

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Godolphin, JD; Warren, HR (2011) Improved conditions for the robustness of binary block designs against the loss of whole blocks. *Journal of statistical planning and inference*, 141 (11). pp. 3498-3505. ISSN 0378-3758 DOI: <https://doi.org/10.1016/j.jspi.2011.05.003>

Downloaded from: <http://researchonline.lshtm.ac.uk/168/>

DOI: [10.1016/j.jspi.2011.05.003](https://doi.org/10.1016/j.jspi.2011.05.003)

Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: <http://creativecommons.org/licenses/by-nc-nd/2.5/>

Improved Conditions for the Robustness of Binary Block Designs Against the Loss of Whole Blocks

J.D. Godolphin⁽¹⁾ and H.R. Warren⁽²⁾

(1) *Dept. of Mathematics, University of Surrey, Guildford, Surrey. GU2 7XH, UK*

(2) *Dept. of NCDE, London School of Hygiene and Tropical Medicine, London. WC1E 7HT, UK*

Abstract

Several authors have investigated conditions for a binary block design, D , to be maximally robust such that every eventual design obtained from D by eliminating $r_{[v]} - 1$ blocks, is connected, where $r_{[v]}$ is the smallest treatment replication. Four new results for the maximal robustness of D with superior properties are given. An extension of these results to widen the assessment of robustness of the planned design is also presented.

Keywords: connectivity, estimability, incomplete block design, information matrix

1. Introduction

The consequences of a disconnected block design may be so severe that the original aims of the experiment will be spoiled: pairwise treatment contrasts will be inestimable and it will not be possible to test the usual null hypothesis that all treatments have the same effect. Researchers have investigated methods for guarding against a disconnected design because of observation loss during the experiment; see the survey of methods given by Godolphin (2004, 2006) which considers cases where it is not possible to anticipate beforehand which observations may be missing. However, there are several experiments where the loss of a whole block is more likely than the loss of an individual observation; for example if the experimental units are the leaves on a plant then any damage to the plant will invalidate the block. Ghosh (1982) considered this problem of robustness against the

Email address: j.godolphin@surrey.ac.uk; Helen.Warren@lshtm.ac.uk (J.D. Godolphin⁽¹⁾ and H.R. Warren⁽²⁾)

unavailability of data for a balanced incomplete block design (BIBD) and Ghosh *et al.* (1983) extended this to m -associate partial BIBDs. Baksalary and Tabis (1987) derived a condition for an arbitrary binary block design to be maximally robust against the unavailability of data which is expressed solely in terms of the design parameters. These authors and also Sathe and Satam (1992) gave further conditions for maximal robustness that require the additional knowledge of how treatments are allocated to blocks. Related work is due to Bhaumik and Whittinghill (1991) and Morgan and Parvu (2008), who considered the optimality and relative efficiency of eventual designs which are obtained after the loss of a fixed number of blocks when the planned design is a BIBD.

The purpose of this paper is to show that the Baksalary-Tabis and the Sathe-Satam approaches do not take full account of the information given by basic design parameters, thus implying that stronger conditions exist. New conditions for a binary block design to be maximally robust are given which either improve on the Baksalary-Tabis and Sathe-Satam conditions or coincide with them. Like the Baksalary-Tabis conditions, the results are easy to apply and can be used before a decision on the design for the experiment is made. Furthermore, these conditions can be widened to assess designs for their robustness against the loss of a specified number of blocks without necessarily being maximally robust. A number of illustrative examples are presented.

2. Baksalary-Tabis and Sathe-Satam conditions

The notation of Baksalary and Tabis (1987) is adopted here. Let $D = \text{BD}(v, b, n, N)$ denote a binary block design on v treatments applied to n experimental units arranged in b blocks with $v \times b$ treatment-block incidence matrix N . The block sizes and the treatment replication numbers are given, respectively, by

$$k_{[1]} \geq k_{[2]} \geq \cdots \geq k_{[b]} \quad \text{and} \quad r_{[1]} \geq r_{[2]} \geq \cdots \geq r_{[v]}. \quad (2.1)$$

Consider an eventual design, $D_{\#} = \text{BD}(v, b - r_{[v]} + 1, n_{\#}, N_{\#})$ which is realized after the loss of $r_{[v]} - 1$ blocks from D . Ghosh (1982) has defined D to be maximally robust against the unavailability of data and with respect to estimability of treatment contrasts if $D_{\#}$ is connected irrespective of the choice of blocks removed from D . In this paper, a

design D having the property detailed by Ghosh will be described as being ‘maximally robust to loss of blocks’: this alternative phrase being considered more descriptive given the numerous approaches to dealing with potential observation loss since 1982. The following theorem, derived by Baksalary and Tabis (1987) gives a sufficient condition for the maximal robustness of a block design in terms of block sizes of the design.

Theorem 1. *For a binary design, D , the condition*

$$k_{[r_{[v]}]} + k_{[b]} > v. \quad (2.2)$$

is sufficient for D to be maximally robust to loss of blocks and with respect to estimability of treatment contrasts.

