Report on Statistical Analysis of the Clutha Agricultural Development Board's Spring 2012 Calf trial (SFF L12-083)

Grant Bennett, Malcolm Deverson May 2013

Executive Summary

Independent analysis showed that if the data is assumed to be random and representative of what occurs on the average farm, then there is not a statistically significant effect of treatment.

However, a more in depth analysis showed that probiotics could improve daily weight gain, with a statistically significant effect of up to 57 grams extra growth a day, and this effect occurred on 2 of the 3 farms in the trial.

Future work is required to identify the conditions in under which a probiotic has a positive effect and how they can be best used to benefit New Zealand farmers.

Introduction

This document explains the statistical work that has been undertaken on the Clutha Agricultural Development Board's probiotic calf trial data collected in the spring of 2012. This was a MPI Sustainable Farming Fund trial (L12-083) co-funded by DairyNZ.

The trial involved approx. 300 calves in 20 pens (10 replications) across three farms in the Clutha district. Control and probiotic supplement groups were matched for birth date and weight. The supplement was a multi-strain probiotic (lacto-bacillus and yeasts) added to the milk replacement feed.

The project trial group worked within the current farm systems and calving procedures on each of the three farms – near Clinton, in West Otago and near Balclutha.

Summary points from the analyses

1) A variety of statistical models were used to analyse the raw data from the SFF L12-083 Ag. Board spring 2012 probiotics calf trial.
2) Two statistical analyses were initially performed, one in house and one by Dairy NZ. These analyses concluded with considerably different results. The confidence in the difference between the means was highly significant in one analysis and not significant in the other.
3) If the data is assumed to be truly random and representative of what occurs on the average of all three farms, there is not a significant effect of treatment.
4) The probiotic treatment has a variable effect on different farms. People cannot guarantee that it will work on any one particular farm or with any random group of calves.
5) Where the probiotic treatment does have an effect on weight gain, it is significant and it is positive.
6) The differences between the in house and Dairy NZ analyses were the result of how they treated "pen" as a factor (either as a fixed or a random effect) and then how these results could be applied to other farms.
7) The in-house analysis applied to just those farms in the trial, and the results were not transferable to other farms.
8) The Dairy NZ analysis did allow for this, but did not consider the implications of interactions, either between treatment and farms, or between treatment and initial weights, i.e. the analysis didn't go far enough.
9) The probiotic treatment had an effect on 2 of the 3 farms in the trial, and these were the farms where the calves moved on to grass at some stage during the trial.
10) Other noticed features during the trial work that were not able to be statistically validated included:
    a. the death rate between control (10 deaths) and treated (2 deaths) groups
    b. indications from the meal intake on the one farm reliably measured
11) The project team suggests that further work is required since the probiotic treatment appeared to improve growth by up to 57 grams a day but this not seen on every farm. Areas that should be investigated include mechanisms of action, why probiotics succeed on some farms but not others, how diet affects probiotics, if probiotics affects nitrogen metabolism of the lower gut or pathogen resistance.

Statistical explanation

Two statistical analyses were initially performed, one in house and one by Dairy NZ. These analyses concluded with considerably different results. The confidence in the difference between the means was highly significant in one analysis and not significant in the other.

Consequently, Richard Sedcole was engaged as an independent statistician to critique these previous two analyses and also conduct an independent analysis.

His conclusions were:

1) The treatment has a variable effect on different farms. People cannot guarantee that it will work on any one particular farm or with any random group of calves.
2) On balance, the treatment does have an effect on some farms, and where it occurs, it's positive.
3) The differences between the in house and Dairy NZ analyses were the result of how they treated "pen" as a factor (either as a fixed or a random effect[1]) and then how these results could be applied to other farms.
4) The in-house analysis applied to just those farms in the trial, and the results were not transferable to other farms.
5) The Dairy NZ analysis did allow for this, but did not consider the implications of interactions, either between treatment and farms, or between treatment and initial weights, i.e. the analysis didn't go far enough.
6) Treatment worked on 2 of the 3 farms in the trial, and these were the farms where the calves moved on to grass at some stage during the trial.

[1] Explanatory note: Any experiment has "factors". Typically, one factor at least is a treatment with two or more "levels". In this case we have one treatment with two levels (some and none). This factor is a "fixed effect" factor, as these levels are specified, and are the levels we want to make inferences about. Another factor may be blocking. We may group "experimental units" together to try to have relative uniformity within each block, and uncontrolled non-uniformity spread "between blocks" This type of factor is a random effect, and we might think of the blocks as being random samples from some larger population and are representative of the typical variation we expect between random samples. So this would be a random effect factor. It depends on how we make inferences. In the current case, we assume that the farms are random samples from all possible similar farms, and we want to make inferences that are applicable to these similar farms. Now we see that the same treatment is used on all farms, so that is a fixed effect, but there are different farms, so that is random. We now turn to pairs. Pairs were constructed on the bases of initial live weight, but we are making inferences about growth rate generally, and we are not specifying a particular live weight grouping as another form of "treatment", but rather using it to deal with non-uniformity. So pairing is random. It's not so clear cut as some, and I prevaricated between considering pairing as a fixed or random effect. In the end, I felt that random would be more appropriate. I should say that statisticians can get into quite intense discussions over fixed vs random effects!

Analysis One: the in-house analysis

The in-house analysis converted weight data growth to growth rate in grams per day (Figure 1) and compared paired pens using a paired mean T-test. This found a highly significant effect of treatment (P < 0.01). Correction for breed and clusters of deaths was attempted and the effect of treatment remained significant.

