Effects of Experimental Design and Its Role in Interpretation of Results

*N. W. Shelton, S. S. Dritz[1], M. D. Tokach, R. D. Goodband,
J. L. Nelssen, J. M. DeRouchey, and L. W. Murray[2]*

Summary

A total of 256 weanling pigs (PIC TR4 × 1050, initially 13.8 lb and 21 d of age) were used in a 28-d growth trial to compare allotment methods of a completely randomized design (CRD) and a randomized complete block design (RCBD). Two treatments were used to compare these designs: a negative control with no antibiotic or growth promoter and a positive control with 35 g/ton of Denagard (Novartis Animal Health), 400 g/ton of chlortetracycline, and zinc from zinc oxide at 3,000 and 2,000 ppm in Phases 1 and 2, respectively. Experimental diets were fed in 2 phases: Phase 1 from d 0 to 14 and Phase 2 from d 14 to 28. Eight replications of each dietary treatment were used for each experimental design. The first statistical model examined dietary treatment, experimental design, and the design × dietary treatment as fixed factors. With the exception of pens in the CRD having a trend for improved ($P < 0.07$) F/G from d 0 to 14 compared with pens in the RCBD, no other design or design × dietary treatment differences were detected ($P > 0.11$) for any responses variables, indicating that treatment means reacted similarly in each of the experimental designs.

In both the CRD and the RCBD, pig weights were increased ($P < 0.003$) with supplementation of growth promoters on d 14 and 28. Variation of weight within pen remained the same in the CRD from d 0 to 28 at approximately 20% but increased from 3% on d 0 to 10% on d 28 for the RCBD. Dietary addition of growth promoters increased ($P < 0.003$) ADG and ADFI and improved F/G ($P < 0.04$) in both the CRD and RCBD from d 0 to 14, with lower P-values for the CRD than the RCBD. From d 14 to 28, the CRD detected an increase ($P < 0.001$) in ADG and ADFI with dietary addition of growth promoters, and the RCBD detected an increase ($P < 0.001$) only in ADFI. Over the entire 28-d trial, growth promoters increased ($P < 0.001$) ADG and ADFI and improved ($P < 0.03$) F/G in the CRD and increased ($P < 0.02$) ADG and ADFI in the RCBD. Lower standard errors for the difference were also estimated for ADG and F/G in the CRD than in the RCBD from d 0 to 28.

The average corrected relative efficiency for each of the three periods was 2.08 for ADG, 5.05 for ADFI, and 0.80 for F/G. The gain and intake values suggest that the added variation explained by blocks in the RCBD was beneficial for achieving a more reduced estimate of σ^2_{error} compared with analyzing that particular data set as a CRD. The variance ratios of the CRD to RCBD from d 0 to 28 depict the different responses well with ADG at 0.67, ADFI at 1.70, and F/G at 0.22. When these ratios were compared with an F-test, they were well below the upper critical limit of 4.60, suggesting that the CRD offered estimates for σ^2_{error} similar to those of the RCBD. With the same estimate for σ^2_{error}, the non-centrality parameter for each design would be similar, and therefore, the increase in degrees of freedom (DF) for the error term would lead to greater power

[1] Food Animal Health and Management Center, College of Veterinary Medicine, Kansas State University.
[2] Department of Statistics, Kansas State University.

to detect differences in the CRD. Additional studies are needed to verify these results and determine whether blocking is an efficient use of error DF.

Key words: allotment, experimental design, data interpretation

Introduction

Experimental design is a major factor that must be considered when planning research trials. The primary designs used in swine production and nutrition research include the completely randomized design (CRD) and the randomized complete block design (RCBD). Modifications or additions to these designs can be performed to generate more complex designs, such as a Latin square, that typically are used in specific instances when experimental units are limited. One of the main functions of the experimental design is to dictate the process of allotting treatments to experimental units (EU). But no matter what design is used, it is important to balance studies by having equal replication of each treatment factor to maximize the power available to detect treatment differences.

