Statistica Sinica **18**(2008), 235-253

UNIFORMLY BALANCED REPEATED MEASUREMENTS DESIGNS IN THE PRESENCE OF SUBJECT DROPOUT

Dibyen Majumdar, Angela M. Dean and Susan M. Lewis

University of Illinois at Chicago, The Ohio State University and University of Southampton

Abstract: Low, Lewis and Prescott (1999) showed that a crossover design based on a Williams Latin square of order 4 can suffer substantial loss of efficiency if some observations in the final period are unavailable. Indeed, if all observations are missing, the design becomes disconnected. We derive the information matrix for the direct effects of a Uniformly Balanced Repeated Measurements Design (UBRMD) in t periods when subjects may drop out before the end of the study, and examine the maximum loss of information. The special case of loss of observations in the final period only is examined in detail. In particular we show that a UBRMD in $t \geq 5$ periods remains connected when some or all observations in the final period are unavailable.

Key words and phrases: Crossover designs, efficiency, missing observations, Williams Latin Squares.

1. Introduction

Cross-over experiments are widely used for comparing the responses to various different stimuli or treatments in areas ranging from psychology and human factor engineering to medical and agricultural applications; see, for example, the books by Jones and Kenward (2003), Ratkowsky, Evans and Alldredge (1992) and Senn (2002). Such experiments extend over a sequence of time periods. Each subject receives one treatment per period and an observation is made at the end of the period. The influence of a treatment on the subject's response may extend (or carry over) into the period following that in which it is administered. This is known as a *first-order carry-over effect* or *first-order residual effect.* In a simple statistical model for crossover studies, the response for a given subject in a given period is regarded as a sum of the effects of the subject, the period, the treatment given in this period (the *direct effect* of the treatment), the carry-over effect from the treatment given in the preceding period, and a random error.

There is an extensive literature that assures us that a carefully designed crossover study can produce a wealth of information and that the parameters of interest can be estimated with high precision; see, for instance, Stufken (1996). This is based on the implicit, but critical, assumption that the experiment yields all planned observations. Yet, in many studies such as clinical trials, there is a substantial probability that some subjects will drop out of the study prior to completion of their treatment sequence. Low, Lewis and Prescott (1999) observed that a dropout rate of between 5% and 10% is not uncommon and, in some areas, can be as high as 25%. They give an example of a design in four periods, based on a Williams Latin square (Williams (1949)), where there is substantial loss of information if some observations are unavailable in period 4. Indeed, if all observations in the final period are not available, the design becomes disconnected, i.e., elementary contrasts are no longer all estimable.

It is important to note that a similar situation may arise in an interim analysis. When interim results on a cross-over experiment are analyzed, the interim design may consist of the planned design without the final several periods.

The loss of connectedness is the most severe consequence of the unavailability of observations. A general study of the loss of connectedness that results from unavailability of observations was done by Ghosh (1979, 1982). For crossover designs, Low, Lewis and Prescott (1999) formulated requirements for a planned design to be robust to dropouts in terms of the properties of the implemented designs that might result under a "completely-at-random" dropout mechanism (Diggle and Kenward (1994)). Godolphin (2004) also studied the problem of loss of connectedness of various designs, including crossover designs.

An experimenter generally starts with a design, the *planned design,* that possesses desirable properties, including high efficiency or optimality. If no subject drops out, the study yields the entire information that was envisioned at the planning stage. Dropouts, however, lead to loss of information. The *implemented design* is the design that corresponds to all *available* observations, and this design can be identified only at the conclusion of the experiment.

For the case of a Williams Latin Square of order 4 the *expected* information loss for various probability distributions of dropouts was studied by Low, Lewis and Prescott (1999). In this article, we focus on the *maximum information loss* that may be anticipated. For instance, in a study with four periods where subjects are expected to remain at least through the first three periods, dropouts, if any, would occur in the final period only. In this case minimal information is attained when *all* subjects drop out in the final period, which gives the *minimal design.*

In this paper we assume that the *planned design* belongs to the class of *Uniform Balanced Repeated Measurement Designs* (UBRMDs). This is an important class of designs that have been studied extensively in the literature and are a popular choice in practice. UBRMDs have elegant *combinatorial balance* and, under the simple model with additive i.i.d. errors with constant variance, possess various optimality properties; see, for example, Hedayat and Afsarinejad (1978), Cheng and Wu (1980), Kunert (1984), Hedayat and Yang (2003), and Hedayat and Yang (2004). (Refer to Stufken (1996) and Hedayat and Yang (2003) for additional references). A design is called *uniform* if (a) for each subject, each treatment is allocated to the same number of periods, and (b) for each period, each treatment is allocated to the same number of subjects. Furthermore, a design is called *balanced for carryover effects* (*balanced,* in short) if, in the order of application, each treatment is preceded by every other treatment the same number of times and is not preceded by itself.

The goal of this research is to study the maximum loss of information and the resulting loss of precision of the estimators that result from subject dropout when the planned design is a UBRMD. Since the maximum loss is attained by the *minimal design*, we study properties of this design, including its information matrix and efficiency. If the maximum loss of information is not deemed to be large, then the experimenter may conclude that no modification of the plan for the experiment is needed. On the other hand, if the loss is large, the experimenter should consider alternative strategies.

We work in the same setup as Low, Lewis and Prescott (1999); in particular we assume a completely-at-random dropout mechanism. Also, we assume throughout that a subject who leaves the study does not re-enter. In Section 2 we derive general formulae for, and study the properties of, the information matrix of the direct effects of the minimal design when the planned design is a UBRMD, with subject dropouts occurring in the final m periods only. We examine the connectedness of the minimal design and, in particular, show that a UBRMD based on $t \geq 5$ treatments remains connected even when all observations in the final period are unavailable. We also develop measures for the maximum loss of precision due to subject dropout and the efficiency of the minimal design. In Section 3 we study the case of one-period dropout in more detail, including the special case when the planned design is based on Williams Latin squares. Also in this section, we identify members of the class of UBRMDs for which the loss of information is small.

The focus of this paper is to study certain properties of UBRMDs in the presence of subject dropout. For the broader problem of designing a crossover experiment in the presence of subject dropout one has to choose a planned design from a class (not necessarily the class of UBRMDs) of highly efficient designs for which the loss of information is small, and the corresponding minimal design is highly efficient.

2. Setup and General Results

Consider a planned design with p periods, s subjects and t treatments. The simple model for the vector of response variables obtained from the implemented design can be written as

$$Y = X_S\beta + X_P\alpha + X_D\tau + X_C\rho + \epsilon, \tag{1}$$

where ϵ is the vector of random error variables, β is a vector of s subject effects, α is a vector of p time period effects, τ is a vector of t direct treatment effects, ρ is a vector of t carry-over effects, and the X matrices are the corresponding design matrices. The treatments are labelled $0, 1, \ldots, t-1$. For the purposes of designing efficient experiments, all effects in the model are assumed to be fixed effects.

