

Who Says Study Effect Doesn't Matter?

O. M. Little, PhD, MBA
OML Associates, Dripping Springs, Texas

INTRODUCTION:

This was an animal feeding study. It was designed to compare two 'Target' rations (Targt1 and Targt2) with two competing products (Compt1 and Compt2) and a control (Contrl). The measures of major interest were weight gained, average daily gain, feed consumed, feed efficiency and coccidial effects of the active rations. The same protocol was used at two sites: One in a plains state, another in a mountain state.

METHODS AND MATERIALS:

Eighty healthy, unweaned, Holstein calves from two studies were acclimated with a milk replacer until weaned. When weaned the calves were then randomized to one of five rations. **The treatments are blinded.** Each site started with 40 calves to compare two 'Target' rations (Targt1 and Targt2) with two competing products (Compt1 and Compt2) and a control (Contrl). The calves were double-weighed and eartagged and individually penned to prevent calf-to-calf contact. At completion of the acclimation, treatment rations were begun. After the study ration was started a coccidiosis challenge was conducted. The medicated ration was fed for 28 days blind to the investigator. Weekly weighings were conducted during the medicated ration feeding period. Double weights were taken at the end of 28 days.

DATA GATHERED:

The majority of variables analyzed came from two data points: Feed consumed and weight gained. Several variables were calculated using these data alone and in combination. Weights gathered in the initial double weighings were analyzed to establish baseline comparability. The double weighings conducted at the conclusion were used for comparison of gains. Feed consumption was measured daily by weighing feed provided to the calves and then 'weighed back' to determine exactly the amount of feed consumed. This daily feed was recorded and used with feed gains to get Feed Efficiency - feed required per pound of gain. The total weight gain was divided by 28, the length of the feeding period, for Average Daily Gain. In addition to feed consumption and weights, two other important variables were collected. These were a daily fecal sample which was scored and examined for oocysts. The daily fecal score was a determination of the consistency of the stool. Calves can dehydrate and die from severe coccidiosis infections manifested by extremely loose stools. Over the course of the feeding period these scores were averaged and reported.

The calves were inoculated with coccidiosis at the start of the medicated ration feeding period. The coccidia were allowed a 14 day feeding period to become established. The oocyst count (egg forming units of coccidia) were then monitored through the course of the study. The oocyst counts were averaged over the last two weeks of the study. Because the oocyst counts were logarithmic a different preparation of the data was used. The empty cells (i.e., zeroes) were replaced by 0.5 so that the logarithmic functions worked. Once the values were established, analysis of variance was applied to the adjusted data points.

ANALYSIS:

This was a pooled analysis of results from a common protocol. The treatments and investigational sites were compared in the model:

$$\text{Dep Var} = \text{Group} + \text{Study} + \text{Group} * \text{Study} + \text{Error}$$

Where: Dep Var = Efficacy measures (baseline weight, feed conversion, weight gains, average daily gain, fecal scores and coccidia infestation.)
Group = Treatment Groups (i= 1, 2, 3, 4, 5)
Study = Investigational Sites (j= 1, 2)

These variables were analyzed using two-way analysis of variance. Analysis of covariance was obviated by the use of the differences in both the weight gained and the feed consumed based on the article by Steel (1).

The Food & Drug Administration in their Guideline for the Format and Content of Clinical Study Results recommends the SAS Type III sums of squares be used for analysis. The Student-Newman-Kuels test followed the ANOVA.

RESULTS AND DISCUSSION:

An element not mentioned so far was the fact that one of the competing products (Compt1) was an inappropriate comparative drug. Compt1 had no coccidicidal claim on the approved label. This meant that the Compt1 group had to be deleted prior to potential publication. This was a simple error; simple, but expensive. Someone did not check all labels before signing the protocol for execution. The remaining data was still usable, but there had been a large expense for the eliminated fifth treatment group. Instead of five groups the dataset was trimmed to four.

