The Distribution of Loss in Two-Treatment Biased-Coin Designs

Anthony C. Atkinson *
The London School of Economics, London WC2A 2AE, UK

February 12, 2002

Abstract

The paper compares randomised rules of the biased-coin type for the sequential allocation of treatments in a clinical trial. An important characteristic is the *loss*, which measures the increase in the variance of parameter estimates due to the imbalance caused by randomisation. Simulations are used to find the small sample distribution of loss. For some rules a simple chi-squared approximation to the asymptotic distribution holds well down to very small sample sizes.

Keywords: balance; Bayesian coin; bias; biased-coin design; chi-squared approximation; D_A-optimum design; loss; minimisation; randomization.

1 Introduction

Patients for a clinical trial arrive sequentially and are each to be given one of t treatments in a parallel group trial. The statistical literature describes many rules for deciding which treatment should be allocated to the patient who has just arrived. Examples are the minimisation rule of Pocock and Simon (1975) and the "biased-coin" rule of Efron (1971). It is however not known exactly how many patients there will be. If recruitment of patients ceases when the trial is unbalanced, the variance of the estimated treatment effects will be larger than if the trial were balanced. Burman (1996) expresses this increase in terms of a "loss". A comparison of several allocation rules is given by Atkinson (2002) who uses simulation to find the expected value

*a.c.atkinson@lse.ac.uk

of the loss and of the allocation bias for each. However, the allocations are restrictedly randomised. Consequently the loss will have a distribution over repetitions of each allocation rule. As Cox (1982) has commented, it is little consolation to an experimenter confronted with an unbalanced randomization to claim that, on average, the randomization scheme used produces good results. Accordingly, this paper studies the distribution of loss for representative allocation schemes. In many cases surprisingly simple results, based on the chi-squared distribution, are obtained.

The context of the design of sequential clinical trials is given, for example by Senn (1997, Chapter 6) and Matthews (2000). A summary is in Senn (2000, §2.1). Smith (1984b) and Burman (1996) study the properties of families of biased-coin designs, with an emphasis on asymptotic properties including the expected loss. Recent references to the clinical use of biased-coin designs are collected by Atkinson (2002, §2). The design problem is that, since it is not known when recruitment to the trial will stop, the sequential allocation of treatments should be made so that the trial is always reasonably balanced. This balance should be achieved for each prognostic factor so that, if these factors are included in the analysis, the effect of the treatments will be estimated with minimum variance. If randomisation is included, the design is not forced to be completely balanced: stopping the trial at an arbitrary point will then lead to a slight increase in the variance of the parameter estimates.

In general, increasing randomisation will reduce the possibility of bias and will increase the variance of estimates. Seven different schemes are compared in this paper, for four and nine prognostic factors. The numerical comparisons are for two treatments. But the simulation methods and analytical results are applicable to any number of treatments.

The linear model theory on which the comparisons of designs are based is introduced in §2, where the increase in the variance of estimated treatment effects due to randomisation is derived as a loss, expressed as the number of patients on whom information is unavailable. The seven allocation schemes are defined in §3. Simulation results for the distribution of loss are in §4. The paper concludes, in §5, with a discussion of the distribution of loss and of the importance of various kinds of bias.

2 Modelling and Design

2.1 A Regression Model

It is assumed that each of n patients receives one of t treatments which is allocated in the knowledge of a vector of prognostic factors, but in ignorance of the outcome of the earlier allocations. In many trials the response will approximately normally distributed, perhaps after transformation, so it is appropriate to use the regression model

$$E(Y_n) = \alpha_j + z_n^T \theta, \tag{1}$$

when the patient receives treatment j and has a vector z_n of $q-1$ prognostic factors: the θ are nuisance parameters. If there are only two treatments, interest is assumed to be solely in the difference between α_1 and α_2.

The responses of the first n patients are therefore modelled by

$$E(Y_n) = G_n \psi = H_n \alpha + Z_n \theta, \tag{2}$$

with H_n the $n \times t$ matrix of indicator variables for the treatments with one non-zero entry per row, and Z_n is the $n \times (q-1)$ matrix of prognostic factors, including any interactions and higher-order terms. The subscript n will only be used when it is necessary to distinguish between quantities for the $(n+1)$st patient for whom a treatment allocation is required and those for the n patients to whom treatments have already been allocated.

For any t, interest is in contrasts between the α, the mean level of response being an additional nuisance parameter making q nuisance parameters in all. Let K^T be a $(t-1) \times t$ matrix of contrasts orthogonal to the mean. An example is given by Atkinson (1982). For $t = 2$ these contrasts reduce to the difference between treatments. Since the volume of the normal theory confidence ellipsoid for least squares estimates of the contrasts is unaffected by non-singular linear transformations of the contrasts, the exact form of K is unimportant, provided the contrasts span the $t-1$ dimensional space orthogonal to the overall mean. Because the θ in (2) are nuisance parameters, the contrasts need augmenting by a $(t-1) \times (q-1)$ matrix of zeroes

$$A^T = (K^T \quad 0) \tag{3}$$

to reflect interest solely in contrasts in the treatment parameters. If specific contrasts are of interest, K can be modified accordingly.

2.2 Estimation and Loss

Estimation of the parameters in the model is by least squares. The volume of the confidence ellipsoid for the linear combinations of the α from the model including the prognostic factors is proportional to the square root of the determinant

$$|A^T(G^TG)^{-1}A| = |K^T\{H^TH - H^TZ(Z^TZ)^{-1}Z^TH\}^{-1}K|. \tag{4}$$

This generalized variance is minimized by a balanced design, in which both an equal number of patients is allocated to each treatment, and there is balance over all prognostic factors so that $H^TZ = 0$. We compare designs using a loss derived from (4).

With two treatments the parameter of interest is $\Delta = \alpha_1 - \alpha_2$. Following Smith (1984a) the model (2) can be written

$$E(Y) = a\Delta + 1\beta_o + Z\theta = a\Delta + F\beta, \tag{5}$$

where a is the $n \times 1$ vector of allocations with elements +1 and -1, and the constant term and covariates are included in the $n \times q$ matrix F. Then (4) becomes

$$\mathrm{var}(\hat{\Delta}) = \sigma^2\{a^Ta - a^TF(F^TF)^{-1}F^Ta\}^{-1}. \tag{6}$$

In (6) $a^Ta = n$, so that we obtain the revealing form

$$\mathrm{var}(\hat{\Delta}) = \frac{\sigma^2}{n - a^TF(F^TF)^{-1}F^Ta} = \frac{\sigma^2}{n - L_n}, \tag{7}$$

where L_n is the loss after n trials.