We define the following incidence and replication matrices: $N_{SD} = X_S'X_D$, $N_{SC} = X_S'X_C$, $N_{PD} = X_P'X_D$, $N_{PC} = X_P'X_C$, $N_{DC} = X_D'X_C$, $r_D = N_{DS}\mathbf{1}_s = N_{DP}\mathbf{1}_p$, $r_C = N_{CS}\mathbf{1}_s = N_{CP}\mathbf{1}_p$, where the "prime" denotes transpose and $\mathbf{1}_a$ is a vector of a unit elements. Also we define $J_{a\times b} = \mathbf{1}_a\mathbf{1}_b'$, $J_a = J_{a\times a}$, $N_{ji} = N_{ij}'$ (for $i, j = S, P, D, C$), $r_D^\delta = diag(r_D)$, and $r_C^\delta = diag(r_C)$. Moreover, I_a denotes an $a \times a$ identity matrix. We order the responses period by period for each subject in turn, so that, $X_P = \mathbf{1}_s \otimes I_p$ and $X_S = I_s \otimes \mathbf{1}_p$.

The joint information matrix for estimating the direct and carry-over (residual) treatment effects is given by

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \tag{2}$$

$$\text{where} \quad C_{11} = r_D^\delta + \frac{1}{ps}r_Dr_D' - \frac{1}{p}N_{DS}N_{SD} - \frac{1}{s}N_{DP}N_{PD}$$

$$C_{22} = r_C^\delta + \frac{1}{ps}r_Cr_C' - \frac{1}{p}N_{CS}N_{SC} - \frac{1}{s}N_{CP}N_{PC}$$

$$C_{12} = N_{DC} + \frac{1}{ps}r_Dr_C' - \frac{1}{p}N_{DS}N_{SC} - \frac{1}{s}N_{DP}N_{PC}.$$

The information matrices for the direct effects and the carry-over effects, respectively, are

$$C_D = C_{11} - C_{12}C_{22}^{-}C_{21}\,, \tag{3}$$

$$C_R = C_{22} - C_{21}C_{11}^{-}C_{12}\,. \tag{4}$$

In this article we focus primarily on C_D.

Throughout, we assume that the *planned design* d_{plan} is a UBRMD with $p = t$ time periods, $s = gt$ subjects, based on t treatments and, in the *implemented design* d_{imp}, all subjects complete their allocated treatment sequence in the first $t - m$ periods $(1 \le m < t - 1)$. After the first $t - m$ periods, subjects may start dropping out of the study completely at random. Since we assume that, once a subject drops out of the study, the subject will not return, the *worst case scenario* occurs when all subjects drop out at period $t - m$. The design $d_{\min}$, composed of the first $t - m$ periods of d_{plan}, is called the *minimal design*.

For even t, a Williams Latin Square gives a UBRMD, as does any sequentially counterbalanced Latin Square (see Isaac, Dean and Ostrom (2001), for a survey). For t odd, a UBRMD cannot be constructed from one Williams Latin square, but such a design with $2t$ subjects can be constructed from a pair of squares. In addition, when t is a composite number, Higham (1998) has shown that there exists a UBRMD in t subjects, t periods and t treatments. The union of UBRMDs (identical or distinct) with the same value of t is a UBRMD. Here are some examples.

Example 1. Three UBRMDs are shown below, where the columns show the treatment sequences and the rows correspond to the time periods. The design d_{2plan} is a Williams Latin Square of order 4, while designs d_{1plan} and d_{3plan} consist of a pair of Williams Latin Squares for $t = 3$ and $t = 5$ treatments, respectively.

d_{1plan}

1	2	0	2	0	1
0	1	2	0	1	2
2	0	1	1	2	0

d_{2plan}

0	1	2	3
1	2	3	0
3	0	1	2
2	3	0	1

d_{3plan}

1	2	3	4	0	3	4	0	1	2
0	1	2	3	4	4	0	1	2	3
2	3	4	0	1	2	3	4	0	1
4	0	1	2	3	0	1	2	3	4
3	4	0	1	2	1	2	3	4	0

If $m = 1$, i.e. subjects may drop out in the final period only, then each array with the last row deleted gives the corresponding minimal design $d_{\min}$. It can be verified that the information matrices of $d_{\min}$ have rank 2, 1 and 4, respectively. So, as noted by Low, Lewis and Prescott (1999), the minimal design corresponding to d_{2plan} is disconnected; indeed the only estimable direct treatment contrast in $d_{2\min}$ is $\tau_0 - \tau_1 + \tau_2 - \tau_3$. On the other hand the minimal designs corresponding to d_{1plan} and d_{3plan} are connected, the former has a nonzero eigenvalue 0.125 with multiplicity 2; the nonzero eigenvalues of the latter are 2.61 and 3.73, each with multiplicity 2.

Since UBRMDs with three periods have been studied in Jones and Kenward (2003) and Low (1995), henceforth we consider $t \geq 4$. The following lemma shows that $d_{\min}$ corresponds to the maximal loss of information. For nonnegative definite matrices A and B we use $A \succeq B$ to denote the fact that $A - B$ is a nonnegative definite matrix, the Löwner order.

Lemma 2. $C_D^{d_{plan}} \succeq C_D^{d_{imp}} \succeq C_D^{d_{\min}}$.

This follows from known general results for linear models; for instance, it is a consequence of Theorem 2.1 of Hedayat and Majumdar (1985). Lemma 2 says that $d_{\min}$ has the "smallest information matrix" among all possibilities for d_{imp}. The matrices, C_{11}, C_{22} and C_{12} for $d_{\min}$ are as given in Theorem 4. First, we need some notation.

The P and U matrices. For $j = 1, \ldots, t$, P_j denotes a $t \times s$ matrix with (h, i) entry 1 if subject i receives treatment h in period j of d_{plan}; it is 0 otherwise. For $j = 0, 1, \ldots, t-1$; $k = 0, 1, \ldots, t-1$, let, $U_{jk} = P_{t-j}P'_{t-k}$. Note that, since d_{plan} is a UBRMD,

$$P_j\mathbf{1}_s = g\mathbf{1}_t, P'_j\mathbf{1}_t = \mathbf{1}_s. \tag{5}$$

Also, the entries of U_{jk} are nonnegative with row and column sums equal to g. Hence $\frac{1}{g}U_{jk}$ is a doubly stochastic matrix; in particular $U_{jj} = gI_t$. The following lemma gives the properties of these matrices that we need.

Lemma 3. *If* $U_1, \ldots, U_M$ (*not necessarily distinct*) *are* $t \times t$ *matrices such that* U_i/g *is doubly stochastic for each* $i = 1, \ldots, M$, *and* $a_1, \ldots, a_M$ *are nonnegative real numbers, then for* $x \in R^t$ *with* $x'x = 1$, $x'U_ix \leq g$ *for* $i = 1, \ldots, M$, *and* $x'(\sum a_iU_i)'(\sum a_iU_i)x \leq g^2(\sum a_i)^2$.

Proof. If we write $W_i = U_i/g$, then it follows from the properties of doubly stochastic matrices (see, for example, Bapat and Raghavan (1997, Chap. 2)) that $x'W_ix \leq 1$. Also, $x'(\sum a_iW_i)'(\sum a_iW_i)x = x'(\sum\sum a_ia_jW'_iW_j)x \leq \sum\sum a_ia_j \sqrt{x'W'_iW_ixx'W'_jW_jx} \leq \sum\sum a_ia_j$. The lemma follows.

