On the robustness of crossover trials against subject drop-out - examples of perpetually connected designs

Journal Title
XX(X):2–18

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

P J Godolphin[1,2] and E J Godolphin[3,4]

Abstract

When performing a repeated measures experiment, such as a clinical trial, there is a risk of subject drop-out during the experiment. If one or more subjects leave the study prematurely a situation could arise where the eventual design is disconnected, implying that very few treatment contrasts for both direct effects and carryover effects are estimable. This paper aims to identify experimental conditions where this problem with the eventual design can be avoided. It is shown that in the class of Uniformly Balanced Repeated Measurement Designs (UBRMDs) consisting of two or more Latin squares, there are planned designs with the following useful property. Provided that all subjects have completed the first two periods of study, such a design will not be replaced by a disconnected eventual design due to drop-out, irrespective of the type of drop-out behaviour that may occur. Designs with this property are referred to as *perpetually connected*. These experimental conditions are identified and examined in the paper and an example of at least one perpetually connected UBRMD design is given in each case. The results improve upon previous contributions in the literature that have been confined largely to cases in which drop-out occurs only in the final periods of study.

Keywords

crossover design, subject drop-out, missing data, clinical trial, uniformly balanced repeated measurement design, perpetually connected

Prepared using sagej.cls *[Version: 2015/06/09 v1.01]*

1 Introduction

The crossover design is regularly implemented in scientific experiments and has a rich history, stretching back to the 19th century (Jones and Kenward[1], p5). In medical research, the most common design is that with two periods and two treatments, with these frequently used in trials of neurology, psychiatry and pain treatment[2]. However, there is often a need for greater comparison and crossover trials with t treatments, where t could be much greater than two, are not uncommon. This is especially the case in the early phases of drug development[3]. Pharmacokinetic studies frequently utilise crossover designs for clinical trials where the number of treatments is three or more, for example see Grattan et al.[4].

Missing data is a problem that occurs in many clinical trials and can have substantial consequences on study quality. Strategies to limit the impact of missing data on the analysis and interpretation of clinical trials are supported by the National Academy of Science Report[5]. The report recommends that 'a more principled approach to design and analysis in the presence of missing data is both needed and possible' and that 'careful design and conduct limit the amount and impact of missing data'[5]. Crossover trials with a large number of treatments under test can become lengthy and as such, missing data is commonly seen through subject drop-out. First introduced by Rubin[6] and also defined on page 32 of Molenberghs and Kenward[7], data are said to be *missing at random* (MAR) if, conditional on the observed data, the probability of missing observations is conditionally independent of the unobserved data. If the MAR assumption holds then conventional likelihood-based methods on the observed data, which ignores the missingness mechanism, will provide valid estimates (Molenberghs and Kenward[7], chapter 12). It is assumed in this paper that all subject drop-out is MAR. For a discussion of crossover trials where this assumption does not hold, consult Rosenkranz[8] or Matthews and Henderson[9].

In crossover trials, subjects may drop out after receiving one or more treatments, which could be due to a positive or negative reaction to a trial intervention (Higgins and Green[10], §16.4.3) or unrelated to treatment[7]. Furthermore, there is increased risk of drop-out in crossover trials because they are usually longer in length when compared to the equivalent parallel group study (Higgins and Green[10], §16.4.2).

[1]Nottingham Clinical Trials Unit, University of Nottingham, UK
[2]Stroke Trials Unit, Division of Clinical Neuroscience, University of Nottingham, UK
[3]Department of Mathematics, University of Surrey, UK
[4]Department of Mathematics, Royal Holloway University of London, London, UK

Corresponding author:
Peter Godolphin, Nottingham Clinical Trials Unit, Queens Medical Centre, University of Nottingham, Nottingham, NG7 2UH, UK
Email: Peter.Godolphin@nottingham.ac.uk

Prepared using sagej.cls

Low *et al.*[11] have suggested that the expected drop-out rate in a crossover trial would be between 5-10%, and in some cases could be as high as 25%. Missing data in crossover trials and some implications are discussed in the literature (Jones and Kenward[1], p202-204). Whilst subject drop-out in parallel group trials may lead to complications with analysis and greater risk of bias[12], there can be a substantial loss of information for crossover studies. With this loss of information, it is possible that the eventual design which remains after drop-out is disconnected[13], which is an unwelcome situation since not all treatment contrasts are estimable. In such circumstances the experiment may be severely compromised.

Due to subjects in a crossover trial receiving more than one treatment, a washout period is often implemented[14], where subjects take no study intervention. The aim of washout is to reduce the effects of a treatment in one study period from carrying over into the next and unduly influence the perceived treatment effect in the following period. Even with the inclusion of washout periods, it is advisable to ensure any crossover design is *balanced*, implying that the impact of carryover effects that may still be present is distributed evenly over the direct treatment effects. Furthermore, the importance of balance for the efficiency of the design is well documented (Jones and Kenward[1], p154-p177). Balance is achieved by ensuring that each treatment is preceded by every other treatment the same number of times, and is never preceded by itself. A design is referred to as *uniform* if each treatment occurs equally often in a period, and for each subject, each treatment occurs in the same number of periods. Many uniform designs require that the number of periods is equal to the number of treatments; furthermore the number of subjects is often taken to be a multiple of the number of treatments. *Uniformly Balanced Repeated Measurement Designs* (UBRMDs) satisfy these constraints and have been a popular choice for researchers and practitioners, dating back for very many years[15]. The UBRMDs have attractive optimality properties, see, for example Hedayat and Afsarinejad[16], Cheng and Wu[17] and Hedayat and Yang[18].