The CRD is the simplest of all designs; treatments are allotted to EU independently of any factors. This design allows for the most degrees of freedom (DF) for the error term in the model to test for treatment differences. However, the CRD can be unreliable if the EU are not homogenous. Non-homogeneity of EU can cause inflated error variance components and can increase the chance of a type 2 error. In the RCBD, treatments are allotted to EU on the basis of some factor, commonly referred to as the blocking factor, which should reduce the error variance if the blocking factor is important. The blocking factor groups EU based on that particular factor into a block, with each treatment having a minimum of one EU in each block. The primary function of blocking is to obtain groups of homogenous EU. Blocking factors vary according to the type of trial and may be different depending on the desired treatment structures. One of the assumptions in this design is that treatments would respond similarly in each block or that there were no true block × treatment interactions because the mean square calculated as the block × treatment source estimates the error variance structure for the model. One way to examine the blocking factor's effectiveness is to determine its relative efficiency (RE). Relative efficiency is a calculation performed after the trial is completed to show the ratio between an estimated error term if the study were conducted as a CRD and the error term for the RCBD. It also describes the increased number of experimental units that are needed in a CRD to achieve the same error variance component term as in a RCBD. For example, if the RE for a particular response variable was calculated to be 2.00, one could assume that the estimate for the error variance component was 2.00 times greater in the CRD than the RCBD, and theoretically, the CRD would need twice as many experimental units to achieve the same estimate error variance component as a RCBD.

It has been a common practice to block nursery studies to achieve a reduced estimate for the error component of an experiment. Often these studies are blocked simultaneously by location in the barn and initial weight. Both of these factors could affect performance and affect the interpretation of results if not equalized across treatments. The main goal of this trial was to determine the impact of blocking by initial BW and location on trial interpretation.

Procedures

The procedures used in this experiment were approved by the Kansas State University Institutional Animal Care and Use Committee. The study was conducted at the Kansas State University Swine Teaching and Research Center in Manhattan, KS.

A total of 256 weanling pigs (PIC TR4 × 1050, initially 13.8 lb and 21 d of age) were used in a 28-d growth trial to compare allotment methods of a CRD and a RCBD. Two treatments were used to compare these designs: a negative control with no antibiotic or growth promoter and a positive control with growth promoting levels of antibiotics and pharmacological levels of zinc. The positive control contained 35 g/ton of Denagard (Novartis Animal Health), 400 g/ton of chlortetracycline, and zinc from zinc oxide at 3,000 and 2,000 ppm in Phases 1 and 2, respectively. Experimental diets were fed in 2 phases: Phase 1 from d 0 to 14 and Phase 2 from d 14 to 28 (Table 1). Phase 1 and 2 diets were fed in meal form and formulated to contain 1.41% and 1.31% standardized ileal digestible lysine, respectively. Phase 1 diets contained 15% spray-dried whey and 3.75% fish meal, and Phase 2 diets were based on corn and soybean meal. Eight replications of each dietary treatment were used for each experimental design.

For the allotting of pens, a group of 4 pens located in the same location were randomized such that 2 pens would be used in the CRD, 2 pens would be used in the RCBD, and the RCBD pens would contain each of the 2 dietary treatments. This was performed throughout barn, and at the conclusion of allotting pens to designs, all pens on the CRD were randomized to treatments with equal replication. For the allotting of pigs to pens, initially weaned pigs were split to each of the 2 designs such that each design would have equal weights and variations of weights for all pigs. In addition, to reduce any bias, both gender and litter were balanced between experimental designs. Pigs assigned to the CRD were allotted to pens so that the average weight and within-pen variation of weight were similar between all pens. Pigs in the RCBD were blocked by weight and put into the location blocks.

Each pen contained a 4-hole dry self feeder and a nipple waterer to provide ad libitum access to feed and water. Pens had wire-mesh floor and allowed for approximately 3 ft^2/pig. Weights and feed disappearance were measured every 14 d to determine ADG, ADFI, and F/G. In addition, variation of pig weight within pen was examined by comparing the CV. After statistics were analyzed for each design, uncorrected and corrected RE were calculated from the RCBD for the growth performance responses. The uncorrected RE was determined by dividing an estimated CRD error variance term (σ^2_{error}) by the σ^2_{error} for the RCBD. The corrected RE was derived by multiplying the uncorrected RE and a correction for DF value. A more detailed description of these calculations and terms is available by Kuehl (2000[3]). In addition to the RE, an F-test was conducted for the ratio of the CRD error variance component to the RCBD error variance component. This F-test was a 2-tailed test and used the CRD error DF for the numerator and the RCBD error DF for the denominator. The lower critical limit was set at 0.30, and the upper critical limit was at 4.60.