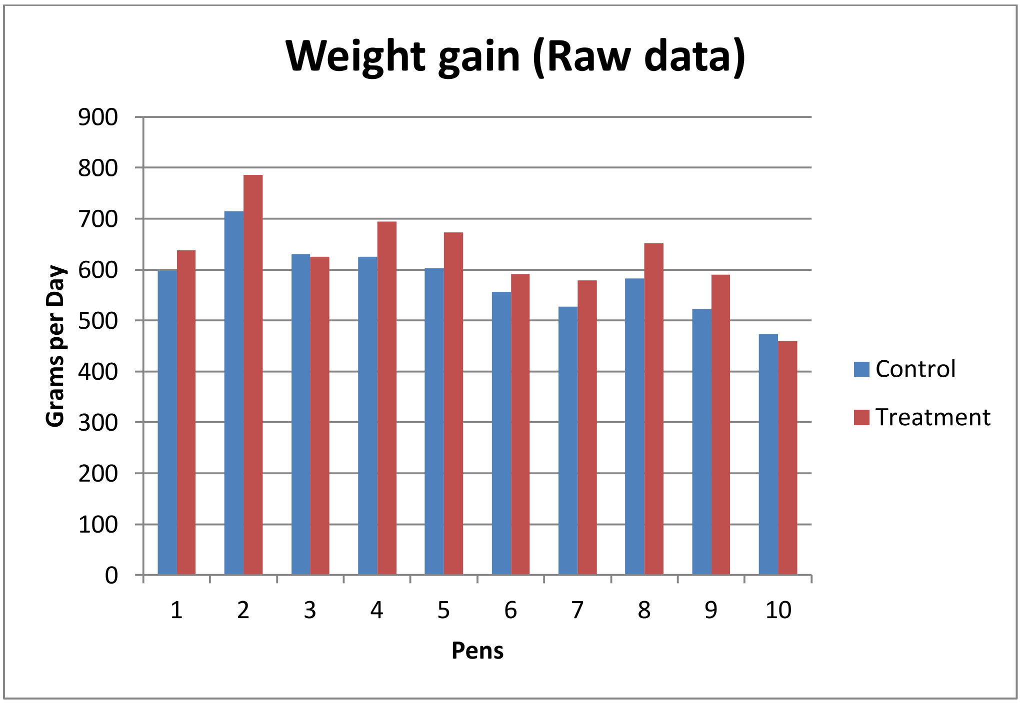


Figure 1, Average weight gain of animals. Calves were blocked into pairs of control and treatment pens attempting to match birth date and birth weight. Pens 1, 2 and 3 are from Farm 1, pens 4, 5, 6 and 7 are from Farm 2 and pens 8, 9 and 10 are from Farm 3.

Analysis Two: the DairyNZ analysis

Weight data was first converted to growth rate in grams per day by linear regression. A mixed liner model based on the restricted maximum likelihood (REML) approach was used to analyse these numbers (see Appendix A).

DairyNZ performed three analyses. Firstly all animals were assessed up to day 15 when all animals were in pens and the pens treated as replicates. No significant treatment effect (p=0.163) was observed over this limited time period.

Secondly, the growth of animals from all farms were assessed up to week 7 (one pair of pens in Farm 2 was removed since controls were treated at the end of the trial by the farmer). While treated animals grew faster (0.70 vs 0.64), this was not significant (p=0.129).

Thirdly, the growth of animals still present on farms 1 and 3 were assessed up to weeks 12-16. Treated animals grew faster, and there was a trend towards significance (P = 0.072) in the observed increase in growth rates of the probiotic treated calves.

Analysis Three: Richard Sedcole's analysis

As a result of conflicting results of the in-house and DairyNZ analyses, Dr Richard Sedcole (Consultant in Statistics & Biometrics, Lincoln University) was engaged to:

1) Critique the work done so far and identify why these analyses drew such discordant results, and
2) If appropriate, conduct an independent analysis

Critique of the in-house and Dairy NZ analyses

Richard critiqued the in-house analysis and found that

- The potential confounding effects of type of stock, breed of stock and deaths were not affecting the result.
- The in house and DairyNZ analyses were discordant due to how they treated the pen as a variable. The in house analysis treated the pen as a fixed variable and the Dairy NZ analysis treated it as a random variable. The results of these two ways of analysing the data were so drastically different since pens were linked to birth weight and birthdate, both of which have large effects on future growth.

Richard Sedcole's independent analysis

Initially a REML analysis found that while treated animals grew faster than the controls (0.5505 vs 0.5068), this effect was within the margin of error (Standard error of differences: 0.05618) and not significant.

Substantial farm to farm differences had been suspected given the literature on probiotics (drastically affected by farm management) and the raw data in Figure 1 from the in-house analysis above. Therefore a separate farm analysis was conducted (Table 1) to generate predicated means of growth (kgs per day) and p values.

This found that a treatment effect could only be seen on some farms, however predicated means should not be used to calculate the difference rather raw data should be referred to (Table 2).

Table 1, Predicated mean of growth by farm

	Treatment	Control	Difference	P value	Conclusion
Farm 1	0.665	0.658	+7g/day	Not significant	No significant effect of treatment on this farm
Farm 2	0.602	0.548	+54g/day	P > 0.01	Treatment had a slightly significant effect on weight gain
Farm 3	0.574	0.527	+47g/day	P > 0.02	Treatment had a significant effect on weight gain

Table 2, Average growth per day in pens from Farm 1, 2 and 3

	Treatment	Control	Difference	P value
Farm 1	0.662	0.644	+18g/day	Not significant
Farm 2	0.634	0.577	+57g/day	P > 0.01
Farm 3	0.573	0.534	+39g/day	P > 0.02

Lastly an ANOVA analysis was performed which assumed the data was a random representative sample. It showed that treatment did not have a significant effect on weight gain.

Conclusions

This series of analyses demonstrated that the probiotic supplement could increase growth by up to 57 grams a day, but that this effect was "patchy" and only seen on some farms in the trial.

Scientifically this is highly plausible, since probiotics would only be expected to have a large effect if an animal was deficient in native probiotic bacteria or under stress and vulnerable to pathogens that probiotics attack.

The fact that the calves in some farms appeared to respond well to probiotics and that others did not should be focus of future work.

Future areas of research that should be considered include

- Teasing out the mechanisms of how probiotics work. These are poorly defined in stock animals and what little science is available is derived from rat and mice models (no rumen!).

  E.g. Do probiotics
    - effect rumen development?
    - change the protein metabolism and ammonia transport in the lower gut by acidifying the gut? This would have important consequences for nitrogen partitioning which in turn effects how much nitrogen ends up in the urine.
    - effect pathogen colonisation of the gut or rumen and allows an animal to resist infection?
    - effect asymptomatic pathogens that colonise the lower gut? Organisms of particular interest are *E. coli* 0157:H7 and paraTB.
        - *E. coli* 0157:H7 is a very dangerous bacterium that attacks humans and is carried by cows. Its colonisation of the rectal anal junction in cows is known to be reduced by probiotics.
        - ParaTB causes Johnes disease and appears to induce inflammatory damage to the gut. Probiotics have a known action of reducing inflammatory damage (aka immunomodulation) and some deer farmers report far fewer paraTB deaths when using probiotics.

- Identifying where probiotics fail and why.
  For example does feed type matter? The penned calves in Farm 1 did not respond to probiotic treatment. Is it a coincidence that it was the only farm that only fed its calves a meal preparation and no grass?