There were two sets of results: (1) The set with all five treatments and (2) the set with four treatment groups. Both sets were examined in detail. The comparison and analysis of the two sets of results was most informative. While the Group*Study (ie, treatment*investigator) interaction appeared after deletion of the fifth treatment group. There was however, a large propensity for interaction when all five treatments were first considered.

It must be kept in mind that the coccidia results were very important in the study. The lack of a coccidial claim was the reason for the deletion of the fifth treatment group. The magnitude of differences in the weight gained and feed consumed means played a prominent role in the results. Calf starter rations (those following weaning) usually provide good weight gains and some level of protection against coccidia. The coccidia activity is important with the weight gains and feed efficiency.

The fifth group (Compt1) provided adequate numbers and sufficient variability to minimize the level of the Study*Group interaction. When the Compt1 group was removed, the Study*Group interaction then became statistically significant.

Table 1, was edited output from Proc Tabulate, summarized the entire study. By deleting frames from Tabulate, editing was accomplished easily. This table summarized all results from the study. In this table, there were some not clearly explained events: The locations of treatment differences between treatment groups - without appropriate probability values.

Table 2 examined the means from both studies separately. Treatment means were ranked with the largest on the left to smallest on the right.

Table 3A presented the results from the two sites stacked. The results from both sites were assembled according to common variables. The order of variables was mixed between the two sites (rows with common variable names). The differences between the means were larger than expected.

Table 3B was the same data with added artwork. The decimals were connected.

Table 4 was the revised analysis - without the fifth group, Compt1. This was edited output from Tabulate. Note the Study effect ($p < 0.001$). This occurred for three of the four weight variables. There was both a statistically significant Group (ie, treatment) effect and a Group*Study interaction for the oocyst counts.

Table 5A was reconfigured with four groups. Substantial differences in magnitude existed between the means. The Study effect was flagged in the rightmost column.

Table 5B presented the crossovers between the means again. This clarified the events that caused the investigator effect and consequently the interaction.

CONCLUSION:

The postmortem examination of the study showed two things had happened. At the time they were not considered important. Detailed analysis revealed that these were major issues:

1. Site 1, in the plains, had held the calves longer before starting the study. When calves were held 15-28 days, these calves were not inclined to stop eating and go into a wait-state, to let the study begin on time. The calves continued to eat. Because they continued to eat, they continued to grow. Because they continued to grow, they ate more (healthy adult cattle will eat 3% of body weight per day). As calves grew, the amount of food consumed grew (ie, 3% of 150 pounds is 5.0 pounds; 3% of 180 pounds is 5.4 pounds). Therefore the calves from the plains site were larger, eating more and this increased the magnitude of the differences between the means, and the excessive variability that caused the difference between the sites and consequently the interaction between Study and Groups.
2. Looking closely at Site 1 daily feed records, there was another more subtle event that had not appeared to be a problem at the time. It turned out that it was a serious problem. Looking closely at the 'weigh-back' of the feed; there were almost always zeroes for the amount of 'weigh-back'. This meant that everything put in front of the calves was consumed. There was an equally good chance that had the calves been given more, they may have eaten more and grown even larger - making the differences even larger.

Obviously, the coccidicidal effect of the ration was a key measure of the impact of the study. The absence of statistically significant treatment differences with the weight related variables was disappointing to the sponsor. The element unmentioned so far was the presence of highly statistically significant ($p < 0.001$) investigator effects for three of the four weight variables. There was the treatment effect for the coccidia oocyst levels which was complicated by a Group*Study interaction. It was therefore not reportable. The interaction indicates that the two investigators handled the treatments substantially differently, and the level of difference between the two sites was extremely inconsistent.

Analysis of variance (ANOVA) remains a powerful tool. The suspicion that something was amiss arose when the treatment groups began to separate without attendant p-values. It took the extra

work of 'studying the means' to understand exactly what was going on. Then the issue became one of locating just what was causing the effects.

REFERENCE:

1. Steel, R.G.D., "Which dependent variable? Y or Y-X?" Mimeo Series BU-54M, Biometrics Unit, Cornell University, Ithaca, NY, 1954.

