11-2016

Evaluating the Efficiency of Treatment Comparison in Crossover Design by Allocating Subjects Based On Ranked Auxiliary Variable

Yisong Huang
Georgia Southern University, yh00049@georgiasouthern.edu

Hani Samawi
Georgia Southern University, hsamawi@georgiasouthern.edu

Robert Vogel
Georgia Southern University, rvogel@georgiasouthern.edu

Jingjing Yin
Georgia Southern University, jyin@georgiasouthern.edu

Worlanyo E. Gato
Georgia Southern University, wgato@georgiasouthern.edu

See next page for additional authors

Communications for Statistical Applications and Methods
2016, Vol. 23, No. 6, 543–553

http://dx.doi.org/10.5351/CSAM.2016.23.6.543
Print ISSN 2287-7843 / Online ISSN 2383-4757

Evaluating the efficiency of treatment comparison in crossover design by allocating subjects based on ranked auxiliary variable

Yisong Huang[a], Hani M. Samawi[1,a], Robert Vogel[a], Jingjing Yin[a], Worlanyo Eric Gato[b], Daniel F. Linder[c]

[a]Department of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, USA; [b]Department of Chemistry, Georgia Southern University, USA; [c]Department of Biostatistics and Epidemiology, Medical College of Georgia, Augusta University, USA

Abstract

The validity of statistical inference depends on proper randomization methods. However, even with proper randomization, we can have imbalanced with respect to important characteristics. In this paper, we introduce a method based on ranked auxiliary variables for treatment allocation in crossover designs using Latin squares models. We evaluate the improvement of the efficiency in treatment comparisons using the proposed method. Our simulation study reveals that our proposed method provides a more powerful test compared to simple randomization with the same sample size. The proposed method is illustrated by conducting an experiment to compare two different concentrations of titanium dioxide nanofiber (TDNF) on rats for the purpose of comparing weight gain.

Keywords: ranked set sampling, ranked auxiliary covariate, experimental design, treatment allocation method, crossover design, Latin square design, TDNF, weight gain

1. Introduction

The goal of designing a statistical experiment is to obtain valid conclusions. The results and conclusions drawn from a statistical experiment depend on the quality of the data collected from the experiment. A well-designed experiment involves randomization, a detailed description of the experiment and its objectives which include the specification of the treatments and controls.

In any experiment, variability arising from a nuisance factor can affect the results. Levels of effects from a nuisance factor may be changing over the duration of the experiment. Under such circumstances, randomization can be introduced to reduce the bias. In other cases, the nuisance factor may be known but is uncontrollable. When the source of variability is predictable and controllable, a blocking design can be introduced to systematically eliminate the effect of the nuisance factor on treatment comparisons. The blocking principle is utilized in several types of designs. One type of design is the Latin square. A Latin square is used to eliminate two nuisance sources of variability.

In this paper, we introduce a method based on the baseline ranked auxiliary variable for treatment allocation in crossover studies which can be analyzed as a set of n replicated Latin squares. When we

[1] Corresponding author: Department of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA, 30458, USA. E-mail: hsamawi@georgiasouthern.edu

Published 30 November 2016 / journal homepage: http://csam.or.kr

Table 1: Crossover design example

Latin square replicates	I		II		III		IV	
Subject	1	2	3	4	5	6	7	8
Period 1	A	B	A	B	B	A	A	B
Period 2	B	A	B	A	A	B	B	A

consider time as one of the blocking factors in a Latin square design, it provides the foundation for a p-period, p-treatment crossover design.

A crossover design is a study in which subjects receive sequences of different treatments. The primary objective of crossover studies is to compare the outcome of interests between p treatment groups on the same subject (Senn, 1993). The main advantage of a crossover design is efficiency with the ability to study individual reaction to treatment. A typical layout of a crossover design using a Latin square is illustrated in Table 1. Table 1 shows a crossover design of four Latin squares, each with 2-periods, 2-treatments, and 2-sequences. A total of 8 subjects are enrolled in the study as shown in Table 1. In the first period, half of the subjects are randomly assigned to treatment A, and the other half to treatment B. At the end of the first period, the response is measured. If needed, a period of time, called wash-out time, is allowed to pass in which any effects of the treatment is eliminated. Then the experimenter switches the subjects who received treatment A to treatment B, and the subjects who received the treatment B to treatment A.