Theorem 4. *Let* d_{plan} *be a UBRMD with* t *treatments,* t *time periods,* $s = gt$ *subjects, and let* $d_{\min}$ *consist of the first* $t - m$ *periods of* d_{plan}. *Then for* $d_{\min}$,

(i) *the information matrix for estimating direct and carry-over treatment effects is given by* (3) *and* (4) *with*

$$C_{11} = \frac{g[(t-m)^2 - m]}{t-m}I_t - \frac{g(t-2m)}{t-m}J_t - \frac{1}{t-m}\sum\sum_{j \neq k = 0, \ldots, m-1} U_{jk}, \tag{6}$$

$$C_{22} = \frac{g}{t-m}[((t-m)^2 - (t+1))I_t - t^{-1}((t-m)^2 - (t+1) - m(m+1))]J_t$$

$$-\frac{1}{t-m}\sum_{j\neq k=0,\dots,m}\sum U_{jk}, \tag{7}$$

$$C_{12}=\frac{g}{t-m}[(m+1)J_t-tI_t]-\sum_{j=0}^{m-1}U_{j(j+1)}-\frac{1}{t-m}\sum_{\substack{j=0\\ j\neq k}}^{m-1}\sum_{k=0}^{m}U_{jk} \tag{8}$$

and (ii) *if*

$$t\geq 2m+2 \tag{9}$$

then a g-inverse of C_{22} *is* $C_{22}^{-}=A^{-1}$, *where*

$$A=\frac{g}{t-m}[(t-m)^2-(t+1)]I_t-\frac{1}{t-m}\sum_{j\neq k=0,\dots,m}\sum U_{jk}. \tag{10}$$

Proof. (i) Since d_{plan} is a *UBRMD*, every treatment appears $s/t=g$ times in every period and s times in total. Also, in the order of application, each treatment is preceded by every other treatment the same number of times and is not preceded by itself. It follows that for the design $d_{\min}$,

$$N_{DS}=J_{t\times s}-\sum_{j=0}^{m-1}P_{t-j},N_{CS}=J_{t\times s}-\sum_{j=0}^{m}P_{t-j},$$
$$N_{DP}=gJ_{t\times(t-m)},\ N_{CP}\ =\ \left[\mathbf{0}_t\ gJ_{t\times(t-m-1)}\right],$$
$$r_D=g(t-m)\mathbf{1}_t,\ r_C=g(t-m-1)\mathbf{1}_t,$$

where $\mathbf{0}_t$ is a vector with t zero elements. Inserting the above formulae into (2) and using (5) yields expressions C_{11}, C_{22} and C_{12} as in the statement of the theorem, after some algebra.

(ii) Using the relation $C_{22}\mathbf{1}_t=\mathbf{0}_t$, it can be verified that $C_{22}A^{-1}C_{22}=C_{22}$, as long as A^{-1} exists; hence A^{-1} is a g-inverse of C_{22}. We now show that (9) guarantees the nonsingularity of A. It follows from the fact that the row sums of $\sum_{j\neq k=0,\dots,m}\sum U_{jk}$ are $gm(m+1)$, and Lemma 3, that the minimum eigenvalue of A is

$$\lambda_{\min}(A)=\frac{g}{t-m}[(t-m)^2-(t+1)-m(m+1)]. \tag{11}$$

This is positive if $(t-m)^2-(t+1)-m(m+1)>0$, which is equivalent to $t\geq 2m+2$.

From the proof of Theorem 4, (9) guarantees that the A in (10) is positive definite, and hence it is sufficient for the nonsingularity of A. This condition plays a critical role in the derivation of bounds for the eigenvalues of $C_D^{d_{\min}}$. Also, it can be shown that (9) is a necessary and sufficient condition for $rank(C_{22}) = t-1$. The next result gives a bound on the eigenvalues of $C_D^{d_{\min}}$ which is used to study the loss of precision for the estimators of the treatment contrasts and the efficiency of $d_{\min}$.

Theorem 5. *Suppose $m \geq 1$ and $t \geq 2m+2$ and d_{plan} is a UBRMD with t treatments, t time periods, and $s = gt$ subjects. Denote the eigenvalues of $C_D^{d_{\min}}$ by $g\theta_0 = 0, g\theta_1, \ldots, g\theta_{t-1}$. For $r = 1, \ldots, t-1$, $\theta_r \geq \theta_L(t,m)$, where*

$$\theta_L(t,m) = \frac{t}{t-m}\Big[(t-2m) - \frac{t(m+1)^2}{(t-m)^2 - (t+1) - m(m+1)}\Big]. \qquad (12)$$

Proof. Suppose $x \in R^t$, with $x'x = 1$, and $x'\mathbf{1}_t = 0$, such that $C_D^{d_{\min}} x = g\theta_r x$, $r = 1, \ldots, t-1$. Then $g\theta_r = x'C_{11}x - x'C_{12}A^{-1}C_{21}x$. The maximum eigenvalue of A^{-1} is $1/(\lambda_{\min}(A))$ where $\lambda_{\min}(A)$ is given by (11). Hence, $g\theta_r \geq x'C_{11}x - (1/\lambda_{\min}(A))x'C_{12}C_{21}x$. If we write,

$$V = (t-m+1)\sum_{j=0}^{m-1} U_{j(j+1)} + \sum_{\substack{j=0 \\ k\neq j,j+1}}^{m-1}\sum_{k=0}^{m} U_{jk},$$

then $-C_{21} = (t-m)^{-1}[gtI_t + V' - g(m+1)J_t]$. So,

$$g\theta_r \geq x'C_{11}x - \frac{1}{\lambda_{\min}(A)[(t-m)^2]}[(gt)^2 + gtx'(V+V')x + x'VV'x]. \qquad (13)$$

The following inequalities can be derived by applying Lemma 3:

$$\begin{aligned} x'C_{11}x &\geq \frac{gt}{t-m}(t-2m) \\ x'(V+V')x &\leq 2g[(t-m+1)m + m(m-1)] = 2gmt \\ x'VV'x &\leq (gmt)^2. \end{aligned} \qquad (14)$$

Inserting them into (13) and using the fact $\lambda_{\min}(A) > 0$, which follows from the condition $t \geq 2m+2$, we get a lower bound to θ_r which, upon simplification, reduces to (12).

The results of Theorem 5 are useful in studying properties of $d_{\min}$. Connectedness is a basic property of a design. A sufficient condition for $d_{\min}$ to be connected for direct treatment effects is $\theta_L(t,m) > 0$. It follows from Lemma 2 that d_{imp} is connected whenever $d_{\min}$ is connected. Corollary 6 follows from (12).