PROTOCOL TITLE LINE 1 -----
 PROTOCOL TITLE LINE 2 -----
 PROTOCOL TITLE LINE 3 -----

TABLE 1

Table B-1: Summary of results - by treatment

----- T R E A T M E N T S -----

			Comptr2			Control			Comptr1			Target1			Target2			
			N-	Mean	Std	N-	Mean	Std	N-	Mean	Std	N-	Mean	Std	N-	Mean	Std	
Table Set	Variable	Treatment P-value																
149	B-2	TlFdCnsd	0.9452	16	165.8	58.6	16	161.0	62.9	16	164.1	64.5	16	168.3	58.0	16	167.7	57.8
	B-3	Tl Gain	0.3575	16	64.13	13.7ab	16	49.36	22.3b	16	61.74	27.3ab	16	64.00	17.5ab	16	70.03	20.6a
	B-4	Fd/lb Gn	0.8866	16	2.53	0.49	16	5.60	9.00	16	3.31	2.40	16	2.67	0.72	16	2.51	0.77
	B-5	Av Da Gn	0.3586	16	2.29	0.49ab	16	1.76	0.80b	16	2.21	0.97ab	16	2.29	0.63ab	16	2.50	0.73a
	B-6	DaFclScr	0.8328	16	1.16	0.16b	16	1.41	0.37a	16	1.22	0.24ab	16	1.21	0.19ab	16	1.12	0.13b
	B-7	MnOocyst	0.0528<	16	1.89	0.55	16	1.80	0.39	16	1.73	0.51	16	1.73	0.31	16	1.56	0.25

< Approaching statistically significant difference (0.05<p<0.10)

** Statistically highly significantly different (p<0.01)

a,b Means/Std groups with different suffixes are statistically significantly different

TABLE 2:

R A N K I N G O F M E A N S B Y S T U D Y

Variable	Largest		Study 1 Plains		Smallest
FdConsmd	Comptr2 220.3	Comptr1 220.1	Control 220.0	Target1 219.6	Target2 217.9
Ttl Gain	Comptr1 82.00	Target2 80.37	Comptr2 74.50	Target1 73.75	Control 63.13
Av Da Gn	Comptr1 2.93	Target2 2.87	Comptr2 2.66	Target1 2.63	Control 2.25
Fd/lb Gn	Control 3.60	Target1 2.99	Comptr2 2.97	Target2 2.74	Comptr1 2.72
DaFclScr	Control 1.40	Comptr2 1.23	Comptr1 1.20	Target1 1.19	Target2 1.15
TlOocyst	Comptr2 2.10	Comptr1 1.77	Target1 1.71	Control 1.68	Target2 1.43
Study 2 Mounts					
FdConsmd	Target2 117.5	Target1 116.9	Comptr2 111.3	Comptr1 108.1	Control 101.9
Ttl Gain	Target2 59.42	Target1 54.33	Comptr2 53.84	Comptr1 41.29	Control 35.79
Av Da Gn	Target2 2.12	Target1 1.94	Comptr2 1.92	Comptr1 1.47	Control 1.28
Fd/lb Gn	Control 7.59	Comptr1 3.91	Target1 2.35	Target2 2.29	Comptr2 2.09
DaFclScr	Control 1.42	Comptr1 1.25	Target1 1.24	Comptr2 1.10	Target2 1.09
TlOocyst	Control 1.91	Target1 1.77	Comptr1 1.70	Target2 1.70	Comptr2 1.68

TABLE 3A:

R A N K I N G O F M E A N S F I V E G R O U P S

Variable	Study	Largest				Smallest
FdConsmd	1	Comprr2	Comprr1	Control	Target1	Target2
	1	220.3	220.1	220.0	219.6	217.9
	2	Target2	Target1	Comprr2	Comprr1	Control
	2	117.5	116.9	111.3	108.1	101.9
Ttl Gain	1	Comprr1	Target2	Comprr2	Target1	Control
	1	82.00	80.37	74.50	73.75	63.13
	2	Target2	Target1	Comprr2	Comprr1	Control
	2	59.42	54.33	53.84	41.29	35.79
Av Da Gn	1	Comprr1	Target2	Comprr2	Target1	Control
	1	2.93	2.87	2.66	2.63	2.25
	2	Target2	Target1	Comprr2	Comprr1	Control
	2	2.12	1.94	1.92	1.47	1.28
Fd/lb Gn	1	Control	Target1	Comprr2	Target2	Comprr1
	1	3.60	2.99	2.97	2.74	2.72
	2	Control	Comprr1	Target1	Target2	Comprr2
	2	7.59	3.91	2.35	2.29	2.09
DaFc1Scr	1	Control	Comprr2	Comprr1	Target1	Target2
	1	1.40	1.23	1.20	1.19	1.15
	2	Control	Comprr1	Target1	Comprr2	Target2
	2	1.42	1.25	1.24	1.10	1.09
TlOocyst	1	Comprr2	Comprr1	Target1	Control	Target2
	1	2.10	1.77	1.71	1.68	1.43
	2	Control	Target1	Comprr1	Target2	Comprr2
	2	1.91	1.77	1.70	1.70	1.68

TABLE 3B:

RANKING OF MEANS FIVE GROUPS

Variable	Study	Largest				Smallest	
FdConsmd	1	Comptr2	Comptr1	Control	Target1	Target2	
	1	220.3	220.1	220.0	219.6	217.9	
	2	Target2	Target1	Comptr2	Comptr1	Control	
	2	117.5	116.9	111.3	108.1	101.9	
Ttl Gain	1	Comptr1	Target2	Comptr2	Target1	Control	
	1	82.00	80.37	74.50	73.75	63.13	
	2	Target2	Target1	Comptr2	Comptr1	Control	
	2	59.42	54.33	53.84	41.29	35.79	
Av Da Gn	1	Comptr1	Target2	Comptr2	Target1	Control	
	1	2.93	2.87	2.66	2.63	2.25	
	2	Target2	Target1	Comptr2	Comptr1	Control	
	2	2.12	1.94	1.92	1.47	1.28	
Fd/lb Gn	1	Control	Target1	Comptr2	Target2	Comptr1	
	1	3.60	2.99	2.97	2.74	2.72	
	2	Control	Comptr1	Target1	Target2	Comptr2	
	2	7.59	3.91	2.35	2.29	2.09	
DaFclScr	1	Control	Comptr2	Comptr1	Target1	Target2	
	1	1.40	1.23	1.20	1.19	1.15	
	2	Control	Comptr1	Target1	Comptr2	Target2	
	2	1.42	1.25	1.24	1.10	1.09	
TlOocyst	1	Comptr2	Comptr1	Target1	Control	Target2	
	1	2.10	1.77	1.71	1.68	1.43	
	2	Control	Target1	Comptr1	Target2	Comptr2	
	2	1.91	1.77	1.70	1.70	1.68	

PROTOCOL TITLE LINE 1 -----
 PROTOCOL TITLE LINE 2 -----
 PROTOCOL TITLE LINE 3 -----

TABLE 4:

Table B-1: Summary of results - by treatment

			- - - - - T R E A T M E N T S - - - - -											
			Comptr2			Control			Target1			Target2		
			N-	Mean	Std	N-	Mean	Std	N-	Mean	Std	N-	Mean	Std
Table Set	Variable	Treatment P-value												
153 B-2	TlFdCnsd	0.8493 +++	16	165.8	58.6	16	161.0	62.9	16	168.3	58.0	16	167.7	57.8
B-3	Tl Gain	0.9176 +++	16	64.13	13.7a	16	49.36	22.3b	16	64.00	17.5a	16	70.03	20.6a
B-4	Fd/lb Gn	0.8240	16	2.53	0.49	16	5.60	9.00	16	2.67	0.72	16	2.51	0.77
B-5	Av Da Gn	0.9166 +++	16	2.29	0.49a	16	1.76	0.80b	16	2.29	0.63a	16	2.50	0.73a
B-6	DaFclScr	0.7768	16	1.16	0.16b	16	1.41	0.37a	16	1.21	0.19b	16	1.12	0.13b
B-7	MnOocyst	0.0143 * #	16	1.89	0.55	16	1.80	0.39	16	1.73	0.31	16	1.56	0.25