The design that is chosen, without the presence of subject drop-out, is referred to as the *planned design* and the design which occurs in practice after any drop-out is the *eventual design*. It is expected that the final analysis is based on the eventual design. In this paper, the planned design is considered to belong to the family of UBRMDs and will consist of m Latin squares where $m \geq 2$. Single square designs are not considered. They are unlikely to be popular for crossover trials or many other crossover studies, except for specialized experiments; for a comprehensive survey of crossover studies where the number of subjects is less than or equal to the number of treatments, see Bate *et al.*[19].

Much of the work in this area has been confined to eventual designs where subject drop-out has occurred towards the end of the study. Majumdar *et al.*[20] showed that if $t \geq 5$ the eventual design for any planned UBRMD is connected if drop-out behaviour is restricted to the last period only. They further showed that if $t \geq 8$ the eventual design for any planned UBRMD is connected if drop-out behaviour is restricted to the last two periods only. Zhou and Majumdar[21] consider a subclass of the UBRMDs and show that such designs are relatively efficient if subject drop-out occurs at random and is limited to

the final period. However, even when drop-out occurs at random it seems likely that some subjects may leave the study prematurely, after just the first few periods. A crossover design is said to be *perpetually connected* if all subjects complete period one and period two, and the eventual design is treatment connected for direct effects and carryover effects, irrespective of drop-out behaviour thereafter. Clearly a perpetually connected design carries a much lower risk for a compromised experiment since it relies formally upon the participation of the subjects only during the first two periods of study.

The purpose of this paper is to show that a UBRMD exists which is perpetually connected for many experimental situations which are likely to arise in practice. Two particular cases which show how designs are at differing risks of becoming disconnected are illustrated in §2. The next section considers a number of commonly chosen UBRMDs for medical and other experiments, with three, four, five, six and seven treatments. It is shown that a design can be chosen for most of these cases that is perpetually connected, thus reducing the risk of inestimable treatment contrasts, even if there is severe and unpredictable information loss. The basis for this is an established theorem which shows the important role of the most extreme eventual design, herein defined as the minimal design. A discussion of the implication of these findings is given in §4 and the theoretical basis for these results is contained in the Appendix.

2 Two Illustrations

2.1 Designs to compare four treatments using eight subjects

To illustrate the ideas in the paper consider an experiment to compare four treatments labelled $1, 2, 3$ and 4 by using eight subjects over four periods. Here and throughout the paper the columns of the design refer to subjects and the rows refer to periods in sequential order. Two UBRMDs for this study have been suggested by Low *et al.* [11] and Godolphin [13], which are specified here as Design $2A$ and Design $2B$.

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

Design 2*A*. Two Williams replicates

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

Design 2*B*. Two distinct squares

In general, there are two distinct balanced Latin squares, both squares proposed by Williams [22] as shown for example by Bate *et al.* [19]. Designs $2A$ and $2B$ are two essentially different UBRMDs to compare four treatments over four periods using eight subjects. Each design is universally optimal, but the general conclusion from both Low *et al.* [11] and Godolphin [13] is that Design $2B$ is preferable to Design $2A$ if there is the likelihood of drop-out in the fourth period.

1	2	3	4	1	2	3	4
2	3	4	1	2	3	4	1
4	1	2	3	4	1	2	3
∗	4	∗	2	∗	4	∗	2

Design 2*C*. Design $2A$ after drop-out

1	2	3	4	1	2	3	4
2	3	4	1	4	1	2	3
∗	1	∗	3	∗	3	∗	1
∗	4	∗	2	∗	4	∗	2

Design 2*D*. Design $2B$ after drop-out

However, both designs are at risk in the presence of subject drop-out. Design $2C$ is the eventual design when four subjects are lost from Design $2A$ in period four and Design $2D$ is the eventual design when four subjects are lost from Design $2B$ in period three. Design $2C$ and Design $2D$ are disconnected. It is shown in the Appendix that in neither case is it possible to formulate the usual treatment sums of squares for direct effects nor for carryover effects, and that several elementary treatment direct contrasts and elementary treatment carryover contrasts are not estimable. An experiment based on either design which suffers the drop-out described would be compromised since little could be achieved from the results which are obtained. Furthermore, the design obtained by taking two copies of the alternative Williams Latin square has properties similar to those of Design $2A$. Therefore we are unable to suggest a perpetually connected UBRMD to compare four treatments over four periods using eight subjects.

2.2 Designs to compare four treatments using twelve subjects

It is interesting to consider an extension of the experiment discussed in §2.1 by using twelve subjects to compare four treatments over four periods. There are several balanced designs to select for this experiment. One UBRMD is Design $2E$ which extends Design $2A$ and consists of three replicates of a Williams balanced square. Another UBRMD for the experiment is Design $2F$ which consists of three mutually orthogonal Latin squares (MOLS). Unlike Design $2E$, none of the individual Latin squares in Design $2F$ is balanced, however the combination of these three squares does yield a UBRMD.

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

Design 2*E*. Three Williams Squares

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

Design 2*F*. Three MOLS

Each of the two designs is universally optimal. However if some subjects leave the experiment early then Design $2E$ is at risk. Design $2G$ represents the eventual design after the termination of the experiment

based on Design $2E$ in which six subjects completed treatments from the first three periods only and then dropped out.