1.1. Treatment allocation based on ranked auxiliary variable

The idea of treatment allocation based on a ranked auxiliary variable is inspired by ranked set sampling. Ranked set sampling (RSS) was first introduced by McIntyre (1952). The principle of RSS is to select a sample systematically from a population. It was originally introduced as an efficient alternative to simple random sampling. The sample mean of RSS is an unbiased estimator of the population mean, and variance based on RSS is smaller than the variance on simple random sampling (SRS) with the same effective sample size. The original method for ranking units is by visual judgment ranking (McIntyre, 1952). Lynne Stokes (1977) proposed concomitant-based ranking which yields a smaller variance than that from SRS. Concomitant-based ranking is ranking based on the covariates rather than the characteristics of interest.

Different ranking schemes have been devised when there are more than one concomitant variable. Samawi *et al.* (1996) proposed extreme ranked set sample (ERSS). This ERSS has important applications in genetics for quantitative trait loci mapping (Chen, 2007). Bivariate RSS was introduced by Al-Saleh and Zheng (2002). Linder *et al.* (2015) proposed stratified bivariate RSS for regression estimators. They showed this estimator is more efficient to stratified bivariate SRS under same effective sample size. Samawi and Al-Sagheer (2001) discuss the relationship between extreme ranked set samples and median ranked set sample for distribution estimation. Since the sample mean from RSS is unbiased estimator of population mean, it is not a problem to draw statistical inference based on the sample mean of RSS. Based on the studies using RSS, we can conclude that RSS can be used in experiments to increase the efficiency of treatment comparison.

The idea proposed in this paper is to evaluate the improvement of the efficiency of treatment comparison by incorporating the subjects selection based on ranked auxiliary variable. Subjects will be allocated to treatment by ranking on an available auxiliary variable in the crossover studies using Latin square design. Section 2 presents the proposed statistical analysis of treatment allocation based on ranked auxiliary variables in crossover design. Section 3 is a simulation study. In Section 4, the proposed method is illustrated by conducting weight gain experiment to compare two different

concentrations of TDNF on rats. Final remarks and discussion are presented in Section 5. In the next section, we will discuss the details of statistical analysis in crossover designs based on incorporating subjects selection based on ranked auxiliary variable.

2. Proposed method

The purpose is to investigate the performance of the proposed treatment allocation methods based on auxiliary variable in crossover studies. The primary aim is to improve the efficiency of treatment comparisons by allocating subjects using ranked auxiliary variable which is correlated with the outcome of interest. The treatment effect is estimated by the fixed effect model and random effect model respectively. The proposed method is described in the next subsection.

2.1. Treatment allocation based on concomitant-based RSS in crossover design

Assume a situation where the outcome of interest (Y) in a crossover study is known be correlated with a baseline auxiliary variable (X), in the biologically sense or from the literature and easy to obtain and rank. In addition, we have the information of X before the study started. Suppose we perform a p-treatment, p-period, p-sequence crossover design using n replicates of Latin squares. Therefore, we need np out of np^2 subjects selected randomly from a population. Then randomize n replicates of Latin squares to each set of p subjects.

Therefore, the protocol of treatment allocation based on a ranked auxiliary variable to crossover studies using Latin squares can be described as follow:

1. Select p sets, each of p subjects randomly from an available pool of subjects.

2. In each set, rank the p subjects based on auxiliary baseline variable (X) which is highly correlated with outcome and easy to rank.

3. From the first set of p subjects, the subjects ranked lowest, in (X), is chosen for actual quantification. From the second set of p subjects, the subject ranked second lowest in (X), is measured. The process is continued until the subject ranked highest is measured from the p^{th} set of p units. At this point, p subjects are selected.

4. Randomly choose a Latin square design to these p subjects.

5. This cycle is repeated n times until the desired sample size (np).

For example, $p = 3$, the process can be illustrated as Table 2.

After the crossover study, we collect the measures of outcome of interest (Y). Two scenarios are considered. If only np^2 subjects are available to the experiment, we consider the subjects effect as fixed effect. Model (A.1) is used for the analysis. If we can select subjects from a large population, we consider the subjects effect as a random effect. Model (B.1) is used for the analysis. See Appendixs A and B for details of analysis.