Suppose that K is the vector of block sizes, K^δ is the diagonal matrix with successive elements of K on the diagonal and $K^{-\delta}$ is the inverse of K^δ . Let κ_* denote the smallest off-diagonal element of $NK^{-\delta}N'$ and let λ_* be the smallest off-diagonal element of NN' . Two further sufficient conditions for the maximal robustness of D are given by Baksalary and Tabis (1987) in terms of κ_* and λ_* . We cite this result in the improved form due to Sathe and Satam (1992), which requires the terms

$$p(k_{[j]}, \rho) = \frac{k_{[j]\{\rho\}}(k_{[j]} - k_{[j]\{\rho\}})}{k_{[j]}} \quad \text{and} \quad q(k_{[j]}, \rho) = k_{[j]}p(k_{[j]}, \rho), \quad (2.3)$$

where $k_{[j]\{\rho\}} = \min\{\text{Int}(\frac{1}{2}k_{[j]}), \rho\}$ and $\text{Int}(\frac{1}{2}k_{[j]})$ denotes the integer part of $\frac{1}{2}k_{[j]}$.

Theorem 2. *Let $D = BD(v, b, n, N)$ be a binary block design and let κ_* and λ_* be the minimal concurrencies defined above. Then each of the two conditions*

$$(i) \quad \kappa_* > \frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, k_{[b]})}{k_{[b]}(v - k_{[b]})} \quad \text{and} \quad (ii) \quad \lambda_* > \frac{\sum_{j=1}^{r_{[v]}-1} q(k_{[j]}, k_{[b]})}{k_{[b]}(v - k_{[b]})} \quad (2.4)$$

is sufficient for D to be maximally robust to loss of blocks and with respect to estimability of treatment contrasts.

3. Conditions for Maximal Robustness in terms of Basic Design Parameters

Two conditions for maximal robustness to loss of blocks are derived that improve on or complement the condition of Theorem 1. This is possible because the Baksalary-Tabis argument leading to condition (2.2) makes use of available information on block sizes but does not take full account of the distribution of treatment replication numbers which also affects the connectivity of $D_\#$. Two preliminary lemmas are required.

Lemma 1. *Let $x_0 = 1$ and for each $m = 1, 2, \dots$ define*

$$x_m = r_{[v-y_m+1]} - r_{[v]} + 1, \quad \text{and} \quad y_m = k_{[b-x_{m-1}+1]}. \quad (3.1)$$

Then x_0, x_1, x_2, \dots and y_1, y_2, \dots are monotonically nondecreasing sequences of integers which terminate at the stop values $x_\#$ and $y_\#$, respectively, where $1 \leq x_\# \leq r_{[1]} - r_{[v]} + 1$ and $k_{[b]} \leq y_\# \leq k_{[1]}$.

Proof: From (2.1) $x_1 - x_0 = r_{[v-y_1+1]} - r_{[v]} \geq 0$ and $y_2 - y_1 = k_{[b-x_1+1]} - k_{[b]} \geq 0$ and a simple induction argument shows that x_0, x_1, x_2, \dots and y_1, y_2, \dots are monotonically nondecreasing sequences of positive integers. Stop values for the sequences are attained if there is a value for m such that $x_{m+1} = x_m$ or $y_{m+1} = y_m$; then $x_\# = x_m$ and $y_\# = y_m$. Otherwise the stop values are $y_\# = k_{[1]}$ and $x_\# = r_{[1]} - r_{[v]} + 1$, which follow because $v - y_m + 1 \geq 1$ for each m , since for a binary design the block-size y_m cannot be greater than v , and $b - x_{m-1} + 1 \geq 1$ follows for any m from the constraint $1 \leq r_{[m]} \leq b$.

Now consider a design $D_\#$ which is obtained after the loss of an arbitrary set of $r_{[v]} - 1$ blocks from D . The next result gives lower bounds for the number of blocks and for the number of treatments occurring in these blocks for any subset of blocks from $D_\#$ that can be regarded as a sub-design of D .

Lemma 2. *Suppose that S is a nonempty subset of the blocks of $D_\#$ with the property that any treatment which occurs in a block belonging to S has all of its replicates occurring in blocks contained in S . Then the number of blocks in S is at least as large as $x_\#$ and the number of treatments occurring in the blocks belonging to S is at least as large as $y_\#$.*

Proof: The largest block in the set S has size at least as large as $y_1 = k_{[b]}$. Since this is a block of D it must contain a treatment replicated at least $r_{[v-k_{[b]}+1]}$ times in D . Therefore this treatment is replicated at least $x_1 = r_{[v-k_{[b]}+1]} - r_{[v]} + 1$ times in S , hence S contains at least x_1 blocks. But this implies that the largest block in S has size at least as large as $y_2 = k_{[b-x_1+1]}$ experimental units; and this shows in turn that S contains at least $x_2 = r_{[v-y_2+1]} - r_{[v]} + 1$ blocks. Continuing this argument iteratively shows that there is a treatment in a block belonging to S which is replicated at least $x_\#$ times and the largest block in S has size at least as large as $y_\# = k_{[b-x_\#+1]}$. Lemma 2 follows.