Corollary 6. *Suppose d_{plan} is a UBRMD with t treatments, t time periods, and $s = gt$ subjects. A sufficient condition for the minimal design $d_{\min}$ to be connected is that*

$$(t-2m)[(t-m)^2 - (t+1) - m(m+1)] - t(m+1)^2 > 0. \qquad (15)$$

For a given m, it follows from (15) that $d_{\min}$ is connected if a polynomial in t of degree 3, which has leading coefficient 1, is positive. For $m = 1$, (15) reduces to $t^3 - 5t^2 + 4 > 0$, i.e., $t \geq 5$, and for $m = 2$ it reduces to $t^3 - 9t^2 + 8t + 12 > 0$, i.e., $t \geq 8$. These observations lead to the following result.

Corollary 7. *Suppose d_{plan} is a UBRMD with t treatments, t time periods, and $s = gt$ subjects. For each $m \geq 1$, there is a positive integer $t^*(m)$ such that the design $d_{\min}$ is connected if $t \geq t^*(m)$. In particular, for $m = 1$, $d_{\min}$ is connected whenever $t \geq 5$; for $m = 2$, $d_{\min}$ is connected whenever $t \geq 8$.*

One way to measure the *goodness* of a connected design d is by the harmonic mean of the eigenvalues of the information matrix C_D^d, $H_d = (t-1)/trace(C_D^d)^+$, where C^+ denotes the Moore-Penrose inverse of C. This is the value of the *A-criterion*; hence H_d is a measure of the precision of estimators of the direct treatment contrasts for the design d. It follows from Lemma 2 that $H_{d_{\min}} \leq H_{d_{\text{imp}}} \leq H_{d_{\text{plan}}}$. Since d_{plan} is the design that was chosen at the start of the experiment on the basis of its desirable properties, especially efficiency, it is of interest to examine the *loss of precision* in the implemented design d_{imp} with respect to d_{plan} due to subject dropout. This loss may be measured by

$$L_{d_{\text{imp}}:d_{\text{plan}}} = \frac{H_{d_{\text{plan}}} - H_{d_{\text{imp}}}}{H_{d_{\text{plan}}}} = 1 - \frac{trace(C_D^{d_{\text{plan}}})^+}{trace(C_D^{d_{\text{imp}}})^+}.$$

Clearly, the *maximum loss of precision due to subject dropout* for d_{plan} is given by

$$ML_{d_{\text{plan}}} = L_{d_{\min}:d_{\text{plan}}} = \frac{H_{d_{\text{plan}}} - H_{d_{\min}}}{H_{d_{\text{plan}}}} = 1 - \frac{trace(C_D^{d_{\text{plan}}})^+}{trace(C_D^{d_{\min}})^+},$$

i.e., $ML_{d_{\text{plan}}} \geq L_{d_{\text{imp}}:d_{\text{plan}}}$.

When d_{plan} is a UBRMD, we get, using Theorem 5,

$$trace(C_D^{d_{\min}})^+ = \sum_{r=1}^{t-1} \frac{1}{g\theta_r} \leq \frac{t-1}{g\theta_L(t,m)}.$$

From Hedayat and Afsarinejad (1978), we obtain for the UBRMD d_{plan},

$$C_D^{d_{\text{plan}}} = \frac{gt(t-2)(t+1)}{t^2-t-1}[I_t - \frac{1}{t}J_{t,t}], \quad trace(C_D^{d_{\text{plan}}})^+ = \frac{(t-1)(t^2-t-1)}{gt(t-2)(t+1)}.$$

Therefore we obtain the following result.

Corollary 8. *Suppose d_{plan} is a UBRMD with t treatments, t time periods, and $s = gt$ subjects. An upper bound to the maximum loss of precision due to subject* *dropout is given by* $ML_{d_{\text{plan}}} \leq UML(t,m)$ *where,*

$$UML(t,m) = 1 - \frac{(t^2-t-1)\theta_L(t,m)}{t(t-2)(t+1)}, \tag{16}$$

with $\theta_L(t,m)$ *given by* (12).

For $m = 1$ and $t \geq 5$, the values of $UML(t,1)$ for selected values of t are given in Table 1, where the planned design d_{plan} is a UBRMD. Similarly for $m = 2$ and $t \geq 8$, the values of $UML(t,2)$ for selected values of t are given in Table 2. As one would expect, for fixed m, the bounds decrease with t and become reasonably small when t is considerably larger than $t^*(m)$. In general, (16) is conservative. Hence, the prospect of subject dropout may not be a big concern when t is much larger than $t^*(m)$.

If the UBRMD d_{plan} is chosen to have certain combinatorial structures, the bound $UML(t,m)$ can be improved. One such structure is considered next.

Type $\mathcal{W}_m$ UBRMD. Suppose the subjects of the UBRMD d_{plan} can be partitioned into g sets of t subjects each such that, within each group, every treatment appears once in each of the periods $t-m, t-m+1, \ldots, t$ for fixed $m \geq 1$. Then for $j,k = 0,\ldots,m$, $j \neq k$,

$$U_{jk} = P_{t-j}P'_{t-k} = \sum_{l=1}^{g}\Pi_{jkl},$$

where each Π_{jkl} is a permutation matrix of order t and P_i is defined in Section 2. If for each $j,k = 0,\ldots,m$, $j \neq k$ and $l = 1,\ldots,g$, the eigenvalue 1 of Π_{jkl} has multiplicity one, then we say the UBRMD d_{plan} is of *type* $\mathcal{W}_m$.

If $m \geq 2$ an UBRMD of type $\mathcal{W}_m$ is also of type $\mathcal{W}_{m-1}$. Examples of UBRMDs of *type* $\mathcal{W}_{t-1}$ are UBRMD's that are cyclically generated, for instance the Williams Latin Squares and pairs of Williams Latin Squares given in Families 1 and 3 of Hedayat and Afsarinejad (1978), and the class of sequentially counterbalanced squares described by Isaac, Dean and Ostrom (2001).

It is known that the eigenvalues of a permutation matrix Π of order t are the roots of unity, $e^{i(2\pi r/t)} = \cos(2\pi r/t) + i\text{Sin}(2\pi\text{r}/\text{t})$, $r = 0,1,\ldots,t-1$, *unless* the permutation can be factored into the *product of two or more disjoint cycles*, in which case the multiplicity of 1 as an eigenvalue of Π is larger than one (see, for example, Davis (1979)). If we set $\psi_r = \cos(2\pi r/t)$ then, for a UBRMD of *type* $\mathcal{W}_m$, the eigenvalues of $\Pi_{jkl} + \Pi'_{jkl}$ are $2\psi_r$, $r = 0,1,\ldots,t-1$. Since $U_{jk}+U_{kj} = \sum_{l=1}^{g}(\Pi_{jkl}+\Pi'_{jkl})$, $\mathbf{1}'_t x = 0$ and $x'x = 1$ implies $x'(U_{jk}+U_{kj})x \leq 2g\psi_1$.

Hence the inequalities (14) may be replaced by

$$x'C_{11}x \geq \frac{gt}{t-m}\Big[(t-2m)+\frac{m(m-1)}{t}(1-\psi_1)\Big]$$
$$x'(V+V')x \leq 2gmt\psi_1$$
$$x'VV'x \leq (gmt)^2\,.$$

If we insert these inequalities into (13) we obtain the following result.