$$Y_{ij[k]l} = \mu + \alpha_i + \tau_j + \beta_{k(l)} + \gamma_l + \epsilon_{ij[k]l}, \quad k = 1, \ldots, p;\ i = 1, \ldots, p;\ j = 1, \ldots, p;\ l = 1, \ldots, n, \quad \text{(A.1)}$$

$$Y_{ij[k]l} = \mu + \alpha_i + \tau_j + \beta_{[k](l)} + \gamma_l + \epsilon_{ij[k]l}, \quad k = 1, \ldots, p;\ i = 1, \ldots, p;\ j = 1, \ldots, p;\ l = 1, \ldots, n. \quad \text{(B.1)}$$

Table 2: Example process of $p = 3$

Replicate	Set	Ranking according to X	Select subject	Treatment allocation
1	1	$X_{11(1)}, X_{11(2)}, X_{11(3)}$	$X_{1(1)}$	Randomly assign a Latin square to the 3 subjects
	2	$X_{12(1)}, X_{12(2)}, X_{12(3)}$	$X_{1(2)}$	
	3	$X_{13(1)}, X_{13(2)}, X_{13(3)}$	$X_{1(3)}$	
2	1	$X_{21(1)}, X_{21(2)}, X_{21(3)}$	$X_{2(1)}$	Randomly assign a Latin square to the 3 subjects
	2	$X_{22(1)}, X_{22(2)}, X_{22(3)}$	$X_{2(2)}$	
	3	$X_{23(1)}, X_{23(2)}, X_{23(3)}$	$X_{2(3)}$	
⋮	⋮	⋮	⋮	⋮
n	1	$X_{n1(1)}, X_{n1(2)}, X_{n1(3)}$	$X_{n(1)}$	Randomly assign a Latin square to the 3 subjects
	2	$X_{n2(1)}, X_{n2(2)}, X_{n2(3)}$	$X_{n(2)}$	
	3	$X_{n3(1)}, X_{n3(2)}, X_{n3(3)}$	$X_{n(3)}$	

Table 3: Empirical nominal value (0.05) for fixed effect model

μ	Period	Treatment	Subject		Corr(X, Y)	$\alpha = 0.05$	
			Latin square 1	Latin square 2		Proposed	SR*
1	0 0 0	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.058	0.040
					0.7	0.049	0.041
					0.5	0.055	0.042
2	0 0 0	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.043	0.054
					0.7	0.046	0.061
					0.5	0.043	0.048
1	0 0.2 0.3	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.052	0.042
					0.7	0.051	0.050
					0.5	0.060	0.047
2	0 0.2 0.3	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.053	0.050
					0.7	0.041	0.052
					0.5	0.052	0.052
1	0 0 0	0 0 0	1 3 6	2 3 4	0.9	0.057	0.047
					0.7	0.056	0.045
					0.5	0.050	0.057
2	0 0 0	0 0 0	1 3 6	2 3 4	0.9	0.047	0.046
					0.7	0.050	0.052
					0.5	0.047	0.046
1	0 0.2 0.3	0 0 0	1 3 6	2 3 4	0.9	0.053	0.054
					0.7	0.044	0.053
					0.5	0.050	0.047
2	0 0.2 0.3	0 0 0	1 3 6	2 3 4	0.9	0.049	0.057
					0.7	0.049	0.050
					0.5	0.046	0.047

SR = simple randomization.

3. Simulation

3.1. Design of simulation study

In this simulation, we compare the power of hypothesis testing using the proposed method and simple randomization. We want to explore the improvement in power of our proposed method compared to simple randomization. We simulate a 3-treatments, 3-periods crossover design using 2 Latin square replicates. The total number of subjects is 6, 3 in each Latin square. Subjects in the proposed method are ranked according to the baseline auxiliary variable denote as X. The corresponding outcome denotes as Y. The correlation between X and Y is set at (0.5, 0.7, 0.9) for each settings. For each simulation settings. Random samples of 2,000 Monte Carlo simulations are generated from bivariate

Table 4: Empirical Power comparison using proposed method and simple random sample (fixed effect model)

