Anisimov, V. V., Yeung, W. Y., and Coad, D. S. (2017) Imbalance properties of centre-stratified permuted-block and complete randomisation for several treatments in a clinical trial. Statistics in Medicine, 36(8), pp. 1302-1318.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

This is the peer reviewed version of the following article: Anisimov, V. V., Yeung, W. Y., and Coad, D. S. (2017) Imbalance properties of centre-stratified permuted-block and complete randomisation for several treatments in a clinical trial. Statistics in Medicine, 36(8), pp. 1302-1318, which has been published in final form at http://dx.doi.org/10.1002/sim.7206. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

http://eprints.gla.ac.uk/133643/

Deposited on: 6 March 2017

Imbalance properties of centre-stratified permuted-block and complete randomisation for several treatments in a clinical trial

Vladimir V. Anisimov[1], Wai Y. Yeung[2] and D. Stephen Coad[3]
[1]School of Mathematics and Statistics, University of Glasgow,
15 University Gardens, Glasgow G12 8QW, U.K.
[2]Roche Products Limited, 6 Falcon Way, Shire Park,
Welwyn Garden City AL7 1TW, U.K.
[3]School of Mathematical Sciences, Queen Mary, University of London,
Mile End Road, London E1 4NS, U.K.

Abstract

Randomisation schemes are rules that assign patients to treatments in a clinical trial. Many of these schemes have the common aim of maintaining balance in the numbers of patients across treatment groups. The properties of imbalance that have been investigated in the literature are based on two treatment groups. In this paper, their properties for $K > 2$ treatments are studied for two randomisation schemes: centre-stratified permuted-block and complete randomisation. For both randomisation schemes, analytical approaches are investigated assuming that the patient recruitment process follows a Poisson-gamma model. When the number of centres involved in a trial is large, the imbalance for both schemes is approximated by a multivariate normal distribution. The accuracy of the approximations is assessed by simulation. A test for treatment differences is also considered for normal responses and numerical values for its power are presented for centre-stratified permuted-block randomisation. To speed up the calculations, a combined analytical/approximate approach is used.

Keywords multi-centre clinical trial, multivariate normal distribution, Poisson-gamma model, power, covariance, simulation

1   Introduction

Randomisation schemes used in clinical trials are considered as essential and of great importance to maintain a balance in the numbers of patients across treatment groups and to gain some randomness in assigning a treatment to a patient to avoid any selection or accidental bias. Properties of various randomisation schemes have been studied by, for example, Rosenberger and Lachin [1] and Senn [2]. Typically, the purpose of trial design and selecting the treatment allocation method is to construct the most efficient estimator of the treatment effect, that is, the one with minimum variance. Under the assumption that the variances of the responses are the same for each treatment group, this can be achieved by equal allocation. In general, the optimal allocation is defined by some linear relation between the numbers of patients on the treatment arms (Brittain

1

and Schlesselman [3]). However, due to the stochastic nature of both patient enrolment and the randomisation process, imbalance in assignments can still occur, even in well-defined randomisation schemes. Since a larger imbalance can reduce the power of a study, the statistical analysis of properties of imbalance is of paramount importance for clinical trial design.

A survey of using various methods of randomisation in clinical trials is provided by McPherson *et al.* [4]. It is stressed that the analysis of imbalance in the numbers of patients across treatment groups becomes more important in clinical trials where fewer patients are recruited and involved in the study, as an imbalance in the numbers of patients in the treatment groups may decrease the probability of correctly detecting a genuine treatment effect. For multi-centre trials, Kundt and Glass [5] emphasise the importance of stratified randomisation in order to reduce the possibility of imbalance across treatment groups.

In the particular case of equal allocation ratios, under complete randomisation for two-treatment trials ignoring centres, each patient is equally likely to be assigned to either of the treatments. Although in this case for n patients the total imbalance is $O(\sqrt{n})$, it is known that imbalance caused by complete randomisation is unlikely to have a strong impact in a large trial with more than 200 patients (Lachin [6]).

Another randomisation scheme that is mainly used in clinical trials by pharmaceutical companies is randomly permuted-block randomisation. There are two basic schemes: unstratified and centre-stratified permuted-block randomisation. Given an assigned proportion of patients for each treatment within a block, a sequence of treatment assignments is generated by randomly listing all of the possibilities for different permuted blocks. Patients are then assigned to the treatment according to this sequence in order upon their arrival. In this paper, unequal allocations are allowed within a block. However, in order to achieve a balance in the numbers of patients across treatment groups, the treatment allocation ratio within a block is usually assumed to be one. Under this assumption, imbalance will not occur in any of the complete blocks, but it may occur in the incomplete block. In other words, under the permuted-block design, the total imbalance in the numbers of patients across treatments is determined by imbalances in the incomplete blocks. A comparison of the pros and cons of unstratified and some variations of centre-stratified randomisation using simulation is considered in Morrissey *et al.* [7].

The treatment imbalance properties of different randomisation schemes have been investigated by Hallstrom and Davis [8] and Kundt and Glass [5] for stratified block randomisation, Lachin [6, 9] for complete randomisation and urn randomisation, and Anisimov [10, 11, 12] for centre-stratified permuted-block randomisation. All of the above authors have considered the imbalance in the number of patients for two treatments. This paper extends these results to more than two treatment groups. In the previous papers, the expectation and the variance of the imbalance are calculated, and it is proved that the distribution of the overall imbalance is approximately normal. The main focus of this paper is to investigate the imbalance properties of two randomisation schemes: centre-stratified permuted-block and complete randomisation. The overall imbalance is represented as a vector of imbalances on different treatments rather than a scalar in the two-treatment case, and asymptotically it has a multivariate normal distribution.

In Section 2, we consider centre-stratified permuted-block and complete randomisation for cases when there are more than two treatments. The imbalance for both randomisation schemes will then be defined for a particular treatment in a particular centre and across all centres. In Section 3, we consider the case where the numbers of patients

recruited in different centres are random variables. This is based on the Poisson-gamma patient recruitment model developed by Anisimov and Fedorov [13]. The expectations, variances and covariances of the imbalance within a centre and across all centres are evaluated. In Section 4, the variances and covariances of the imbalance for the two randomisation schemes are compared. A simulation study is considered for a particular scenario in Section 5. The values for the expectations and the covariance matrices are given, together with the empirical distributions of the imbalance, which confirm the theoretical results produced in Section 3. In Section 6, a test will be described for all pairs of treatment differences with a control group and hence how the power can be obtained. Numerical results for the power and the sample size are also provided by simulation for different scenarios. By fixing a particular level of power to be achieved for the balanced case, the sample size can be found using simulation. A similar, but more complicated, approach is used in the imbalanced case. Finally, there is a discussion in Section 7. Derivations of the properties of the imbalance are given in an appendix.

2   Randomisation schemes for more than two treatments

Consider a multi-centre clinical study, where in total n patients have to be recruited by N clinical centres. Patients have to be assigned to one of $K \geq 3$ treatments. Assume that n_i patients will be recruited in centre i. Then $\sum_{i=1}^{N} n_i = n$.

There are two types of permuted-block randomisation schemes that are commonly used in clinical trials: unstratified and centre-stratified permuted-block randomisation. Unstratified permuted-block randomisation means that patients are randomised to treatment according to independent randomly permuted blocks of a fixed size without regard to centres. Centre-stratified randomisation means that each of the centres has a separate permuted-block randomisation scheme. Patients for the study are randomised to treatments according to the independent randomly permuted blocks of a fixed size within each centre.

It is obvious that unstratified permuted-block randomisation will minimise the imbalance in the number of patients on different treatments for the whole study and increase the imbalance in each centre compared to centre-stratified permuted-block randomisation.