- Do probiotics decrease feed intake and improve feed efficiency in non-pastoral systems and does this vary by weight class?

On Farm 1 a group of heavy calves treated with probiotics appeared to eat 21% less meal than control calves, but still grew as fast as the controls. Note: this effect was not seen in the lighter groups of animals on Farm 1 and it is unknown if this was the result of the probiotic supplement or an aberration.

- Do probiotics affect calf survival? During the trial untreated calves died much more often (10 vs 2). However, it is unknown if this is the result of treatment. To test this would either require a very large on farm survey measuring the odds of survival or a challenge experiment that deliberately tries to make calves sick and tests if probiotics improves their survival.

Appendix A, DairyNZ analysis

A statistical analysis of the BioBrew trial

The data supplied were formatted for analysis by coding for 'farm' (Balclutha, 1; Clinton, 2; West Otago, 3), treatment (BioBrew, 1; Control, 2) and weights were entered according to date of measurement. Bulls, steers and dead animals were removed and data from pens 4 and D on the Clinton farm was not used after the calves were released to pasture, because individuals were some Control calves were placed in the 'treatment' group. All analyses were for live-weight gain (determined by regression, within periods), rather than live weight, which would have been affected by age at trial commencement and breed.

DATA ANALYSIS

1. Daily Lwt gain for 1st 15 days from start (all calves on all farms were in pens)
2. Daily Lwt gain for 7 weeks from start (a mixture of pens or pens and pasture, depending on farm)
3. Daily Lwt gain to 10th Dec for Farms 1 and 3. Calves that were not measured on 10th Dec were not included in this analysis.

Daily gains were determined for each calf for each of these 3 periods (2 periods for Farm 2) periods using regression analysis.

These **3 periods** of liveweight gains were then each analysed using mixed models including Farm and Treatment as fixed effects and penpair, pen within penpair and calf within pen as random effects.

Pens 4 and d from Farm 2 were omitted for the analysis of Daily Lwt gain for 7 weeks

Validity.

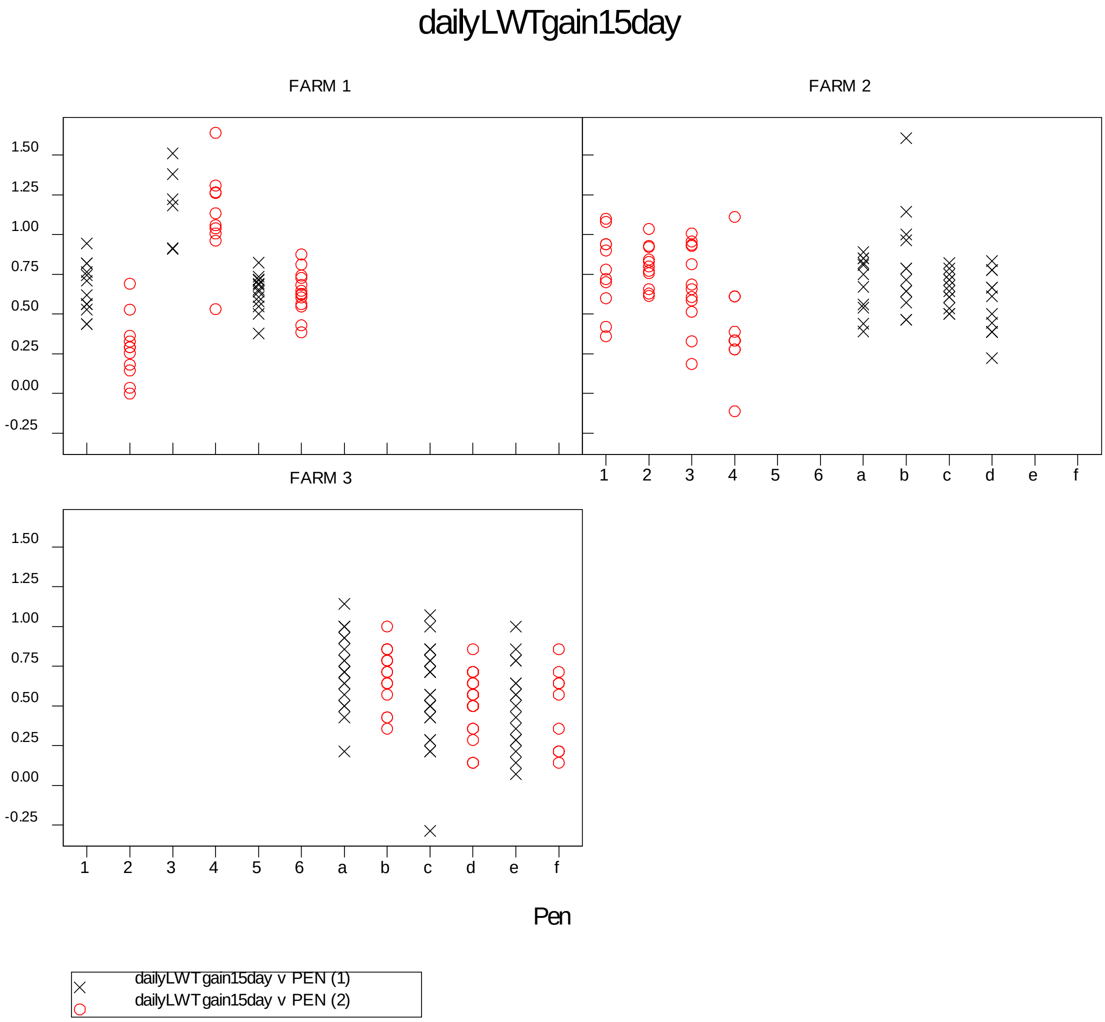
The analysis of Daily Liveweight gain for 1st 15 days is valid.

However statistical analysis of daily liveweight gain for 7 weeks and Daily liveweight gain to 10th December have assumed that the paddocks used when the calves were put on pasture were randomly chosen. **This analysis not valid** if there were differences in quality and amount of pasture available during grazing.

Through the analyses, the term 'FARM' refers to effects of the three farms; TRT, refers to the effects of treatment (BioBrew vs. Control) and the FARM.TRT is the interaction of farm by treatment. When the interaction is not statistically significant it has been dropped from the model (increasing the analytical power).