μ	Period	Treatment	Subject		Corr(X, Y)	$1-\beta$		Relative
			Latin square 1	Latin square 2		Proposed	SR	power
1	0 0 0	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.713	0.411	1.735
					0.7	0.336	0.193	1.745
					0.5	0.221	0.135	1.637
2	0 0 0	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.690	0.460	1.500
					0.7	0.308	0.169	1.825
					0.5	0.237	0.149	1.593
1	0 0.2 0.3	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.716	0.419	1.709
					0.7	0.306	0.179	1.709
					0.5	0.238	0.128	1.859
2	0 0.2 0.3	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.689	0.425	1.621
					0.7	0.301	0.176	1.710
					0.5	0.219	0.145	1.512
1	0 0 0	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.566	0.327	1.734
					0.7	0.246	0.160	1.534
					0.5	0.152	0.113	1.347
2	0 0 0	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.560	0.329	1.703
					0.7	0.247	0.143	1.727
					0.5	0.190	0.155	1.230
1	0 0.2 0.3	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.545	0.313	1.740
					0.7	0.241	0.145	1.668
					0.5	0.185	0.156	1.190
2	0 0.2 0.3	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.540	0.326	1.655
					0.7	0.256	0.161	1.587
					0.5	0.192	0.129	1.494
1	0 0 0	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.655	0.360	1.819
					0.7	0.304	0.174	1.747
					0.5	0.219	0.129	1.699
2	0 0 0	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.693	0.410	1.690
					0.7	0.315	0.175	1.805
					0.5	0.214	0.127	1.681
1	0 0.2 0.3	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.702	0.410	1.713
					0.7	0.311	0.183	1.699
					0.5	0.244	0.135	1.807
2	0 0.2 0.3	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.700	0.400	1.751
					0.7	0.325	0.186	1.752
					0.5	0.214	0.142	1.507

SR = simple randomization.

normal distributions. Throughout this simulation, we assume there is no carry over effect. So we can drop this effect from the model. Because the subject effect can be considered as fixed or random, we perform the simulation on the fixed effect model and the mixed effect model respectively.

We specify the period effect (α), treatment effect (τ), and subject effect (β) in the simulation. Calculate the empirical power using proposed method and simple randomization. Because the sample size is small, we check if the significant level is hold for null hypothesis under all parameter settings.

Similar assumptions apply to the mixed effect model. In the mixed effect model using simple randomization, the subject effect ($\beta_{k(l)}$) is considered as a random effect. It is assumed to be *i.i.d* with mean 0 and variance $\sigma^2_{\beta[k]}$.

3.2. Result of simulation

Table 3 presents the estimated probability of type I error under the null hypothesis of no mean differ-

Table 5: Empirical nominal value (0.05) for random effect model

μ	Period	Treatment	Subject		Corr(X, Y)	$\alpha = 0.05$	
			Latin square 1	Latin square 2		Proposed	SR
1	0 0 0	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.049	0.043
					0.7	0.047	0.053
					0.5	0.059	0.053
2	0 0 0	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.059	0.048
					0.7	0.048	0.055
					0.5	0.056	0.055
1	0 0.2 0.3	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.050	0.060
					0.7	0.052	0.050
					0.5	0.047	0.055
2	0 0.2 0.3	0 0 0	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.043	0.049
					0.7	0.046	0.056
					0.5	0.050	0.050
1	0 0 0	0 0 0	1 3 6	2 3 4	0.9	0.060	0.050
					0.7	0.050	0.052
					0.5	0.059	0.041
2	0 0 0	0 0 0	1 3 6	2 3 4	0.9	0.047	0.054
					0.7	0.050	0.046
					0.5	0.049	0.048
1	0 0.2 0.3	0 0 0	1 3 6	2 3 4	0.9	0.054	0.045
					0.7	0.052	0.050
					0.5	0.055	0.047
2	0 0.2 0.3	0 0 0	1 3 6	2 3 4	0.9	0.049	0.054
					0.7	0.053	0.049
					0.5	0.057	0.053

SR = simple randomization.

ences of treatment effects, for proposed method and simple randomization using fixed effect model, respectively. In general, both methods give close estimates to the nominal value ($\alpha = 0.05$) for all suggested correlations, period effects and subject effects.

Table 4 shows the estimated empirical power under different alternative hypothesis for treatment comparison using fixed effect model. The table shows that using the proposed method is more powerful than using simple randomization for all parameter settings. Also, the power increases as the correlation coefficient increases. The relative power shows the power of proposed method is increased by an average of 65%. When the correlation between X and Y is 0.9, our proposed method increases the power by an average of 70%. When the correlation between X and Y is 0.7, our proposed method increases the power by an average of 70%. When the correlation between X and Y is 0.5, our proposed method increases the power by an average of 54%.

Table 5 presents the estimated probability of type I error under the null hypothesis of no mean differences of treatment effects, for proposed method and simple randomization using mixed effect model, respectively. In general, both models give close estimates to the nominal value ($\alpha = 0.05$) for all suggested period effects and subject effect.