* Statistically significant treatment effect (p<0.05)

Statistically significant different interaction (p<0.05)

+++ Statistically very highly significant study effect (p<0.001)

a,b Means/Std groups with different suffixes are statistically significantly different

TABLE 5A: R A N K I N G O F M E A N S F O U R G R O U P S

Variable	Study	Largest		Smallest		ANOVA Result
FdConsmd	1	Comptr2	Control	Target1	Target2	
	1	220.3	220.0	219.6	217.9	
	2	Target2	Target1	Comptr2	Control	+++
	2	117.5	116.9	111.3	101.9	
Ttl Gain	1	Target2	Comptr2	Target1	Control	
	1	80.37	74.50	73.75	63.13	
	2	Target2	Target1	Comptr2	Control	+++
	2	59.42	54.33	53.84	35.79	
Av Da Gn	1	Target2	Comptr2	Target1	Control	
	1	2.87	2.66	2.63	2.25	
	2	Target2	Target1	Comptr2	Control	+++
	2	2.12	1.94	1.92	1.28	
Fd/lb Gn	1	Control	Target1	Comptr2	Target2	
	1	3.60	2.99	2.97	2.74	
	2	Control	Target1	Target2	Comptr2	
	2	7.59	2.35	2.29	2.09	
DaFclScr	1	Control	Comptr2	Target1	Target2	
	1	1.40	1.23	1.19	1.15	
	2	Control	Target1	Comptr2	Target2	
	2	1.42	1.24	1.10	1.09	
TlOocyst	1	Comptr2	Target1	Control	Target2	
	1	2.10	1.71	1.68	1.43	
	2	Control	Target1	Target2	Comptr2	*
	2	1.91	1.77	1.70	1.68	#

* Statistically significant treatment effect ($p < 0.05$)# Statistically significant tmt*study interaction ($p < 0.05$)+++ Statistically significant investigator effect ($p < 0.001$)

TABLE 5B:

RANKING OF MEANS FOUR GROUPS

Variable	Study	Largest		Smallest		ANOVA Result
FdConsmd	1	Comprr2	Control	Target1	Target2	+++
	1	220.3	220.0	219.6	217.9	
	2	Target2	Target1	Comprr2	Control	
	2	117.5	116.9	111.3	101.9	
Ttl Gain	1	Target2	Comprr2	Target1	Control	+++
	1	80.37	74.50	73.75	63.13	
	2	Target2	Target1	Comprr2	Control	
	2	59.42	54.33	53.84	35.79	
Av Da Gn	1	Target2	Comprr2	Target1	Control	+++
	1	2.87	2.66	2.63	2.25	
	2	Target2	Target1	Comprr2	Control	
	2	2.12	1.94	1.92	1.28	
Fd/lb Gn	1	Control	Target1	Comprr2	Target2	
	1	3.60	2.99	2.97	2.74	
	2	Control	Target1	Target2	Comprr2	
	2	7.59	2.35	2.29	2.09	
DaFc1Scr	1	Control	Comprr2	Target1	Target2	
	1	1.40	1.23	1.19	1.15	
	2	Control	Target1	Comprr2	Target2	
	2	1.42	1.24	1.10	1.09	
TlOocyst	1	Comprr2	Target1	Control	Target2	* #
	1	2.10	1.71	1.68	1.43	
	2	Control	Target1	Target2	Comprr2	
	2	1.91	1.77	1.70	1.68	

* Statistically significant treatment effect ($p < 0.05$)# Statistically significant tmt*study interaction ($p < 0.05$)+++ Statistically significant investigator effect ($p < 0.001$)