Assume that there are in total $K \geq 3$ treatments with the allocation in the block $(k_1, k_2, ..., k_K)$, where k_j is the number of patients within a block that are allocated to treatment j for $j = 1, ..., K$. Let $B = \sum_{j=1}^{K} k_j$ be the block size. As a simple example, suppose that $K = 3$ and that our block is $(2, 2, 2)$. Then we have $B = 6$.

Consider centre-stratified permuted-block randomisation. The total number of patients n_i recruited in centre i may not be a multiple of the block size B. This may lead to incomplete blocks in some of the centres. In the incomplete block, the numbers of patients on different treatments may not be proportional to the allocations within the block $(k_1, k_2, ..., k_K)$, and this will cause imbalance in this block. Many incomplete blocks may lead to larger imbalance across all centres. Therefore, we can study the imbalance properties of a randomisation scheme by studying the properties of incomplete blocks. For some centre i, denote by r the size of the incomplete block, where $r = 1, ..., B - 1$. In the above example, we could have $r = 3$.

Let $\xi_j(r)$ be the number of patients on treatment j in an incomplete block of size

3

r. Then the variable $\xi_j(r)$ has a multivariate hypergeometric distribution, that is,

$$\mathbf{P}[\xi_1(r) = s_1, \ldots, \xi_K(r) = s_K] = \frac{\prod_{j=1}^{K} \binom{k_j}{s_j}}{\binom{B}{r}}$$

for $s_j = 0, 1, \ldots, \min(k_j, r)$ and $j = 1, \ldots, K$ with

$$\mathbf{E}[\xi_j(r)] = \frac{k_j r}{B}, \quad \mathbf{Var}[\xi_j(r)] = \frac{k_j r (B - k_j)(B - r)}{B^2 (B - 1)} \tag{2.1}$$

and

$$\mathbf{Cov}[\xi_j(r), \xi_m(r)] = -\frac{k_j k_m r (B - r)}{B^2 (B - 1)}.$$

If $r = 0$, we set $\xi_j(r) = 0$.

Let n_{ij} be the total number of patients in centre i on treatment j. Denote by $\lfloor z \rfloor$ the integer part of z and $mod(z, B) = z - \lfloor z/B \rfloor B$. Then

$$n_{ij} = \left\lfloor \frac{n_i}{B} \right\rfloor k_j + \xi_j \left[mod(n_i, B) \right]. \tag{2.2}$$

The centre-stratified permuted-block design is studied and referred to as permuted-block randomisation in this paper.

Consider now complete randomisation. Suppose that each patient is assigned to treatment j with probability p_j without regard to clinical centres. Then the total number of patients in centre i on treatment j, n_{ij}, has a binomial distribution with parameters n_i and p_j. Typically, p_j is chosen as the proportion of patients that are allocated to treatment j in a complete block of size B, and so $p_j = k_j/B$. Note that blocks are only introduced here for comparison purposes as the same allocation ratios apply for the different randomisation schemes, and the size of a block is not used in subsequent calculations. Therefore, for complete randomisation, we use for convenience $p_j = k_j/B$. For equal treatment proportions within each complete block, we have $k_1 = k_2 = \ldots = k_K$ with block size $B = K k_1$ and $p_j = 1/K$. Thus, in the above example, we see that $k_j = 2$ and $p_j = 1/3$.

Define for centre i the imbalance on treatment j as

$$\Delta_{ij} = n_{ij} - \frac{n_i}{B} k_j \tag{2.3}$$

for $j = 1, \ldots, K$ and $i = 1, \ldots, N$. Now define the overall imbalance on a particular treatment j for all centres by

$$\Delta_j = \sum_{i=1}^{N} \Delta_{ij}. \tag{2.4}$$

In what follows, Δ_{ij} and Δ_j with and without a superscript * will refer to the imbalances under complete randomisation and the permuted-block design, respectively.

3 Imbalance properties of randomisation schemes

3.1 Patient Recruitment Model

Consider a Poisson-gamma model for the patient recruitment process developed by Anisimov and Fedorov [13]. This model accounts for the stochasticity in patient arrival and also the variation in recruitment rates between different centres. The model

has been validated for many large clinical trials. It provides the opportunity to model analytically patient recruitment at the centre level, and therefore it can be used as a convenient framework for our purposes.

Assume that the patient recruitment process in centre i follows a Poisson process with rate λ_i. The rates $\{\lambda_i\}$ are viewed as independent gamma distributed random variables with known shape and rate parameters (α, β), respectively. Assume for simplicity that all centres start to recruit patients at the same time. Then, for given rates, the total number of patients n_i recruited in centre i has a mixed binomial distribution with parameters n and p_i, where n is the total number of patients recruited in all centres and p_i is the probability of recruiting a patient in centre i given by $p_i = \lambda_i / \sum_{k=1}^{N} \lambda_k$. Since $\{\lambda_i\}$ are gamma distributed random variables, p_i has a beta distribution with parameters $(\alpha, \alpha(N-1))$. Therefore, n_i has a beta-binomial distribution with

$$\mathbf{P}[n_i = l] = \mathbf{P}(n, N, \alpha, l) = \binom{n}{l} \frac{\mathbf{B}(\alpha + l, \alpha(N-1) + n - l)}{\mathbf{B}(\alpha, \alpha(N-1))}, \qquad (3.1)$$

where $\mathbf{B}(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ denotes the beta function, and

$$\mathbf{E}[n_i] = \frac{n}{N}, \qquad \mathbf{Var}[n_i] = \frac{n(N-1)(\alpha N + n)}{N^2(\alpha N + 1)}.$$

3.2 Permuted-Block Design

Under the permuted-block design with random $\{n_i\}$, consider the properties of the imbalance defined in (2.3).

Denote by $R_i = mod(n_i, B)$ the variable which represents the size of the incomplete block in a particular centre i. As $n_i = \lfloor n_i/B \rfloor B + R_i$, then using (2.2), $\Delta_{ij} = \xi_j(R_i) - R_i k_j/B$. Thus, by conditioning on R_i and using (2.1), we obtain $\mathbf{E}[\Delta_{ij}] = 0$ and

$$\mathbf{Var}[\Delta_{ij}] = \frac{k_j(B - k_j)}{B(B-1)}\mathbf{E}[R_i] - \frac{k_j(B - k_j)}{B^2(B-1)}\mathbf{E}[R_i^2], \qquad (3.2)$$

where expressions for $\mathbf{E}[R_i]$ and $\mathbf{E}[R_i^2]$ are given in the appendix.

We now consider the covariance of the imbalances on two different treatments j and m in a trial. Let i and l be two different centres. Then, due to the independence of the numbers of patients on a particular treatment j in different centres, the imbalances on a particular treatment j in different centres are also independent, so that $\mathbf{Cov}[\Delta_{ij}, \Delta_{lm}] = 0$. By again conditioning on R_i, the covariance for two different treatments j and m at the same centre i is

$$\mathbf{Cov}[\Delta_{ij}, \Delta_{im}] = -\frac{k_j k_m}{B(B-1)}\mathbf{E}[R_i] + \frac{k_j k_m}{B^2(B-1)}\mathbf{E}[R_i^2]. \qquad (3.3)$$

For calculating characteristics of R_i, we can use (3.1) and (A.1). However, these relations simplify if we consider the case where $n/N > 2B$, which can be achieved in large clinical trials. In this case, the random variable R_i has an approximate discrete uniform distribution on the integers between 0 and $B-1$ with probabilities $1/B$. Therefore, assuming that $n/N > 2B$ we can use the approximations

$$\mathbf{E}[R_i] = \frac{B-1}{2}, \qquad \mathbf{Var}[R_i] = \frac{B^2 - 1}{12}.$$

After simple algebra, (3.2) and (3.3) become

$$\mathbf{Var}[\Delta_{ij}] \quad = \quad \frac{k_j(B-k_j)(B+1)}{6B^2} \tag{3.4}$$

and

$$\mathbf{Cov}[\Delta_{ij}, \Delta_{im}] = -\frac{k_j k_m (B+1)}{6B^2}. \tag{3.5}$$

