprisingly little effect of the value of σ on the distribution. This impression of normality is supported by the normal QQ plots of Figure 9 for $n = 50$ and 100 when σ has the higher value of two. There is no evidence of any strong departure from normality: standard normal, or t, tests could be used straightforwardly for inference. The adaptive nature of the design does not affect inference.

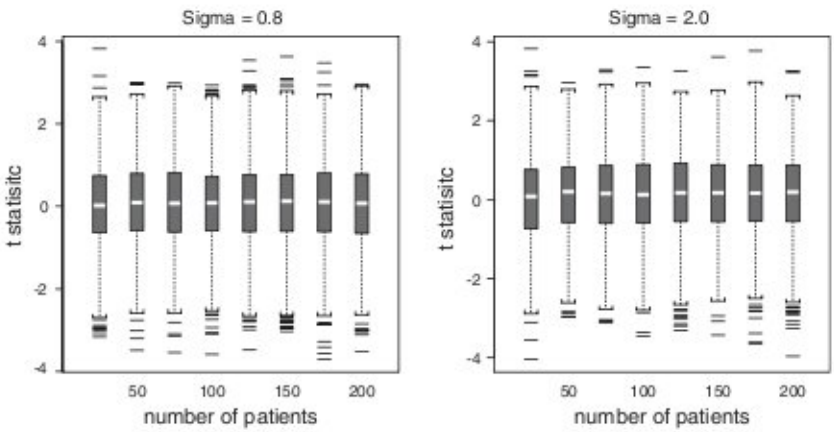

Figure 8. One thousand individual regularized adaptive designs, rule A: box plots of pseudo-null distribution of the test statistic t_Δ. Left-hand panel, $\sigma = 0.8$; right-hand panel, $\sigma = 2.0$.

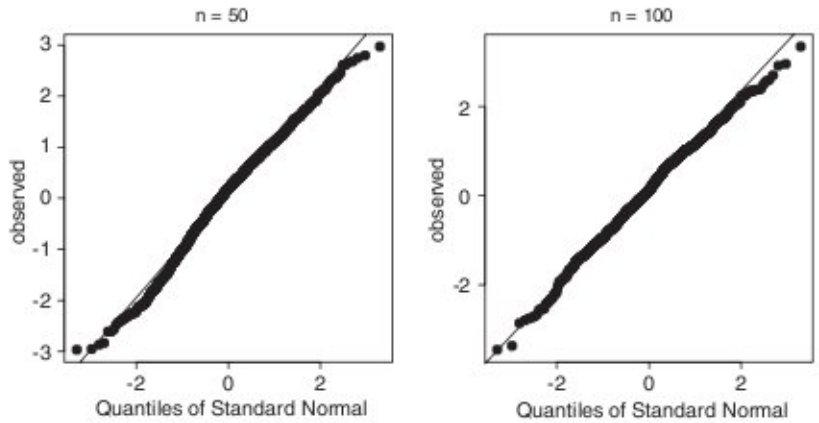

Figure 9. One thousand individual regularized adaptive designs, rule A: normal QQ plots of the pseudo-null distribution of the test statistic t_Δ. Left-hand panel, $n = 50$; right-hand panel, $n = 100$.

8. EXTENSIONS

We have assumed that the responses for all n patients are known when allocation is to be made to patient $n+1$. However, the responses may be available on only some number m of the patients where $m < n$. Allocation with such delayed responses is considered by Bai *et al.* [22] for binary responses. The extension of our method for normal observations to delayed response is straightforward. The estimates $\hat{\alpha}_1$ and $\hat{\alpha}_2$ required in Section 5.1 for calculation of $\hat{p}$ are based on the m observations for which responses are available. But, the calculation of the variances $d_A(.)$ in (A2) uses the covariate values for all n patients. The

properties of the resulting designs will therefore be intermediate between those of the response-adaptive designs introduced here and the biased-coin covariate-adaptive designs compared in Atkinson [18].

A family of designs with good properties in the comparisons of Atkinson [18] are Bayesian biased-coin designs derived from a utility proposed by Ball *et al.* [24]. Atkinson and Biswas [25] generalize the derivation of this utility to give designs when unequal allocations are required and apply the results to the construction of adaptive-biased coin designs similar to those given here. However, the right-hand panel of Figure 4 shows that adaptive designs have appreciably higher average loss than skewed designs, unless the error variance is negligible. The reduction in loss from use of the Bayesian designs in a response-adaptive setting is consequently less important than it is when covariate-adaptive designs, such as those in Atkinson [18], are appropriate.

Finally, Atkinson [26] describes the extension of the adaptive designs of this paper to more than two treatments. For three treatments the linear combination (4) is replaced by

$$a^{\mathrm{T}} = \{p_1 \ -p_2 \ 1-p_1-p_2 \ \cdots \ 0\}, \quad 0 \leqslant p \leqslant 1$$

Again, $a^{\mathrm{T}}\widehat{\omega}$ is to be estimated with minimum variance. Numerical examples in which the target proportions for the three treatments are 0.8, 0.15 and 0.05 illustrate the importance of regularization when the designs are highly skewed.

APPENDIX A

With t treatments up to $s \leqslant t-1$ independent linear combinations similar to a in (2) can be estimated. Atkinson [12] gives an example and calls the matrix of combinations A.