Theorem 3. *The design D is maximally robust to loss of blocks and with respect to estimability of treatment contrasts provided that*

$$k_{[r_{[v]}]} + k_{[b-x_\#+1]} > v. \quad (3.2)$$

Proof: Assume that $D_{\#}$ is a disconnected design. It follows from the P -process of Godolphin (2004) that the blocks of $D_{\#}$ can be arranged in two nonempty sets S_1 and S_2 such that no treatment allocated to blocks in S_1 occurs in any of the blocks in S_2 . Let S_1 contain the largest block of $D_{\#}$ which has block-size no smaller than $k_{[r_{[v]}]}$ so the blocks in S_1 contain at least $k_{[r_{[v]}]}$ treatments. By Lemma 2, the blocks of S_2 contain at least $y_{\#} = k_{[b-x_{\#}+1]}$ treatments, which implies that $k_{[r_{[v]}]} + k_{[b-x_{\#}+1]} \leq v$. Condition (3.2) ensures that the original assumption is invalid, i.e. $D_{\#}$ is a connected design.

That the condition (3.2) can provide a significant improvement on condition (2.2) is apparent from the following examples.

Example 1 Twelve treatments are allocated to 39 units arranged in six blocks with sizes $k_{[1]} = 9$, $k_{[2]} = 8$, $k_{[3]} = 7$, $k_{[4]} = 6$, $k_{[5]} = 5$ and $k_{[6]} = 4$; treatment replications are $r_{[i]} = 4$ ($i = 1, \dots, 6$), $r_{[i]} = 3$ ($i = 7, 8, 9$) and $r_{[i]} = 2$ ($i = 10, 11, 12$). Then $k_{[r_{[v]}]} + k_{[b]} = k_{[2]} + k_{[6]} = 12$ so condition (2.2) fails. However $x_1 = r_{[9]} - r_{[12]} + 1 = 2$, $y_2 = k_{[5]} = 5$, $x_2 = r_{[8]} - r_{[12]} + 1 = 2$, therefore $x_{\#} = 2$ so $k_{[r_{[v]}]} + k_{[b-x_{\#}+1]} = k_{[2]} + k_{[5]} = 13 > v$, i.e. condition (3.2) is satisfied and we conclude that the design is maximally robust.

Example 2 A design D for six treatments allocated to 120 units arranged in 40 blocks consists of a combination of four designs: a BIBD(5, 10, 20, N_1) design for five treatments in ten blocks of size 2 with incidence matrix N_1 (10 blocks of size 2); BIBD(5, 10, 30, N_2) and BIBD(5, 10, 30, N_3) designs (20 blocks of size 3); a reinforced BIBD in the sense of Notz *et al.* (1994), i.e. a BIBD(5, 10, 30, N_4) plus an added sixth treatment in each block (10 blocks of size 4). The replication numbers are $r_{[i]} = 22$ for $1 \leq i \leq 5$ and $r_{[6]} = 10$. Thus $k_{[r_{[v]}]} + k_{[b]} = k_{[10]} + k_{[40]} = 6$ so condition (2.2) fails. However $y_1 = k_{[40]}$ hence $x_1 = r_{[6-k_{[40]}+1]} - r_{[6]} + 1 = r_{[5]} - r_{[6]} + 1 = 13$; also $x_2 = 13 = x_{\#}$ so condition (3.2) is satisfied, i.e. $k_{[10]} + k_{[40-13+1]} = 7 > v$. The design is shown to be maximally robust although the Baksalary-Tabis condition does not detect this.

It is clear that the argument based on the P -process in the proof of Lemma 2 can be amended so that the roles of block sizes and treatment replication numbers are interchanged. The following result provides a criterion which complements condition (3.2).

Theorem 4. *The design D is maximally robust to loss of blocks and with respect to*

estimability of treatment contrasts provided that

$$r_{[1]} + r_{[v-y_{\#}+1]} > b + r_{[v]} - 1. \quad (3.3)$$

Proof: Assume that $D_{\#}$ is disconnected and, as in Theorem 3, let the blocks of $D_{\#}$ be arranged in two sets S_1 and S_2 such that there is no overlap of treatments occurring in blocks of S_1 with those occurring in blocks of S_2 . Let the treatment with largest replication number, which is at least as large as $r_{[1]} - r_{[v]} + 1$, occur in a block of S_1 . Thus S_1 contains at least $r_{[1]} - r_{[v]} + 1$ blocks. By Lemma 2 S_2 contains at least $r_{[v-y_{\#}+1]} - r_{[v]} + 1$ blocks. This gives a contradiction if the sum of blocks in these two sets is larger than the total number of blocks in $D_{\#}$, i.e.

$$(r_{[1]} - r_{[v]} + 1) + (r_{[v-y_{\#}+1]} - r_{[v]} + 1) > b - r_{[v]} + 1,$$

which is condition (3.3).

Although the inequalities (3.2) and (3.3) are obtained from the same information in the form of design parameters consisting of block sizes and replication numbers, these two conditions do sometimes perform differently. In particular, condition (3.3) fails to show that the designs in Examples 1 and 2 are maximally robust, unlike condition (3.2). The following two examples show that (3.3) can succeed where (3.2) does not.

Example 3 Eight treatments are allocated to 52 units arranged in ten blocks with sizes $k_{[i]} = 8$ ($i = 1, 2, 3$) and $k_{[i]} = 4$ ($4 \leq i \leq 10$); treatment replications are $r_{[1]} = 9$, $r_{[i]} = 7$ ($2 \leq i \leq 4$), $r_{[i]} = 6$ ($5 \leq i \leq 7$) and $r_{[8]} = 4$. Condition (3.2) fails for this design D . However $y_1 = k_{[10]} = 4$, $x_1 = r_{[v-y_1+1]} - r_{[v]} + 1 = r_{[5]} - r_{[8]} + 1 = 3$ and $y_2 = k_{[b-x_1+1]} = k_{[8]} = 4$; therefore $y_{\#} = 4$ so that $r_{[1]} + r_{[v-y_{\#}+1]} = 15$ and $b + r_{[v]} - 1 = 13$; i.e. condition (3.3) is valid in this case and D is maximally robust.