Table 6 shows the estimated empirical power under different alternative hypothesis for treatment comparison using mixed effect model. The table shows that using the proposed method is more powerful than using simple randomization for all parameter settings. Also, the power increases as the correlation coefficient increases. The relative power is calculated as the power of proposed method divided by the power of simple randomization. The relative power shows the power of proposed method is increased by an average of 63%. When the correlation between X and Y is 0.9, our proposed method increases the power by an average of 66%. When the correlation between X and Y is 0.7, our

Table 6: Empirical power comparison using proposed method and simple random sample (random effect model)

μ	Period	Treatment	Subject		Corr(X, Y)	$1-\beta$		Relative
			Latin square 1	Latin square 2		Proposed	SR	
1	0 0 0	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.698	0.426	1.639
					0.7	0.322	0.182	1.766
					0.5	0.227	0.135	1.681
2	0 0 0	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.712	0.431	1.653
					0.7	0.301	0.176	1.707
					0.5	0.233	0.132	1.772
1	0 0.2 0.3	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.696	0.422	1.648
					0.7	0.324	0.189	1.719
					0.5	0.218	0.132	1.652
2	0 0.2 0.3	0.1 0.12 0.15	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.677	0.410	1.651
					0.7	0.282	0.176	1.599
					0.5	0.217	0.129	1.689
1	0 0 0	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.557	0.310	1.797
					0.7	0.267	0.142	1.883
					0.5	0.170	0.127	1.340
2	0 0 0	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.565	0.329	1.717
					0.7	0.246	0.151	1.626
					0.5	0.187	0.128	1.467
1	0 0.2 0.3	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.566	0.327	1.732
					0.7	0.243	0.157	1.553
					0.5	0.194	0.110	1.759
2	0 0.2 0.3	1 0.8 1.2	0.1 0.12 0.15	0.12 0.15 0.1	0.9	0.533	0.338	1.581
					0.7	0.252	0.152	1.656
					0.5	0.168	0.112	1.504
1	0 0 0	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.698	0.420	1.661
					0.7	0.318	0.187	1.699
					0.5	0.238	0.137	1.741
2	0 0 0	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.690	0.448	1.540
					0.7	0.308	0.174	1.770
					0.5	0.080	0.054	1.481
1	0 0.2 0.3	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.689	0.417	1.651
					0.7	0.325	0.500	0.650
					0.5	0.233	0.123	1.894
2	0 0.2 0.3	0.1 .0.5 0.7	1 3 6	2 3 4	0.9	0.693	0.420	1.649
					0.7	0.306	0.193	1.592
					0.5	0.216	0.138	1.565

SR = simple randomization.

proposed method increases the power by an average of 60%. When the correlation between X and Y is 0.5, our proposed method increases the power by an average of 62%. This results lead to similar conclusions for fixed effect model.

4. Illustration on the effect of gaining weight associated with short term-term ingestion of TiO_2 nanofiber (TDNF)

The main objective of this experiment study was to investigate the adverse effect associated with short term-term ingestion of TiO_2 nanofiber (TDNF). Specifically the lungs will be assessed for global differential gene expression changes. The experiment was conducted by the authors to illustrate the proposed method of the first author Dr.PH dissertation. However, for the illustration propose in this paper we will focus on the effect of weight gaining associated with short term-term ingestion of TiO_2 nanofiber (TDNF), when subject is allocated to treatments based on baseline ranked weight as

Table 7: Latin square design layout for the matched paired rats

Latin square replicates	I		II	
Rank matched pairs	1	2	3	4
Period 1 (rat1 in the matched pair)	A	B	A	B
Period 2 (rat2 in the matched pair)	B	A	B	A

Table 8: ANOVA table for weight gain in rats using crossover design

Source	df	Type I SS	Mean square	F-value	p-value
Replicates	1	544.50	544.50	22.75	0.0412
Rats (rep)	2	250.82	125.41	5.24	0.1602
Period	1	2.00	2.00	0.08	0.7997
Treatment	1	220.50	220.50	9.21	0.0935

Table 9: ANOVA table for weight gain in rats using complete randomized design

Parameter	Estimate	Standard error	t-value	p-value
Intercept	100.45	5.930	16.93	<0.0001
treat 1	10.50	8.392	1.25	0.2575
treat 2	0.00			

explained above.