1	2	3	4	1	2	3	4	1	2	3	4
2	3	4	1	2	3	4	1	2	3	4	1
4	1	2	3	4	1	2	3	4	1	2	3
$*$	4	$*$	2	$*$	4	$*$	2	$*$	4	$*$	2

Design 2*G*. Design $2E$ with six missing observations in period four

Design $2G$ is disconnected and is seriously compromised. Although 42 of the 48 possible measurements are available, only one of the six elementary treatment direct effects is estimable and only two of the six elementary treatment carryover effects are estimable

On the other hand, this situation does not arise with Design $2F$. It is an assumption of the paper that all subjects persevere with the experiments for periods one and two. If drop-out occurs in periods three and/or four the experiment will be less efficient, however both sets of treatment sums of squares can be formed and all elementary contrasts between each set of treatment contrasts can be estimated, regardless of eventual drop-out behaviour. Design $2F$ has the perpetually connected property which is clearly a useful property of the design. It is evident that Design $2F$ is the more sensible choice of planned UBRMD since it does not have the risk associated with Design $2E$.

3 Perpetually Connected UBRMDs

3.1 Criterion for identifying perpetually connected designs

The crossover experiment is designed to utilize mt subjects to compare t treatments over t different periods, where $m \geq 2$, and the planned design is a UBRMD that consists of m Latin Squares of order $t \times t$. Such a design is universally optimal among a wide class of competing designs (c.f. Jones and Kenward[1], §4.3 and references therein). Any two planned designs with the same values of m and t will be equally efficient. However the two eventual designs which remain after a given pattern of subject drop-out may have markedly different properties, particularly with regard to connectivity, as illustrated for particular values $t = 4$ and $m = 3$ by the planned designs $2E$ and $2F$ in §2.2.

For definiteness the concept of connectivity employed throughout the paper is given formally.

Definition 1. A design is said to be *connected* if the corresponding design matrix has maximal rank.

When a crossover design is connected then all elementary contrasts in the treatment direct effects are estimable and, furthermore, all elementary contrasts in the treatment carryover effects are estimable; see

Appendix, (p14-15). If subject drop-out occurs during the course of the experiment it is clearly desirable to ensure that the ensuing eventual design is connected. Since the number of eventual designs that could result from possible subject drop-out is considerable, even for modest values of t and m, it is obviously useful if a procedure can be found which bypasses the need to check each eventual design individually.

It is assumed that no subject drop-out occurs in the first two periods. For any given planned design the assessment of risk due to drop-out should take account of the number of possible eventual designs, and the structure of these designs, which are connected according to Definition 1. Evidently, the preferred position is that where none of the potential eventual designs are disconnected, which is the motivation behind the formal definition of a perpetually connected design given here.

Definition 2. A planned design is said to be *perpetually connected* if all subjects complete period one and period two, and the eventual design is connected irrespective of drop-out behavior thereafter.

Perpetually connected designs do exist for some values of t and m; an example for $t = 4$ and $m = 3$ is given by the three MOLS of Design $2F$. In general, to assist in identifying perpetually connected designs it is useful to note the eventual design that is realized after the most extreme form of drop-out that can arise, subject to the assumption of no drop-out in period 1 or period 2.

Definition 3. An eventual design is said to be a *minimal design* if it consists solely of the first two rows of the planned design, i.e. all mt subjects complete the first two periods and then drop out.

For a given planned UBRMD, the significant role played by the minimal design in checking for perpetual connectivity follows from the assertion that the planned design is perpetually connected if and only if the corresponding minimal design is connected. This assertion is stated as a formal theorem which is proved in the Appendix. It suggests the following procedure which can be used for checking the given planned design.

Criterion To check for perpetual connectivity of a planned design it is required to ascertain the rank of the design matrix for the corresponding minimal design. If the design matrix for the minimal design has maximal rank then the planned design is perpetually connected; otherwise the planned design is not perpetually connected.

In the remainder of this section, several examples of perpetually connected crossover designs are given for various values of t and m such that $mt^2 \leq 100$, by using this Criterion.

3.2 Designs for three treatments

A common situation with clinical trials is to compare three treatments, labelled $1, 2,$ and 3, using $3m$ subjects over three periods. No balanced 3×3 Latin square exists, indeed Newcombe [23] has pointed out

that no UBRMD exists for this experiment when m, the number of Latin squares in the design, is an odd integer. Attention is confined to designs where m is even: such a design is perpetually connected if no type of drop-out behaviour in the final period causes a disconnected design, provided that all subjects complete the first two periods of the study.

Design $3A$ is a set of two mutually orthogonal Latin squares which gives a UBRMD when $m = 2$.

$$\begin{array}{cccccc} 1 & 2 & 3 & 1 & 2 & 3 \\ 2 & 3 & 1 & 3 & 1 & 2 \\ 3 & 1 & 2 & 2 & 3 & 1 \end{array}$$

Design 3*A*. Two MOLs of order 3×3

Design $3A$ is also obtained by the construction method of Williams[22]. Although it would usually be considered too small for practical use it is interesting to note that Design $3A$ is perpetually connected. Furthermore, it is straightforward to show that when m is any one of the even integers $m = 2, 4, 6, 8, 10$ or 12, a perpetually connected design exists for three treatments using $3m$ subjects over three periods simply by taking $\frac{1}{2}m$ replicates of Design $3A$.