In all figures the black X are BioBrew animals, and the red O are controls

Analysis over the first 15 days in pens was undertaken because all animals received a similar treatment, enabling a valid comparison for the first 15 days of BioBrew treatment. The figure shows the distribution of daily gains in pens for the three farms:



REML variance components analysis

Response variate:	**dailyLWTgain15day**
Fixed model:	Constant + FARM + TRT + FARM.TRT
Random model:	PENPAIR + PENPAIR.FARMPEN + PENPAIR.FARMPEN.FARMCALF
Number of units:	**273** (3 units excluded due to zero weights or missing values)

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
FARM	0.82	2	0.41	6.9	0.680
TRT	**2.38**	**1**	**2.38**	**6.2**	**0.172**
FARM.TRT	2.14	2	1.07	6.4	0.398

Table of predicted means for FARM

FARM	1	2	3
	0.7537	0.6732	0.5890

Standard errors of the difference for FARM	
Average	0.171
Maximum	0.178
Minimum	0.167

Table of predicted means for TRT

TRT	1	2
	0.705	0.639

Standard error of differences for TRT: 0.0378

Table of predicted means for FARM.TRT

TRT	1	2
FARM		
1	0.8268	0.6806
2	0.6791	0.6673
3	0.6089	0.5692

Standard errors of differences for FARM	
Average:	0.065
Maximum:	0.072
Minimum:	0.061

This analysis shows that for the first 15 days of the trial (in pens, indoors) the differences in growth rate of calves on the three farms was not significant (P= 0.680) and there were no significant differences in daily gain (kg) of calves receiving BioBrew (0.70) or Controls (0.64), (P = 0.172) and there was no interaction (P = 0.398).

When the interaction was omitted from the analysis, for the first 15 days of treatment:

REML variance components analysis

Response variate:	dailyLWTgain15day
Fixed model:	Constant + FARM + TRT
Random model:	PENPAIR + PENPAIR.FARMPEN + PENPAIR.FARMPEN.FARMCALF
Number of units:	273 (3 units excluded due to zero weights or missing values)

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
FARM	0.83	2	0.42	6.9	0.675
TRT	2.36	1	2.36	8.1	**0.163**

Predicted means for FARM

FARM	1	2	3
	0.752	0.673	0.589

Standard errors of differences for FARM	
Average:	0.169
Maximum:	0.177
Minimum:	0.165

Predicted means for TRT

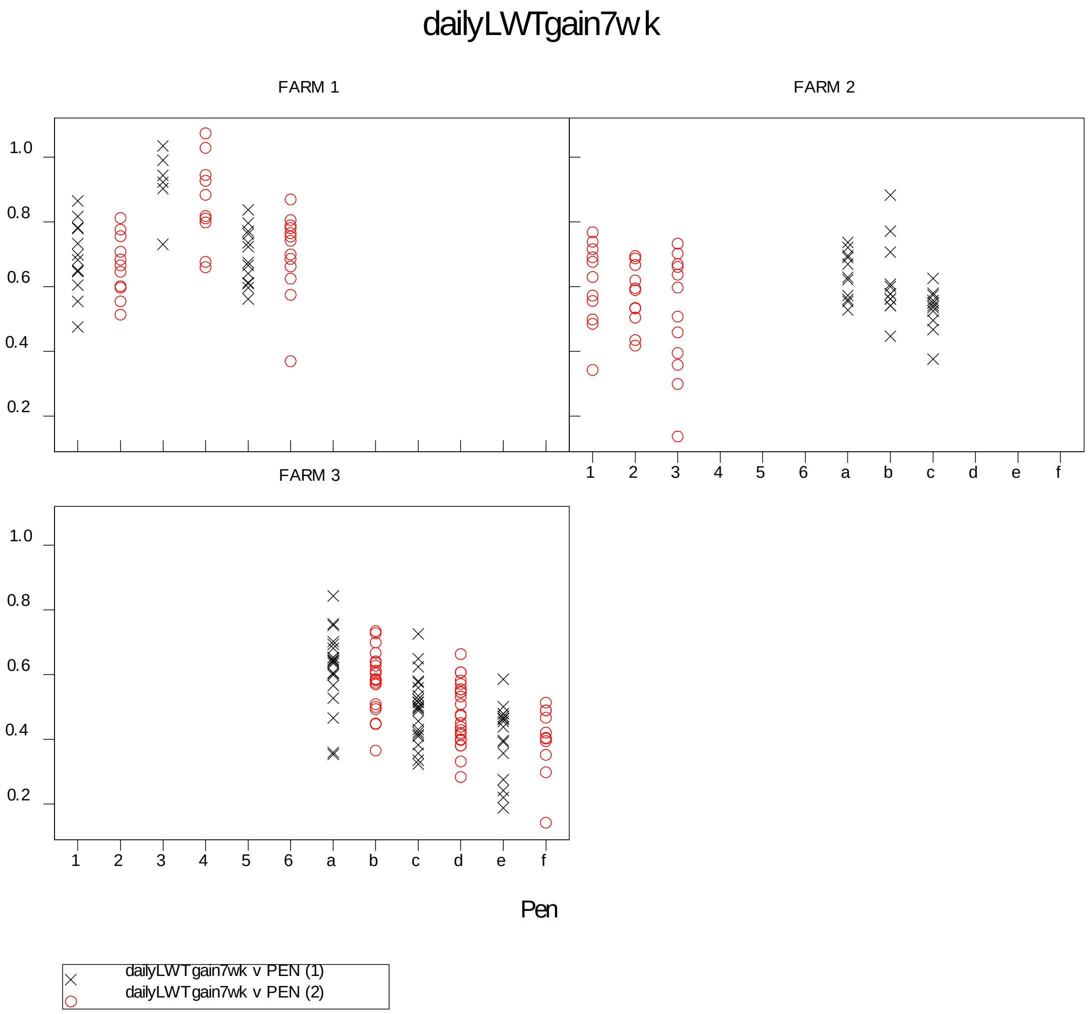
TRT	BioBrew	Control
	0.700	0.642

Standard error of differences: 0.0375

This analysis of daily **liveweight gains over the 2 weeks** of treatment when all calves were in pens showed that average rates of gain did not differ significantly (P=0.675) between farms and the probability of differences **between treatments was not statistically significant (P = 0.163).**

Q. How can daily gains of 0.70 and 0.64 be non-significant?
A. Because, either the variation between individuals, or the small number of replicates (pens) resulted in high level of uncertainty (0.16) that the difference was real.

Analysis over the first 7 weeks of the trial has been based on the original pen allocations, even though pens have been combined within treatment groups. The analysis also assumes that pastures were similar for all calves. Replicates have been reduced because animals from pens D and 4 (Clinton) have been omitted from the analysis.