If the design is exactly balanced, L_n is zero. Otherwise the loss of information is expressed in terms of number of patients. For the randomized designs studied here L_n is a random variable. There are theoretical and simulation results on the expectation $E(L_n) = \mathcal{L}_n$. The results of Smith (1984a) and of Smith (1984b) provide asymptotic values $\mathcal{L}_\infty$ for the expected value of the loss. Burman (1996) focused attention on non-asymptotic expected values $\mathcal{L}_n$, using simulation to study small sample properties. Simulations, which are an extension of those of Burman are used by Atkinson (1999b) to study the progress of the expected value of the loss towards its asymptotic value both for $t = 2$ and $t = 3$. He also gives the extension of (7) to the comparison of t treatments. Here we are interested in the distribution of the values L_n around $\mathcal{L}_n$ when $t = 2$.

2.3 Biased-Coin Designs and Optimality

Optimum design theory is helpful in constructing and describing many of the designs which have been suggested for sequential clinical trials.

If the vector of allocation and prognostic factors for the $(n+1)$st patient is g_{n+1}, G_{n+1} in (2) is formed by adding the row g_{n+1}^T to G_n. This vector includes both the indicator for the allocated treatment and also z_{n+1}, the vector of prognostic factors for the new patient, which are given, not chosen. The experimental region consists solely of points corresponding to allocation of the jth treatment, $j = 1, \ldots, t$. In the iterative construction of D_A-optimum designs to minimize the generalized variance (4), the next trial would be added where the variance

$$d_A(j, n, z_{n+1}) = g_{n+1}^T (G_n^T G_n)^{-1} A \{A^T (G_n^T G_n)^{-1} A\}^{-1} A^T (G_n^T G_n)^{-1} g_{n+1} \quad (8)$$

is a maximum over the design region. In §3 we call this the deterministic design rule. To provide a randomised form of this iterative construction, Atkinson (1982) suggests allocating treatment j with probability

$$\pi_A(j|z_{n+1}) = \frac{d_A(j, n, z_{n+1})}{\sum_{j=1}^{t} d_A(j, n, z_{n+1})}. \quad (9)$$

In (9) the variances $d_A(.)$ could be replaced by any monotone function $\psi\{d_A(.)\}$. In Atkinson (1999a) it is shown that the version of the general Bayesian biased-coin procedure of Ball, Smith, and Verdinelli (1993) which uses D_A-optimality leads to

$$\psi(u) = (1 + u)^{1/\gamma},$$

with γ a parameter to be elucidated from the experimenter. The simulations in Atkinson (2002) show how the bias and loss for this family of rules depends on the value of γ.

3 Allocation Rules

For those allocation rules depending on the variances $d_A(j, n, z_{n+1})$, balance is more nearly achieved by allocating the treatment for which $d_A(j, n, z_{n+1})$, $j = (1, 2)$ is larger. In the absence of prognostic factors, treatment 1 would be allocated if $n_2 > n_1$. By extension of this result, treatment one can be described as being "under represented" if $d_A(1, n, z_{n+1}) > d_A(2, n, z_{n+1})$.

The allocation rules are expressed in terms of probabilities

$$\pi(1) = \text{ prob } [a_{n+1} = 1 | \{d_A(1, n, z_{n+1}) > d_A(2, n, z_{n+1})\}],$$

that is the probability of allocating the "under-represented" treatment one.

D: Deterministic (Sequential Design Construction)

$$\pi_D(1) = 1.$$

The treatment with larger variance $d_A(j, n, z_{n+1})$ is always selected. Asymptotically, for any reasonable distribution over time of prognostic factors, the design will be balanced over the factors and there will be no loss: $\mathcal{L}_\infty = 0$.

R: Completely Randomized

$$\pi_R(1) = 0.5,$$

with $\mathcal{L}_\infty = q$, the number of nuisance parameters, including the constant. The result that randomisation over q variates causes an expected increase in variance of q goes back at least to Cox (1951).

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

E: Efron's Biased-Coin

The "under-represented" treatment is allocated with a probability greater than one half. In particular Efron (1971) elucidated the properties of the rule

$$\pi_E(1) = 2/3,$$

although without covariates. Again, as for the deterministic rule with covariates, $\mathcal{L}_\infty = 0$. Values other than 2/3 will give a different rate of convergence to $\mathcal{L}_\infty$ and a different probability that the clinician can guess correctly which treatment will be allocated next. Some investigations of the properties of schemes with values other than 2/3 are given by Burman (1996). A biased-coin randomization of a'peculiar' deterministic allocation rule for several factors is presented by Efron (1980).

A: $\mathbf{D_A}$-Optimality

With two treatments the biased-coin allocation of Atkinson (1982) according to (9) becomes

$$\pi_A(1) = \frac{d_A(1, n, z_{n+1})}{\sum_{j=1}^{2} d_A(j, n, z_{n+1})}.$$

Burman (1996) shows that $\mathcal{L}_\infty = q/5$.

C: Balanced Covariates

This rule does not depend on the variances $d_A(j, n, z_{n+1})$. The values of the $q-1$ covariates are dichotomised about their individual medians, giving 2^{q-1} possible cells in which the value of z_{n+1} could lie. The under-represented treatment in the cell indicated by z_{n+1} is then allocated, the probability being 0.5 if the numbers of the two treatments are equal. Even if the numbers receiving treatment 1 and treatment 2 are equal in all cells, $\mathcal{L}_n$ will not be identically zero, since the criterion takes no account of the observed values of the z, merely categorising them as above or below the median. A potential practical problem is that the value of the median of each covariate is assumed known. The value of the loss will be inflated if the median is incorrect or if the distribution of the covariates is skewed. Some numerical illustrations of the effect of the resulting inflation of loss are in Atkinson (2002).

A randomised version of the rule could have a biased coin within each cell. If q is not small, the large number of cells may be sparsely filled. There will then be a lack of balance over the margins of the table.

M: Minimisation - Pocock and Simon

The family of rules introduced by Pocock and Simon (1975) are concerned with marginal balance. For the ith element of the covariate vector z_{n+1} let $k(i, n+1)$ be 1 if the element is below the median and 2 if it is above. The marginal totals $m(i, k, j)$ record the number of times a patient with level k of the ith covariate is allocated treatment j. A simple, non-randomised rule exemplified by Pocock and Simon calculates the effect on the balance of allocating treatment 1 as

$$C_1 = \sum_{i=2}^{q} |m\{i, k(i, n+1), 2\} - m\{i, k(i, n+1), 1\} - 1|,$$

with the effect of allocating treatment 2 being

$$C_2 = \sum_{i=2}^{q} |m\{i, k(i, n+1), 2\} - m\{i, k(i, n+1), 1\} + 1|.$$

Treatment 1 is allocated if $C_1 < C_2$ and vice versa, with random allocation when $C_1 = C_2$. A biased-coin version of this rule, in which the treatments allocated with some probability such as $2/3$, rather than certainty, is widely used. Recent examples are surveyed in Atkinson (2002).