The imbalance Δ_j on treatment j for all centres is defined in (2.4) and $\mathbf{E}[\Delta_j] = 0$. Given $\{n_i\}$, the variables $\{\Delta_{ij}\}$ for a particular treatment j in different centres are independent. Thus,

$$\mathbf{Var}[\Delta_j] \quad = \quad \sum_{i=1}^{N} \mathbf{Var}[\Delta_{ij}].$$

Using (3.2), we obtain

$$\mathbf{Var}[\Delta_j] \quad = \quad \sum_{i=1}^{N} \left[\frac{k_j(B-k_j)}{B(B-1)} \mathbf{E}[R_i] - \frac{k_j(B-k_j)}{B^2(B-1)} \mathbf{E}[R_i^2] \right]. \tag{3.6}$$

Similarly, the covariance of Δ_j and Δ_m is

$$\mathbf{Cov}[\Delta_j, \Delta_m] = \sum_{i=1}^{N} \mathbf{Cov}[\Delta_{ij}, \Delta_{im}].$$

Using (3.3), we obtain

$$\mathbf{Cov}[\Delta_j, \Delta_m] = \sum_{i=1}^{N} \left\{ -\frac{k_j k_m}{B(B-1)} \mathbf{E}[R_i] + \frac{k_j k_m}{B^2(B-1)} \mathbf{E}[R_i^2] \right\}. \tag{3.7}$$

For $n/N > 2B$, the random variable R_i has an approximate discrete uniform distribution. Using (3.4), the variance of Δ_j is

$$\mathbf{Var}[\Delta_j] = \frac{N k_j(B-k_j)(B+1)}{6B^2}. \tag{3.8}$$

Finally, using (3.5), the covariance of Δ_j and Δ_m is

$$\mathbf{Cov}[\Delta_j, \Delta_m] = -\frac{N k_j k_m (B+1)}{6B^2}. \tag{3.9}$$

3.3 Complete Randomisation

Under complete randomisation, the imbalance on treatment j in centre i is defined in (2.3).

Denote by $\text{Bin}(n, p)$ a binomial random variable with parameters (n, p). Then we know that $n_{ij} \mid n_i \sim \text{Bin}(n_i, k_j/B)$. Thus, by conditioning on n_i, we obtain $\mathbf{E}[\Delta_{ij}^*] = 0$ and

$$\mathbf{Var}[\Delta_{ij}^*] = \frac{n k_j(B-k_j)}{NB^2}. \tag{3.10}$$

6

Since Δ_{ij}^* and Δ_{lm}^* are independent variables for different centres i and l, their covariance is zero. By again conditioning on n_i, the covariance of Δ_{ij}^* and Δ_{im}^* for a particular centre i is

$$\mathbf{Cov}[\Delta_{ij}^*, \Delta_{im}^*] = -\frac{nk_jk_m}{NB^2}. \tag{3.11}$$

The imbalance over all centres on treatment j defined in (2.4) has expectation $\mathbf{E}[\Delta_j^*] = 0$. Given $\{n_i\}$, the variables $\{\Delta_{ij}^*\}$ for a particular treatment j are independent for different centres, so that

$$\mathbf{Var}[\Delta_j^*] = \sum_{i=1}^N \mathbf{Var}[\Delta_{ij}^*] = \frac{nk_j(B-k_j)}{B^2}. \tag{3.12}$$

Using (3.11), we get that the covariance of Δ_j^* and Δ_m^* is

$$\mathbf{Cov}[\Delta_j^*, \Delta_m^*] = \sum_{i=1}^N \mathbf{Cov}[\Delta_{ij}^*, \Delta_{im}^*] = -\frac{nk_jk_m}{B^2}. \tag{3.13}$$

Remark 1 For equal treatment allocation, we have $k_1 = k_2 = ... = k_K$ and $B = Kk_1$, so that

$$\mathbf{Var}[\Delta_{ij}^*] = \frac{n(K-1)}{NK^2}$$

and

$$\mathbf{Cov}[\Delta_{ij}^*, \Delta_{im}^*] = -\frac{n}{NK^2}.$$

It follows that

$$\mathbf{Var}[\Delta_j^*] = \frac{n(K-1)}{K^2}$$

and

$$\mathbf{Cov}[\Delta_j^*, \Delta_m^*] = -\frac{n}{K^2}.$$

3.4 Overall Imbalance

When the number of centres N involved in a clinical trial is large, the imbalance defined in (2.4) is approximated by a normal distribution with mean zero, and variance (3.6) and (3.12), respectively, for the permuted-block design and complete randomisation. Let the overall imbalances for the permuted-block design and complete randomisation be $\boldsymbol{\Delta}$ and $\boldsymbol{\Delta}^*$, respectively. Then each of these is a vector of imbalances on different treatments for all centres. In other words, $\boldsymbol{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_K)^\intercal$ and $\boldsymbol{\Delta}^* = (\Delta_1^*, \Delta_2^*, \ldots, \Delta_K^*)^\intercal$.

As each of the Δ_j and Δ_j^* are asymptotically normally distributed, both $\boldsymbol{\Delta}$ and $\boldsymbol{\Delta}^*$ are asymptotically multivariate normal. Under the permuted-block design, the asymptotic multivariate normal distribution of $\boldsymbol{\Delta}$ has zero mean vector and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1K} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{K1} & \sigma_{K2} & \ldots & \sigma_K^2 \end{pmatrix}, \tag{3.14}$$

where $\sigma_j^2 = \mathbf{Var}[\Delta_j]$ in (3.8) and $\sigma_{jm} = \mathbf{Cov}[\Delta_j, \Delta_m]$ in (3.9).

Similarly, for complete randomisation, the overall imbalance $\mathbf{\Delta}^*$ has zero mean vector and covariance matrix

$$\mathbf{\Sigma}^* = \begin{pmatrix} \sigma_1^{*2} & \sigma_{12}^* & \cdots & \sigma_{1K}^* \\ \sigma_{21}^* & \sigma_2^{*2} & \cdots & \sigma_{2K}^* \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{K1}^* & \sigma_{K2}^* & \cdots & \sigma_K^{*2} \end{pmatrix},$$

where $\sigma_j^{*2} = \mathbf{Var}[\Delta_j^*]$ in (3.12) and $\sigma_{jm}^* = \mathbf{Cov}[\Delta_j^*, \Delta_m^*]$ in (3.13).

Remark 2 For $K = 2$, the overall imbalance for permuted-block randomisation was defined by Anisimov [10] as the difference between the numbers of patients on treatments 1 and 2. We denote this overall imbalance by $\tilde{\mathbf{\Delta}} = n_{.1} - n_{.2}$, where $n_{.1} = \sum_{i=1}^N n_{i1}$ and $n_{.2} = \sum_{i=1}^N n_{i2}$, which can be written in terms of our notation as

$$\tilde{\mathbf{\Delta}} = \Delta_1 - \Delta_2 + \left(\frac{n}{B}k_1 - \frac{n}{B}k_2 \right).$$

When there is an equal proportion of patients for the two treatments within a complete block, we have $\tilde{\mathbf{\Delta}} = \Delta_1 - \Delta_2$. Now, the overall imbalance in our notation is $\mathbf{\Delta} = (\Delta_1, \Delta_2)^\mathsf{T}$ with expectation $\mathbf{E}[\mathbf{\Delta}] = (0,0)^\mathsf{T}$ and covariance matrix Σ with entries σ_j^2 in (3.8) and σ_{jm} in (3.9).