3.3 Designs for four treatments

The discussion of §2 shows some difficulties in selecting a UBRMD for an experiment in which four treatments $1, 2, 3$ and 4 are compared using $4m$ subjects over four periods. When $m = 2$ there appears to be no perpetually connected UBRMD. When $m = 3$ a perpetually connected UBRMD does exist and is given by Design $2F$; this design has the interesting property that it is balanced for second-order and third-order carryover effects as well as first order carryover effects.

When $m = 4$, a perpetually connected UBRMD for comparing four treatments using sixteen subjects over four periods is obtained by combining Design $2F$ with the Williams Latin square common to Design $2A$ and Design $2B$. Another perpetually connected UBRMD with the same dimensions is obtained by combining Design $2F$ with the other Williams Latin square that forms part of Design $2B$.

When $m = 5$, a perpetually connected UBRMD for comparing four treatments using twenty subjects over four periods is obtained by combining Design $2F$ with Design $2B$.

When $m = 6$, a perpetually connected UBRMD for comparing four treatments using twenty four subjects over four periods is obtained by combining Design $2F$ with itself, i.e. taking two replicates of the specified set of mutually orthogonal Latin squares.

3.4 Designs for five treatments

This section considers designs to cover experiments in which five treatments labelled $1, 2, 3, 4$ and 5 are compared using $5m$ subjects over five periods. Attention is confined to the cases $m = 2, 3$ and 4. A design is perpetually connected if all subjects complete the first two periods of the study and no type of drop-out behaviour in the final three periods can cause the eventual design to be treatment disconnected, with regard to direct effects or carryover effects or both.

The following three UBRMDs, designated Designs $3B$, $3C$ and $3D$, are perpetually connected. Design $3B$ is the familiar balanced design in two squares due to Williams[22]. Design $3C$ is the combination of three Latin squares due to Newcombe[23]. Design $3D$ consists of four Latin squares; this design is obtained from the construction argument of Bate *et al.*[19] such that half of it is Design $3B$ and the other half is isomorphic to Design $3B$.

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

Design 3*B*. Two 5×5 Latin squares

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

Design 3*C*. Three 5×5 Latin squares

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

Design 3*D*. Four 5×5 Latin squares

3.5 Designs for six and seven treatments

Suppose that six treatments labelled $1, 2, 3, 4, 5$ and 6 are to be compared using $6m$ subjects over six periods, where $m = 2$ or $m = 3$. When twelve subjects are available the Design $3E$ displayed here is perpetually connected.

Design $3E$ consists of two distinct Williams Latin squares which are obtained, for example, by the Balanced Cyclic Square algorithm of Bate *et al.*[19]. A design that consists of two replicates of just one of

these balanced Latin squares is not perpetually connected, i.e. the two squares need to be distinct, as is the case with Design $3E$.

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

Design 3*E*. Two 6×6 Williams Latin squares

Furthermore, by adding either one of these balanced Latin squares to Design $3E$ a design consisting of three Latin Squares is obtained, two of which are the same. This is a combined design for eighteen subjects to compare six treatments over six periods and is a UBRMD which is perpetually connected.

A perpetually connected UBRMD for comparing seven treatments labelled $1, 2, 3, 4, 5, 6$ and 7 using fourteen subjects over seven periods is the familiar Williams design displayed as Design $3F$.

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

Design 3*F*. Williams design for two 7×7 Latin squares

4 Discussion

It is shown in this paper that perpetually connected UBRMDs exist for many of the experimental situations that may be encountered which require up to one hundred measurements. In each of these cases, at least one perpetually connected UBRMD is specified. The examples which are given involve treatment ranges from $t = 3$ to $t = 7$, where the number of periods is t and the number of subjects is mt, where $m \geq 2$. These results imply that practitioners who wish to use any of the suggested UBRMDs can be confident that they will be able to process their results as usual, provided that all subjects complete the first two periods of study and regardless of drop-out in subsequent periods. The concept of forming

conditions on designs which permit any pattern of missing values mimics the situation with incomplete block designs as described, for example, by Godolphin and Godolphin [24 25].

The requirements that there should be at least $2t$ subjects in the planned design and that all subjects should complete the first two periods of study appear to be essential conditions, on theoretical grounds. For any value of t, there is only one degree of freedom available for the worst possible drop-out case, which is that the eventual design is the minimal design with only $2t$ observations recorded. In general, when mt subjects are employed then $1 + (m - 2)t$ degrees of freedom are available for estimating the residual error variance in the minimal design. Attempting to modify these requirements, for example by including the possibility of drop-out in period two, may require a restriction condition on subject drop-out behaviour in the third and possibly higher periods. This appears to be less desirable than the basic requirement specified above. In practice, it seems highly likely that every subject will complete the first period, but for some trials there is a doubt that all subjects will complete the second. Despite these misgivings, the perpetually connected property of a crossover design appears to be a useful concept when considering robustness of UBRMDs.

The assumption that data are MAR is not realistic for all cases, as subject drop-out could be closely related to treatment success or failure in trials utilising a crossover design. However, without exploring the missingness mechanism directly, it is a plausible and general assumption which permits a likelihood based analysis, thus allowing the issues of connectedness to be investigated. Whilst this assumption is not suitable for all possibilities, highlighting this as a limitation of our approach, this paper presents the building blocks of a topic area which is yet to be extensively investigated.