B: A Bayesian Biased-Coin

Ball, Smith, and Verdinelli (1993) suggest that the probabilities of treatment selection $\pi_B(j|z_{n+1})$ be chosen to maximize a utility which combines both the variance of parameter estimates and randomness in proportions determined by a parameter γ. The rule depends upon a function of the information matrix. If this is that for D_A-optimality, Atkinson (1999a) shows that

$$\pi_B(j|z_{n+1}) = \frac{\{1 + d_A(j, n, z_{n+1})\}^{1/\gamma}}{\sum_{k=1}^{t}\{1 + d_A(k, n, z_{n+1})\}^{1/\gamma}}. \tag{10}$$

When $\gamma = 0$, the sequential allocation is non-random, reducing to the deterministic rule. As $\gamma \to \infty$, the procedure tends towards the random allocation rule R. These two values of γ thus provide procedures which respectively minimize variance by maximising balance and minimize potential bias by maximising randomness. Provided $\gamma > 0$, the procedure leads, with a speed depending on the value of γ, to random allocation and a loss of q as $n \to \infty$.

4 The Distribution of Loss

4.1 Expected Loss

The distribution of loss is investigated by simulating series of clinical trials for up to 200 trials using these seven allocation rules. Table 1 gives the average loss for each rule over 1,000 trials when $n = 200$. Also given are the values of $\mathcal{L}_\infty$, when these are known. The covariates are taken to be independently and identically normally distributed with zero mean.

The results show that the losses are mostly close to their asymptotic values, especially when $q = 5$. The deterministic rule has forced balance, with a loss close to zero. For randomisation and D_A-optimality the losses are near to q and $q/5$ and, for the Bayes rule, γ is sufficiently large that the rule is approaching random allocation.

The plots of these losses as functions of n are in Fig. 1, both for $q = 5$ and for $q = 10$. They show that the losses for most rules have levelled out after an initially high start when there are so few observations that the design is unbalanced. The arbitrary rules C and M have reached stable values for $q = 5$, but not for $q = 10$. Particularly for C, most designs are far from balance. Efron's rule at $n = 200$ is still decreasing towards its asymptotic value of zero. The Bayes rules are the only ones for which the loss increases with n, after an initial decline.

These results divide the allocation rules into three groups - those for which the asymptotic value is non-zero, those for which it is zero and the rules C

Allocation Rule	Average Loss $q = 5$	Average Loss $q = 10$	Asymptotic Value $\mathcal{L}_\infty$
D_A-optimality A	1.028	2.0937	$q/5$
Covariate Balance C	1.634	8.015	?
Deterministic D	0.054	0.211	0
Efron's Biased Coin E	0.542	1.913	0
Bayes B ($\gamma = 0.1$)	3.573	7.229	q
Minimisation M	1.522	3.598	?
Random R	4.898	9.886	q

Table 1: Average losses from 1,000 simulations of 200 patient clinical trials for five and ten nuisance parameters - four and nine covariates. Also given, where known, are the asymptotic values of the loss $\mathcal{L}_\infty$

and M, dependent on the categorisation of the covariates, for which $\mathcal{L}_\infty$ is not known.

4.2 A Chi-squared Approximation

We now consider the distribution of the losses summarized in Table 1 and Fig. 1. The left-hand panel of Fig. 2 gives boxplots of the distributions of loss for rule A at eight values of n from 25 to 200. The means of these distributions initially decrease gradually, but comparison with Fig. 1 shows that the large decrease in the mean has already occurred by $n = 25$. The shape of the eight distributions appears similarly skewed.

Since loss is a non-negative quantity we try a standard approximation to non-negative skewed distributions which is a scaled chi-squared distribution on ν degrees of freedom. The scaling is estimated so that the distribution has the correct mean, that is we assume

$$L_n \sim (\mathcal{L}_n/\nu)\chi^2_\nu. \tag{11}$$

An idea of ν can be found by QQ plots of the empirical distribution of loss against a selection of χ^2 distributions. The right-hand panel of Fig. 2 shows the results for $n = 200$ and $\nu = 5$. There are 1,000 observations. The plot is acceptably straight and a little straighter than those for ν equal to four or six, although in all cases the eye is drawn to the 1% of the trials in the top right-hand corner of the plot. In later sections we estimate ν by maximum likelihood.

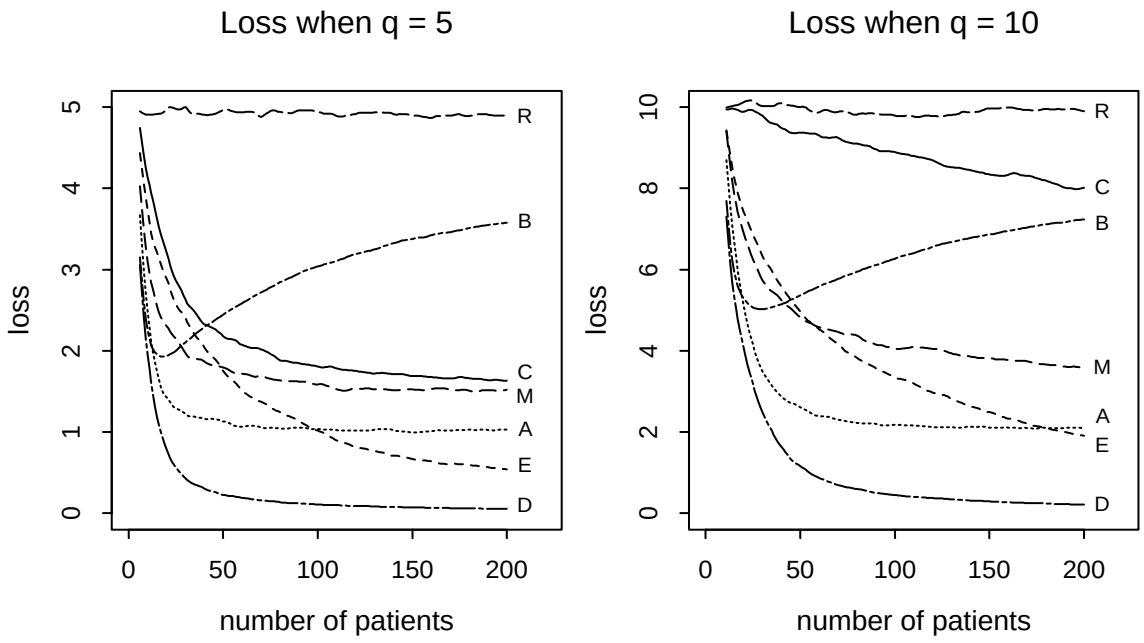

Figure 1: Loss $\mathcal{L}_n$ for seven strategies for sequential allocation of treatments with $t = 2$ and $q = 5$ and 10: A, D_A-optimality; B, Bayesian biased coin, $\gamma = 0.1$; C, Covariate Balance; D, Deterministic; E, Efron's Biased Coin; M Minimisation and R, Random. Means of 1,000 simulations

Fig. 3 shows similar results for Efron's biased coin. The boxplots show how the losses again decrease, more rapidly than those in Fig. 2. The right-hand panel of the figure indicates that $\nu = 3$ is a good approximation to the distribution of loss when $n = 200$. The results, again for $q = 5$, for the Bayes rule with $\gamma = 0.1$ in Fig. 4 show how, for this rule, the average loss increases with n. Again $\nu = 5$ provides a good chi-squared approximation to the distribution when $n = 200$.