Let $u = (1, -1)^\mathsf{T}$. Then the expectation of $\tilde{\mathbf{\Delta}} = u^\mathsf{T}\mathbf{\Delta}$ is zero and its variance can be written as

$$\begin{aligned} \mathbf{Var}[\tilde{\mathbf{\Delta}}] &= u^\mathsf{T}\mathbf{\Sigma}u \\ &= (1, -1) \begin{pmatrix} \frac{Nk_1(B-k_1)(B+1)}{6B^2} & -\frac{Nk_1k_2(B+1)}{6B^2} \\ -\frac{Nk_1k_2(B+1)}{6B^2} & \frac{Nk_2(B-k_2)(B+1)}{6B^2} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{2Nk_1k_2(B+1)}{3B^2}. \end{aligned}$$

For equal proportions within a complete block, we have $\mathbf{Var}[\tilde{\mathbf{\Delta}}] = N(B+1)/6$.

For large N, Δ_1 and Δ_2 are approximately normal. Therefore, the imbalance $\tilde{\mathbf{\Delta}}$ as a linear combination of Δ_1 and Δ_2 is also approximately normal. The result calculated here in terms of our notation for the expectation and variance of $\tilde{\mathbf{\Delta}}$ matches with those of the overall imbalance calculated by Anisimov [10, 11].

4 Comparison of randomisation schemes

One of the main aims of a randomisation scheme is to balance the numbers of patients across treatment groups. This ensures that an adequate level of power can be achieved. The higher the power, the more likely a test will detect a genuine treatment difference. The variance of the imbalance is always considered to be a good indicator of which randomisation schemes provide better balance. The greater the variance of the imbalance of a randomisation scheme, the less efficient is the design for balancing the numbers of patients across groups.

By using the calculated variance of the imbalance in (3.2) and (3.10), comparisons can be made between the two randomisation schemes on their effectiveness in balancing the numbers of patients across treatment groups. Assume that $n_i = sB + r$ and

$s \geq 0$. Then the ratio of the variances of the imbalances for the two randomisation schemes is

$$\frac{\mathbf{Var}[\Delta_{ij} \mid n_i]}{\mathbf{Var}[\Delta_{ij}^* \mid n_i]} = \frac{r(B-r)}{(sB+r)(B-1)}.$$

If $s = 0$ and $r = 0$, there will be no patients in the trial and we cannot calculate the ratio. When $s = 0$, there is only one incomplete block, so that the ratio becomes

$$\frac{B-r}{B-1} = \begin{cases} 1, & r = 1, \\ < 1, & r > 1. \end{cases}$$

When $r = 1$, there will be only one patient in centre i and the variance of the imbalance is the same for complete randomisation and the permuted-block design. When $r > 1$, there is only one incomplete block in centre i. The variance of that imbalance under complete randomisation is greater than that of the permuted-block design for a non-empty incomplete block, which means the permuted-block design is more efficient than complete randomisation.

Now consider the case $s \geq 1$. When $r = 0$, centre i contains only complete blocks and the ratio becomes zero for the permuted-block design. Now suppose $r \geq 1$, which implies that $B - r \leq B - 1$, and, for $s \geq 1$ and $B \geq 1$, implies that $sB + r \geq r$. Therefore, $r(B - r) \leq (sB + r)(B - 1)$. If in centre i, there are more than two blocks which include at least one complete block and one incomplete block, the variance of the imbalance for complete randomisation is greater than that of the permuted-block design.

To conclude, the variance of the imbalance in a particular centre i on a particular treatment j is greater under complete randomisation than the permuted-block design in all cases. In other words, the permuted-block design is more efficient than complete randomisation, except when there is only one patient in each centre.

We will now study the overall imbalance on treatments under these two randomisation schemes. We can look at the covariance matrices for the randomisation schemes to see which randomisation scheme performs better.

The imbalances on a particular treatment j are independent for different centres. The variance of the imbalance on a particular treatment j for all centres is just the sum of all the imbalances on treatment j for each centre i for $i = 1, ..., N$. Therefore, the variance of the imbalance on treatment j for all centres under permuted-block randomisation will be less than that under complete randomisation. Similarly, consider the covariances for two different treatments j and m in centre i under these two randomisation schemes. Using (3.3) and (3.11), the ratio of the two covariances is

$$\frac{\mathbf{Cov}[\Delta_{ij}, \Delta_{im} \mid n_i]}{\mathbf{Cov}[\Delta_{ij}^*, \Delta_{im}^* \mid n_i]} = \frac{r(B-r)}{(sB+r)(B-1)},$$

where $n_i = sB + r$.

The covariance ratio of the imbalance is exactly what we obtained for the variance ratio. Therefore, we can draw the same conclusions here as earlier. The covariance of the imbalance in a particular centre i on two different treatments j and m is greater under complete randomisation than under permuted-block randomisation in all cases. Therefore, all the entries in the covariance matrix Σ^* for complete randomisation have values greater than those in the covariance matrix Σ under permuted-block randomisation. For the overall imbalance across treatments, the covariance under permuted-block

randomisation is less than that under complete randomisation, which implies that, under permuted-block randomisation, a more balanced trial can be achieved across treatments for all centres than complete randomisation.

As proved above,

$$\mathbf{Var}[\Delta_{ij} \mid n_i] = \mathbf{Var}[\Delta_{ij}^* \mid n_i]$$

if $n_i = 1$ and

$$\mathbf{Var}[\Delta_{ij} \mid n_i] < \mathbf{Var}[\Delta_{ij}^* \mid n_i]$$

if $n_i > 1$. Now, when n_i is random, if $n > 1$, then $\mathbf{P}[n_i > 1] > 0$. Thus, it can be shown that

$$\mathbf{Var}[\Delta_{ij}] \quad < \quad \mathbf{Var}[\Delta_{ij}^*].$$

Therefore, if we have more than one patient in a trial, $\mathbf{Var}[\Delta_{ij}] < \mathbf{Var}[\Delta_{ij}^*]$ and permuted-block randomisation always has lower variability in the imbalance than complete randomisation. Similarly, $\mathbf{Cov}[\Delta_{ij}, \Delta_{im}] < \mathbf{Cov}[\Delta_{ij}^*, \Delta_{im}^*]$.

For the overall imbalance, we look at the covariance matrix under the two randomisation schemes. The conclusion is the same as when n_i is fixed. Permuted-block randomisation is better than complete randomisation, and thus permuted-block randomisation provides less imbalance in the number of patients on a particular treatment for all centres than complete randomisation in the sense that the entries in the covariance matrix for permuted-block randomisation are all less than those for complete randomisation. Therefore, using permuted-block randomisation should lead to better quality estimators of treatment effect than using complete randomisation.

5 Simulation results

Results of simulation support the theoretical conclusions of Section 3. Consider centre-stratified randomisation and assume that the patient recruitment process is modelled by the Poisson-gamma model (Anisimov and Fedorov [13]). Then, as shown in Section 3.1, under the assumption that all centres are initiated at the same time, the total number of patients recruited in each centre has a beta-binomial distribution. Within each centre, patients are allocated to a treatment according to some randomly permuted blocks. Assume that an equal proportion of patients is to be allocated to each treatment within each complete block. Some randomly permuted blocks of size B are generated in the simulation and the patients are allocated to treatments according to the sequence formed by these blocks. As the imbalance on treatments can be caused only by the incomplete block in each centre, the treatment allocation to patients in these incomplete blocks is simulated from a multivariate hypergeometric distribution. The imbalance on treatments can be calculated by subtracting the simulated number of patients allocated to each of the treatments from the expected number of patients on each treatment. The imbalance on treatments is calculated within each centre and for all centres. Finally, the vector for the overall imbalance on treatments is calculated.