Theorem 9. *Suppose* d_{plan} *is a UBRMD of type* W_m *for fixed* $m \geq 1$, $t \geq 2m+2$. *Denote the eigenvalues of* C_D^d *by* $g\theta_0 = 0, g\theta_1, \ldots, g\theta_{t-1}$. *For* $r = 1, \ldots, t-1$, $\theta_r \geq \theta_L^*(t,m)$, *where*

$$\theta_L^*(t,m) = \frac{t}{t-m}\Big[(t-2m)+\frac{m(m-1)}{t}(1-\psi_1)-\frac{t(1+2\psi_1 m+m^2)}{(t-m)^2-(t+1)-m(m+1)}\Big]$$

with $\psi_1 = \cos 2\pi/t$.

Since $\psi_1 < 1$, $\theta_L^*(t,m) > \theta_L(t,m)$. Hence replacing $\theta_L(t,m)$ by $\theta_L^*(t,m)$ in (16) gives a sharper upper bound to the *maximum loss of precision due to subject dropout* when d_{plan} is a UBRMD of *type* $\mathcal{W}_m$, i.e., $ML_{d_{\text{plan}}} \leq UML^*(t,m) < UML(t,m)$, where

$$UML^*(t,m) = 1 - \frac{(t^2-t-1)\theta_L^*(t,m)}{t(t-2)(t+1)}.$$

For $m = 1$ and $t \geq 5$, the values of the upper bound $UML^*(t,1)$ to the maximum loss of precision due to subject dropout when d_{plan} is a UBRMD of *type* $\mathcal{W}_m$ for selected values of t are given in Table 1. For $m = 2$ and $t \geq 8$ the values of $UML^*(t,2)$ for selected values of t are given in Table 2.

Table 1. Upper bounds to the maximum loss of precision due to subject dropout and lower bound to the efficiency of the minimal design when $m = 1$.

t	5	6	7	8	9	10
$UML(t,1)$	0.87	0.48	0.33	0.25	0.20	0.17
$UML^*(t,1)$	0.64	0.40	0.30	0.23	0.19	0.16
$EL(t,1)$	0.18	0.66	0.81	0.88	0.92	0.93
$EL^*(t,1)$	0.49	0.76	0.85	0.90	0.93	0.94

Table 2. Upper bounds to the maximum loss of precision due to subject dropout and lower bound to the efficiency of the minimal design when $m = 2$.

t	8	9	10	11	12	16
$UML(t,2)$	0.90	0.63	0.48	0.39	0.33	0.20
$UML^*(t,2)$	0.81	0.59	0.46	0.38	0.32	0.21
$EL(t,2)$	0.15	0.50	0.67	0.77	0.82	0.93
$EL^*(t,2)$	0.27	0.55	0.69	0.78	0.83	0.94

We now consider the efficiency of $d_{\min}$, the design that corresponds to the *worst case scenario*. Let $D(t, gt, t-m)$ denote the class of all connected crossover designs (not necessarily uniform or balanced) in $t-m$ periods and gt subjects, based on t treatments. $d_{\min}$ belongs to this class. Since for an arbitrary design $d \in D(t, gt, t-m)$, $trace(C_D^d)^+ \geq (t-1)^2/traceC_D^d$, a lower bound for $trace(C_D^d)^+$ may be obtained from an upper bound of $traceC_D^d$. The latter bound can be obtained from Theorem 3 of Hedayat and Yang (2004) (which generalized a bound of Stufken (1991)) as follows,

$$\underset{d\in D(t,gt,t-m)}{Max} traceC_D^d = gt(t-m-1) - \frac{2(gt-\delta^*)}{t-m} - \frac{(t-m-1)\delta^{*2}}{g(t-m)(t(t-m)-t-1)},$$

where δ^* is the nearest integer to $[g(t(t-m)-t-1)]/(t-m-1)$. Since the choice $\delta^* = [g(t(t-m)-t-1)]/(t-m-1)$ gives an upper bound to the maximum, it can be shown that $traceC_D^d \leq gMTr(t, m)$, where

$$MTr(t, m) = t(t-m-1) - \frac{t(t-m-1)+1}{(t-m)(t-m-1)}.$$

A measure of the efficiency of $d_{\min}$ in $D(t, gt, t-m)$ is

$$EFF^{d_{\min}}_{D(t,gt,t-m)} = \frac{\underset{d\in D(t,gt,t-m)}{Min} trace(C_D^d)^+}{trace(C_D^{d_{\min}})^+} \geq \frac{(t-1)^2}{gMTr(t,m)(trace(C_D^{d_{\min}})^+)}.$$

It follows from Theorems 5 and 9 that, if we define

$$EL(t, m) = \frac{(t-1)\theta_L(t, m)}{MTr(t, m)} \quad \text{and} \quad EL^*(t, m) = \frac{(t-1)\theta_L^*(t, m)}{MTr(t, m)},$$

then the inequalities $EFF^{d_{\min}}_{D(t,gt,t-m)} > EL(t, m)$ and $EFF^{d_{\min}}_{D(t,gt,t-m)} > EL^*(t, m)$ give *lower bounds to the efficiency* of $d_{\min}$ when d_{plan} is a general UBRMD and a UBRMD of *type* $\mathcal{W}_m$, respectively. Note that both $EL(t, m)$ and $EL^*(t, m)$ take values in $(0, 1)$.

For $m = 1$ and $t \geq 5$, the values of the lower bounds $EL(t, 1)$ and $EL^*(t, 1)$ to the efficiency of $d_{\min}$ in $D(t, gt, t-1)$ for selected values of t are given in Table 1. For $m = 2$ and $t \geq 8$, the values of $EL(t, 2)$ and $EL^*(t, 2)$ for selected values of t are given in Table 2. Since $EL(t, m)$ (or $EL^*(t, m)$) measures the efficiency of $d_{\min}$ over *all designs* in $D(t, gt, t-m)$, not just those that are derived from UBRMDs, high values of this efficiency bound suggest that a different starting design, instead of the UBRMD d_{plan}, would not have resulted in a substantially better $d_{\min}$. An UBRMD d_{plan} that has a small value of $UML(t, m)$ (or $UML^*(t, m)$) and a large value of $EL(t, m)$ (or $EL^*(t, m)$) for $d_{\min}$ clearly is a good design for use when there is a possibility of subject dropout.

3. Further Results for the Case $m = 1$

In the previous section we derived upper bounds $UML(t,m)$ and $UML^*(t, m)$ to the maximum loss of precision due to subject dropout $ML_{d_{\text{plan}}}$. In this section, we first establish formulae for $ML_{d_{\text{plan}}}$for two special families of UBRMDs. Then we indicate how to select a starting design UBRMD d_{plan} for which $ML_{d_{\text{plan}}}$ is small. For simplicity, we focus on the case $m = 1$, i.e., subjects remain in the study at least through period $t-1$. We start with definitions of the families of UBRMDs that we will study.