---

[3] Kuehl, R. O. 2000. Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA. pp. 272-275.

Three different SAS (SAS Institute Inc., Cary, NC) models were used to describe the effects of experimental design on trial interpretation. The first model used data combined from the CRD and RCBD and was analyzed as a 2×2 factorial design with the 2 experimental designs (CRD or RCBD) and the 2 dietary treatments treated as fixed factors with no random effects. The remaining models were used to analyze each of the 2 designs independently. The model for the CRD used the dietary treatment as a fixed effect with a random effect of pen within dietary treatment. For the RCBD, dietary treatment was again used as a fixed effect, block was used as a random effect, and the block × dietary treatment was used as a random effect to estimate the error variance component. For each model, pen was used as the experimental unit and analysis of variance (ANOVA) was conducted using the MIXED procedure in SAS.

Results and Discussion

The results from the first model (Table 2) used data sets from both designs. This model examined dietary treatment, experimental design, and the design × dietary treatment as fixed factors with no blocking factors. Equal variance was assumed for both experimental designs; however, it could be that these 2 designs have unequal variances. The main focus of this model was to determine if the treatments means behaved similarly in each design and if overall performance differed in each experimental design. With the exception of pens in the CRD having a trend for improved ($P < 0.07$) F/G from d 0 to 14 compared with pens in the RCBD, no other design or design × dietary treatment differences were detected ($P > 0.11$) for any responses variables. On the basis of these results, it appears that treatment means were similar in each of the experimental designs.

After determining that performance was similar between treatments in each of the experimental designs, models were generated to evaluate the effects of each design separately. Examples of the ANOVA tables for both the CRD and RCBD are shown for overall ADG (d 0 to 28) in Tables 3 and 4, respectively. The variance term used to test for treatment effects is labeled as Pen (Treatment) in the CRD and Treatment × Block in the RCBD. It is also important to determine the difference in DF for the error term of each design. The error term for the CRD has 14 DF, and that for the RCBD design has 7 DF. This difference will affect the power of the F-test in the ANOVA model for each design. The error DF are used as the denominator DF in the ANOVA F-test, and decreasing the DF will decrease the power to detect differences, all things being equal. However, if blocking decreases the estimate of σ^2_{error}, power will increase by increasing the non-centrality parameter. Typically, the loss of DF is more than compensated by the increase in the non-centrality parameter, thereby making the block design an advantageous use of those DF.

In both the CRD and the RCBD, pig weights were increased ($P < 0.003$) with supplementation of growth promoters on d 14 and 28 (Table 5). Variation of pig weight within pen did not differ ($P > 0.52$) on d 0, 14, or 28 with the addition of growth promoters in either experimental design. However, in the CRD, variation of weight within pen remained the same from d 0 to 28 at approximately 20% but increased from 3% on d 0 to 10% on d 28 for the RCBD. The difference in within-pen variation between the 2 designs is reflective of the allotment of pigs to EU. The increase in within-pen variation when pigs begin with more uniform weight variation (RCBD) is in agreement with other studies.

Dietary addition of growth promoters increased ($P < 0.003$) ADG and ADFI and improved F/G ($P < 0.04$) in both the CRD and RCBD from d 0 to 14 (Table 6). The P-values were lower in the CRD than the RCBD because of the increase in denominator DF used in the ANOVA model and similar standard error for difference in means (SED). From d 14 to 28, the CRD detected an increase ($P < 0.001$) in ADG and ADFI with dietary addition of growth promoters, and the RCBD detected an increase ($P < 0.001$) only in ADFI. The reason why the RCBD did not detect ($P > 0.10$) an improvement in ADG with promoters was an increase in the SED compared with that for the CRD. Over the entire 28-d trial, growth promoters increased ($P < 0.001$) ADG and ADFI and improved ($P < 0.03$) F/G in the CRD. However, for the RCBD, only ADG and ADFI were increased ($P < 0.02$). For the entire trial, reduced SED were also estimated for ADG and ADFI in the CRD compared with the RCBD.