The above procedure is repeated s times. The sample mean and the sample covariance matrix of the overall imbalance are obtained after s runs. For K treatments, let $\widehat{\boldsymbol{\Delta}}_p$ be the vector for the overall imbalance obtained in the pth simulation run. The sample mean vector and covariance matrix are calculated as

$$\bar{\widehat{\boldsymbol{\Delta}}} = \frac{1}{s} \sum_{p=1}^{s} \widehat{\boldsymbol{\Delta}}_p, \qquad \hat{\boldsymbol{\Sigma}} = \frac{1}{s-1} \sum_{p=1}^{s} (\widehat{\boldsymbol{\Delta}}_p - \bar{\widehat{\boldsymbol{\Delta}}})(\widehat{\boldsymbol{\Delta}}_p - \bar{\widehat{\boldsymbol{\Delta}}})^{\mathsf{T}}.$$

These values are compared with the theoretical mean vector and covariance matrix using scenarios based on realistic assumptions. Consider the scenario where $n = 640$, $N = 80$, $K = 4$ and $B = 8$, with $\alpha = 1.2$ and $\beta = 2$ for the patient recruitment process. The theoretical mean vector and covariance matrix of the overall imbalance can be calculated by using results in Section 3. Clearly, the theoretical expectation of the overall imbalance is the vector of zeros. The theoretical covariance matrix has diagonal entries in (3.6) and off-diagonal entries in (3.7). If the size of the incomplete block R_i in centre i has an approximate discrete uniform distribution, the theoretical covariance matrix has entries (3.8) and (3.9).

The number of runs is taken as $s = 100,000$. The sample mean vector of the overall imbalance is

$$\bar{\bar{\boldsymbol{\Delta}}} = \begin{pmatrix} 0.001, & -0.005, & -0.008, & 0.012 \end{pmatrix}^{\mathsf{T}},$$

which is very close to the theoretical mean vector. The theoretical covariance matrix of the overall imbalance is

$$\boldsymbol{\Sigma} = \begin{pmatrix} 21.548 & -7.183 & -7.183 & -7.183 \\ -7.183 & 21.548 & -7.183 & -7.183 \\ -7.183 & -7.183 & 21.548 & -7.183 \\ -7.183 & -7.183 & -7.183 & 21.548 \end{pmatrix}. \tag{5.1}$$

When R_i, the size of the incomplete block in centre i, has an approximate discrete uniform distribution, we have

$$\boldsymbol{\Sigma} = \begin{pmatrix} 22.5 & -7.5 & -7.5 & -7.5 \\ -7.5 & 22.5 & -7.5 & -7.5 \\ -7.5 & -7.5 & 22.5 & -7.5 \\ -7.5 & -7.5 & -7.5 & 22.5 \end{pmatrix}. \tag{5.2}$$

The sample covariance matrix of the overall imbalance is

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 21.458 & -7.149 & -7.188 & -7.121 \\ -7.149 & 21.657 & -7.238 & -7.270 \\ -7.188 & -7.238 & 21.569 & -7.143 \\ -7.121 & -7.270 & -7.143 & 21.533 \end{pmatrix},$$

which has values very close to the theoretical values in (5.1), and also rather close to those in (5.2).

For the above scenario, the results from simulation are consistent with the numerical values for the theoretical mean vector and covariance matrix in (5.1). When the size of the incomplete block R_i in centre i has an approximate discrete uniform distribution, the relative error of approximation of the covariance matrix compared to the theoretical covariance matrix is about 4%. Note that the relative error is larger when we have fewer patients on average per centre. For example, simulations were also carried out for the scenarios where $n = 232$ and $N = 100$, and $n = 496$ and $N = 80$. In the former scenario, the relative error of approximation of the covariance matrix by using a discrete uniform distribution was about 30%, whereas, in the latter, it was about 6%.

The theoretical covariance matrix in (5.1) is calculated under the assumption that the patient recruitment process follows a Poisson-gamma model (Anisimov [12]). On the other hand, (5.2) is calculated under the assumption that R_i has an approximate discrete uniform distribution. This approximation becomes more accurate as the ratio

n/N increases. In practical problems, it is recommended by Anisimov [12] that the uniform distribution for R_i be used when $n/N > 2B$. However, in the above scenario, we have $n/N = 8$ and $2B = 16$, but the discrete uniform approximation is still quite good. This means that it should be sufficiently accurate to apply the approximate expressions using a uniform distribution in large phase III trials, where n/N is typically more than the block size, as these expressions are simple for calculation purposes and do not require any specific information about patient recruitment. The approximation is less reliable in phase II trials or smaller phase III trials for rare diseases, where patient recruitment may be slow. Nevertheless, given information about the parameters of the recruitment process, we can use the expressions derived in Section 3.2 and calculate the theoretical expressions for comparing different scenarios at the design stage of the trial.

The imbalance Δ_j on a particular treatment j defined in (2.4) for $j = 1, ..., K$ for all centres is approximated by a normal distribution with mean zero and variance (3.6), or (3.8) if the size of the incomplete block has an approximate discrete uniform distribution. The results of the simulations support this. Consider the scenario with $K = 4$ treatments and block size $B = 8$. The values of Δ_j for $j = 1, ..., 4$ are calculated for $100,000$ simulations. Figure 1 shows the histogram of the values of Δ_j for each treatment $j = 1, ..., 4$.

Figure 1. Histograms for simulated values of Δ_j for $j = 1, \ldots, 4$.
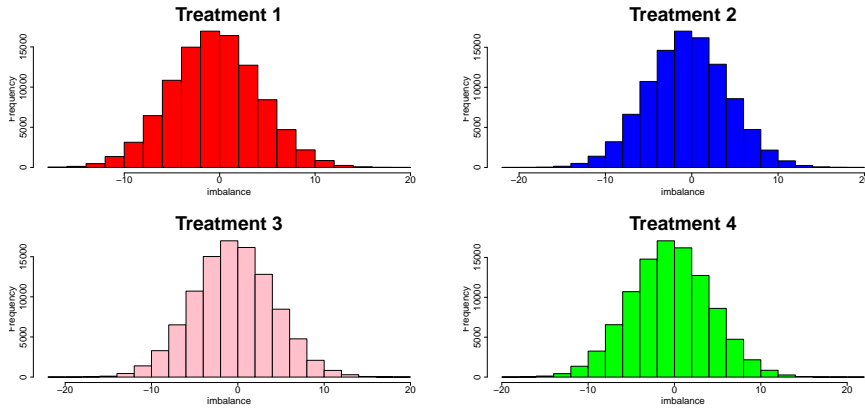
The curves for the empirical density functions of $\Delta_j / \sqrt{\mathbf{Var}[\Delta_j]}$ for $j = 1, ..., 4$ are shown in Figure 2. It is visible that these curves practically coincide with the standard normal density function.

Figure 2. Empirical density functions of $\Delta_j / \sqrt{\mathbf{Var}[\Delta_j]}$ for $j = 1, \dots, 4$.
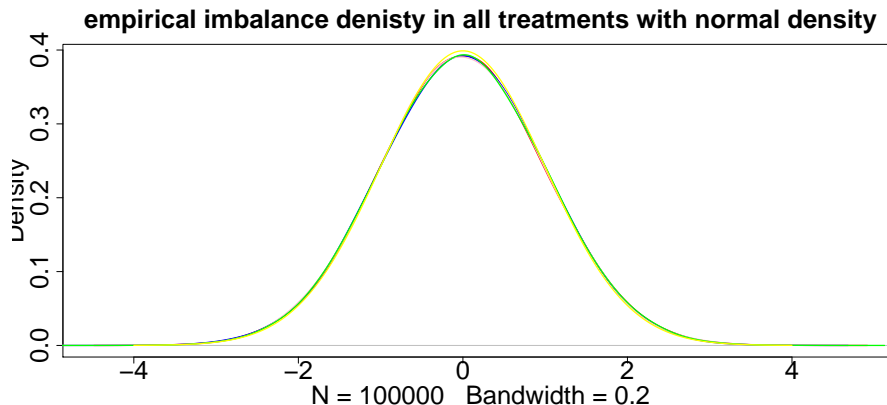
The yellow line represents the standard normal density function. The red, blue, pink and green lines represent the empirical density functions of $\Delta_j / \sqrt{\mathbf{Var}[\Delta_j]}$ for $j = 1, 2, 3, 4$, respectively.