The examples of perpetually connected designs in the paper are limited to small trials with less than 100 observations. Experiments that utilise crossover designs with multiple treatments are often early phase trials, animal experiments or pharmacokinetic studies, which historically recruit few subjects. The majority of the designs presented consider cases for 12 or more subjects, which follows guidelines from the Committee for Medicinal Products for Human Use [26] and guidance from the Food and Drug Administration [27]. It is hoped that these results assist researchers to plan experiments under conditions which are robust to the consequences of subject drop-out.

Acknowledgements

The authors would like to thank an anonymous referee for helpful comments and some additional references.

Declaration of conflicting interests

The authors declare no potential conflicts of interests with respect to the research, authorship and/or publication of this article. Both authors contributed to the study equally.

Funding

P J Godolphin was funded for this summary of independent research by the National Institute for Health Research (NIHR)'s Research Methods Fellowship and Internship Programme (RMFI-2014-05-13). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, NHS or the Department of Health.

References

1. Jones B, Kenward GM. *Design and Analysis of Cross-Over Trials*, 3rd edn. Monographs on Statistics and Applied Probability 138. CRC Press. 2015.
2. Wellek S, Blettner M. *On the Proper Use of the Crossover Design in Clinical Trials: Part 18 of a Series of Evaluation of Scientific Publications*, 2nd edn. Deutsches Arzteblatt International. Addison-Wesley. 2012;109(15):276-281.
3. Senn S, D'Angelo G, Potvin D. *Carry-over in cross-over trials in bioequivalence: theoretical concerns and empirical evidence*, Pharmaceutical Statistics. 2004. 3(2):133-142.
4. Grattan T, Hickman R, Darby-Dowman A. *A five way crossover human volunteer study to compare the pharmacokinetics of paracetamol following oral administration of two commercially available paracetamol tables and three development tablets containing paracetamol in combination with sodium bicarbonate or calcium carbonate*, European Journal of Pharmaceutics and Biopharmaceutics. 2000. 49:225-229.
5. National Research Council. *The Prevention and Treatment of Missing Data in Clinical Trials*. Panel on Handling Missing Data in Clinical Trials. Committee of National Statistics, Division of Behavioral and Social Sciences and Education. The National Academies Press: Washington, DC; 2010.
6. Rubin DB. *Inference and Missing Data*. Biometrika. 1976. 63(3):581-592.
7. Molenberghs G, Kenward GM. *Missing Data in Clinical Studies*, Chichester, UK: Wiley. 2007.
8. Rosenkranz GK. *Analysis of cross-over studies with missing data*, Statistical Methods in Medical Research. 2015. 24(4):420-433.
9. Matthews JNS, Henderson R. *Two-period, two-treatment crossover designs subject to non-ignorable missing data*, Biostatistics. 2013. 14(4):626-638.
10. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*, The Cochrane Collaboration, 2011.
11. Low JL, Lewis SM, Prescott P. *Assessing robustness of crossover designs to subjects dropping out*, Statistics and Computing. 1999. 9(3):219-227.
12. Shih WJ. *Problems in dealing with missing data and informative censoring in clinical trials*, Current Controlled Trials in Cardiovascular Medicine. 2002. 3(1):4.

13. Godolphin JD. *Simple pilot procedures for the avoidance of disconnected experimental designs*, Applied Statistics. 2004. 53:133-147.
14. Mills EJ, Chan A, Wu P et al. *Design, analysis and presentation of crossover trials*, Trials. 2009. 10:1-6.
15. Cochran WG, Autrey KM, Cannon CY. *A double change-over design for dairy cattle feeding experiments*, Jour. Dairy Sci. 1941. 24:937-951.
16. Hedayat A, Afsarinejad K. *Repeated Measurements Designs 2*, Annals of Statistics. 1978. 6(3):619-628.
17. Cheng CS, Wu J. *Balanced repeated measurement designs*, Annals of Statistics. 1980. 8:1272-1283. (Corrigendum. Annuals of Statistics (1983), 11:349)
18. Hedayat AS, Yang M. *Universal optimality of balanced uniform crossover designs*, Annals of Statistics. 2003. 31:978-983.
19. Bate ST, Godolphin EJ, Godolphin JD. *Choosing cross-over designs when few subjects are available*, Computational Statistics and Data Analysis. 2008. 52:1572-1586.
20. Majumdar D, Dean AM, Lewis SM. *Uniformly balanced repeated measurements designs in the presence of subject dropout*, Statistica Sinica. 2008. 18:235-253.
21. Zhou S, Majumdar D. *On uniformly balanced crossover designs efficient under subject dropout*, Journal of Statistical Theory and Practice. 2012. 6(1):178-189.
22. Williams EJ. *Experimental designs balanced for the estimation of residual effects of treatments*, Australian Journal of Scientific Research. 1949. 2(2):149-168.
23. Newcombe RG. *Sequentially balanced three-squares cross-over designs*, Statistics in Medicine. 1996. 15(20):2143-2147.
24. Godolphin JD and Godolphin EJ. *The robustness of resolvable block designs against the loss of whole blocks or replicates*, Journal of Statistical Planning and Inference. 2015. 163:34-42.
25. Godolphin JD and Godolphin EJ. *The use of treatment concurrencies to assess robustness of binary block designs against the loss of whole blocks*, Australian and New Zealand Journal of of Statistics. 2015. 57:225-239.
26. CHMP. *Guideline on the investigation of bioequivalence*, European Medicines Agency. 2010.
27. *Guideline for industry: statistical approaches to establishing bioequivalence*, Food and Drug Administration. 2001.
28. Srivastava JN and Anderson DA. *Some basic properties of multidimensional partially balanced designs*, Annals of Mathematical Statistics. 1970. 41:1438-1445.
29. Godolphin JD. *On the connectivity problem for m-way designs*, Journal of Statistical Theory and Practice. 2013. 7:732-744.
30. Harville DA. *Matrix algebra from a Statisticians's perspective*, Springer-Verlag, New York. 1997.