Example 4 Let D consist of b blocks of size k_0 containing $v-1$ treatments with common replication number r_0 together with a control treatment which occurs in every block. It is evident from the P -process that if blocks are lost from D then the eventual design will be connected, as long as v treatments occur in it, since all remaining blocks contain the control. In particular, the design is maximally robust although this property is not established from condition (2.2) or from (3.2) if $k_0 \leq \frac{v}{2}$. For this design D condition (3.3) is $r_{[1]} + r_{[v-k_0+1]} = b + r_0 > b + r_0 - 1 = b + r_{[v]} - 1$, confirming D is maximally

robust.

An immediate corollary of Theorem 4 states that the condition

$$r_{[1]} + r_{[v-k_{[b]}+1]} > b + r_{[v]} - 1. \quad (3.4)$$

is sufficient for D to be maximally robust to loss of blocks and with respect to estimability of treatment contrasts. Conditions (3.3) and (3.4) do not appear to be mentioned in the literature although they provide useful and simple preliminary checks for maximal robustness by using just the information on block sizes and replication numbers. Naturally, (3.3) is preferred to (3.4) in general.

4. Conditions for Maximal Robustness using Minimal Concurrence

Baksalary and Tabis (1987) and Sathe and Satam (1992) have given further conditions for maximal robustness to loss of blocks and with respect to estimability of treatment contrasts which take account of the design configuration for D in terms of the incidence matrix N , as summarized by Theorem 2. Conditions (2.4) (i) and (ii) focus attention on the weighted minimal concurrence and the minimal concurrence by utilizing κ_* and λ_* , the smallest off-diagonal elements of $NK^{-\delta}N'$ and NN' respectively. In this section it is shown, firstly, that improvements to the conditions of Theorem 2 are made by effectively reducing the bounds which κ_* and λ_* need to attain in order to establish that the design is maximally robust. It is helpful to note the following lemma which is a stand-alone result of considerable interest in its own right.

Lemma 3. *Let x_m and y_m be defined by (3.1) with stop values $x_{\#}$ and $y_{\#}$ respectively. If either (i) $x_{\#} > \frac{1}{2}(b - r_{[v]} + 1)$, or (ii) $y_{\#} > \frac{1}{2}v$, then D is maximally robust to loss of blocks and with respect to estimability of treatment contrasts.*

Proof: Lemma 2 shows that either of conditions (i) or (ii) is sufficient to deny the existence of two sets of blocks of $D_{\#}$ such that treatments occurring in blocks from one set do not overlap with treatments occurring in blocks from the other.

It is clear that further investigation of conditions for maximal robustness of D is only required if the stop values $x_{\#}$ and $y_{\#}$ for this design are sufficiently small so that both conditions of Lemma 3 fail. Such designs are not uncommon, however; for instance,

neither of the conditions of the Lemma are satisfied by any of Examples 1, 2 or 3. In what follows it is assumed that conditions (i) and (ii) of Lemma 3 do not apply.

Theorem 5. *Let $D = BD(v, b, n, N)$ be a binary block design and let κ_* and λ_* be defined as in Theorem 2. Let $x_0 = 1$ and for each $m = 1, 2, \dots$ define x_m and y_m by (3.1) and let $x_\#$ and $y_\#$ be the respective stop values. Then any one of the two conditions*

$$(i) \quad \kappa_* > \frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, y_\#)}{y_\#(v - y_\#)} \quad \text{and} \quad (ii) \quad \lambda_* > \frac{\sum_{j=1}^{r_{[v]}-1} q(k_{[j]}, y_\#)}{y_\#(v - y_\#)}, \quad (4.1)$$

is sufficient for D to be maximally robust to loss of blocks and with respect to estimability of treatment contrasts, with $p(\cdot, \cdot)$ and $q(\cdot, \cdot)$ defined by (2.3). Furthermore, the bounds in (i) and (ii) coincide with or are lower than the corresponding bounds in Theorem 2.

Proof: Assume that $D_\#$ is disconnected and, therefore, that the blocks of $D_\#$ are arranged in nonempty sets S_1 and S_2 such that the treatments which occur in the blocks in S_1 do not occur in any of the blocks in S_2 . From the proof of Theorem 3 it follows that the blocks of S_1 contain at least $k_{[r_{[v]}]}$ treatments and that the blocks of S_2 contain at least $y_\# = k_{[b-x_\#+1]}$ treatments. Thus suppose that v_0 treatments occur in blocks in one set and $v - v_0$ treatments occur in blocks in the other, where $0 < v_0 \leq \frac{1}{2}v$. Then

$$0 \leq k_{[b]} \leq y_\# \leq v_0 \leq \frac{1}{2}v. \quad (4.2)$$

Furthermore it follows easily from Result 2 of Sathe and Satam (1992, p.97) that

$$\frac{p(k_{[j]}, m_2)}{m_2(v - m_2)} \leq \frac{p(k_{[j]}, m_1)}{m_1(v - m_1)}, \quad (4.3)$$

whenever m_1, m_2 are integers satisfying $0 < m_1 < m_2 \leq \text{Int}(\frac{1}{2}v)$. Applying (4.3) to the inequalities (4.2), we have

$$\frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, v_0)}{v_0(v - v_0)} \leq \frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, y_\#)}{y_\#(v - y_\#)} \leq \frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, k_{[b]})}{k_{[b]}(v - k_{[b]})}. \quad (4.4)$$