6 Power and sample size

6.1 Description of Test

Here, we study the impact of imbalance across treatments on the power for centre-stratified randomisation. We also consider the imbalanced case and investigate how an increase in the sample size can compensate for the loss in power. This will be shown by numerical results from simulation.

Let n be the total number of patients to be randomised to $K > 2$ treatments at N clinical centres. The centre-stratified randomisation has blocks of size B. Let n_j^* be the number of patients randomised to treatment j for $j = 1, ..., K$ and let $\bar{X}_j$ be the sample mean of the patient responses on treatment j. Assume that the observations are independent and normally distributed random variables with unknown means m_j for $j = 1, ..., K$ and known variance σ^2. Suppose that patients in the first group receive the standard treatment and let this group be the control group. We will be carrying out $K-1$ tests assuming that there is no centre effect in the treatment responses. According to the ICH E9 guideline [14], there is no need to include a centre effect in the analysis model. For example, in large mortality trials with not so many patients per centre, there may be no reason to expect the centres to have any influence on the primary or secondary outcomes.

Consider testing the null hypothesis $H_0 : m_1 = m_2 = ... = m_K = m$ for some non-negative constant m against H_1 : at least one $m_j - m_1 = h_j > 0$ for some $j = 2, ..., K$. Consider the test statistics

$$S_j = \frac{\bar{X}_j - \bar{X}_1}{\sigma \sqrt{1/n_j^* + 1/n_1^*}}$$

for $j = 2, ..., K$. Under H_0, these are dependent standard normal random variables. Given γ as the significance level of one test, denote by Φ the standard normal distribution function and let z_γ satisfy $1 - \Phi(z_\gamma) = \gamma$. We reject H_0 if, for at least one j with

13

$j = 2, ..., K, S_j > z_\gamma$. Denote by γ^* the significance level of the overall test, that is,

$$\mathbf{P}\left\{\cup_{j=2}^{K}(S_j > z_\gamma)|H_0\right\} = \gamma^*. \tag{6.1}$$

Consider now the hypothesis H_1. Denote for $j = 2, ..., K$,

$$\eta_j = \frac{\bar{X}_j - \bar{X}_1 - h_j}{\sigma\sqrt{1/n_j^* + 1/n_1^*}}.$$

Under H_1, the variables $\{\zeta_j = \sqrt{n_j^*}(\bar{X}_j - m_j)/\sigma, j = 1, ..., K\}$ are mutually independent standard normal random variables. So, for any $j = 2, ..., K$, η_j can be represented as a linear combination of two independent normal random variables,

$$\eta_j = \zeta_j c_{j1} - \zeta_1 c_{j2}, \tag{6.2}$$

where

$$c_{j1} = \sqrt{n_1^*}/\sqrt{n_1^* + n_j^*}, \qquad c_{j2} = \sqrt{n_j^*}/\sqrt{n_1^* + n_j^*}. \tag{6.3}$$

Thus, under H_1,

$$S_j = \eta_j + \frac{h_j}{\sigma\sqrt{1/n_j^* + 1/n_1^*}}.$$

Let β^* be the probability of a type II error. For a given level of significance γ, the power can be expressed as

$$\mathbf{P}\left\{\cup_{j=2}^{K}\left(\eta_j + \frac{h_j}{\sigma\sqrt{1/n_j^* + 1/n_1^*}} > z_\gamma\right)\Big| H_1\right\} = 1 - \beta^*, \tag{6.4}$$

where the variables η_j are given in (6.2). Correspondingly, relation (6.1) can be rewritten in the form

$$\mathbf{P}\left\{\cup_{j=2}^{K}(\eta_j > z_{\gamma^*}^*)|H_0\right\} = \gamma^*, \tag{6.5}$$

where, for any particular γ^*, we can find by simulation the value $z_{\gamma^*}^*$ satisfying (6.5) and then compute the power in (6.4) for this $z_{\gamma^*}^*$.

6.2   Impact of Imbalance on Power and Sample Size

For the balanced case, since $n_1^* = n_2^* = ... = n_K^*$, by (6.3), $c_{j1} = c_{j2} = 1/\sqrt{2}$. Thus, relation (6.5) has the form

$$\mathbf{P}\left\{\cup_{j=2}^{K}\left(\zeta_j - \zeta_1 > z_{\gamma^*}^*\sqrt{2}\right)|H_0\right\} = \gamma^*,$$

which can be rewritten as

$$\mathbf{P}\left\{\cap_{j=2}^{K}\left(\zeta_j - \zeta_1 \le z_{\gamma^*}^*\sqrt{2}\right)|H_0\right\} = 1 - \gamma^*. \tag{6.6}$$

Now, since the ζ_j are independent, (6.6) can be expressed as

$$\mathbf{E}[\Phi^{K-1}(z_{\gamma^*}^*\sqrt{2} + \zeta_1)|H_0] = 1 - \gamma^*. \tag{6.7}$$

So it is necessary to find $z_{\gamma^*}^*$ as a solution of (6.7), which can be achieved by simulating only a sequence of values for ζ_1. Using this approach for $\gamma^* = 0.05$ and any K, the values $z_{0.05}^*(K)$ can be calculated. In Table 1, these values are given for $K = 3, ..., 8$ and are based on 100,000 simulations.

Table 1. Critical values when the overall significance level is 5%.

K	3	4	5	6	7	8
$z_{0.05}^*$	1.919	2.063	2.160	2.236	2.294	2.350

Now we study the power of the test using simulation for four different scenarios when $\gamma^* = 0.05$, $\sigma = 1$ and $\beta^* = 0.05$. Some theoretical derivations simplify and accelerate the calculations. Using (6.2) and (6.3), relation (6.4) can be rewritten as

$$\mathbf{P}\left\{\cap_{j=2}^K \left(\zeta_j c_{j1} \leq \zeta_1 c_{j2} - \frac{h_j}{\sqrt{1/n_j^* + 1/n_1^*}} + z_{0.05}^*(K) \right) \Big| H_1 \right\} = \beta^*. \quad (6.8)$$

First consider the balanced case where we have the same number of patients on each of the treatments. Then, since $n_1^* = n_2^* = ... = n_K^* = n/K$, by (6.3), $c_{j1} = c_{j2} = 1/\sqrt{2}$. Thus, (6.8) has the form

$$\mathbf{P}\left\{\cap_{j=2}^K \left(\zeta_j \leq \zeta_1 - h_j\sqrt{\frac{n}{K}} + \sqrt{2}z_{0.05}^*(K) \right) \Big| H_1 \right\} = \beta^*, \quad (6.9)$$

and, since the ζ_j are independent, (6.9) can be expressed as

$$\mathbf{E}\left[\prod_{j=2}^K \Phi\left(\zeta_1 - h_j\sqrt{\frac{n}{K}} + \sqrt{2}z_{0.05}^*(K) \right) \Big| H_1 \right] = \beta^*.$$

So, for any particular values h_j, the sample size n in the balanced case can be evaluated by simulating only the variable ζ_1 and finding the minimum n such that the left-hand side of (6.9) is less than or equal to β^*.

Consider now the imbalanced case. Here, the values n_j^* for $j = 1, ..., K$ are in general unequal random variables defined by the patient recruitment. Therefore, using (6.3), we can write relation (6.8) in the form

$$\mathbf{P}\left\{\cap_{j=2}^K \left(\zeta_j \leq \sqrt{\frac{n_j^*}{n_1^*}}\left(\zeta_1 - h_j\sqrt{n_1^*} + z_{0.05}^*(K)\sqrt{\frac{n_j^* + n_1^*}{n_j^*}} \right) \right) \Big| H_1 \right\} = \beta^*, \quad (6.10)$$

or, since the ζ_j are independent, for any given scenario of recruitment $n_1^*, ..., n_K^*$, (6.10) can be expressed as

$$\mathbf{E}\left[\prod_{j=2}^K \Phi\left(\sqrt{\frac{n_j^*}{n_1^*}}\left(\zeta_1 - h_j\sqrt{n_1^*} + z_{0.05}^*(K)\sqrt{\frac{n_j^* + n_1^*}{n_j^*}} \right) \right) \Big| H_1 \right] = \beta^*,$$

where expectation is with respect to ζ_1 and different values of $n_1^*, ..., n_K^*$.