Class A type $\mathcal{W}_1$ UBRMD. Start with a $t\times t$ square W that is a UBRMD with columns denoting treatment sequences and rows denoting periods. Let $P_t(W)$ and $P_{t-1}(W)$ be the $t \times t$ matrices defined in Section 2 corresponding to periods t and $t-1$, respectively, for square W, i.e., the (h,i) entry of $P_j(W)$ is 1 if the (j,i) entry of W is h; it is 0 otherwise. Let $\Pi = P_t(W)P_{t-1}(W)'$. If 1 is an eigenvalue of Π of multiplicity one, then the design d_{plan} that assigns g subjects to each sequence (column) of W is called a Class A type $\mathcal{W}_1$ UBRMD.

Class B type $\mathcal{W}_1$ UBRMD. We start with two $t \times t$ squares W_1 and W_2 such that the $t \times 2t$ design $(W_1\ W_2)$ is a UBRMD with columns denoting treatment sequences and rows denoting periods. For $\delta = 1, 2$, let $P_t(W_\delta)$ and $P_{t-1}(W_\delta)$ be the $t \times t$ matrices defined as in the previous paragraph, and take $\Pi_\delta = P_t(W_\delta)P_{t-1}(W_\delta)'$. Suppose 1 is an eigenvalue of Π_δ of multiplicity one for each $\delta = 1, 2$. Suppose also that W_1 and W_2 are *complementary* in the sense $\Pi_2 = \Pi_1'$. Then the design d_{plan} that assigns $g/2$ subjects to each sequence (column) of W is called a Class B type $\mathcal{W}_1$ UBRMD.

Note that, for ease of implementation of the study, the experimenter will generally assign several subjects to each of a small number of treatment sequences (see Jones and Kenward (2003, p.159)). Examples of Class A type $\mathcal{W}_1$ UBRMD when t is even are the Williams Latin squares given in Family 1 of Hedayat and Afsarinejad (1978) with g subjects assigned to each sequence, and examples of Class B type $\mathcal{W}_1$ UBRMD when t is odd are the pair of William squares given in Family 3 of Hedayat and Afsarinejad (1978) with $g/2$ subjects assigned to each sequence.

Theorem 10 *Suppose $t \geq 4$.* (i) *If d_{plan} is a Class A type W_1 UBRMD then, for the minimal design $d_{\min}$ that consists of the first $t-1$ periods of d_{plan}, the eigenvalues of $C_D^{d_{\min}}$ are $g\theta_0 = 0$ and $g\theta_r$, where*

$$\theta_r = \frac{t}{t-1}\Big[t - 2 - \frac{2t(1+\cos(\frac{2\pi r}{t}))}{t(t-3) - 2\cos(\frac{2\pi r}{t})}\Big], \quad r = 1, \ldots, t-1. \qquad (17)$$

(ii) *If d_{plan} is a Class B type W_1 UBRMD then, for the design $d_{\min}$ that consists of the first $t-1$ periods of d_{plan}, the eigenvalues of $C_D^{d_{\min}}$ are $g\theta_0 = 0$*

and $g\theta_r$*, where*

$$\theta_r = \frac{t}{t-1}\Big[t-2-\frac{t(1+\cos(\frac{2\pi r}{t}))^2}{t(t-3)-2\cos(\frac{2\pi r}{t})}\Big], \quad r=1,\ldots,t-1. \tag{18}$$

Proof. Write $U = U_{01} = P_tP'_{t-1}$. Then $C_{11} = (gt(t-2)/(t-1))(I_t - J_t/t)$, $C_{12} = -(1/(t-1))(gtI_t - 2gJ_t + tU)$, and $C_{22} = A - (g(t^2-3t-2)/(t(t-1)))J_t$, where $A = ((gt(t-3)/(t-1))I_t - (1/t-1)(U+U')$. Consider the spectral decomposition, $U+U' = \sum_{r=0}^{t-1}\alpha_r h_r h'_r$, with $h'_r h_r = 1$, $r = 0,1,\ldots,t-1$, $h'_r h_q = 0$, for $r \neq q$; $\alpha_0 = 2g$, $h_0 = \mathbf{1}_t/\sqrt{t}$. For $r = 1,\ldots,t-1$, $h'_r\mathbf{1}_t = 0$. Let $\gamma_r = (I_t + U/g)h_r$, $r = 0,1,\ldots,t-1$. It can be shown that

$$C_D^{d_{\min}} = \frac{gt}{t-1}\Big[(t-2)\Big(I_t - \frac{1}{t}J_{t,t}\Big) - gt\sum_{r=1}^{t-1}(gt(t-3)-\alpha_r)^{-1}\gamma_r\gamma'_r\Big].$$

Note that, by Lemma 3, $\alpha_r \leq 2g$. Hence for $t \geq 4$, $gt(t-3)-\alpha_r \geq gt(t-3)-2g > 0$.

(i) In this case, $U = g\Pi$. For $r = 1,\ldots,t-1$, $(U+U')h_r = \alpha_r h_r$ implies $\alpha_r = 2g\psi_r$, where $\psi_r = \cos(2\pi r/t)$. Since $\gamma_r = (I_t + \Pi)h_r$, $\gamma'_l\gamma_r = h'_l(2I_t + \Pi + \Pi')h_r = (2+2\psi_r)h'_l h_r$. Therefore, if we write $\gamma_0^* = 1/\sqrt{t}\mathbf{1}_t$, $\gamma_r^* = (2+2\psi_r)^{-1/2}\gamma_r = (2+2\psi_r)^{-1/2}(I_t+\Pi)h_r$, $r = 1,\ldots,t-1$, then $\{\gamma_0^*, \gamma_1^*, \ldots, \gamma_{t-1}^*\}$ is an orthogonal and normalized basis of $\mathcal{R}^t$, and for $r = 1,\ldots,t-1$,

$$C_D^{d_{\min}}\gamma_r^* = \frac{gt}{t-1}\Big[t-2-\frac{t(2+2\psi_r)}{t(t-3)-2\psi_r}\Big]\gamma_r^* = \frac{gt}{t-1}\Big[t-2-\frac{2t(1+\cos(\frac{2\pi r}{t}))}{t(t-3)-2\cos(\frac{2\pi r}{t})}\Big]\gamma_r^*,$$

which establishes (i).

(ii) Here $U = g(\Pi_1 + \Pi'_1)/2 = U'$. For $r = 1,\ldots,t-1$, $(U+U')h_r = \alpha_r h_r$ implies $\alpha_r = 2g\psi_r$; $\gamma_r = (I_t + U/g)h_r = (1+\psi_r)h_r$. It follows that, $h_0, h_1, \ldots, h_r$ are orthogonal and normalized eigenvectors of $C_D^{d_{\min}}$, and

$$C_D^{d_{\min}}h_r = \frac{gt}{t-1}\Big[t-2-\frac{t(1+\cos(\frac{2\pi r}{t}))^2}{t(t-3)-2\cos(\frac{2\pi r}{t})}\Big]h_r,$$

for $r = 1,\ldots,t-1$. This establishes (ii).

For $t = 4$ it can be shown that $\theta_1 = 0$, $\theta_2 = 2.67$, and $\theta_3 = 0$. Hence $d_{\min}$ is not connected. For $t \geq 5$ however, it follows from Corollary 7 that $d_{\min}$ is connected. The following Corollary is immediate.