REML variance components analysis

Response variate:	**dailyLWTgain7wk**
Fixed model:	Constant + FARM + TRT + FARM.TRT
Random model:	PENPAIR + PENPAIR.FARMPEN + PENPAIR.FARMPEN.FARMCALF
Number of units:	250 (2 units excluded due to zero weights or missing values)

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
FARM	11.75	2	5.88	6.0	**0.039**
TRT	2.30	1	2.30	238.2	**0.130**
FARM.TRT	0.16	2	0.08	238.2	0.924

Table of predicted means for FARM

FARM	1	2	3
	0.7514	0.5781	0.4954

Standard errors of differences for FARM

Average:	0.0759
Maximum:	0.0762
Minimum:	0.0757

Table of predicted means for TRT

TRT	1	2
	0.619	0.598

Standard error of differences for TRT: 0.0144

Table of predicted means for FARM.TRT

TRT	1	2
FARM		
1	0.7594	0.7434
2	0.5930	0.5631
3	0.5048	0.4860

Standard errors of differences for FARM

Average:	0.025
Maximum:	0.027
Minimum:	0.021

The analysis showed for the first 7 weeks of measurements, there were significant (P = 0.039) differences between farms in calf daily gain (0.75, 0.58 and 0.50 kg/d), but no significant differences in daily gain (kg) of calves receiving BioBrew (0.62) or Controls (0.60), (P = 0.130) and there was no interaction (P = 0.924) between farm and treatment.

When the FARM.TRT interaction was omitted from the analysis of the 7 weeks of treatment

REML variance components analysis

Response variate:	**dailyLWTgain7wk**
Fixed model:	Constant + FARM + TRT
Random model:	PENPAIR + PENPAIR.FARMPEN + PENPAIR.FARMPEN.FARMCALF
Number of units:	250 (2 units excluded due to zero weights or missing values)

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
FARM	11.70	2	5.85	6.0	**0.039**
TRT	2.32	1	2.32	240.2	**0.129**

Table of predicted means for FARM

FARM	1	2	3
	0.752	0.578	0.495

Standard errors of differences for FARM	
Average:	0.0761
Maximum:	0.0764
Minimum:	0.0759

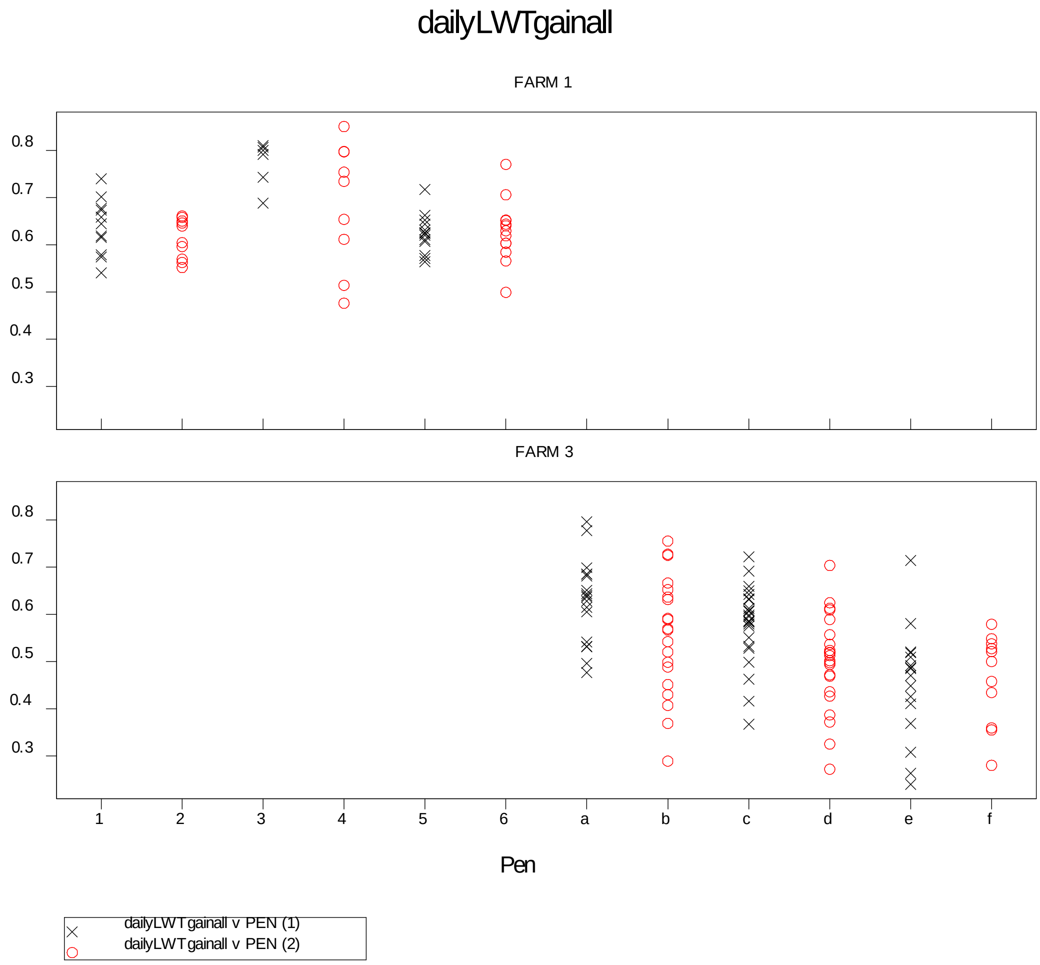
Table of predicted means for TRT

TRT	BioBrew	Control
	0.619	0.598

Standard error of differences: 0.0139

This analysis of daily **liveweight gains over the 7 weeks** of treatment showed that average rates of gain differed significantly (P=0.039) between farms but the probability of **differences between treatments was not significant (P = 0.129)**

Comparison of liveweight gains were made over a 12-16 week period from calves on two farms (1, Balclutha and 3, West Otago) that were weighed on 10th December, well after the BioBrew administration was completed.