With the parameters ω estimated by least squares

$$\mathrm{var}\{A^{\mathrm{T}}\widehat{\omega}\} = \sigma^2 A^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}A \tag{A1}$$

D-optimum experimental designs for the linear regression model $E(Y) = G\omega$ maximize the determinant $|G^{\mathrm{T}}G|$. Such optimum designs can be constructed sequentially. In Section 3.2 (6) and (7) described augmentation of G_n to G_{n+1} on the allocation of treatment j to patient $n+1$. A useful matrix result is that

$$\begin{aligned} |G_{n+1}^{\mathrm{T}}G_{n+1}| &= \{1 + g_{n+1}^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}g_{n+1}\}|G_n^{\mathrm{T}}G_n| \\ &= \{1 + d(g_{n+1}, n)\}|G_n^{\mathrm{T}}G_n| \end{aligned}$$

The optimum treatment for the $(n+1)$st patient is therefore that for which $d(g_{n+1}, n)$ is a maximum.

In the clinical trials considered in this paper, where interest is in the matrix of coefficients A, D-optimality is replaced by $\mathrm{D_A}$-optimality with designs being found to minimize $|A^{\mathrm{T}}(G_{n+1}^{\mathrm{T}}G_{n+1})^{-1}A|$. These designs can again be constructed iteratively, the variance $d(g_{n+1}, n)$ being replaced by

$$d_{\mathrm{A}}(j, n, x_{n+1}) = g_{n+1}^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}A\{A^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}A\}^{-1}A^{\mathrm{T}}(G_n^{\mathrm{T}}G_n)^{-1}g_{n+1},$$

$$(j = 1, \ldots, t) \tag{A2}$$

where g_{n+1} combines the allocation indicator h_{n+1} for the $(n+1)$st patient and z_{n+1}, the extended vector of prognostic factors, known for the new patient. In the numerical examples of this paper, $t = 2$ and A is the vector (4).

REFERENCES

1. Senn S. Consensus and controversy in pharmaceutical statistics (with discussion). *The Statistician* 2000; **49**: 135–176.
2. Zelen M. Play the winner and the controlled clinical trial. *Journal of the American Statistical Association* 1969; **64**:131–146.
3. Wei LJ, Durham SD. The randomized play-the-winner rule in medical trials. *Journal of the American Statistical Association* 1978; **73**:840–843.
4. Durham SD, Flournoy N, Li W. Sequential designs for maximizing the probability of a favorable response. *Canadian Journal of Statistics* 1998; **3**:479–495.
5. Rosenberger WF, Stallard N, Ivanova A, Harper CN, Ricks ML. Optimal adaptive designs for binary response trials. *Biometrics* 2002; **57**:909–913.
6. Rosenberger WF, Lachin JL. *Randomization in Clinical Trials: Theory and Practice*. Wiley: New York, 2002.
7. Bartlett RH, Roloff DW, Cornell RG, Andrews AF, Dillon PW, Zwischenberger JB. Extracorporeal circulation in neonatal respiratory failure: a prospective randomized study. *Pediatrics* 1985; **76**:479–487.
8. Tamura RN, Faries DE, Andersen JS, Heiligenstein JH. A case study of an adaptive clinical trial in the treatment of out-patients with depressive disorder. *Journal of the American Statistical Association* 1994; **89**:768–776.
9. Rout CC, Rocke DA, Levin J, Gouws E, Reddy D. A reevaluation of the role of crystalloid preload in the prevention of hypotension associated with spinal anesthesia for elective cesarean section. *Anesthesiology* 1993; **79**:262–269.
10. Bandyopadhyay U, Biswas A. Adaptive designs for normal responses with prognostic factors. *Biometrika* 2001; **88**:409–419.
11. Efron B. Forcing a sequential experiment to be balanced. *Biometrika* 1971; **58**:403–417.
12. Atkinson AC. Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika* 1982; **69**:61–67.
13. Smith RL. Properties of biased coin designs in sequential clinical trials. *The Annals of Statistics* 1984; **12**: 1018–1034.
14. Smith RL. Sequential treatment allocation using biased coin designs. *Journal of the Royal Statistical Society, Series B* 1984; **46**:519–543.
15. Burman C-F. *On Sequential Treatment Allocations in Clinical Trials*. Department of Mathematics: Göteborg, 1996.
16. Matthews JNS. *An Introduction to Randomized Controlled Clinical Tials*. Edward Arnold: London, 2000.
17. Senn S. Unbalanced claims for balance. *Applied Clinical Trials* 2004; **13**(4):15–16.
18. Atkinson AC. The comparison of designs for sequential clinical trials with covariate information. *Journal of the Royal Statistical Society, Series A* 2002; **165**:349–373.
19. Atkinson AC. The distribution of loss in two-treatment biased-coin designs. *Biostatistics* 2003; **4**:179–193.
20. Begg CB. On inference from Wei's biased coin design for clinical trials (with discussion). *Biometrika* 1990; **77**:467–484.
21. Cox DR. A note on design when response has an exponential family distribution. *Biometrika* 1988; **75**:161–164.
22. Bai ZD, Hu F, Rosenberger WF. Asymptotic properties of adaptive designs for clinical trials with delayed response. *Annals of Statistics* 2002; **30**:122–139.
23. Baldi A, Antognini, Giovagnoli A. On the large sample optimality of sequential designs for comparing two treatments. *Sequential Analysis* 2005; **24**:205–217. DOI: 10.1081/SQA-200056200.
24. Ball FG, Smith AFM, Verdinelli I. Biased coin designs with a Bayesian bias. *Journal of Statistical Planning and Inference* 1993; **34**:403–421.
25. Atkinson AC, Biswas A. Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics* 2005; **6**(1):118–125.
26. Atkinson AC. Adaptive biased-coin designs for clinical trials with several treatments. *Discussiones Mathematicae Probability and Statistics* 2004; **24**:85–108.