4.1. Experimental design

Male Sprague Dawley rats were purchased from Taconic Bioscience Inc., Hudson, NY. The animals were 6–7 weeks old at the time of treatment. To apply the proposed method of selecting the rats to this experiment, we asked Taconic Bioscience Inc. to divide the rats into four weight groups ranging from low weight group to maximum weight group. Four rats were randomly selected from each ordered group. For this illustration, we used two sets of selected rats. Each set has four rats with four different baseline weights ranked from minimum weight to maximum weight. We used two treatments: (1) Treatment A = total ingestion concentration of 0 ppm (control − FC), (2) Treatment B = 60 ppm (medium concentration − FMC). We used a Latin square model to test for the difference in treatment weight gain after two weeks of feeding the rats with the two concentrations. TDNF was delivered to animals via oral gavage. We formed four pairs of rats matching them based on their baseline weights. Our Latin square design is shown in Table 7: we treated each pair of rats as a blocking factor (as Period 1 and Period 2). Animal weight and dietary ingestion was measured during the course of the study

4.2. Data analysis and results

To compare the difference between treatment A and B effect on weight gain, we analyzed the data using a nested effect design as follows:

$$\text{weight gain} = \text{overall mean} + \text{period} + \text{treatment} + \text{rats (replicate)} + \text{replicates} + \text{error}.$$

Table 8 shows that there is a marginal significance different between Treatment A = total ingestion concentration of 0 ppm (control − FC), and Treatment B = 60 ppm (medium concentration − FMC) in weight gain after 2 weeks of feeding. Treatment B has a reduced weight gain in 2 weeks feeding, since the average weight gain in treatment A is 110.95 (104.17, 117.73, 95% C.I.) gm while in treatment B is 100.45 (93.62, 107.23, 95% C.I.) gm. However, even with very small number of rats, by using our

proposed approach, we achieved marginal significance of different treatment effect. The post power analysis showed that we need at least 16 rates.

Table 9 shows the ANOVA analysis for weight gain in rats using complete randomized design. Using the Latin square approach for a crossover design is more efficient for treatment groups comparisons.

5. Final remarks and discussion

The use of an auxiliary variable in the treatment allocation is introduced in many experiments. In the literature, studies show that the use of the auxiliary variable is beneficial to control for nuisance factor. If we want to control for two nuisance factors, the Latin square design is useful. The Latin square design provides the statistical foundation for p by p crossover design. In this paper, we focused on crossover designs using a Latin squares model.

We propose a treatment allocation method to improve the efficiency of treatment comparisons in a crossover design using Latin squares model based on one ranked auxiliary variable, which is known to be correlated with the outcome of interest. In our proposed protocol, a set of p subjects are selected based on one ranked auxiliary variable. Then we randomize one of the n Latin squares to these p subjects. This process is repeated n times. Because subject effect can be considered as fixed effect or random effect, two models are proposed for making inference with the ANOVA approach. The proposed method is shown to have greater power for all illustrated parameters than SRS in our simulation study, as shown in Section 3.

There are many advantages to using a crossover design. It is attractive to researchers, that subjects are their own control in the studies, which implies elimination of between subjects variation. A crossover design increases the efficiency for estimating and testing treatment effects, when there is no carry-over effect. It requires a smaller sample size than is needed in a parallel group design. However, not all types of studies can be applied to crossover design. For example, if subjects (e.g. rats) need to be sacrificed after the treatment, it does not apply. In addition, if we consider the presence of a carryover effect, the crossover design is not applicable in this settings.

In this paper we propose a method by incorporating a ranked auxiliary variable in a crossover design to increase the efficiency of treatment comparison. The main strength of this proposed method is that it provides an alternative method of treatment allocation. Treatment allocation based on the ranked auxiliary variable provides for increased power. In the simulation study, the proposed method shows higher power compared to simple randomization. Our proposed approach is illustrated on weight gain in rats using baseline ranked weight for rats allocation.

Appendix A: Analysis using fixed effect model

For the scenario that subjects are considered as fixed effect, we use model (A.1) and assume the following:

$$Y_{ij[k]l} = \mu + \alpha_i + \tau_j + \beta_{k(l)} + \gamma_l + \epsilon_{ij[k]l}, \quad k = 1, \ldots, p;\ i = 1, \ldots, p;\ j = 1, \ldots, p;\ l = 1, \ldots, n, \quad \text{(A.1)}$$

where $Y_{ij[k]l}$ is the response of k^{th} ranked subject in l^{th} replicate in j^{th} treatment at i^{th} period.