Similar support for an approximating chi-squared distribution is obtained from the simulations when $n = 10$. Fig. 5 shows the boxplots and QQ plot for random allocation when $n = 10$. There is no noticeable trend of the expected loss with n - it remains close to ten - and the QQ plot for $\nu = 10$ indicates that the distribution fits well. The greatest contrast to this is the pair of plots in Fig. 6 for the deterministic rule: the average loss decreases sharply with n as does the spread of the distribution. When $n = 200$, a good fit is obtained with $\nu = 13$.

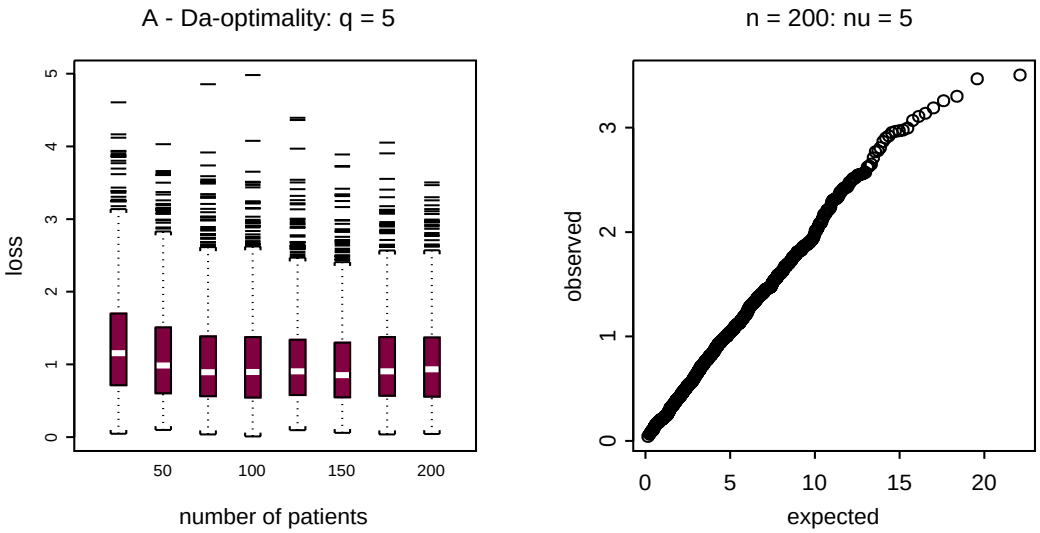

Figure 2: Distribution of loss for D_A-optimality when $q = 5$: (a) boxplots of the distribution of loss L_n; (b) QQ plot of L_{200} against χ^2_5. 1,000 simulations

4.3 Five Nuisance Parameters

These preliminary results suggest that for some, if not all, of the rules, the degrees of freedom ν may be equal to q, the number of nuisance parameters. This would match the suggestion of Cox for random allocation. We now determine for which allocation rules $\nu = q$ provides a good approximation to the distribution of loss.

The QQ plots shown in the figures are each for one sample of 1,000 trials from which the value of ν can be estimated. These simulations were repeated 100 times, giving 100 estimates of ν for selected n and the seven rules. The estimates were found by maximizing the likelihood and tested for equality to q using a likelihood ratio test, which will have an asymptotic chi-squared distribution on one degree of freedom. This maximum likelihood test is more complicated than that which arises in the gamma family of generalized

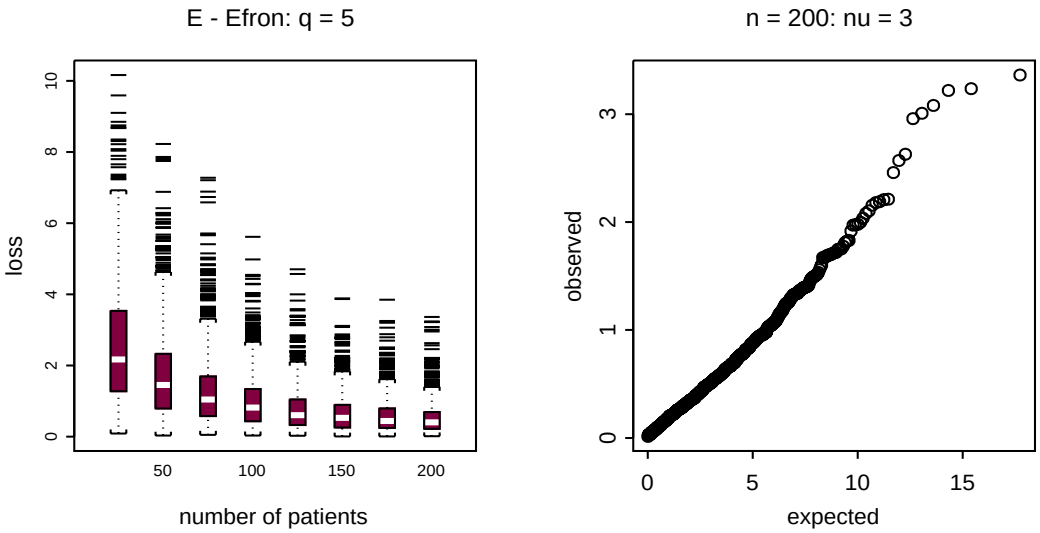

Figure 3: Distribution of loss for Efron's biased coin when $q = 5$: (a) boxplots of the distribution of loss L_n; (b) QQ plot of L_{200} against χ^2_3. 1,000 simulations

linear models, when comparisons are between nested models with the same dispersion parameter, that is, value of ν.

The mean values of the 100 estimates of ν when $n = 200$ and $q = 5$ are given in Table 2. Plots of the mean values for eight values of n, together with asymptotic 95% confidence intervals for ν, are plotted in Fig. 7. These results extend those implied in the QQ plots of the earlier figures. Above $n = 50$ the deterministic rule D has a value of ν around 6. The values for rules C and M decrease to around 4, whereas Efron's biased coin decreases steadily to around three. However, for three rules, A, B and R, the estimates seem to have stabilized around $\nu = 5$.