Thus, to simulate the power and to calculate the sample size n, for a given n, we need to simulate in each run the value of a normal random variable ζ_1 and one recruitment scenario to obtain the values $n_1^*, ..., n_K^*$. After running simulations, we need to find the sample mean of the values

$$\prod_{j=2}^K \Phi\left(\sqrt{\frac{n_j^*}{n_1^*}}\left(\zeta_1 - h_j\sqrt{n_1^*} + z_{0.05}^*(K)\sqrt{\frac{n_j^* + n_1^*}{n_j^*}} \right) \right) \quad (6.11)$$

15

and check if this is less than or equal to β^*. By repeating the simulations for different values of n, we can find the minimum sample size for a given β^* when the sample mean of the values in (6.11) becomes less than or equal to β^* and then compare it with the corresponding value for the balanced case. Table 2 shows the power obtained by the above approach for the balanced and imbalanced cases for 100,000 simulations with $N = 50$.

Table 2. Simulated powers for four scenarios using the direct approach.

Number of treatments	$K = 4$	$K = 5$	$K = 6$	$K = 8$
Size of block	$B = 8$	$B = 10$	$B = 12$	$B = 8$
Balanced case				
n	220	240	252	264
Power	0.951	0.953	0.953	0.953
Imbalanced case				
n	220	240	252	264
Power	0.950	0.951	0.951	0.952
n	221	241	253	265
Power	0.951	0.952	0.952	0.952
n	222	242	254	266
Power	0.952	0.952	0.952	0.953
n	223	243	255	267
Power	0.952	0.953	0.952	0.954
n	224	244	256	268
Power	0.953	0.954	0.954	0.955

For each scenario, the power obtained in the imbalanced case is less than that in the balanced case for the same value of n. As we can see from Table 2, we need at most four more patients in the imbalanced case to obtain the same level of power as in the balanced one. These results are in agreement with those of Anisimov [11, 12], where the case of two treatments was considered.

Since simulating a large number of different recruitment scenarios and calculating the values $n_1^*, ..., n_K^*$ may take a reasonable amount of time in practice, we propose a combined analytical/approximate approach using the asymptotic approximations derived in Section 3. This approach is computationally about twice as fast as the direct one.

Using relations (2.3) and (2.4), we obtain that, in the case of equal treatment allocations, $n_j^* \approx \Delta_j + n/K$ and the vector $\boldsymbol{\Delta} = (\Delta_1, \Delta_2, \dots, \Delta_K)^\mathsf{T}$ is well approximated by a multivariate normal vector which can be written in the form

$$\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\xi},$$

where the matrix $\boldsymbol{\Sigma}$ is defined in (3.14) by relations (3.8) and (3.9), and $\boldsymbol{\xi}$ is a multivariate normal vector with independent components such that

$$\mathbf{E}[\boldsymbol{\xi}] = \mathbf{0}, \qquad \mathbf{E}[\boldsymbol{\xi}\boldsymbol{\xi}^\mathsf{T}] = \mathbf{I}.$$

This means that we can approximate the vector $\mathbf{n}^* = (n_1^*, ..., n_K^*)^\mathsf{T}$ as

$$\mathbf{n}^* \approx \frac{n}{K}\mathbf{1} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{\xi}. \tag{6.12}$$

Using this approach, instead of simulating the patient recruitment and calculating the values $\{n_j^*\}$, we can simulate in each run the independent normal random variables $(\xi_1, ..., \xi_K)$ and one normal variable ζ_1, and then calculate the vector $(n_1^*, ..., n_K^*)$ using (6.12) and hence the values (6.11).

Table 3 shows the results of the power calculations obtained by using the above approximation. As we can see, for $K = 4, 5, 6$, the results of direct simulation of the recruitment as it was carried out in the calculations in Table 2 are practically the same, up to three decimal places and simulation error. However, for $K = 8$, the results differ slightly in the second decimal place. This is because the normal approximation uses a covariance matrix of the form (3.14), where relations (3.8) and (3.9) are based on a discrete uniform approximation for the size of the incomplete block. Recall that, from Section 5, this approximation works well if $n/N > B$. However, for $B = 8$, the ratio $n/N \approx 5$, which is less than 8. Therefore, we can conclude that, for $n/N > B$, we can substantially simplify and accelerate the computations by using the normal approximation for the vector $(n_1^*, ..., n_K^*)$ given by (6.12).

Table 3. Simulated powers for four scenarios using a normal approximation.

| Number of treatments | $K = 4$ | $K = 5$ | $K = 6$ | $K = 8$ |
Size of block	$B = 8$	$B = 10$	$B = 12$	$B = 8$
n	220	240	252	264
Power	0.950	0.953	0.952	0.963
n	221	241	253	265
Power	0.951	0.953	0.952	0.964
n	222	242	254	266
Power	0.952	0.955	0.953	0.965
n	223	243	255	267
Power	0.953	0.954	0.954	0.965
n	224	244	256	268
Power	0.953	0.955	0.954	0.966

7 Discussion

7.1 Conclusions

In this paper, imbalances are defined for complete randomisation and the permuted-block design for clinical trials with more than two treatments, and their properties are investigated. Most of the previous papers deal with imbalance properties of different randomisation schemes for two treatment groups only. The imbalance for more than two treatments is no longer the difference in the numbers of patients on two treatments, but is defined as a vector whose components are the differences between the number of patients in each treatment group and the expected number of patients in this group.

The imbalance is defined within centres and globally. The calculations of the expectations, variances and covariances of the imbalances in a centre or for all centres are provided for the two randomisation schemes. It is assumed that the patient recruitment process follows the Poisson-gamma model (Anisimov and Fedorov [13]). The variances for the two randomisation schemes are compared. For trials with several

treatments, in general, centre-stratified permuted-block randomisation performs better than complete randomisation in balancing the numbers of patients across treatment groups. In other words, complete randomisation provides more uncertainty in the numbers of patients in different treatment groups. A test is also considered for comparing each treatment to a control group. The numerical values for the power of the test in the balanced and imbalanced cases indicate that any imbalance has little effect on the power.

We have concentrated on the imbalance in the form considered here, since it is more reflected in the existing literature. However, the methods developed in the paper allow us to investigate the distribution of the numbers of patients on different treatment arms, and thus to potentially analyse other measures of imbalance. This would help statisticians to decide whether the probability of undesirable imbalance is notable in practice. Note that, in Anisimov [12], the impact of random patient dropout was also investigated on the sample size for both randomisation schemes for $K = 2$ when using the test that compares means. Potentially, the techniques can be used to analyse the impact of dropout for several treatment arms as well. Such extensions will be addressed in a separate paper.

7.2 Implications of Results

With the increasing pace of drug development, it is not unusual for several treatments to be ready simultaneously for testing in a randomised phase III setting; see, for example, Royston *et al.* [15] and Wason and Jaki [16]. This reduces costs and time to recruit patients in a trial, and expedites earlier availability of an effective treatment to the general public. The imbalance properties for the several treatments case are of particular importance as they affect the power of the test for treatment differences. The larger the imbalance is in a trial, the less power it may have to detect a genuine treatment difference. In this paper, the focus is on rather large trials where the randomisation may play a significant role.