μ is an effect common to all response. α_i is the effect of the i^{th} period. $\beta_{k(l)}$ is the effect of the k^{th} ranked subject nested in l^{th} replicate. τ_j is the effect of the j^{th} treatment. γ_l is the effect of the l^{th} replicate. $\epsilon_{ij[k]l}$ is the error term (within subject) in observing response $Y_{ij[k]l}$. It is assumes to be independent with mean 0 and variance $\sigma^2_{[k]}$. The goal of this study is to test the treatment effect.

Table A.1: The ANOVA table for crossover design using Latin squares, proposed method, fixed effect model

Source of variation	Sum of squares (SS)	Expected value of SS	Degree of freedom
Periods	$np\sum_{i=1}^{p}\left(\bar{Y}_{i\cdots}-\bar{Y}_{\cdots\cdot}\right)^2$	$np\sum_{i=1}^{p}\alpha_i^2+\frac{p-1}{p}\sum_{k=1}^{p}\sigma^2_{[k]}$	$p-1$
Treatments	$np\sum_{j=1}^{p}\left(\bar{Y}_{\cdot j\cdot}-\bar{Y}_{\cdots\cdot}\right)^2$	$np\sum_{j=1}^{p}\tau_j^2+\frac{p-1}{p}\sum_{k=1}^{p}\sigma^2_{[k]}$	$p-1$
Subjects (nested in replications)	$p\sum_{k=1}^{p}\sum_{l=1}^{n}\left(\bar{Y}_{\cdot\cdot k(l)}-\bar{Y}_{\cdot\cdot l}\right)^2$	$p\sum_{k=1}^{p}\sum_{l=1}^{n}\beta^2_{k(l)}+\frac{n(p-1)}{p}\sum_{k=1}^{p}\sigma^2_{[k]}$	$n(p-1)$
Replicates	$p^2\sum_{l=1}^{n}\left(\bar{Y}_{\cdots l}-\bar{Y}_{\cdots\cdot}\right)^2$	$p^2\sum_{l=1}^{n}\gamma_l^2+\frac{n-1}{p}\sum_{k=1}^{p}\sigma^2_{[k]}$	$n-1$
Error	$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}\sum_{l=1}^{n}\left(Y_{ij[k]l}-\bar{Y}_{i\cdots}-\bar{Y}_{\cdot j\cdot}-\bar{Y}_{\cdot\cdot[k]l}+2\bar{Y}_{\cdots\cdot}\right)^2$	$\frac{(p-1)(np-2)}{p}\sum_{k=1}^{p}\sigma^2_{[k]}$	$(p-1)(np-2)$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}\sum_{l=1}^{n}\left(Y_{ij[k]l}-\bar{Y}_{\cdots\cdot}\right)^2$	$np\sum_{i=1}^{p}\alpha_i^2+np\sum_{j=1}^{p}\tau_j^2+p^2\sum_{l=1}^{n}\gamma_l^2+p\sum_{k=1}^{p}\sum_{l=1}^{n}\beta^2_{k(l)}+\frac{np^2-1}{p}\sum_{k=1}^{p}\alpha_i^2$	np^2-1

Therefore, the hypothesis testing can be written as:

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_p = 0,$$
$$H_1 : \text{at least one } \tau_p \text{is not equal to 0.}$$

Let $\bar{Y}_{\cdots\cdot}$ denote the average of all observations, $\bar{Y}_{i\cdots}$ denote the average of observations from period i, $\bar{Y}_{j\cdots}$ denote the average of observations in j^{th} treatment, $\bar{Y}_{\cdot\cdot k(l)}$denote the average of k^{th} ranked subject nested in replicate l and $\bar{Y}_{\cdots l}$ denote the average of observations in replicates l. The analysis of variance sums of squares, degrees of freedom, and expected mean squares can be obtained in the appendix. We show the test statistic is approximately F distribution with degree of freedom $[(p-1),(p-1)(np-2)]$. Table A.1 presents the ANOVA table for analyzing the model in A.1 with sum of squares expectations.

Appendix B: Analysis using mixed effect model

If we consider the subject effect as random effect, the following model (B.1) is used for analysis.

$$Y_{ij[k]l} = \mu + \alpha_i + \tau_j + \beta_{[k](l)} + \gamma_l + \epsilon_{ij[k]l}, \quad k = 1,\ldots,p;\ i = 1,\ldots,p;\ j = 1,\ldots,p; l = 1,\ldots,n. \quad \text{(B.1)}$$

Similar assumptions as in model (A.1) are needed except that we need, $\beta_{[k]l}$ the effect of the k^{th} subject nested in l^{th} replicate, to be independent with mean 0 and variance $\sigma^2_{[k]l}$ and $\epsilon_{ij[k]l}$ is independent of $\beta_{[k]l}$.