The narrowness of the confidence intervals in the figure shows that the conclusions are clear of random error from the simulations. This conclusion is confirmed by the values of the statistics for testing $\nu = 5$, which are summarized in Table 3 and plotted in Fig. 8. This figure chiefly shows the

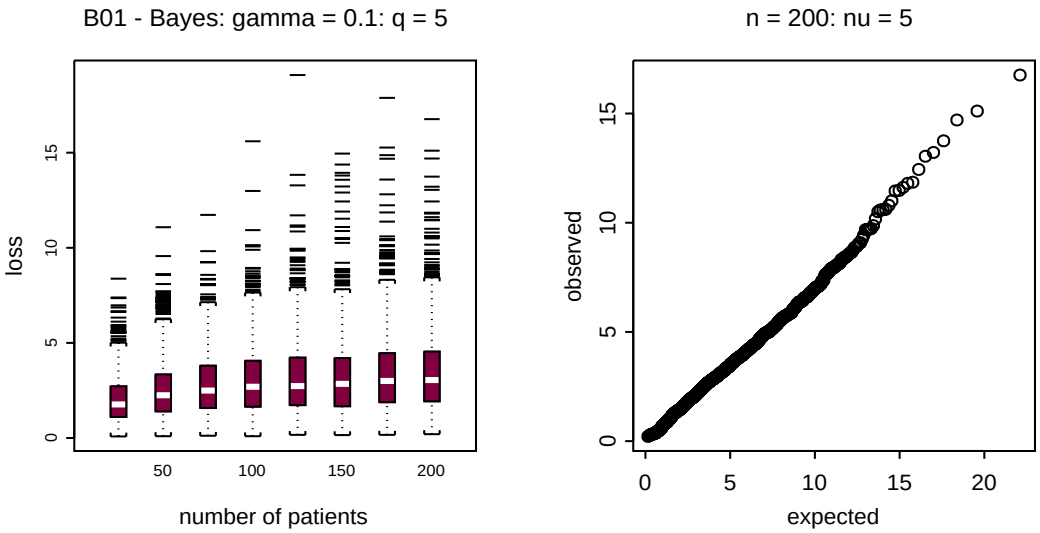

Figure 4: Distribution of loss for the Bayesian biased coin with $\gamma = 0.1$ when $q = 5$: (a) boxplots of the distribution of loss L_n; (b) QQ plot of L_{200} against χ^2_5. 1,000 simulations

value for E increasing steadily to around 160. The values for the rules M, D and C are also significantly high when compared with the null χ^2_1 distribution. Fig. 9 shows the evolution of the statistics that are not significant at $n = 200$. Both rules B and A settle down early on to being well approximated by a chi-squared on five degrees of freedom. Random allocation, R, initially is not so well approximated, due to the lack of balance the rule sometimes gives when there are few patients. The major conclusion is that for three rules, A, B and R, the degrees of freedom are very close to five.

4.4 Ten Nuisance Parameters

The results with $q = 10$ are similar to those for $q = 5$, but less sharp, since the increase in q requires an increase in n for asymptotic results to start to hold. The values for the estimates of ν in Table 2 show that for five out of

Allocation Rule	Average d.f. $q = 5$	Average d.f $q = 10$
D_A-optimality A	5.08	10.28
Covariate Balance C	4.28	10.36
Deterministic D	6.04	12.74
Efron's Biased Coin E	3.10	6.15
Bayes B ($\gamma = 0.1$)	5.11	10.26
Minimisation M	4.05	9.16
Random R	5.14	10.51

Table 2: Average degrees of freedom for the chi-squared approximation to the distribution of loss when $n = 200$. Average of 100 repetitions of 1,000 simulations of 200 patient clinical trials for five and ten nuisance parameters - four and nine covariates

Allocation Rule	Average l.r. $q = 5$	Average l.r $q = 10$
D_A-optimality A	0.99	1.23
Covariate Balance C	15.87	1.54
Deterministic D	20.00	29.20
Efron's Biased Coin E	160.64	154.97
Bayes B ($\gamma = 0.1$)	1.60	1.24
Minimisation M	28.53	5.36
Random R	1.92	2.36

Table 3: Average likelihood ratio test for testing that the degrees of freedom of the chi-squared approximation to the distribution of loss are equal to q. Average of 100 repetitions of 1,000 simulations of 200 patient clinical trials for five and ten nuisance parameters - four and nine covariates

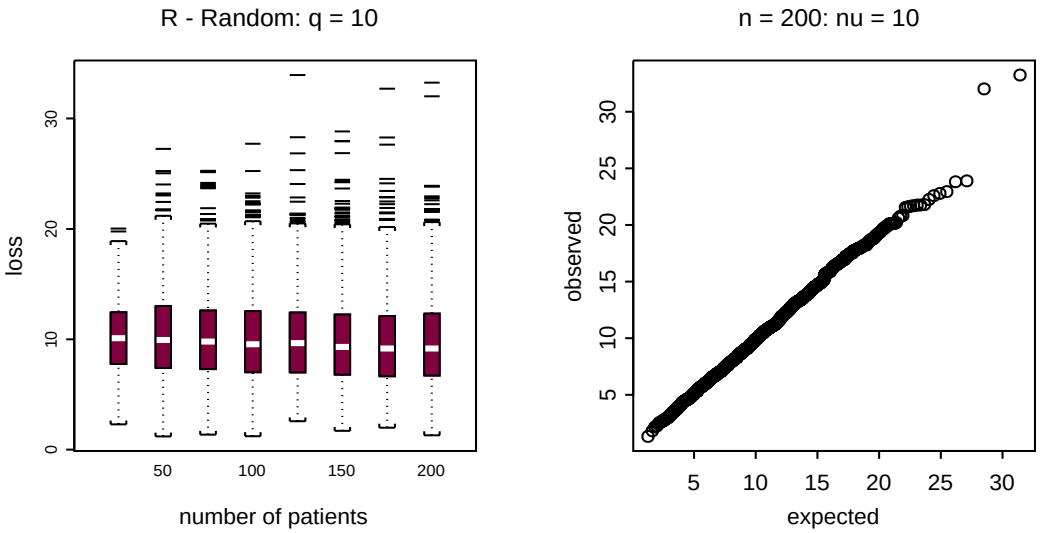

Figure 5: Distribution of loss for random allocation when $q = 10$: (a) boxplots of the distribution of loss L_n; (b) QQ plot of L_{200} against χ^2_{10}. 1,000 simulations

seven of the rules the estimate has almost exactly doubled. Only for rules C and M is the increase rather more than twice. The values for the test statistics in Table 3 also show that doubling the number of covariates affects procedures C and M differently from the others, for which the test statistics are little changed in value. In particular the distribution of loss for rule C is now well approximated by the χ^2_{10} distribution.

The plot of estimates of ν in Fig. 10 shows that rules D and E have estimates far from ten. That for M, just above 9, is perhaps slowly increasing. The other four rules A, B, C and R all have similar approximations. At $n = 200$, the value of the test statistic for R is 2.36, the largest for these four rules. However, the final plot, Fig. 11, shows that the ratio for R is still decreasing at $n = 200$. An asymptotic chi-squared distribution for the loss on q degrees of freedom is acceptable for these four rules when $q = 10$.