The effects of experimental design on the variance components and RE for each of the performance responses are shown in Table 7. It should be noted that the σ^2_{error} and σ^2_{block} are estimates of the true variation components for the entire population of EU. On the basis of these estimates in the RCBD, the RE as well as a ratio of the variance components between the 2 experimental designs were calculated. The uncorrected RE ranged from 0.65 to 10.63, and the corrected RE ranged from 0.59 to 9.64 for each of the growth responses. Each of the three response criteria seemed to follow a pattern for RE regardless of the time period. The average corrected RE for each of the 3 periods was 2.08 for ADG, 5.05 for ADFI, and 0.80 for F/G. The gain and intake values suggest that the added variation explained by blocks in the RCBD was beneficial for achieving a more reduced estimate of σ^2_{error} compared to analyzing that particular data set as a CRD. However, when a different allotment scheme was performed in the CRD, the variance ratio of the CRD to the RCBD ranged from 0.22 to 3.50. The ratios from d 0 to 28 depict the different responses well, with ADG at 0.67, ADFI at 1.70, and F/G at 0.22. These suggest that under a CRD allotment performed in this manner, an estimate for σ^2_{error} was obtained that was similar to that for the RCBD.

The variance ratio between the 2 designs indicated that the CRD estimated σ^2_{error} values for each response variable similar to those for the RCBD. Compared with the critical limits of 0.30 and 4.60 for an F-test between the 2 variance components, the lack of difference becomes even clearer. Observed values greater than the upper limit would suggest that the RCBD had a reduced estimate for σ^2_{error}. No values were near in proximity to the upper limit. However, ratios for F/G from d 14 to 28 and d 0 to 28 were below the lower limit, suggesting the CRD had reduced estimates for σ^2_{error} compared with the RCBD. If blocking had been effective, it should be expected to observe the variance ratios above the upper critical limit.

This experiment also suggests that using a generalized block design, which has more than 1 replication per block, may be a strategy to increase homogeneity of EU but reduce the number of DF assigned to blocks. This generalized block design would also allow for testing of interactions between treatments and blocking factors. Research has shown that various products may behave differently among different weight groups of pigs. To estimate this response, a weight × treatment interaction term is needed in the statistical model, and the generalized block design would accommodate that particular term.

In conclusion, researchers who typically block pigs by weight or some other factor can use RE to determine whether blocking offers better estimates for σ^2_{error} than a CRD. Relative efficiency is a quick method of quantifying the benefit received from a blocking factor. This single study suggests that for this nursery facility in which researchers can control the homogeneity of the average pen pig weight, the CRD estimates for σ^2_{error} are similar to those in a RCBD. With the same estimate for σ^2_{error}, the non-centrality parameter for each design would be similar, and therefore, the increase in DF for the error term would lead to a greater power to detect differences among treatments. Additional studies are needed to verify these results as well as to compare designs in different facilities and stages of production to determine whether blocking is an efficient use of error DF.

Table 1. Composition of diets[1]

Growth promoters[4]	Phase 1[2]		Phase 2[3]	
	No	Yes	No	Yes
Ingredient, %				
Corn	49.19	48.15	61.07	60.17
Soybean meal (46.5% CP)	28.98	29.06	34.97	35.03
Spray-dried whey	15.00	15.00	---	---
Select menhaden fish meal	3.75	3.75	---	---
Monocalcium P (21% P)	1.05	1.05	1.60	1.60
Limestone	0.70	0.70	1.10	1.10
Salt	0.33	0.33	0.33	0.33
Vitamin premix	0.25	0.25	0.25	0.25
Trace mineral premix	0.15	0.15	0.15	0.15
Lysine HCl	0.30	0.30	0.30	0.30
DL-methionine	0.175	0.175	0.125	0.125
L-threonine	0.125	0.125	0.110	0.110
Zinc oxide	---	0.384	---	0.256
Denagard	---	0.175	---	0.175
Chlortetracycline	---	0.400	---	0.400
Total	100.00	100.00	100.00	100.00
Calculated analysis				
SID[5] amino acids, %				
Lysine	1.41	1.41	1.31	1.31
Isoleucine:lysine	60	60	63	63
Leucine:lysine	120	120	129	129
Methionine:lysine	36	36	33	33
Met & Cys:lysine	58	58	58	58
Threonine:lysine	62	62	62	62
Tryptophan:lysine	17	17	18	18
Valine:lysine	65	65	69	69
Total lysine, %	1.55	1.55	1.45	1.45
ME, kcal/lb	1,495	1,495	1,495	1,495
SID lysine:ME, g/Mcal	4.28	4.28	3.97	3.97
CP, %	22.3	22.3	21.9	21.9
Ca, %	0.88	0.88	0.85	0.85
P, %	0.78	0.78	0.75	0.75
Available P, %	0.50	0.50	0.42	0.42
Available P:calorie, g/Mcal	1.51	1.51	1.26	1.26