Now since $D_\#$ is disconnected by assumption it follows from a result of Sathe and Satam (1992, p.95) that the term κ_* satisfies the inequality

$$\kappa_* \leq \frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, v_0)}{v_0(v - v_0)}. \quad (4.5)$$

However, v_0 is unknown so a bound for the right hand side of (4.5) is required. The upper bound of Sathe and Satam (1992) is the final term on the right side of (4.4); but the upper bound given by the central term in (4.4) will either coincide with or be smaller than the Sathe-Satam bound. The result (i) of the Theorem follows. The result (ii) of the Theorem is derived in a similar way.

When determining the properties of a suggested design D , the evaluation of the stop values $x_{\#}$ and $y_{\#}$ from (3.1) is a recommended preliminary step. If D is maximally robust but $x_{\#}$ and $y_{\#}$ are such that conditions (i) and (ii) of Lemma 3 do not apply, it is possible that either condition of Theorem 5 may detect this property even though both conditions of Theorem 2 do not. This is illustrated by the following example.

Example 5 Let D be a cyclic design consisting of 10 treatments, arranged in 12 blocks of size five and two further blocks of size four:

1	2	3	4	5	6	7	8	9	10	1	2	3	4
2	3	4	5	6	7	8	9	10	1	2	3	4	5
3	4	5	6	7	8	9	10	1	2	3	4	5	6
6	7	8	9	10	1	2	3	4	5	6	7	8	9
7	8	9	10	1	2	3	4	5	6	7	8		

where columns show the blocks. In this case the terms $x_{\#} = x_2 = r_{[6]} - r_{[10]} + 1 = 3$ and $y_{\#} = y_2 = k_{[12]} = 5$. The upper bound on the right hand side of equation (4.1) (i) is given by $\{4p(5, 5)\}/(5 \times 5) = \frac{24}{125}$. Since $k_{[b]} = 4$ then the corresponding bound on the right hand side of equation (2.4) (i) is $\{4p(5, 4)\}/(4 \times 6) = \frac{1}{5}$. The value for the weighted minimal concurrence κ_* is $\frac{1}{5}$ so condition (4.1) (i) shows that design D is maximally robust but the Sathe-Satam condition (2.4) (i) does not.

The bound on the right side of (4.1) (ii) is given by $\{4q(5, 5)\}/25 = \frac{24}{25}$ and the bound on the right side of (2.4) (ii) is $\{4q(5, 4)\}/24 = 1$. The minimal concurrence λ_* is unity which confirms that D is maximally robust but, as in the previous case, this is not picked up by the Sathe-Satam condition (2.4) (ii).

If for some design D the x, y sequence (3.1) does not advance beyond $y_{\#} = k_{[b]}$ then it may be possible to strengthen conditions (4.1) which coincide with conditions (2.4).

Sathe and Satam (1992, Corollary 1) point out that inequalities (2.4) can be improved if λ^* , the maximum concurrence between any pair of treatments, is suitably small, i.e.

$$\lambda^* < r_{[v]} - \omega + 1, \quad (4.6)$$

where ω is the maximum number of identical blocks of size $k_{[b]}$. However, it is often the case that (4.6) itself can be sharpened by using an argument which is similar to that of Lemma 2. To see this we require the following result.

Lemma 4. *Suppose that S is a nonempty subset of the blocks of $D_{\#}$ with the property that any treatment which occurs in a block belonging to S has all of its replicates occurring in blocks belonging to S . If the condition*

$$\lambda^* < r_{[v-k_{[b]}+1]} + r_{[v-k_{[b]}+2]} - r_{[v]} - \omega + 1, \quad (4.7)$$

is satisfied then the blocks of S contain at least $k_{[b]} + 1$ treatments.

Proof: The blocks of S must contain at least $k_{[b]}$ treatments. Assume that they contain exactly $k_{[b]}$ treatments; then S must consist of identical blocks of size $k_{[b]}$, hence no more than ω blocks are contained in S .

Consider any two treatments belonging to the blocks of S ; note that two such treatments exist since $k_{[b]} \geq 2$. The total number of blocks in D containing one or both of these treatments is at least equal to $r_{[v-k_{[b]}+1]} + r_{[v-k_{[b]}+2]} - \lambda^*$ and the number of blocks in S containing one or both of these treatments cannot be smaller than this quantity by an amount bigger than $r_{[v]} - 1$. Therefore a necessary condition for the blocks of S to contain exactly $k_{[b]}$ treatments is given by

$$r_{[v-k_{[b]}+1]} + r_{[v-k_{[b]}+2]} - \lambda^* - r_{[v]} + 1 \leq \omega,$$

and the contradiction of (4.7) establishes the lemma.

The sensitivity of condition (4.7) compared to (4.6) is measured by the difference

$$r_{[v-k_{[b]}+1]} + r_{[v-k_{[b]}+2]} - 2r_{[v]},$$

which is strictly positive when at most $k_{[b]} - 1$ treatments have the minimal replication number $r_{[v]}$ and may be large for some designs compared to the maximum concurrence λ^* . The following result is an immediate consequence of Lemma 4 and Theorem 5.