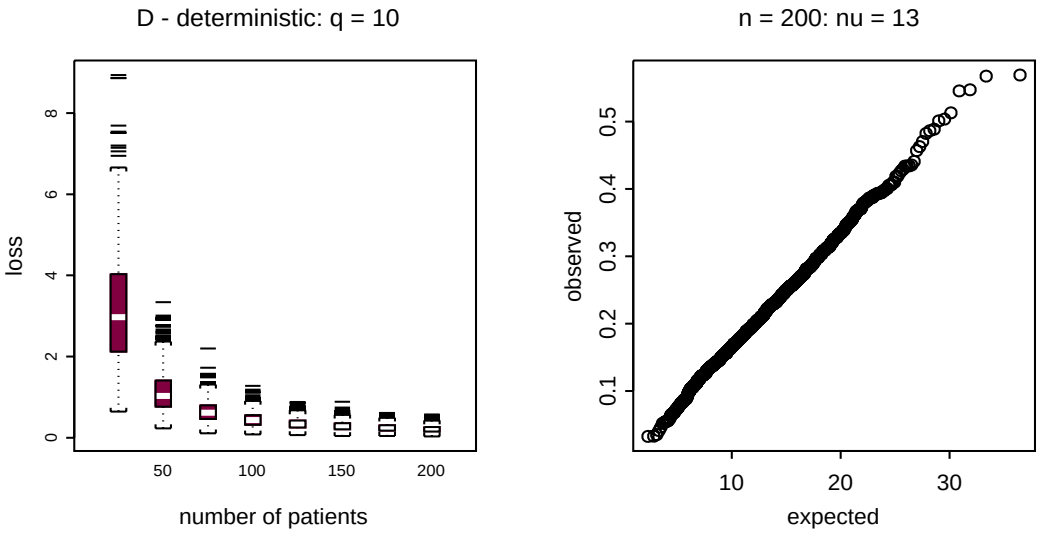

Figure 6: Distribution of loss for deterministic allocation when $q = 10$: (a) boxplots of the distribution of loss L_n; (b) QQ plot of L_{200} against χ^2_{13}. 1,000 simulations

5 Discussion

The conclusion from this study is clear. The distribution of loss divides the rules into three groups.

The first group contains the rules A, B and R for which $\mathcal{L}_\infty$ is either $q/5$ or q. Then we can be explicit about the degrees of freedom in (11) and state that

$$L_n \sim (\mathcal{L}_n/q)\chi^2_q = \mathcal{L}_n F_{q,\infty}. \tag{12}$$

The results of the likelihood ratio tests when $q = 5$ in Fig. 9 show that this approximation holds for B when n is at least 25, for A when $n \geq 50$ and for R when $n \geq 75$. Larger minimum sample sizes are indicated by the results for $q = 10$ in Fig. 11: 50 for B, 75 for A and 150 for R, roughly twice the sample sizes required for $q = 5$. These results illustrate the rate of convergence to the asymptotic results implicit in Smith (1984a).

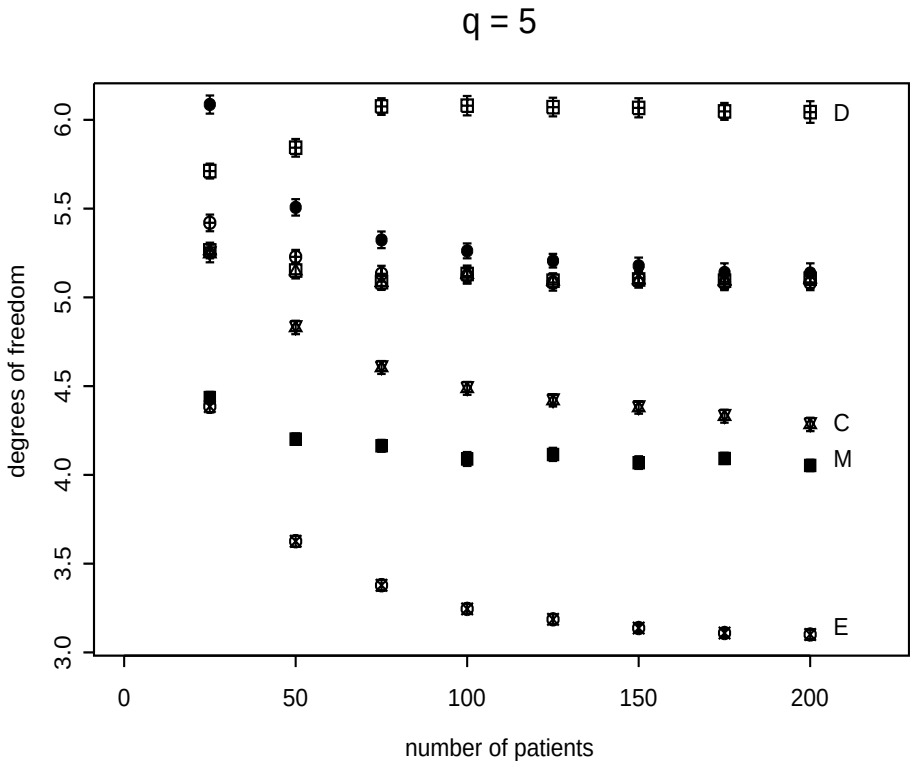

Figure 7: Mean estimates of ν and 95% confidence intervals from 100 estimates based on simulations of 1,000 trials with $q = 5$. C, Covariate Balance; D, Deterministic; E, Efron's Biased Coin and M Minimisation

The second group of rules, D and E, are those for which $\mathcal{L}_\infty = 0$. Over the range of values of n studied the results show that the distribution of loss is well approximated by a chi-squared distribution, but that the degrees of freedom depend on n, especially for E. Figures 7 and 10 show this dependence. For $n = 200$ and $q = 5$, ν is around 6 for D when $q = 5$ and around 3 for E. As Table 2 indicates, these values roughly double when $q = 10$.

Finally there are the rules C and M, which are not based on the linear model (1). Theoretical results for these rules have not yet been obtained, so that the values of $\mathcal{L}_\infty$ are not known. The results here establish chi-squared approximations to the distributions of loss for these rules. For $q = 5$ they behave similarly, with a value of ν around 4. But doubling the value of q has

a very different effect on the two rules: in particular for C the expected loss and the degrees of freedom of the chi-squared distribution for $n = 200$ much more than double.

The deterministic strategy D minimizes loss while excluding randomization. The effect of randomization is most easily quantified using "selection bias", introduced by Blackwell and Hodges (1957). This is the probability that the allocation of the treatment to the next patient can be correctly guessed. It is one for D, 0 for complete randomization and has intermediate values for the other rules - for example, one third for rule E. An asymptotic expression for the mean selection bias which includes rules A and B is given by Smith (1984a, p.1033). Simulations of expected selection bias for the non-asymptotic sample sizes studied in this paper are in Atkinson (2002).

Of course, in a double-blind trial the clinician is not able to guess the next treatment to be allocated and so is unable to influence which patient receives which treatment. This is even more so in the case of a multi-centre trial in which treatments are allocated centrally. Selection bias should therefore be considered as a calculable surrogate for all the reasons for which randomness is required. Efron (1971) and Steele (1980) show that biased-coin designs give good protection against smooth trends and short-range cyclical patterns. Smith (1984b) extends this work and includes correlated errors. Taves (2001) gives a further discussion of reasons for randomisation in clinical trials and Matthews (2000, §2.2) considers the various biases which may arise.