Appendix

The notation of Bate *et al.*[19] is largely adopted here. The design selected for the experiment, the planned design, is a uniform balanced repeated measures design which is used throughout the experiment if no drop-out occurs. In this case, the $n \times 1$ observation vector Y, where $n = mt^2$, is assumed to follow the additive fixed-effects model, described in matrix form as

$$E[Y] = \mu 1_n + X_1\tau + X_2\rho + X_3\alpha + X_4\beta, \tag{1}$$

where μ is a parameter, $\tau, \rho, \alpha, \beta$ are vectors of treatment direct, treatment carryover, row (period) and column (subject) effects of sizes $t \times 1$, $t \times 1$, $t \times 1$ and $mt \times 1$ respectively. Here X_1, X_2, X_3 and X_4 are components of the design matrix for the planned design and 1_n is the $n \times 1$ vector, all of whose elements are unity. The design matrix $X = [1_n\ X_1\ X_2\ X_3\ X_4]$ has dimension $n \times \{(m+3)t + 1\}$.

In practice some observations may be lost from the planned experiment through subject drop-out. In this case there are fewer than mt^2 observations, i.e. the observation vector Y has size n where $n < mt^2$. Under the assumptions of the paper $E[Y]$ has the same form as (1), i.e. in this eventual design all components on the right side of (1) are retained in the model although the size of the period parameter vector α may be reduced to $t - q$, where q denotes the number of the final periods, if any, where no measurements are made. Of course $q > 0$ only if all subjects drop out from the study after $t - q$ periods rather than the planned quota of t periods. Since no subject drops out in the second period, by assumption, then the most extreme case of subject drop-out occurs when $q = t - 2$, i.e. all subjects receive two treatments only, then leave the study. This extreme case, termed the minimal design, is specified by Definition 3 in §3.1. It is noted that this same term is used by Majumdar *et al*[20] to describe a milder form of the definition, where all subjects are assumed to complete $t - q$ periods and then drop out, but these authors assume that q is typically one or two. However for perpetual connectivity it is necessary to consider the extreme form given here. Clearly every planned design has a minimal design associated with it.

Connectivity Criterion

It is a basic requirement that all elementary contrasts between direct treatment effects are estimable and this is sometimes the criterion of connectivity that is considered by researchers. For example, Majumdar *et al*[20] consider conditions for $t - 1$ eigenvalues of the information matrix for direct effects to be strictly positive; to avoid ambiguity we say that direct effects are *singly connected* when these conditions are satisfied. The general specification given in Definition 1 is that the planned design is connected if the rank of the design matrix is maximal, e.g. $\text{rank}(X) = (m+3)t - 3$. This definition corresponds to the criterion of complete connectivity due to Srivastava and Anderson[28] which is considered in some detail

by Godolphin[29]. Although these authors confine their results to the multi-way classification design, the approach extends easily to the change-over planned design. Indeed for this planned design the kernel of the design matrix is the column space of the $\{(m+3)t+1\} \times 4$ matrix Π_{plan}, given by

$$\Pi_{\text{plan}} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ 1_t & 0_t & 0_t & 0_t \\ 0_t & 1_t & 0_t & 0_t \\ 0_t & \phi_t & 1_t & 0_t \\ 0_{mt} & 0_{mt} & 0_{mt} & 1_{mt} \end{bmatrix} \tag{2}$$

where ϕ_t is the $t \times 1$ vector with unity in the first position and zero in the other $t-1$ positions, and 0_t is the $t \times 1$ null vector. Any vector x orthogonal to Π_{plan}, i.e. any x such that $x'\Pi_{\text{plan}} = 0$, is contained in the estimable space, i.e. the parametric combination $[\mu\ \tau'\ \rho'\ \alpha'\ \beta']\,x$ is estimable. In particular, if $x = [0\ \lambda'\ 0'_{(m+2)t}]'$ then $x'\Pi_{\text{plan}} = 0$ if and only if $\lambda' 1_t = 0$, i.e. any linear contrast $\lambda'\tau$ in treatment direct effects is estimable; similarly, a linear contrast, $\lambda'\rho$, in treatment carryover effects is estimable when the the rank of the design matrix is maximal; confer Theorems 1,3 and Corollary 1 of Godolphin[29].

Effect of Subject Drop-out

Let $\mathcal{R}_{\text{plan}} = \mathcal{C}(X')$ be the column space of X', i.e. the row space for the planned design which is the space spanned by the mt^2 vectors consisting of the rows of the design matrix. Thus $\mathcal{R}_{\text{plan}}$ is orthogonal to $\mathcal{C}(\Pi_{\text{plan}})$. Each row of X corresponds to an observation from the planned design, hence the effect of a single subject dropping out from the final q periods of the experiment is, effectively, to generate an alternative space spanned by vectors which are the rows of X excepting those q rows that correspond to the missing observations. This alternative space is a subspace of $\mathcal{R}_{\text{plan}}$. In order to investigate the effect of an arbitrary number of subjects dropping out of the experiment in the third period or later, let $\mathcal{R}$ denote a subspace of $\mathcal{R}_{\text{plan}}$ spanned by vectors which are rows of X and include all $2mt$ rows that correspond to the first two periods of the experiment. Also let $\mathcal{R}_*$ denote the subspace which is formed after the removal of *all* rows of X except the $2mt$ vectors corresponding to the first two periods; and let X_* be the $2mt \times \{(m+3)t+1\}$ matrix with these rows. Then $\mathcal{R}_* = \mathcal{C}(X'_*)$ and $\mathcal{R}_* \subseteq \mathcal{R} \subseteq \mathcal{R}_{\text{plan}}$.