Theorem 6. Let $D = BD(v, b, n, N)$ be a binary block design such that (4.7) is satisfied and let κ_* and λ_* be defined as in Theorem 2. Then any one of the two conditions

$$(i) \quad \kappa_* > \frac{\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, k_{[b]} + 1)}{(k_{[b]} + 1)(v - k_{[b]} - 1)} \quad \text{and} \quad (ii) \quad \lambda_* > \frac{\sum_{j=1}^{r_{[v]}-1} q(k_{[j]}, k_{[b]} + 1)}{(k_{[b]} + 1)(v - k_{[b]} - 1)}, \quad (4.8)$$

is sufficient for D to be maximally robust to loss of blocks and with respect to estimability of treatment contrasts, with $p(.,.)$ and $q(.,.)$ defined by (2.3).

We complete this section with examples to illustrate a number of points.

Example 6 To see that conditions (i) and (ii) are not equivalent for each of Theorems 2, 5 and 6, consider an equi-replicate variance balanced design D with $r_{[v]} = 5$. This design has eight treatments arranged in eight blocks of size four and four blocks of size two, given in Table 1 of Gupta and Jones (1983, p. 436):

1	1	1	1	2	2	3	4	1	2	3	4
2	2	3	4	3	5	5	5	5	6	7	8
3	7	6	6	4	7	6	6				
4	8	8	7	5	8	8	7				

where columns show the blocks. The concurrencies are $\kappa_* = \frac{1}{2}$, $\lambda_* = 1$ and $\lambda^* = 2$. The design D is maximally robust since condition (i) of Theorems 2 and 5 is satisfied, i.e. $\kappa_* > \frac{1}{3}$, although condition (ii) is $\lambda_* > \frac{4}{3}$ which fails for this design. In this case $y_{\#} = 2 = k_{[b]}$ and condition (4.7) is $\lambda^* < 5$ so Theorem 6 applies; however conditions (i) and (ii) are $\kappa_* > \frac{4}{15}$ and $\lambda_* > \frac{16}{15}$, i.e. again one condition is satisfied but not the other.

Example 7 A regular group divisible design D for six treatments consists of twenty blocks of size three, given by

1	1	1	1	2	2	2	3	3	1	1	1	1	1	1	2	2	2	2	2
2	2	4	5	3	4	5	4	4	3	3	3	3	3	3	4	4	4	4	4
3	4	6	6	6	5	6	5	6	5	5	5	5	5	5	6	6	6	6	6

where columns show the blocks. This design is cited as R53 on page 190 of Clatworthy (1973). None of the conditions of Theorems 1 to 5 are satisfied for design D , although several bounds are attained but not surpassed. For instance:

condition (3.2) is not valid since $k_{[r_{[v]}]} + k_{[b-x_{\#}+1]} = 3 + 3$ and $v = 6$;

Lemma 2 (ii) is not valid since $y_{\#} = 3$ and $\frac{1}{2}v = 3$;

condition (4.1) (i) is not valid since $\sum_{j=1}^{r_{[v]}-1} p(k_{[j]}, y_{\#}) / \{y_{\#} (v - y_{\#})\} = \frac{2}{3}$ and $\kappa_{*} = \frac{2}{3}$;

condition (4.1) (ii) is not valid since $\sum_{j=1}^{r_{[v]}-1} q(k_{[j]}, y_{\#}) / \{y_{\#} (v - y_{\#})\} = 2$ and $\lambda_{*} = 2$.

Furthermore the maximum concurrence is $\lambda^{*} = 7$ but the bound given by inequality (4.7) of Lemma 4 has value 5 so Theorem 6 cannot apply. Therefore maximal robustness is not established for this design. Notwithstanding the lack of evidence, however, Sathe and Satam (1992) do infer that the design cited as R53 is maximally robust.

To see by inspection that the Sathe-Satam conjecture is incorrect it suffices to lose $r_{[v]} - 1 = 9$ blocks and leave an eventual design $D_{\#}$ which is disconnected. Let the first nine blocks in the displayed design be removed. The remaining blocks split into two sets S_1 and S_2 such that the six blocks of S_1 contain all odd-numbered treatments and the five blocks of S_2 contain all even-numbered treatments. Hence R53 is not maximally robust. This example illustrates that the conditions of the theorems in this note appear to be sensitive to detecting maximal robustness, at least for some designs.

Examples 8a and 8b Let the design D_a consist of ten treatments arranged in twenty blocks of size four:

1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
2	3	4	5	1	3	4	5	1	2	2	3	4	5	1	3	4	5	1	2
6	7	8	9	10	6	7	8	9	10	6	8	9	10	7	6	8	9	10	7
7	8	9	10	6	8	9	10	6	7	8	9	10	7	6	9	10	7	6	8

where columns show the blocks. This design represents two isomorphic copies of the singular group divisible design cited as S9 on page 108 of Clatworthy (1973). It is not surprising to find that D_a is maximally robust, indeed both conditions of Theorem 5 are satisfied. Suppose, however, that another design D_b is used which differs from D_a in that the penultimate block $(4, 1, 10, 6)'$ is replaced by $(4, 2, 10, 8)'$. It is interesting to ask if this minor amendment implies a different robustness status for D_b .