In general, designs with small expected loss have high expected selection bias, and vice versa. Atkinson (2002) uses simulation to study the admissibility of allocation rules measured by expected loss and expected selection bias. A rule with higher loss and higher bias than another rule is inadmissible. The need for a systematic comparison of rules, to which this paper is intended to make a contribution, is highlighted by the claim of Treasure and MacRae (1998) that minimisation forms a "platinum standard". Unfortunately, the form of minimisation given here, rule M, is inadmissible.

The calculations described in this paper assume that the observations are normally distributed, perhaps after transformation. However, provided the treatment effects are moderate, the results should extend to generalized linear models in the same manner as Cox (1988) argued that factorial designs extend to such models. The technical point is that, if the treatment effects are moderate, the variation in response will not be large. Then the iterative weights used in least squares fitting will have similar values for all observations. As a consequence, maximum likelihood estimation is close to unweighted least squares. Whatever the appropriate model it is however necessary to take account of the distribution of loss as well as its average value.

References

Atkinson, A. C. (1982). Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika 69*, 61–67.

Atkinson, A. C. (1999a). Bayesian and other biased-coin designs for sequential clinical trials. *Tatra Mountains Mathematical Publications 17*, 133–139.

Atkinson, A. C. (1999b). Optimum biased-coin designs for sequential treatment allocation with covariate information. *Statistics in Medicine 18*, 1741–1752.

Atkinson, A. C. (2002). The comparison of designs for sequential clinical trials with covariate information. *Journal of the Royal Statistical Society, Series A 165*. (In Press).

Ball, F. G., A. F. M. Smith, and I. Verdinelli (1993). Biased coin designs with a Bayesian bias. *Journal of Statistical Planning and Inference 34*, 403–421.

Blackwell, D. and J. L. Hodges (1957). Design for the control of selection bias. *Annals of Mathematical Statistics 28*, 449–460.

Burman, C.-F. (1996). *On Sequential Treatment Allocations in Clinical Trials*. Göteborg: Department of Mathematics.

Cox, D. R. (1951). Some systematic experimental designs. *Biometrika 38*, 312–323.

Cox, D. R. (1982). A remark on randomization in clinical trials. *Utilitas Mathematica 21A*, 245–252.

Cox, D. R. (1988). A note on design when response has an exponential family distribution. *Biometrika 75*, 161–164.

Efron, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika 58*, 403–417.

Efron, B. (1980). Randomizing and balancing a complicated sequential experiment. In R. J. Miller, B. Efron, B. W. Brown, and L. E. Moses (Eds.), *Biostatistics Casebook*, pp. 19–30. New York: Wiley.

Matthews, J. N. S. (2000). *An Introduction to Randomized Controlled Clinical Tials*. London: Edward Arnold.

Pocock, S. J. and R. Simon (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics 31*, 103–115.

Senn, S. (2000). Consensus and controversy in pharmaceutical statistics (with discussion). *The Statistician 49*, 135–176.

Senn, S. J. (1997). *Statistical Issues in Drug Development*. Chichester: Wiley.

Smith, R. L. (1984a). Properties of biased coin designs in sequential clinical trials. *The Annals of Statistics 12*, 1018—1034.

Smith, R. L. (1984b). Sequential treatment allocation using biased coin designs. *Journal of the Royal Statistical Society, Series B 46*, 519—543.

Steele, J. M. (1980). Efron's conjecture on vulnerability to bias in a method for balancing sequential trials. *Biometrika 67*, 503–504.

Taves, D. R. (2001). Letter to the editor about Atkinson 1999b. *Statistics in Medicine 20*, 813–816.

Treasure, T. and K. D. MacRae (1998). Minimisation: the platinum standard for trials? *British Medical Journal 317*, 362–363.

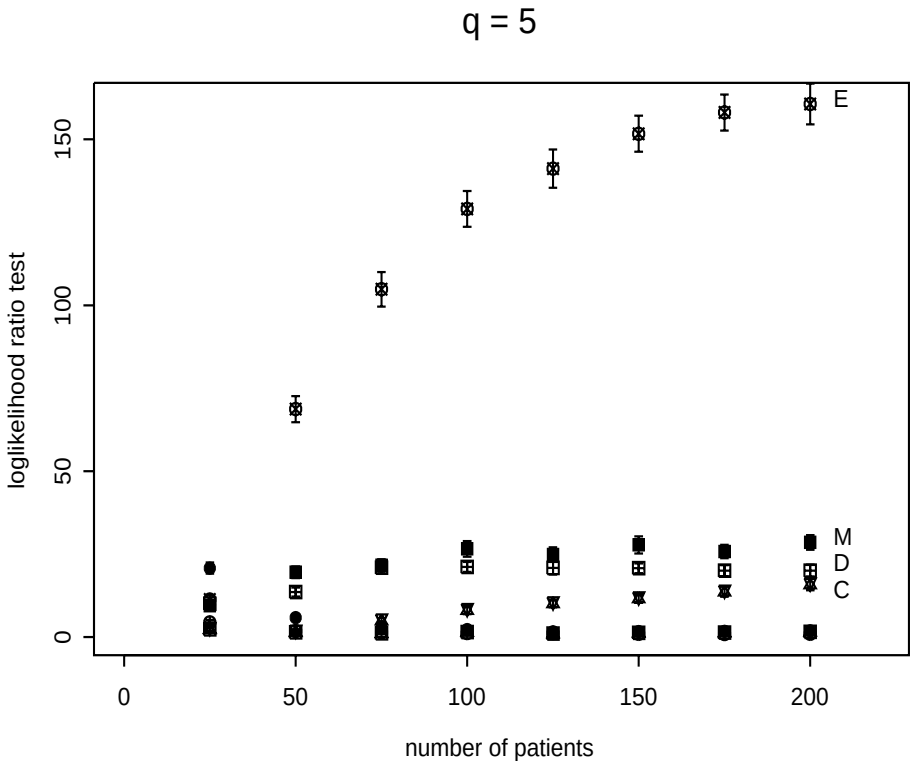

Figure 8: Loglikelihood ratio test for $\nu = 5$: means and 95% confidence intervals from 100 estimates based on simulations of 1,000 trials with $q = 5$. C, Covariate Balance; D, Deterministic; E, Efron's Biased Coin and M Minimisation