The results provide the analytical methodology for investigating the properties of imbalance in the case of multiple treatments and for comparing various scenarios. Since asymptotically the overall imbalance for permuted-block randomisation depends only on the number of strata and the number of treatments, and not the sample size, the imbalance should not have a significant impact on the power and sample size of the study. Several examples support this conclusion and show that the loss in power due to the imbalance can be compensated for by adding only a few patients.

From another standpoint, using stratified randomisation in general may improve different operational characteristics. In particular, it may substantially reduce drug supply overages. Therefore, in the cases when the choice of randomisation is not dictated by the type of data, these results support wider use of stratified randomisation in clinical trials.

Appendix: Derivations of Properties of Imbalance

First consider the permuted-block design. Then

$$\mathbf{E}[\Delta_{ij}] = \mathbf{E}[\mathbf{E}[\Delta_{ij} \mid R_i]] = 0,$$

and using (2.1) we obtain

$$
\begin{aligned}
\mathbf{Var}[\Delta_{ij}] &= \mathbf{E}[\mathbf{E}[\Delta_{ij}^2 \mid R_i]] \\
&= \mathbf{E}\left[\mathbf{E}[\xi_j^2(R_i) \mid R_i] - R_i^2 \left(\frac{k_j}{B}\right)^2 \right] \\
&= \mathbf{E}\left[\mathbf{Var}[\xi_j(R_i) \mid R_i] \right] = \mathbf{E}\left[\frac{k_j R_i (B - k_j)(B - R_i)}{B^2(B-1)} \right],
\end{aligned}
$$

which yields (3.2).

As for all i, j, $\mathbf{E}[\Delta_{ij}] = 0$, then the covariance for two different treatments j and m at the same centre i is calculated as

$$
\mathbf{Cov}[\Delta_{ij}, \Delta_{im}] = \mathbf{E}[\Delta_{ij}\Delta_{im}].
$$

Now, we have

$$
\begin{aligned}
\mathbf{E}[\Delta_{ij}\Delta_{im}] &= \mathbf{E}\left[\mathbf{E}\left[\left(\xi_j(R_i) - R_i\frac{k_j}{B}\right)\left(\xi_m(R_i) - R_i\frac{k_m}{B}\right) \mid R_i \right] \right] \\
&= \mathbf{E}[\mathbf{Cov}[\xi_j(R_i), \xi_m(R_i) \mid R_i]] = \mathbf{E}\left[-\frac{k_j k_m R_i(B - R_i)}{B^2(B-1)} \right].
\end{aligned}
$$

Thus, we obtain (3.3).

In Anisimov [10, 12], the probability $\mathbf{P}[mod(n_i, B) = r] = \mathbf{P}[R_i = r]$ is calculated as

$$
q_r(n, N, \alpha, B) = \sum_{s=0}^{n/B-1} \mathbf{P}(n, N, \alpha, r + sB) \tag{A.1}
$$

for $r = 0, 1, ..., B - 1$, where $\mathbf{P}(n, N, \alpha, r + sB)$ is defined in (3.1). So $\mathbf{E}[R_i]$ can be written as

$$
\begin{aligned}
\mathbf{E}[R_i] &= \sum_{r=0}^{B-1} r\mathbf{P}[R_i = r] \\
&= \sum_{r=0}^{B-1} \sum_{s=0}^{n/B-1} r \binom{n}{r + sB} \frac{\mathbf{B}(\alpha + r + sB, \alpha(N-1) + n - r - sB)}{\mathbf{B}(\alpha, \alpha(N-1))}.
\end{aligned}
$$

Similarly,

$$
\mathbf{E}[R_i^2] = \sum_{r=0}^{B-1} \sum_{s=0}^{n/B-1} r^2 \binom{n}{r + sB} \frac{\mathbf{B}(\alpha + r + sB, \alpha(N-1) + n - r - sB)}{\mathbf{B}(\alpha, \alpha(N-1))}.
$$

Now consider complete randomisation. Then, since $n_{ij} \mid n_i \sim \text{Bin}(n_i, k_j/B)$, we have $\mathbf{E}[n_{ij} \mid n_i] = n_i k_j / B$ and $\mathbf{E}[\Delta_{ij}^*] = 0$. Further,

$$
\mathbf{Var}[\Delta_{ij}^* \mid n_i] = \mathbf{Var}[n_{ij} \mid n_i] = n_i \frac{k_j}{B}\left(1 - \frac{k_j}{B}\right).
$$

As $\mathbf{E}[n_i] = n/N$, we obtain (3.10).

The covariance of Δ_{ij}^* and Δ_{im}^* for a particular centre i is

$$
\begin{aligned}
\mathbf{Cov}[\Delta_{ij}^*, \Delta_{im}^*] &= \mathbf{E}[\mathbf{E}[\Delta_{ij}^* \Delta_{im}^* \mid n_i]] \\
&= \mathbf{E}\left[\mathbf{E}\left[\left(n_{ij} - \frac{n_i}{B} k_j\right)\left(n_{im} - \frac{n_i}{B} k_m\right) \mid n_i\right]\right] \\
&= \mathbf{E}\left[\mathbf{Cov}[n_{ij}, n_{im} \mid n_i]\right].
\end{aligned}
$$

The conditional covariance of n_{ij} and n_{im} can be obtained from the multinomial distribution and is calculated as $-n_i k_j k_m / B^2$. So the covariance of Δ_{ij}^* and Δ_{im}^* is given by (3.11).

Acknowledgements

References

[1] Rosenberger WF, Lachin JM. *Randomization in Clinical Trials: Theory and Practice.* Wiley: New York, 2002.

[2] Senn S. *Statistical Issues in Drug Development.* Wiley: Chichester, 1997.

[3] Brittain E, Schlesselman JJ. Optimal allocation for the comparison of proportions. *Biometrics* 1982; **38**: 1003-1009.

[4] McPherson GC, Campbell MK, Elbourne DR. Use of randomisation in clinical trials: a survey of UK practice. *Trials* 2012; **13**: 198.

[5] Kundt G, Glass A. Evaluation of imbalance in stratified blocked randomization: some remarks on the range of validity of the model by Hallstrom and Davis. *Methods of Information in Medicine* 2012; **51**: 55-62.

[6] Lachin J. Properties of simple randomization in clinical trials. *Controlled Clinical Trials* 1988; **9**: 312-326.

[7] Morrissey M, McEntegart D, Lang M. Randomisation in double-blind multicentre trials with many treatments. *Contemporary Clinical Trials* 2010; **31**: 381-391.

[8] Hallstrom A, Davis K. Imbalance in treatment assignments in stratified blocked randomization. *Controlled Clinical Trials* 1988; **9**: 375-382.

[9] Lachin J. Statistical properties of randomization in clinical trials. *Controlled Clinical Trials* 1988; **9**: 289-311.

[10] Anisimov V. Effect of imbalance in using stratified block randomization in clinical trials. *Proceedings of the 56th Session of the International Statistical Institute* 2007; 1-4.

[11] Anisimov V. Impact of stratified randomization in clinical trials. In *mODa 9 - Advances in Model-Oriented Design and Analysis*, Giovagnoli A, Atkinson AC, Torsney B, May C (eds.). Physica-Verlag HD: Berlin, 2010; 1-8.

[12] Anisimov V. Effects of unstratified and centre-stratified randomization in multi-centre clinical trials. *Pharmaceutical Statistics* 2011; **10**: 50-59.

[13] Anisimov V, Fedorov V. Modelling, prediction and adaptive adjustment of recruitment in multicentre trials. *Statistics in Medicine* 2007; **26**: 4958-4975.

[14] ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Statistics in Medicine* 1999; **18**: 1905-1942.

[15] Royston P, Parmar MKB, Qian W. Novel designs for multi-arm clinical trials with survival outcomes with an application in ovarian cancer. *Statistics in Medicine* 2003; **22**: 2239-2256.

[16] Wason JMS, Jaki T. Optimal design of multi-arm multi-stage trials. *Statistics in Medicine* 2012; **31**: 4269-4279.