It turns out that none of the conditions of Theorems 1 to 5 are satisfied for D_b . In particular the bounds for conditions (4.1) (i) and (ii) of $\frac{1}{4}$ and 1 are attained, but not

surpassed, by concurrencies κ_* and λ_* respectively. Moreover, the maximum concurrence $\lambda^* = 7$ and the bounds given by inequalities (4.6) and (4.7) are 7 and 9 respectively, so Theorem 6 applies. The bounds for conditions (4.8) (i) and (ii) of $\frac{6}{25}$ and $\frac{24}{25}$ are surpassed by concurrencies κ_* and λ_* respectively, implying that design D_b is maximally robust. Note that condition (4.7) of Lemma 4 has a key part to play in this conclusion since the Sathe-Satam condition (4.6) on its own does not show that Theorem 6 is admissible.

5. Further Developments

Key issues requiring decisions by the statistical team responsible for design selection should include robustness considerations if, as is often the case, there is a possibility of observation loss during the experiment. Block sizes and treatment replication numbers affect the vulnerability of a planned experiment where the loss of whole blocks is thought to be a possibility. Theorems 3 and 4 and Lemma 3 provide basic conditions for maximal robustness which depend, in general, on the evaluation of $x_{\#}$ and $y_{\#}$ from Lemma 2 using only the information which is available on block sizes and replication numbers. In particular, if the block sizes are suitably large then any binary design which is arranged in blocks of these sizes will be maximally robust, whatever the configuration.

If experimental constraints do not allow for block sizes or treatment replications which satisfy these basic conditions, then other factors are required to discriminate between possible competing designs for the experiment. Godolphin (2006) pointed out that block designs that are equally efficient are not necessarily equally vulnerable to observation loss, and a similar conclusion can be anticipated if there is a tendency for whole blocks to be invalidated. It is sensible to concentrate on the concurrence or weighted concurrence between any pair of treatments and compare κ_* or λ_* with the bounds given in Theorems 5 or 6. For a proper design, where all block sizes are the same, conditions (i) and (ii) of Theorems 5 or 6 are equivalent and this situation is conveniently summarized in Table 1. For values of v ranging from 4 to 12 and for all block sizes satisfying $2 \leq k \leq \frac{v}{2}$, Table 1 specifies the minimal concurrence sufficient to ensure maximal robustness of the design for a range of values of the smallest treatment replication from $r_{[v]} = 2$ up to $r_{[v]} = 20$.

Note that Table 1 can also be used for designs with variable block sizes. If a design D

v	k	2	3	4	5	6	7	8	9	10	11	$r_{[v]}$ 12	13	14	15	16	17	18	19	20
4	2	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5
5	2	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4
6	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3
	3	1	1	1	1	2	2	2	2	3	3	3	3	3	4	4	4	4	5	5
7	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
	3	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4
8	2	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2
	3	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3
	4	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5
9	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
	3	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3
	4	1	1	1	1	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4
10	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2
	3	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
	4	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4
	5	1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5
11	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2
	3	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
	4	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3
	5	1	1	1	1	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4
12	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
	4	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3
	5	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4
	6	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5

Table 1: Minimum values of λ_* to ensure maximal robustness of a proper design; v = number of treatments, k = common block size and $r_{[v]}$ = smallest replication number.

has block sizes given by (2.1) then it is straightforward to see that D will be maximally robust if the minimum concurrence λ_* is at least as large as the value indicated in Table 1, with $k_{[1]}$ replacing k in the table. This is particularly helpful if the entries in the table are the same for each block-size; for example, if D has nine treatments arranged in blocks of size 2, 3 and 4, with minimal replication number $r_{[v]} = 4$, and if every pair of treatments occurs in a block at least once then D is maximally robust. Otherwise, the entry in the table for the largest block-size applies; for example, if $v = 8$, $r_{[v]} = 5$ and the two block sizes are $k_{[1]} = 4$ and $k_{[8]} = 2$ then the table shows that a value of 2 for the smallest concurrence λ_* is sufficient for maximal robustness. However, in these circumstances an improved condition may be identified by other results of section 4; for example, the variance balanced design D of Example 6 has parameters $v = 8$, $r_{[v]} = 5$, $k_{[1]} = 4$ and $k_{[8]} = 2$ but D is maximally robust, even though $\lambda_* = 1$, as verified by the condition on weighted concurrencies.

A useful extension of the theorems established in this paper is that they can be widened to provide conditions for assessing robustness of designs to the loss of b_* blocks, where b_* is smaller than $r_{[v]} - 1$. These conditions will have particular value in experimental situations where the smallest replication number, $r_{[v]}$, is relatively large and the loss of as many as $r_{[v]} - 1$ blocks would be considered a remote possibility; in such cases it may be more relevant to consider the consequences of the loss of smaller numbers of blocks. The authors are grateful to a referee for the suggestion that the results in this paper can be generalized to cover the concept of the robustness of a design against the loss of b_* blocks, for any specified b_* such that $1 \leq b_* \leq r_{[v]} - 1$. This generalization requires an adjustment to the definitions of x_m and y_m given in Lemma 1. Let $x_{0,b_*} = 1$ and for each $m = 1, 2, \dots$ define

$$x_{m,b_*} = r_{[v-y_{m,b_*}+1]} - b_*, \quad \text{and} \quad y_{m,b_*} = k_{[b-x_{m-1,b_*}+1]}. \quad (5.1)$$