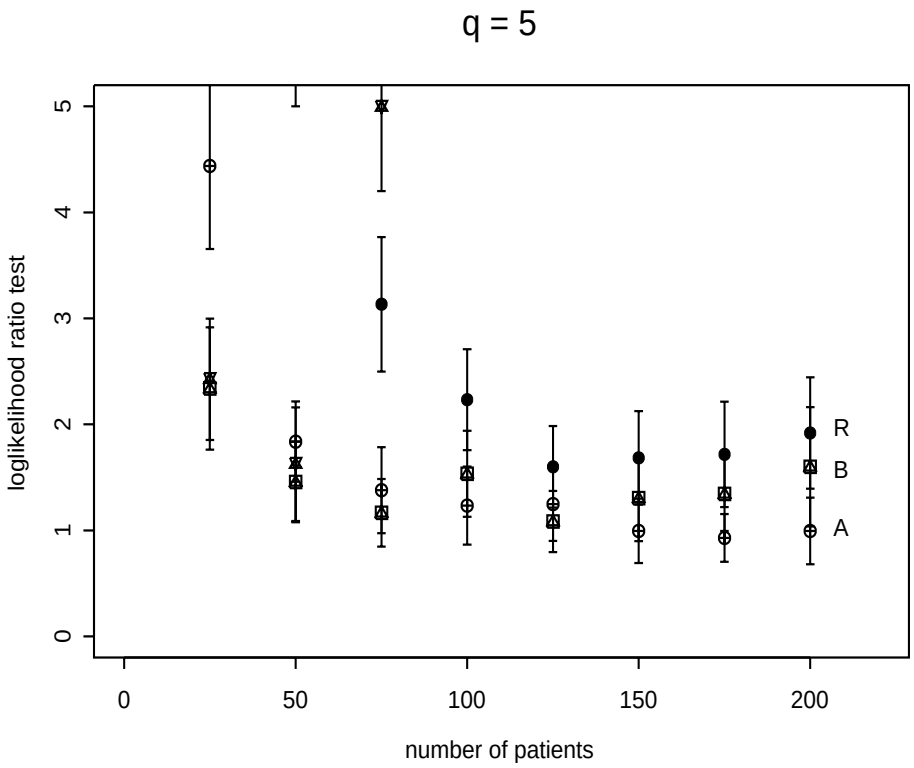

Figure 9: Loglikelihood ratio test for $\nu = 5$: means and 95% confidence intervals from 100 estimates based on simulations of 1,000 trials with $q = 5$. A, D_A-optimality; B, Bayesian biased coin and R, Random

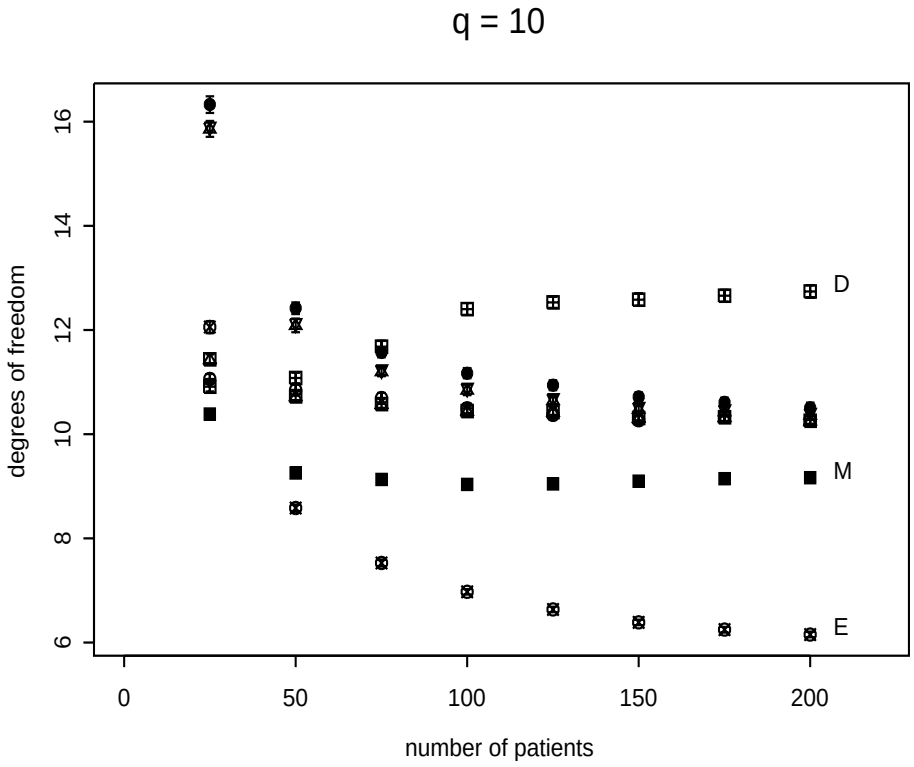

Figure 10: Mean estimates of ν and 95% confidence intervals from 100 estimates based on simulations of 1,000 trials with $q = 10$. D, Deterministic; E, Efron's Biased Coin and M Minimisation

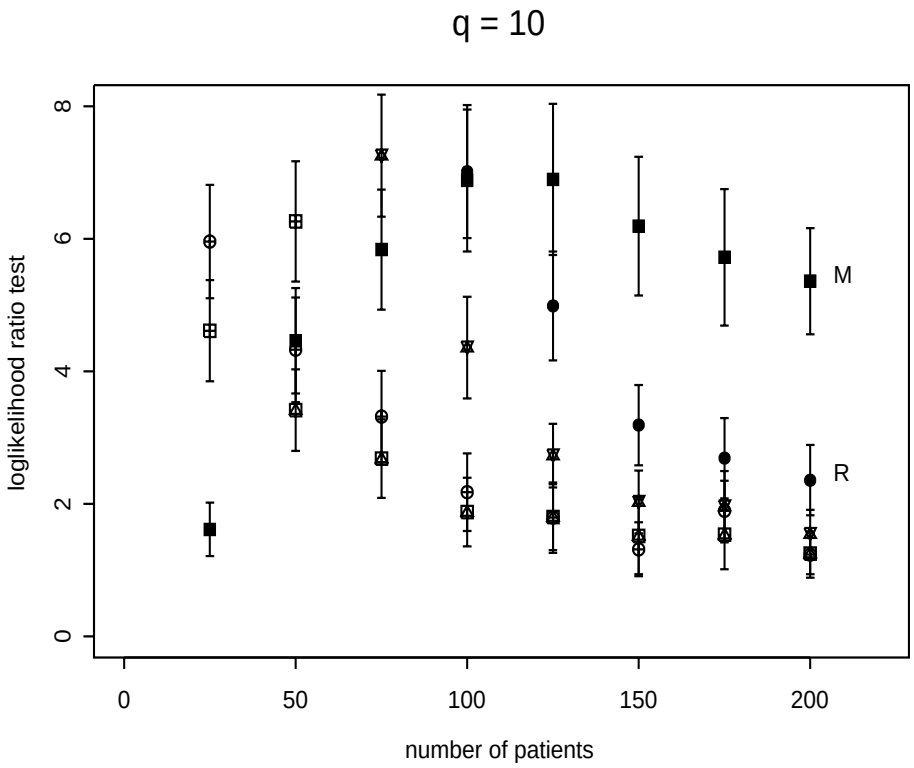

Figure 11: Loglikelihood ratio test for $\nu = 10$: means and 95% confidence intervals from 100 estimates based on simulations of 1,000 trials with $q = 10$. M Minimisation and R, Random