Corollary 11 *Suppose* $t \geq 5$*. If* d_{plan} *is a Class A or Class B type* $\mathcal{W}_1$ *UBRMD then the loss of precision due to subject dropout is*

$$L_{d_{\text{imp}}:d_{\text{plan}}} \leq L_{d_{\min}:d_{\text{plan}}} = ML_{d_{\text{plan}}} = 1 - \frac{(t-1)(t^2-t-1)}{t(t-2)(t+1)}\Big(\sum_{r=1}^{t-1}\frac{1}{\theta_r}\Big)^{-1},$$

where the θ_r*'s are given by* (17) *and* (18) *for Class A and B, respectively.*

Theorem 10 also gives a lower bound to the efficiency of $d_{\min}$ in $D(t, gt, t-1)$, $EFF^{d_{\min}}_{D(t,gt,t-1)} > EL^{AB}(t)$, where

$$EL^{AB}(t) = \frac{(t-1)^2}{MTr(t,1)} \left(\sum_{r=1}^{t-1} \frac{1}{\theta_r} \right)^{-1}$$

and where $MTr(t,1) = t(t-2) - (t-1)/(t-2)$. The difference between the bounds $EL^{AB}(t)$ and $EL^*(t,1)$ defined in Section 2 is that in the former (which applies to Class A or Class B type $\mathcal{W}_1$ UBRMD) exact values of the eigenvalues θ_r are used while in the latter (which applies to any type $\mathcal{W}_1$ UBRMD) these are replaced by the lower bound $\theta^*_L(t,1)$.

Table 3 gives $ML_{d_{\text{plan}}}$ and $EL^{AB}(t)$ for Class A and Class B type $\mathcal{W}_1$ UBRMD for selected values of t. Note that the values of the maximum loss due to subject dropout $ML_{d_{\text{plan}}}$ are substantially lower than the upper bounds given in Table 1, while the values of the efficiency bound $EL^{AB}(t)$ are higher.

Table 3. Maximum loss of precision due to subject dropout and lower bound to the efficiency of the minimal design for Class A and Class B type $\mathcal{W}_1$ UBRMD d_{plan} when $m = 1$.

t	5	6	7	8	9	10
Class	B	A	B	A	B	A
$ML_{d_{\text{plan}}}$	0.35	0.30	0.20	0.18	0.14	0.13
$EL^{AB}(t)$	0.90	0.89	0.97	0.97	0.98	0.98

For $g > 1$, a Class A type $\mathcal{W}_1$ UBRMD consists of g replications of a square while a Class B type $\mathcal{W}_1$ UBRMD consists of $g/2$ replications of a pair of squares. The next result implies that the loss due to subject dropout may be smaller if distinct squares (or pair of squares) are used instead of replications.

Corollary 12. *Suppose* $t \geq 5$. *If* d_{plan} *is the union of* g $t \times t$ *Class A type* $\mathcal{W}_1$ *UBRMDs or* $g/2$ $t \times 2t$ *Class B type* W_1 *UBRMDs that are not necessarily identical, then*

$$L_{d_{\text{imp}}:d_{\text{plan}}} \leq L_{d_{\min}:d_{\text{plan}}} = ML_{d_{\text{plan}}} \leq 1 - \frac{(t-1)(t^2-t-1)}{t(t-2)(t+1)} \left(\sum_{r=1}^{t-1} \frac{1}{\theta_r} \right)^{-1},$$

where the θ_r*'s are given by* (17) *and* (18) *for Class A and B, respectively.*

Proof. We sketch a proof for Class A type $\mathcal{W}_1$ UBRMDs; the proof for Class B type $\mathcal{W}_1$ UBRMDs is identical. Suppose $d_{\min} = \bigcup_{i=1}^{g} d_i$, where d_i is the minimal design for a Class A type $\mathcal{W}_1$ UBRMD design based on t subjects for

$i = 1, \ldots, g$. It follows from Theorem 2.1 of Hedayat and Majumdar (1985) that, $C_D^{d_{\min}} \succeq \sum_{i=1}^{g} C_D^{d_i}$. Take $B = C_D^{d_{\min}} + J_t/t$, and $B_i = C_D^{d_i} + J_t/(gt)$, for $i = 1, \ldots, g$. Since each d_i is connected, the matrices $B_1, \ldots, B_g$ and B are positive definite. Clearly, $B \succeq \sum_{i=1}^{g} B_i$. It follows that

$$B^{-1} \preceq \Big(\sum_{i=1}^{g} B_i\Big)^{-1} \preceq \frac{1}{g^2}\Big(\sum_{i=1}^{g} B_i^{-1}\Big),$$

where the first inequality is well known in matrix theory and the second is given in Bapat and Raghavan (1997, Thm. 3.11.1). It can be shown that, $B^{-1} = (C_D^{d_{\min}})^{+} + J_t/t$, and $B_i^{-1} = (C_D^{d_i})^{+} + (g/t)J_t$, for $i = 1, \ldots, g$. Hence, $(C_D^{d_{\min}})^{+} \preceq g^{-2}\sum_{i=1}^{g}(C_D^{d_i})^{+}$. This implies,

$$trace(C_D^{d_{\min}})^{+} \leq \frac{1}{g^2}\sum_{i=1}^{g} trace(C_D^{d_i})^{+} = \frac{1}{g^2}\sum_{i=1}^{g}\sum_{r=1}^{t-1}\frac{1}{\theta_r} = \frac{1}{g}\sum_{r=1}^{t-1}\frac{1}{\theta_r}.$$

The result follows.

For the setup of Corollary 12, it is clear that a lower bound to the efficiency of $d_{\min}$ in $D(t, gt, t-1)$ is

$$EFF_{D(t,gt,t-1)}^{d_{\min}} > \frac{(t-1)^2}{MTr(t,1)\Big(\sum_{r=1}^{t-1}\frac{1}{\theta_r}\Big)}.$$

Corollary 12 indicates that the use of distinct Class A or Class B type $\mathcal{W}_1$ UBRMDs instead of replications of the same design will not increase the maximum loss of precision due to subject dropout $ML_{d_{\rm plan}}$. There are examples where $ML_{d_{\rm plan}}$ actually decreases. In their Example 2, Low, Lewis and Prescott (1999) studied the case $t = 4$, $s = 24$, $m = 1$ and showed that the use of distinct William Squares instead of replications of the same square reduces $ML_{d_{\rm plan}}$. An example for $t = 6$ is given below. The implication is that a UBRMD with more distinct sequences is likely to perform better under subject dropout.

Consider the "extreme" design $d_{\rm plan}^e$ that consists of one subject assigned to each of the $t!$ possible sequences ($s = t!$). For the case $m = 1$, it can be shown that $U_{01} = ((t-2)!)[J_t - I_t]$ and the information matrix of the minimal design $d_{\min}^e$ is

$$C_D^{d^e} = \frac{at(t-2)[(t-2)!]}{t-1}\Big(I_t - \frac{1}{t}J_t\Big), \quad \text{where } a = \frac{t^4 - 5t^3 + 6t^2 + t - 2}{t^3 - 4t^2 + 3t + 2}.$$

For $d_{\rm plan}^e$ the maximum loss of precision due to subject dropout is,

$$ML_{d_{\rm plan}^e} = 1 - \frac{a(t^2 - t - 1)}{(t-1)^2(t+1)}.$$

Numerical studies indicate that this is the smallest value of $ML_{d_{\rm plan}}$ among all UBRMDs with $t!$ or fewer subjects. We are currently investigating the nature of planned designs that attain the minimum and maximum values of $ML_{d_{\rm plan}}$, as well as designs that fall between in these extremes. However, as mentioned earlier in this section, a small number of treatment sequences is generally preferred, so it is doubtful that crossover designs with a large number of distinct sequences will be used widely in practice.