The argument of Lemma 1 shows that $\{x_{m,b_*}\}$ and $\{y_{m,b_*}\}$ are monotonically nondecreasing sequences of integers which terminate at the stop values $x_{\#,b_*}$ and $y_{\#,b_*}$, where $1 \leq x_{\#,b_*} \leq r_{[1]} - b_*$ and $k_{[b]} \leq y_{\#,b_*} \leq k_{[1]}$. Let $D_{\#,b_*}$ denote an eventual design obtained by the loss of any b_* blocks from D . The argument of Lemma 2 establishes that if S is a nonempty subset of the blocks of $D_{\#,b_*}$ with the property that any treatment which

occurs in a block belonging to S has all of its replicates occurring in blocks contained in S , then S contains at least $x_{\#,b_*}$ blocks and at least $y_{\#,b_*}$ treatments. Theorems 3 to 6 and Lemma 3 can then be extended to cover the loss of b_* blocks, with $b_* = r_{[v]} - 1$ as a particular case, by replacing $x_{\#}$, $y_{\#}$ with $x_{\#,b_*}$, $y_{\#,b_*}$ and $r_{[v]} - 1$ with b_* , as necessary. For example, Theorem 5 can be generalized as follows:

Theorem 7. *Let $D = BD(v, b, n, N)$ be a binary block design and let κ_* and λ_* be defined as in Theorem 2. Let $x_{0,b_*} = 1$ and for each $m = 1, 2, \dots$ define x_{m,b_*} and y_{m,b_*} by (5.1) and let $x_{\#,b_*}$ and $y_{\#,b_*}$ be the respective stop values. Then any one of the two conditions*

$$(i) \quad \kappa_* > \frac{\sum_{j=1}^{b_*} p(k_{[j]}, y_{\#,b_*})}{y_{\#,b_*} (v - y_{\#,b_*})} \quad \text{and} \quad (ii) \quad \lambda_* > \frac{\sum_{j=1}^{b_*} q(k_{[j]}, y_{\#,b_*})}{y_{\#,b_*} (v - y_{\#,b_*})}, \quad (5.2)$$

is sufficient for D to be robust against the loss of b_ blocks and with respect to the estimability of treatment contrasts, with $p(\cdot, \cdot)$ and $q(\cdot, \cdot)$ defined by (2.3).*

Example 7 (revisited) Consider again the regular group divisible design for six treatments arranged in twenty blocks of size three, cited as R53 by Clatworthy (1973). In this case $r_{[v]} = 10$ and it is shown that this design is not robust to the loss of 9 blocks, i.e. it is not maximally robust. Suppose however that $b_* = 8$ and that $x_{\#,8}$, $y_{\#,8}$ are defined by (5.1). It follows that

$$\frac{\sum_{j=1}^8 p(k_{[j]}, y_{\#,8})}{y_{\#,8} (v - y_{\#,8})} = \frac{16}{27} < \frac{2}{3} = \kappa_*.$$

Hence from Theorem 7 (i) it is established that any eventual design, $D_{\#,8}$, formed by the removal of 8 blocks is connected. The same conclusion is also reached after examining condition (ii) of Theorem 7. It follows that the Clatworthy design R53 is robust to the loss of up to 8 blocks, although this design is not maximally robust.

Acknowledgements

The authors are grateful to the associate editor and to a referee for suggestions for improving the paper.

Helen Warren (née Thornewell)'s PhD research at the University of Surrey was supported by EPSRC grant KB9021.

References

- [1] Baksalary, J.K. and Tabis, Z. (1987) Conditions for the robustness of block designs against the unavailability of data. *J. Statist. Plannng Inf.* 16, 49-54.
- [2] Bhaumik, D.K. and Whittinghill, D.C. (1991) Optimality and robustness to the unavailability of blocks in block designs. *J. R. Statist. Soc. B* 53, 399-407.
- [3] Clatworthy, W.H. (1973) Tables of two-associate-class partially balanced designs. *Nat. Bureau Stand. Appl. Math. Ser.* 63.
- [4] Ghosh, S. (1982) Robustness of designs against the unavailability of data. *J. Statist. Plannng Inf.* 6, 29-32.
- [5] Ghosh, S., Rao, S.B. and Singhi, N.M. (1983) On a robustness property of PBIBD. *J. Statist. Plannng Inf.* 8, 355-363.
- [6] Godolphin, J.D. (2004) Simple pilot procedures for the avoidance of disconnected experimental designs. *Applied Statistics* 53, 133-147.
- [7] Godolphin, J.D. (2006) The specification of rank reducing observation sets in experimental design. *Comp. Statist. Data Anal.* 51, 1862-1874.
- [8] Gupta, S.C. and Jones, B. (1983) Equireplicate balanced block designs with unequal block sizes. *Biometrika* 70, 433-440.
- [9] Morgan, J.P. and Parvu, V. (2008) Most robust BIBDs. *Statistica Sinica* 18, 689-707.
- [10] Notz, W.I., Whittinghill, D.C. and Zhu, Y. (1994) Robustness to the unavailability of data in block designs. *Metrika* 41, 263-275.
- [11] Sathe, Y.S. and Satam, M.R. (1992) Some more robust block designs against the unavailability of data. *J. Statist. Plannng Inf.* 30, 93-98.