[1] A total of 256 weanling pigs (PIC, initially 13.3 lb and 21 d of age) were used in a 28-d trial to compare the effects of experimental design on data interpretation.
[2] Pigs were fed Phase 1 from d 0 to 14.
[3] Pigs were fed Phase 2 from d 14 to 28.
[4] Growth promoters included zinc from zinc oxide at 3,000 ppm in Phase 1 and 2,000 ppm in Phase 2, Denagard at 35 g/ton, and chlortetracycline at 400 g/ton.
[5] Standardized ileal digestible.

Table 2. Effects of experimental design on nursery performance[1]

	Design			Probability, $P <$		
Item	CRD[2]	RCBD[3]	SED	Design × Treatment	Design	Treatment
d 0 to 14						
ADG, lb	0.49	0.47	0.027	0.45	0.44	0.001
ADFI, lb	0.58	0.58	0.030	0.65	1.00	0.001
F/G	1.20	1.24	0.023	0.70	0.07	0.001
d 14 to 28						
ADG, lb	1.07	1.07	0.045	0.44	0.99	0.006
ADFI, lb	1.56	1.55	0.058	0.85	0.81	0.001
F/G	1.46	1.45	0.021	0.16	0.68	0.14
d 0 to 28						
ADG, lb	0.78	0.77	0.033	0.39	0.73	0.001
ADFI, lb	1.07	1.06	0.042	0.72	0.83	0.001
F/G	1.38	1.38	0.016	0.12	0.67	0.38
Weights, lb						
d 0	13.8	13.8	1.00	1.00	0.99	0.99
d 14	20.7	20.4	1.26	0.80	0.79	0.04
d 28	35.6	35.5	1.89	0.70	0.92	0.02

[1] A total of 256 weanling pigs (PIC TR4 ×1050, initially 13.8 lb) were used in a 28-d study with 8 pigs per pen to determine the effect of experimental design on trial interpretation.
[2] Completely randomized design.
[3] Randomized complete block design.

Table 3. Analysis of variance table for the completely randomized design for ADG from d 0 to 28

Source	DF	Sum of squares	Mean square	F value	Pr > F
Treatment	1	0.090671	0.090671	31.1	< 0.0001
Pen (treatment)	14	0.040849	0.002918		
Corrected total	15	0.131520			

Table 4. Analysis of variance table for the randomized complete block design for ADG from d 0 to 28

Source	DF	Sum of squares	Mean square	F Value	Pr > F
Treatment	1	0.042007	0.042007	9.7	0.0171
Block	7	0.096222	0.013746		
Treatment × Block	7	0.030423	0.004346		
Corrected total	15	0.168151			

Table 5. Effects of experimental design and the addition of growth promoters on pig weights and variation in pig weight within pens[1]

	Completely randomized design				Randomized complete block design			
Growth promoter[2]:	No	Yes	SED	Probability, $P <$	No	Yes	SED	Probability, $P <$
d 0								
Avg. wt, lb	13.8	13.8	0.03	0.87	13.8	13.8	0.01	0.64
Avg. pen CV for pig wt, %[3]	20.3	20.8	0.72	0.52	3.1	3.1	0.12	0.99
d 14								
Avg. wt, lb	19.5	21.8	0.39	0.001	19.5	21.3	0.27	0.001
Avg. pen CV for pig wt, %[3]	20.4	20.7	0.80	0.67	9.5	10.4	1.67	0.64
d 28								
Avg. wt, lb	33.5	37.7	0.75	0.001	33.9	37.1	1.87	0.003
Avg. pen CV for pig wt, %[3]	18.6	18.4	1.21	0.89	10.2	9.6	1.20	0.63

[1] A total of 256 weanling pigs (PIC TR4 × 1050, initially 13.8 lb 21 d of age) were used in a 28-d study with 8 pigs per pen to determine the effect of experimental design on trial interpretation.