REML variance components analysis

Response variate:	**dailyLWTgainall**
Fixed model:	Constant + FARM13 + TRT + FARM13.TRT
Random model:	PENPAIR + PENPAIR.FARMPEN + PENPAIR.FARMPEN.FARMCALF
Number of units:	174 (7 units excluded due to zero weights or missing values)

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
FARM13	6.16	1	6.16	4.0	0.069
TRT	4.96	1	4.96	3.6	**0.098**
FARM13.TRT	0.39	1	0.39	4.0	0.568

Table of predicted means for FARM13

FARM13	1	3
	0.659	0.529

Standard error of differences: 0.05168

Table of predicted means for TRT

TRT	BioBrew	Control
	0.614	0.574

Standard error of differences: 0.01916

Table of predicted means for FARM13.TRT

TRT	1	2
FARM13		
1	0.6726	0.6445
3	0.5550	0.5031

Standard errors of differences for FARM13	
Average:	0.0270
Maximum:	0.0293
Minimum:	0.0247

The analysis showed that by the 10th December there was a trend toward significant (P = 0.069) differences between farms in calf daily gain (0.66 and 0.53 kg/d), and the effects of the BioBrew treatment on daily gain (kg) (0.61 vs. 0.57, respectively; (P = 0.098). There was no interaction (P = 0.568) between farm and treatment.

When the FARM13.TRT interaction was omitted from the analysis over 12-16 weeks

REML variance components analysis

Response variate:	**dailyLWTgainall**
Fixed model:	Constant + FARM13 + TRT
Random model:	PENPAIR + PENPAIR.FARMPEN + PENPAIR.FARMPEN.FARMCALF
Number of units:	174 (7 units excluded due to zero weights or missing values)

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
FARM13	6.14	1	6.14	4.0	0.069
TRT	5.30	1	5.30	4.8	**0.072**

Table of predicted means for FARM1 3

FARM13	1	3
	0.659	0.529

Standard error of differences: 0.05176

Table of predicted means for TRT

TRT	BioBrew	Control
	0.615	0.573

Standard error of differences: 0.0184

This analysis of daily **liveweight gains over 12-16 weeks** using data from 2 farms showed that average rates of gain differed (P=0.067) between farms and there was a **trend toward significance for treatment (P = 0.072)**

Overall, the effect of treatment on gains were not large (20g/day until 7 weeks, and 50 g/day over 12-16 weeks), however, **these effects will be invalid** if there were differences in the amount and quality of pastures grazed by the calves.

Appendix. Summaries of pens, dates and calves

Table 1. Pens and mean age and dates when calves entered the trial on 3 farms

	Pen ID		Date & Age at start (d)		
	Suppl	Cont		Supp	Con
West Otago (F 3)					
	A	B	13/8	5	7
	C	D	27/8	8	8
	E	F	10/9	8	7
Clinton (F 2)					
	A	1	24/8	1	1
	B	2	27/8	3	1
	C	3	3/9	3	3
	D	4	10/9	3	4
Balclutha (F 1)					
	1	2	25/8	7	6
	3	4	28/8	12	18
	5	6	30/8	10	6

Table 2. Summary of dates and times where calves were released to pasture

	Pens	Start date	Date to pasture	Weeks in pens	Configuration on pasture
West Otago (F 3)	A, B	13/8	13/9	4	A with C; B with D
	C, D	27/8	13/9	2	
	E, F	10/9	27/9	2.5	E with A, C; F with B, D
Clinton (F 2)					
	A, 1	24/8	18/9	3.5	Calves from pens A, B and C were combined, as were pens 1, 2 and 3
	B, 2	27/8	18/9	3	
	C, 3	3/9	18/9	2	
	D, 4	10/9	1/10	3	Not to be used at pasture

Appendix B. Emails between Richard Sedcole and Grant Bennett

On Wed, Apr 10, 2013 at 8:03 AM, Richard Sedcole <Richard.Sedcole@scorch.co.nz> wrote:

Hi Grant,
Attached is the output from analysing each farm separately.
Again, there doesn't seem to be much one can hang a satisfactory response.
Your clients did say that they had some responses and some non-responses to the product, and the interpretation could be that even if you did nothing, there would be some animals that grew faster than other animals. In some ways it's the old reason for the double-blind experiment.
Basically, at the end, a person buying the product needs to be reasonably satisfied that it'll work for them, although, to be fair your clients say that it sometimes works and sometimes doesn't - you just have to try it out and see how it goes in your particular case.
There is the problem of the animals that died. Perhaps we should select only animals above a certain initial weight. I'm not comfortable with the approach of removing the stuff that went wrong, so you keep just the stuff that worked! We'll have to carefully consider what we're doing in that case.
Anyway, let's know what the next step is.
Regards, R.

BENNET3_OUT.TXT
6K View Download

Grant Bennett <bennetg2@gmail.com> Apr 10 (8 days ago)

to Richard

Thanks for that Richard
In response to this
"There is the problem of the animals that died. Perhaps we should select only animals above a certain initial weight. I'm not comfortable with the approach of removing the stuff that went wrong, so you keep just the stuff that worked! We'll have to carefully consider what we're doing in that case."
I'm not comfortable only selecting animals above certain weight for the analysis since we end deliberately filtering our data by weight-it seems like loading the deck to me. This is why I originally removed animals born at the same time as those that died regardless of their weight. While its biologically reasonable to correct for a deaths in your lightest animals since this skews the average weight gain, it does bring in a dangerous source of bias.
In terms of what next, the key stuff is going to be
1) How do we talk about the 1st and 2nd analysis
2) The pens were blocked on farm by birth weight/birth date since these have by far the biggest effect on future growth, have the current analyses incorporated this blocking effect? (As far as I can tell, dairyNZ did not, but it is not clear)

3) Why if there is a 50-60 grams a day weight difference do we not see a significant effect? E.g. high variance, leads to a high standard error which in turn masks any potential effects.
Ultimately where all this is going is to the writeup and publication, for the board who ran the trial they have seen the 1st analysis (attached for your reference), and from looking at the raw data something appears to be superficially happening. What I need to know why the first analysis is flawed and what the reasonable conclusions that should be drawn from the 1st, 2nd and your analysis.
If you are going to out near Lincoln over the next few days it might be worth going over this stuff in person, or feel free to give me a ring on 021 0245 6702.
Grant

Richard Sedcole Apr 11 (7 days ago)

to me

I've played around with this to see how to get the results you got taking each "pair" of pens at a time. As I thought might happen, there turned out to be some, although very minor, negative variances within Farms, and within pairs. If larger, I would be concerned as these things lead to misleading results.
However, I believe that I have a result consitent with yours. Depending on the way you arrange things the p value changes, but very slightly.
I have left some of the initial attempts in the out file, owing to the time necessary to edit it, but also to show that there is some opportunity for different analyses, with quite different conclusions.

Here, the INFERENCE is the tricky thing - under what conditions are the results applicable to another farm/case/whatever...?

Point one is that Balclutha was removed from the analysis!