Suppose further that the minimal design is connected. Let $X_{\min}$ be the design matrix for the minimal design which has maximal rank $(m+2)t-1$ by assumption. This is smaller than the rank of the planned design matrix since the minimal design has $t-2$ fewer parameters, i.e. the row parameters $\alpha_3, \ldots, \alpha_t$ are absent as they represent missing periods that are not defined for the minimal design. It follows that the kernel of $X_{\min}$ is the column space of the $\{(m+2)t+3\} \times 4$ matrix $\Pi_{\min}$, given by

$$\Pi_{\min} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ 1_t & 0_t & 0_t & 0_t \\ 0_t & 1_t & 0_t & 0_t \\ 0_2 & \phi_2 & 1_2 & 0_2 \\ 0_{mt} & 0_{mt} & 0_{mt} & 1_{mt} \end{bmatrix} \tag{3}$$

where ϕ_2 is the 2×1 matrix with unity in the first position and zero in the other position. However, $\alpha_3, \ldots, \alpha_t$ are necessarily defined for matrix X_* so the columns of X_* corresponding to these parameters will be null vectors 0_{2mt}, i.e. $X_* = [X_{\min}\, 0^*_{2mt}]$, where 0^*_{2mt} denotes the $2mt \times (t-2)$ null matrix. Hence the kernel of X_* is the column space of a matrix with $(m+3)t+1$ rows, $\Pi_{**} = [\Pi_*, \xi]$ say, which contains a component matrix Π_*, where Π_* is given by:

$$\Pi_* = \begin{bmatrix} -1 & -1 & -1 & -1 & 0'_{t-2} \\ 1_t & 0_t & 0_t & 0_t & 0^*_t \\ 0_t & 1_t & 0_t & 0_t & 0^*_t \\ 0_2 & \phi_2 & 1_2 & 0_2 & 0^*_2 \\ 0_{t-2} & 0_{t-2} & 0_{t-2} & 0_{t-2} & I_{t-2} \\ 0_{mt} & 0_{mt} & 0_{mt} & 1_{mt} & 0^*_{mt} \end{bmatrix}, \tag{4}$$

such that $0^*_t, 0^*_2, 0^*_{mt}$ denote $t \times (t-2)$, $2 \times (t-2)$, $mt \times (t-2)$ null matrices. Here ξ is a matrix component of Π_{**} such that Π_{**} has rank $\{(m+3)t+1\} - \{(m+2)t-1\} = t+2$. But the final $t-2$ columns of Π_* is a matrix component which is orthogonal to the remaining matrix component consisting of the first four columns, i.e. these two components are essentially disjoint so that $\text{rank}(\Pi_*) = 4 + (t-2) = t+2$ (Harville[30], Theorem 17.2.4). The columns of Π_{**} are linearly independent so the ξ vector is absent and $\Pi_{**} = \Pi_*$. Therefore the kernel of X_* is $\mathcal{C}(\Pi_*)$.

The following theorem relates to the perpetual connectivity property of a design.

Theorem *A planned design is perpetually connected if and only if the corresponding minimal design is connected.*

To prove the theorem, let q be an integer satisfying $0 \le q < t-2$. The design matrix for the planned design can be expressed as $X = [X'_{\dagger 1}\, X'_{\dagger 2}]'$ where $X_{\dagger 1}$ contains n rows, including all rows for the first two periods $(2mt < n \le mt^2)$, whilst the corresponding rows for all replicates of q of the remaining periods are assumed to be contained in $X_{\dagger 2}$. We put $\mathcal{R} = \mathcal{C}(X'_{\dagger 1})$. Possibly after some redistribution of the columns of $X_{\dagger 1}$ we can write

$$X_{\dagger 1} = \begin{bmatrix} X_{\min} & 0^*_{2t} \\ X_{\dagger 11} & X_{\dagger 12} \end{bmatrix}, \tag{5}$$

where 0^*_{2t} is a $2t \times (t-2-q)$ null matrix.

Assume that the minimal design is connected so that $X_{\min}$ has maximal rank $(m+2)t-1$. The matrix component $X_{\dagger 12}$ is that part of the period component X_3 in model (1) which relates to the $t-2-q$ periods for the eventual design. The matrix $X_{\dagger 12}$ has the dimensions $(n-2t) \times (t-2-q)$, with value unity occurring once in each row and zeros occurring elsewhere, and all $t-2-q$ columns have at least one non-zero entry. Hence rank$(X_{\dagger 12}) = t-2-q$, i.e. $X_{\dagger 12}$ has full row rank. Using Theorem 8.5.3 of Harville[30] it follows from equation (5) that $X_{\dagger 1}$ has rank equal to

$$\text{rank}(X_{\min}) + \text{rank}(X_{\dagger 12}) = \big((m+2)t-1\big) + (t-2-q) = (m+3)t-3-q.$$

This is the maximal rank of $X_{\dagger 1}$. Therefore the eventual design with design matrix $X_{\dagger 1}$ is connected. The partitioning of X into components $X_{\dagger 1}$ and $X_{\dagger 2}$ is arbitrary, subject to the inclusion of $X_{\min}$ in $X_{\dagger 1}$, which shows that no eventual design can be disconnected due to subject drop-out if the associated minimal design is connected. It follows that the planned design is perpetually connected.