[2] Growth promoters included zinc from zinc oxide at 3,000 ppm in Phase 1 and 2,000 ppm in Phase 2, Denagard at 35 g/ton, and chlorrtetracycline at 400 g/ton.

[3] Depicts the in-pen variation in pig weight for each design and treatment combination.

Table 6. Effects of experimental design on interpretation of the growth effects of addition of growth promters[1]

	Completely randomized design				Randomized complete block design			
Growth promoter[2]:	No	Yes	SED	Probability, $P <$	No	Yes	SED	Probability, $P <$
d 0 to 14								
ADG, lb	0.41	0.57	0.029	0.001	0.41	0.54	0.019	0.003
ADFI, lb	0.51	0.65	0.034	0.001	0.52	0.64	0.028	0.003
F/G	1.24	1.15	0.029	0.007	1.28	1.20	0.029	0.04
d 14 to 28								
ADG, lb	1.00	1.14	0.030	0.001	1.03	1.11	0.044	0.11
ADFI, lb	1.46	1.67	0.044	0.001	1.46	1.65	0.024	0.001
F/G	1.46	1.46	0.018	0.91	1.42	1.48	0.037	0.14
d 0 to 28								
ADG, lb	0.70	0.85	0.027	0.001	0.72	0.82	0.033	0.02
ADFI, lb	0.98	1.16	0.037	0.001	0.99	1.14	0.029	0.002
F/G	1.40	1.36	0.016	0.03	1.38	1.39	0.026	0.68

[1] A total of 256 weanling pigs (PIC TR4 × 1050, initially 13.8 lb 21 d of age) were used in a 28-d study with 8 pigs per pen to determine the effect of experimental design on trial interpretation.
[2] Growth promoters included zinc from zinc oxide at 3,000 ppm in Phase 1 and 2,000 ppm in Phase 2, Denagard at 35 g/ton, and chlortetracy-cline at 400 g/ton.

Table 7. Effects of experimental design on the variance components and estimation of the error terms[1]

	Design: CRD[2]	RCBD[3]		Uncorrected	Corrected	Variance ratio
Variance components:	σ^2_{error}	σ^2_{block}	σ^2_{error}	RE[4]	RE[5]	CRD:RCBD[6]
d 0 to 14						
ADG, lb	0.0033	0.0027	0.0015	2.67	2.42	2.20
ADFI, lb	0.0047	0.0036	0.0031	2.07	1.87	1.51
F/G	0.0033	0.0008	0.0033	1.23	1.11	1.00
d 14 to 28						
ADG, lb	0.0036	0.0099	0.0076	2.21	2.01	0.47
ADFI, lb	0.0079	0.0233	0.0023	10.63	9.64	3.50
F/G	0.0013	-0.0019	0.0075	0.76	0.69	0.17
d 0 to 28						
ADG, lb	0.0029	0.0047	0.0043	2.01	1.82	0.67
ADFI, lb	0.0055	0.0105	0.0033	4.01	3.64	1.70
F/G	0.0010	-0.0016	0.0044	0.65	0.59	0.22

[1] A total of 256 weanling pigs (PIC TR4 × 1050, initially 13.8 lb 21 d of age) were used in a 28-d study with 8 pigs per pen to determine the effect of experimental design on trial interpretation.
[2] Completely randomized design.
[3] Randomized complete block design.
[4] Uncorrected relative efficiency = estimated σ^2_{error} for CRD / σ^2_{error} for RCBD and estimated σ^2_{error} for CRD = (SSblock+r(t-1) MSE)/(rt-1) where r = the number of blocks and t = the number of treatments.
[5] Corrected relative efficiency = uncorrected relative efficiency × degrees of freedom correction, and the degrees of freedom correction = (df for RCBD + 1)(df for CRD + 3) / (df for RCBD + 3)(df for CRD + 1).
[6] Variance ratio CRD: RCBD = σ^2_{error} for CRD / σ^2_{error} for RCBD.