So, one might say at the begining, this works (if at all) in 2/3rds of farms! At the least, I wouldn't back down from that statement, and I think that your clients wouldn't be too dismayed at that.

Point two is more difficult: Another analysis (not given here) looked at using the initial wt as a covariate. Overwhelmingly this was responsible for the final growth rate. Because the pairs are based on initial wt (more or less) this may mean that initial wt would swamp out any response due to the treatment. I am still not sure if I have really sorted out the effect here - I would be happier if the result was more clear cut.
However, you have means and SEDs from which you can draw comclusions, but these should be given out with suitable label warnings - perhaps.
Let's know what next.
Regards, Richard.

Richard Sedcole Apr 11 (7 days ago)

to me

Hi Grant,
I've thought about it & I figured that really, "Pair" is a random term within Farm, which should also be random. The only fixed effect is Treatment.
Now, with all farms in there is no effect of treatment, however, just removing Balcluthat we get this:
\
restrict wtgain ; cond=(Farm .ne. 1)

92.. ....

***** REML Variance Components Analysis *****

Response Variate : wtgain

Random model : Farm+Farm.Pair+Farm.Pair.Pen
Fixed model : Constant+Treatmen

Number of units : 209
No absorbing factor

* Analysis is subject to the restriction on wtgain

*** Estimated Variance Components ***

Random term Component S.e.

Farm 0.00097 0.00262
Farm.Pair 0.00257 0.00186
Farm.Pair.Pen -0.00010 0.00037
units 0.01138 0.00115

*** Approximate stratum variances ***

Effective d.f.
Farm 0.18314 1.00
Farm.Pair 0.08058 4.99
Farm.Pair.Pen 0.00975 5.38
units 0.01138 195.62

* Matrix of coefficients of components for each stratum *

Farm 100.95 29.46 14.85 1.00
Farm.Pair 0.00 27.56 13.86 1.00

Farm.Pair.Pen 0.00 0.00 16.23 1.00
units 0.00 0.00 0.00 1.00

*** Wald tests for fixed effects ***

Fixed term Wald statistic d.f.

Treatmen 13.7 1 <- p{F(1,5) = 0.014}

* All Wald statistics are calculated ignoring terms fitted later in the model

* Message: Negative variance components present:

* Fitting of fixed model terms is not sequential: effects and means for
any aliased fixed model terms may therefore be misleading. Wald tests,
likelihood tests and fitted values are unaffected.

*** Table of predicted means for Constant ***

1
0.5581

Table has only one entry: standard error 0.03079

*** Table of predicted means for Treatmen ***

Treatmen Supplement Control
0.5835 0.5327

Standard error of differences: 0.01365

%%%%%%%%%%%%%%%%%%%
So I think that we now have a result that is consistent with yours, but we do have to state, as before, that it works for 2/3rd of the farms, the effect is small, but significant.
When do you want a bill?
Regards, Richard.

Richard Sedcole Apr 11 (7 days ago)

to me

Hi Grant,
I thought that I would see about putting the data together & looking at Farm * Treatment interaction, that I always felt should show up somewhere.

So here we are:
The probs have been estimated from considering what the appropriate DF would be given the stratum variance table.
So. Farm x Treatment very highly significant, which means that the treatment will respond differently on different farms.
Now, the tests are sequential - Treatment, then T X F so it could well be in a balanced data set the Treatment could be NS if the interaction term is in the model. So, for you, the interaction means table is the thing to look at. Farms vary in their responses, and the responses can be of the order shown in the table. NS worse, marginally better, significantly better.
Note the table of SEDs - it's triangular, but text wrapping does make it a bit messy.
You should be able to write all that up.
Regards, Richard.

***** REML Variance Components Analysis *****

Response Variate : wtgain

Random model : Farm+Farm.Pair+Farm.Pair.Pen
Fixed model : Constant+Treatmen+Farm.Treatmen

Number of units : 299
No absorbing factor

******** Warning (Code VC 20). Statement 1 on Line 84
Command: reml [pr=mo,me,co,st,wa] wtgain
Aliased random term

Random term Farm is aliased completely with terms earlier in the model.
This term will be deleted and the model will be re-analysed.

***** REML Variance Components Analysis *****

Response Variate : wtgain

Random model : Farm.Pair+Farm.Pair.Pen
Fixed model : Constant+Treatmen+Farm.Treatmen

Number of units : 299
No absorbing factor

*** Estimated Variance Components ***

Random term Component S.e.

Farm.Pair 0.00318 0.00192
Farm.Pair.Pen -0.00008 0.00037
units 0.01148 0.00097

*** Approximate stratum variances ***

Effective d.f.
Farm.Pair 0.10060 7.00
Farm.Pair.Pen 0.01023 6.51
units 0.01148 279.50

* Matrix of coefficients of components for each stratum *

Farm.Pair 28.14 14.13 1.00
Farm.Pair.Pen 0.00 15.66 1.00
units 0.00 0.00 1.00

*** Wald tests for fixed effects ***

Fixed term Wald statistic d.f.

Treatmen 5.7 1 <- p = {F(1,7)} 0.048
Farm.Treatmen 19.8 4 <- p = {F(4,7)} 0.0008

* All Wald statistics are calculated ignoring terms fitted later in the model

* Message: Negative variance components present:

* Fitting of fixed model terms is not sequential: effects and means for any aliased fixed model terms may therefore be misleading. Wald tests, likelihood tests and fitted values are unaffected.

*** Table of predicted means for Constant ***

1
0.5994

Table has only one entry: standard error 0.03602

*** Table of predicted means for Treatmen ***

Treatmen Supplement Control

0.6120 0.5869

Standard error of differences: 0.01180

*** Table of predicted means for Farm.Treatmen ***

Treatmen Supplement Control
Farm
Balclutha 0.6731 0.6979
Clinton 0.6100 0.5635
West Otago 0.5531 0.4992

Standard error of differences: Average 0.04310
Maximum 0.05053
Minimum 0.01849

Average variance of differences: 0.001988

Standard error of differences for same level of factor:

Farm Treatmen
Average 0.02039 0.04878
Maximum 0.02138 0.05053
Minimum 0.01849 0.04756
Average variance of differences:
0.0004177 0.002381

vdisplay [pr=me;pse=alld]

90.. ....

* Message: Negative variance components present:

* Fitting of fixed model terms is not sequential: effects and means for any aliased fixed model terms may therefore be misleading. Wald tests, likelihood tests and fitted values are unaffected.