Similarly, the hypotheses can be written as:

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_p = 0,$$
$$H_1 : \text{at least one } \tau_p \text{ is not equal to 0.}$$

Table B.1 presents the ANOVA table for analyzing the model in (B.1) with sum of squares expectations.

Table B.1: The ANOVA table for crossover design using Latin squares, proposed method, mixed effect model

Source of variation	Sum of squares (SS)	Expected value of SS	Degree of freedom
Treatments	$np\sum_{i=1}^{p}\left(\bar{Y}_{i\cdots}-\bar{Y}_{\cdots\cdot}\right)^2$	$np\sum_{j=1}^{p}\tau_j^2+\frac{p-1}{p}\sum_{k=1}^{p}\sigma_{[k]}^2$	$p-1$
Periods	$np\sum_{j=1}^{p}\left(\bar{Y}_{\cdot j\cdot}-\bar{Y}_{\cdots\cdot}\right)^2$	$np\sum_{i=1}^{p}\alpha_i^2+\frac{p-1}{p}\sum_{k=1}^{p}\sigma_{[k]}^2$	$p-1$
Subjects	$p\sum_{k=1}^{p}\sum_{l=1}^{n}\left(\bar{Y}_{\cdot\cdot k(l)}-\bar{Y}_{\cdots l}\right)^2$	$\frac{n(p-1)}{p}\left(p\sum_{k=1}^{p}\sigma_{[k]}^2+\sum_{k=1}^{p}\sigma_{\beta[k]}^2\right)$	$n(p-1)$
Replicates	$p^2\sum_{l=1}^{n}\left(\bar{Y}_{\cdots l}-\bar{Y}_{\cdots\cdot}\right)^2$	$p^2\sum_{l=1}^{n}\gamma_l^2+\frac{n-1}{p}\left(p\sum_{k=1}^{p}\sigma_{\beta[k]}^2+\sum_{k=1}^{p}\sigma_{[k]}^2\right)$	$n-1$
Error	$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}\sum_{l=1}^{n}\left(Y_{ij[k]l}-\bar{Y}_{i\cdots}-\bar{Y}_{\cdot j\cdot}-\bar{Y}_{\cdot\cdot[k]l}+2\bar{Y}_{\cdots\cdot}\right)^2$	$\frac{(p-1)(np-2)}{p}\sum_{k=1}^{p}\sigma_{[k]}^2$	$(p-1)(np-2)$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}\sum_{l=1}^{n}y_{ijkl}^2-\frac{y_{\cdots\cdot}^2}{np}$	$np\sum_{i=1}^{p}\alpha_i^2+np\sum_{j=1}^{p}\tau_j^2+p^2\sum_{l=1}^{n}\gamma_l^2$ $+(np-1)\sum_{k=1}^{p}\sigma_{[k]}^2+\frac{np^2-1}{p}\sum_{k=1}^{p}\sigma_{[k]}^2$	np^2-1

References

Al-Saleh MF and Zheng G (2002). Estimation of bivariate characteristics using ranked set sampling, *Australian & New Zealand Journal of Statistics*, **44**, 221–232.

Chen Z (2007). Ranked set sampling: its essence and some new applications, *Environmental and Ecological Statistics*, **14**, 355–363.

Linder D, Samawi H, Yu L, Chatterjee A, Huang Y, and Vogel R (2015). On stratified bivariate ranked set sampling for regression eestimators, *Journal of Applied Statistics*, **42**, 2571–2583.

Lynne Stokes S (1977). Ranked set sampling with concomitant variables communications in statistics theory and methods, *Communications in Statistics - Theory and Methods*, **6**, 1207–1211.

McIntyre GA (1952). A method of unbiased selective sampling using ranked sets, *Australian Journal of Agricultural Research*, **3**, 385–390.

Samawi HM, Ahmed MS, and Abu-Dayyeh W (1996). Estimating the population mean using extreme ranked set sampling, *Biometrical Journal*, **38**, 577–586.

Senn S (1993). *Cross-Over Trials in Clinical Research*, John Wiley & Sons, New York.

Received July 29, 2016; Revised October 6, 2016; Accepted October 7, 2016