Conversely, suppose the minimal design is disconnected so that $X_{\min}$ has rank less than $(m+2)t-1$. We assert that there are eventual designs which occur as a result of subject drop-out which are not the minimal design but they are also disconnected. Indeed each of the mt eventual designs, which occur after $mt-1$ subjects drop out after two periods and the other subject drops out after three periods, cannot have maximal rank if $X_{\min}$ does not have maximal rank. This is because each of the corresponding row spaces is spanned by one more vector than the row space for the minimal design, but each eventual design also has an additional parameter which would account for any increase in rank.

As a consequence of this theorem it is possible to choose from among the UBRMD designs and select a perpetually connected design, when one exists, by exploring the connectivity of the associated minimal design. This strategy has been applied throughout §3 in the main body of the paper.

Designs $2A$ and $2B$

To illustrate the remarks in §2.1, Design $2C$ and Design $2D$ are examined further. The two planned UBRMDs Design $2A$ and Design $2B$ are connected but the associated minimal designs are found to be disconnected, The rank of the minimal design associated with Design $2A$ is rank$(X_{\min}) = 12$, rather than the full rank of $(m+2)t-1 = 15$, and the rank of the minimal design associated with Design $2B$ is rank$(X_{\min}) = 14$. Consequently for each planned design there will be some eventual designs that are also disconnected. It turns out that disconnected eventual designs occur even when several subjects do not drop out but complete their full sequence of treatments. Design $2A$ and Design $2B$ incurred drop-out through the loss of just four subjects, resulting in Design $2C$ and Design $2D$ respectively. The kernel of

the row space of the design matrix can be represented in each case by

$$\Pi = \begin{bmatrix} -1 & -1 & -1 & -1 & 0 \\ 1_4 & 0_4 & 0_4 & 0_4 & \xi_\tau \\ 0_4 & 1_4 & 0_4 & 0_4 & \xi_\rho \\ 0_4 & \phi_4 & 1_4 & 0_4 & \xi_\alpha \\ 0_8 & 0_8 & 0_8 & 1_8 & \xi_\beta \end{bmatrix}, \tag{6}$$

where $\xi_\tau, \xi_\rho, \xi_\alpha$ are vectors with four elements and ξ_β is a vector with eight elements, which have different formulations for Design $2C$ and Design $2D$.

Consider the estimation of the treatment direct effects for Design $2C$. The vector ξ_τ is given by $\xi_\tau = [2\ 1\ -1\ 0]'$, therefore it is evident that there is no 4×1 vector x_τ such that both $x'_\tau \xi_\tau = 0$ and $x'_\tau 1_4 = 0$. Hence no pairwise contrasts in the direct effects are estimable. A similar conclusion is reached for the estimation of the treatment carryover effects for Design $2C$. The estimation of the treatment direct effects for Design $2D$ is similar. The vector ξ_τ is given by $\xi_\tau = [-1\ 0\ -1\ 0]'$, therefore there are two pairwise contrasts in the direct effects which are estimable, namely $\tau_1 - \tau_3$ and $\tau_2 - \tau_4$, but the other four pairwise contrasts in the direct effects are not estimable. There are two pairwise contrasts in the carryover effects which are estimable, given by $\rho_1 - \rho_3$ and $\rho_2 - \rho_4$. For Design $2C$ and Design $2D$ the null hypotheses $H_0^{(\tau)} : \tau_1 = \tau_2 = \tau_3 = \tau_4$ and $H_0^{(\rho)} : \rho = \rho_2 = \rho_3 = \rho_4$ are not testable.

Design $2G$

Design $2G$ is disconnected as the design matrix has rank 20, rather than the maximal $(m+3)t - 3 = 21$, and the kernel of the row space of the eventual design matrix is $\mathcal{C}(\Pi)$, where Π has the form

$$\Pi = \begin{bmatrix} -1 & -1 & -1 & -1 & 0 \\ 1_4 & 0_4 & 0_4 & 0_4 & \xi_\tau \\ 0_4 & 1_4 & 0_4 & 0_4 & \xi_\rho \\ 0_4 & \phi_4 & 1_4 & 0_4 & \xi_\alpha \\ 0_{12} & 0_{12} & 0_{12} & 1_{12} & \xi_\beta \end{bmatrix}, \tag{7}$$

where $\xi_\tau, \xi_\rho, \xi_\alpha$ are vectors with four elements and ξ_β is a vector with twelve elements. The vector ξ_τ is given by $\xi_\tau = [2\ 1\ -1\ 0]'$, therefore there is just one estimable pairwise contrast in the treatment direct effects, namely $\tau_2 - \tau_3$. Also $\xi_\rho = [1\ 2\ -1\ -2]'$ so $\rho_1 - \rho_3$ and $\rho_2 - \rho_4$ are the only estimable pairwise contrasts in the treatment carryover effects. The hypotheses $H_0^{(\tau)} : \tau_1 = \tau_2 = \tau_3 = \tau_4$ and $H_0^{(\rho)} : \rho_1 = \rho_2 = \rho_3 = \rho_4$ are not testable.