*** Table of predicted means for Constant ***

1
0.5994

Table has only one entry: standard error 0.03602

*** Table of predicted means for Treatmen ***

Treatmen Supplement Control

0.6120 0.5869

Standard errors of differences between pairs

Treatmen Supplement *
Treatmen Control 0.01180 *

Treatmen Supplement Treatmen Control

*** Table of predicted means for Farm.Treatmen ***

Treatmen Supplement Control
Farm
Balclutha 0.6731 0.6979
Clinton 0.6100 0.5635
West Otago 0.5531 0.4992

Standard errors of differences between pairs

Farm Balclutha.Treatmen Supplement *
Farm Balclutha.Treatmen Control 0.02138 *
Farm Clinton.Treatmen Supplement 0.04820 0.04820 *
Farm Clinton.Treatmen Control 0.04826 0.04826 0.02131
*
Farm West Otago.Treatmen Supplement 0.05033 0.05033 0.04756
0.04762 *
Farm West Otago.Treatmen Control 0.05053 0.05053 0.04777
0.04782 0.01849

Farm Balclutha.Treatmen Supplement Farm Balclutha.Treatmen
Control Farm Clinton.Treatmen Supplement Farm Clinton.Treatmen Control
Farm West Otago.Treatmen Supplement

Farm West Otago.Treatmen Control *

Farm West Otago.Treatmen Control

Standard error of differences: Average 0.04310
Maximum 0.05053
Minimum 0.01849

Average variance of differences: 0.001988

Standard error of differences for same level of factor:

Farm Treatmen
Average 0.02039 0.04878
Maximum 0.02138 0.05053
Minimum 0.01849 0.04756

Average variance of differences:
0.0004177 0.002381

Apr 12 (6 days ago)
Richard Sedcole

to me

Hi Grant,
Just briefly, I am not altogether happy about these results: I'm a bit of a belt & braces fellow, and like robust data: data that are robust against methods!
Here we have conclusions very much dependent on the method, and we need to look at the implications. Note that if we put Pair as a random term, the effect of treatment vanishes, so having Pair as fixed or random matters.
On balance, I would say fixed, as the pairs were constructed on the basis of initial weights, but this has to be considered when inferring to the general "population" whatever it is.
Generally, the responses are variable - sometimes none, sometimes some, and the conditions under which responses can be expected are yet to be adequately determined.
Note that the two sample t-test is a fixed effect model.
Regards, Richard.

---

to Richard

Thanks for that Richard,
As you have pointed out earlier its all about the inferences you can make,
I'm happy, you have,
- explained why the 1st and 2nd analyses were so different
- highlighted the massive effect farm and birth weight have on the models
- actual managed to put some numbers around the statement "for 2/3rd of the farms, the effect is small, but significant". I.e. treatment varies greatly by farm-from the science of probiotics this is more than reasonable.

My only Question left is,
In the previous email you referred to the "interaction means table", describing how farms vary by responses, can you confrim exactly which bit of the gen stat output are you referring to here since its the key piece of info.
My plan is to think about this a bit, write it up and confirm with you I have not mis-interpreted anything you gave me.
Grant

Richard Sedcole Apr 12 (6 days ago)

to me

Hi Grant,
This from my new tablet, which may explain some funny words.
The interaction table is the table of means for Farms and Treatment. And below that are the SEDs in a triangular table. I looked a bit further into the initial wt as a covariate. Because the pairs were based on the initial wt I thought this might give a clue abt why we got such a change in significance with different models. Treatment x initial wt is significant. Difficult to explain what's really going on, but our conclusion remains: response varies with farms.
Regards, Richard

Apr 14 (4 days ago)
Richard Sedcole

to me

Hi Grant,
Still thinking about the analyses: I cannot feel that Pair is a fixed term. It doesn't make sense, and putting that as random makes Treatment NS. And I think that that is right. As I mentioned before, the paired t-test is, in effect, a fixed effect model. And when we put in Pair as a fixed effect, we got a significant effect of treatment.
Here are two analyses: REML which you already have, and ANOVA using Genstat's algorithm. The latter actually manages to handle it, identifying confounding as it goes - brilliant! Treatment appears in all three strata, but it is the lowest one, with farm.pair.pen as the stratum, that is appropriate, and there we have a clear NS result.
Not what your clients wanted, but my conclusion now is treatment will not have a significant effect on the population of which the data is from a random sample. That sounds a but like a cop out, but it is the statistical inference. The random terms infer that these are random samples from a (perhaps hypothetical) population to which the results apply.
I think that I've cleared up, at least for myself, the problem about the different analyses.
Regards,
Richard.

Hi Grant,
A bit belated, but here a response. Mostly, it is the pairing that affects the results depending on whether they are considered to be fixed or random effects. Pens are random. A further difference occurs if Farms are taken as random effects - which I believe they should, and I've included some analyses looking at initial weights along with commentaries to help explain.
One has to give a simple presentation, so

The treatment has a variable effect on different farms. People cannot guarantee that it will work on any one particular farm.

On balance, the treatment does have an effect, and where it does, it's positive.
The differences in the previous analyses can be explained by considering the inferences that one can make from the analyses, i.e., how are these results applied to other farms. The in-house analysis applied to just those farms in the trial, and the results were not transferable to other farms. The Dairy NZ analysis did allow for this, but did not consider the implications of interactions, either between treatment and farms, or between treatment and initial weights, i.e. the analysis didn't go far enough.
I hope this is satisfactory.
Regards,
Richard.

Hi Grant,
Attached is an account for the 4 hour's work (although you may guess I spent somewhat longer one it, see a bit more below).
You wanted some wording about the method. The data was statistically analysed by REML (Restricted maximum likelihood, or sometime called residual maximum likelihood) in the Genstat implementation. The seminal reference to this is Patterson HD & Thompson R, (1971) Biometrika 58:545-554. This is a pretty tricky paper for the tyro, and a more approachable reference would be the Genstat Manual. I don't have an up-to-date manual at hand, but you would have had something when you wrote your thesis. Failing that, get back to me & I'll download a trial version of Genstat to read off the on-line manual reference.
I explained the concepts of fixed and random effects in a previous email, which you may use in your report. I think that explains enough of the "why" of mixed modelling, but not the "how" of course.
I took the data this morning & padded it out to a balanced design with missing values to try to use the ANOVA algorithm in Genstat.
I got a result, very similar to using REML - a significant, but not highly significant effect of Treatment. But overwhelmingly an effect of initial weight. Get those lbs on early!
Anything else, please ask.
Regards,
Richard Sedcole.