Example 13. Let $t = 6$ and $s = 12g_0$. The array below consists of two Williams Latin squares.

Square 1						Square 2					
1	2	3	4	5	0	2	5	1	3	0	4
0	1	2	3	4	5	4	2	5	1	3	0
2	3	4	5	0	1	5	1	3	0	4	2
5	0	1	2	3	4	0	4	2	5	1	3
3	4	5	0	1	2	1	3	0	4	2	5
4	5	0	1	2	3	3	0	4	2	5	1

Suppose $d^1_{\rm plan}$ is a design that assigns $2g_0$ subjects to the first six columns of the array and $d^2_{\rm plan}$ a design that assigns g_0 subjects to each of the twelve columns. If $m = 1$, then the maximum loss of precision due to subject dropout are $ML_{d^1_{\rm plan}} = 0.30$ and $ML_{d^2_{\rm plan}} = 0.24$. The design $d^e_{\rm plan}$ can be constructed when $g_0 = 60$. For this design, $ML_{d^e_{\rm plan}} = 0.20$.

Since estimation of the residual effects of the treatments is sometimes at least a secondary focus of experiments, we conclude this section with a brief consideration of the information matrix for the residual treatment effects of the minimal design $d_{\rm min}$ obtained from a UBRMD $d_{\rm plan}$ when $m = 1$. Note that, $C_R^{d_{imp}} \succeq C_R^{d_{\rm min}}$. Also, $C_R^{d_{\rm plan}} = C_{22} - C_{21}C_{11}^{-}C_{12}$. Suppose $t \geq 3$. Then, it can be shown that $((t-1)/(gt(t-2)))I_t$ is a generalized inverse of C_{11}. From (2) and (4) we obtain,

$$C_R^{d_{\rm min}} = \Big(\frac{gt(t^2-5t+5)}{(t-1)(t-2)}\Big)\Big[I_t - \frac{1}{t}J_{t,t}\Big] - \Big(\frac{2}{t-2}\Big)[U+U'] \\ -\Big(\frac{t}{g(t-1)(t-2)}\Big)U'U + \Big(\frac{g(5t-4)}{t(t-1)(t-2)}\Big)J_{t,t}.$$

Using this, it can be shown that if the design $d_{\rm plan}$ is a Class A or Class B type $\mathcal{W}_1$ UBRMD then $d_{\rm min}$ is connected for residual treatment effects whenever $t = 3$, or $t \geq 5$. For $t = 4$, Low, Lewis and Prescott (1999) have shown that if $d_{\rm plan}$ is a Williams Latin square then $d_{\rm min}$ is disconnected for the residual treatment effects.

Acknowledgement

The authors would like to thank two referees, an associate editor and the Editor for insightful comments that improved the manuscript substantially. Dibyen Majumdar's research is sponsored by NSF grant DMS-0204532 and Angela Dean's research is sponsored by NSF grant SES-0437251.

References

Bapat, R. B. and Raghavan, T. E. S. (1997). *Nonnegative Matrices and Applications*. Cambridge University Press.

Cheng, C.-S. and Wu, J. (1980). Balanced repeated measurement designs. *Ann. Statist.* **8**, 1272-1283. (Corrigendum. *Ann. Statist.* (1983), **11**, 349).

Davis, P. J. (1979). *Circulant Matrices*. Wiley, New York.

Diggle, P. D. and Kenward, M. G. (1994). Informative dropout in longitudinal data analysis (with discussion). *J. Roy. Statist. Soc. Ser. C* **43**, 49-93.

Ghosh, S. (1979). On robustness of designs against incomplete data. *Sankhyā* **40**, 204-208.

Ghosh, S. (1982). Robustness of BIBD against the unavailability of data. *J. Statist. Plann. Inference* **6**, 29-32.

Godolphin, J. D. (2004). Simple pilot procedures for the avoidance of disconnected experimental designs. *Appl. Statist.* **53**, 133-147.

Hedayat, A. S. and Afsarinejad, K. (1978). Repeated measurements designs, II. *Ann. Statist.* **6**, 619-628.

Hedayat, A. S. and Majumdar, D. (1985). Combining experiments under Gauss-Markov models. *J. Amer. Statist. Assoc.* **80**, 698-703.

Hedayat, A. S. and Yang, M. (2003). Universal optimality of balanced uniform crossover designs. *Ann. Statist.* **31**, 978-983.

Hedayat, A. S. and Yang, M. (2004). Universal optimality of selected crossover designs. *J. Amer. Statist. Assoc.* **99**, 461-466.

Higham, J. (1998). Row-complete latin squares of every composite order exist. *J. Combin Design* **6**, 63-77.

Isaac, P. D., Dean, A. M. and Ostrom, T. (2001). Generating pairwise balanced Latin Squares. *Stat. Appl.* **3**, 25-46.

Jones, B. and Kenward, M. G. (2003). *Design and Analysis of Cross-over Trials*. CRC Press, London.

Kunert, J. (1984). Optimality of balanced repeated measurements designs. *Ann. Statist.* **12**, 1006-1017.

Low, J. L. (1995). The design of crossover experiments subject to dropout. PhD thesis, School of Mathematics, University of Southampton.

Low, J. L., Lewis, S. M. and Prescott, P. (1999). Assessing robustness of crossover designs to subjects dropping out. *Statist. Comput.* **9**, 219-227.

Ratkowsky, D. A., Evans, M. A. and Alldredge, J. R. (1992). *Crossover Experiments: Design, Analysis and Application*. Marcel Dekker, New York.

Senn, S. (2002). *Cross-over Trials in Clinical Research*. Wiley, New York.

Stufken, J. (1991). Some families of optimal and efficient repeated measurements designs, *J. Statist. Plann. Inference* **27**, 75-83.

Stufken, J. (1996). Optimal crossover designs. In *Design and Analysis of Experiments. Handbook of Statistics* **13** (Edited by S. Ghosh and C. R. Rao), 63-90. North Holland, Amsterdam.

Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Austral. J. Statist.* **A2**, 149-168.

Mathematics, Statistics and Computer Science, , University of Illinois at Chicago, 851 S. Morgan, Chicago, Il 60607-7045, USA.

E-mail: dibyen@uic.edu

Department of Statistics, The Ohio State University, 404 Cockins Hall, 1959 Neil Avenue, Columbus, OH 43210, USA.

E-mail: dean.9@osu.edu

School of Mathematics, University of Southampton, Highfield, Southampton S017 1BJ, UK.

E-mail: S.M.Lewis@maths.soton.ac.uk

(Received May 2005; accepted April 2006)