# *APPENDIX*

## 1. *Overview*

The purpose of this appendix to provide additional details related to the statistical methods utilized in *Soy isoflavones do not affect bone resorption in postmenopausal women: A dose response study using a novel approach with $^{41}Ca$.* In setting up a multi-subject model to assess treatment effectiveness, there are several things that one must keep in mind. These include:

- **Time:** The urinary $^{41}Ca/Ca$ ratio decreases over time due to the fact that the amount of available $^{41}Ca$ is continually decreased by excretion. For a given subject, we will assume (as suggested by visual inspection of data) that this relationship is linear over any 2-3 year period that begins at least six months after the initial dose.
- **Dose:** The amount of the initial dose was not the same for all subjects. The amount of dose was not random. Neither was the reduction in dose planned at the beginning of the study. Rather, the dose was reduced as a matter of standard safety protocol when it was determined that measurements could still be easily obtained at a smaller initial dosage level. This will lead to differences in the $^{41}Ca/Ca$ ratio across subjects.
- **Subject:** Even in the case where dose sizes are the same, subjects are not. Subjects who receive the same initial dose will still incorporate different amounts of $^{41}Ca$ into bone due to differences in formation rate. Resorption rates would also be expected to differ by subject. Hence a statistical model must be arranged in such a way that each subject acts as her own control.

## 2. *Adjusting for Dose, Subject, and Time*

Graphical evidence indicates that the urinary $^{41}Ca/Ca$ ratio curves for subjects who received a different initial dose were not fundamentally different. The difference in dosage level simply affects the magnitude of the responses, while the shape of the curve is approximately the same. This is illustrated in Figure 1 below. Furthermore, smaller differences in magnitude exist even for subjects given the same initial dose. These occur due to the simple fact that bone formation rates also differ among the subjects and so the amount of $^{41}Ca$ incorporated into bone will vary by subject even when the initial doses are identical. For this reason subject will need to be included as a blocking variable [1,2] in any model that we wish to use for treatment comparison.

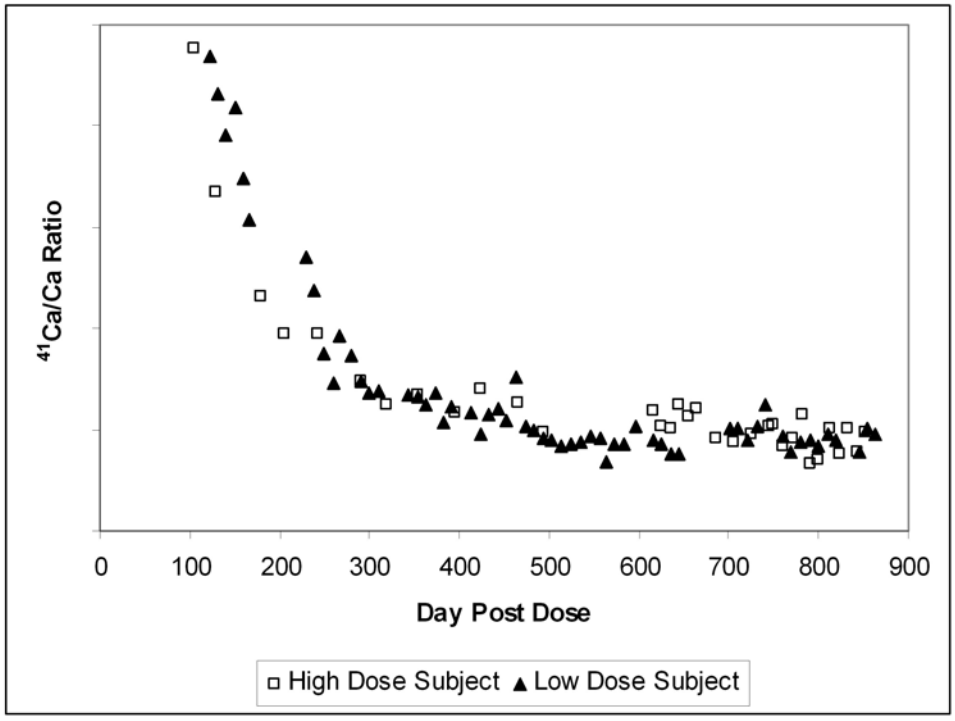


Figure 1. Subjects on different doses – the vertical axis for the low-dose subject has been rescaled to illustrate similarity of form.

By including subject as a variable, we have allowed for differences in the starting amount of $^{41}$Ca in the bone. It is also necessary to include time as a covariate since clearly the urinary $^{41}$Ca/Ca ratio is a decreasing function of time. By including a subject-by-time interaction, we allow for the fact that $^{41}$Ca is naturally resorbed at different rates for different subjects.

The inclusion of these covariates (subject and time) helps to remove a large component of the variation in the urinary $^{41}$Ca/Ca ratio. This makes it possible to observe treatment effects that would otherwise be concealed by that excess variation.

## 3. *Initial Model Results*

Before adding treatments, consider the model including subject and time as described above. The value of $R_k(t_j)$, the urinary $^{41}$Ca/Ca ratio of subject $k$ at time $t_j$ is:

$$R_k(t_j) = b_{0,k} + b_{1,k}t_j + e_k(t_j)$$

where $b_{0,k}$ and $b_{1,k}$ are parameters to be estimated for each subject, and where we assume that the model deviations are independent normal random variables with variance $s^2$. This model allows that each subject be represented by a different line. The expected value of $R_k(t_j)$ depends only on the two parameters for each subject:

$$E\{R_k(t_j)\} = b_{0,k} + b_{1,k}t_j$$

The GLM procedure in SAS 9.1 is used for calculations (coding is given in Section 8). Analysis of variance indicates that all model parameters are significant and hence useful in

decreasing the variation in the response. The model removes 82% of the total variation in the urinary $^{41}Ca/Ca$ ratio, as is required to be able to see treatment effects across subjects and time.

If desired, a nested model can be used to further divide the subject effect into a dose effect and a subject within dose effect. This is unnecessary when analyzing treatments, but included to satisfy the reader that the dose effect is indeed accounted for by the inclusion of subject in the model. The analysis of variance table for this model is shown in Table 1. It illustrates that a good portion of the subject-to-subject variation (about 80%) is in fact due to the initial dose size. Any remaining subject-to-subject variation would thus be due to the absorption rate (subject alone) and resorption rate (subject interacting with time).

| Source | DF | SS | MS | F-Value | Pr > F |
|---|---|---|---|---|---|
| Dose | 1 | 930 | 930 | 44.9 | < 0.0001 |
| Subject(Dose) | 11 | 228 | 20.7 | 33.1 | < 0.0001 |
| Time | 1 | 483 | 483 | 187 | < 0.0001 |
| Dose*Time | 1 | 28 | 28 | 44.7 | < 0.0001 |
| Subject(Dose)*Time | 11 | 11 | 1 | 1.53 | 0.1153 |
| Error | 569 | 357 | 0.628 | | |
| Total | 594 | 2037 | | | |

Table 1. Analysis of Variance Table for Nested Model

Another method of considering the analysis of variance is to simply consider the percentage of total variation explained by each source. The percentages (shown in Table 2) have some fairly straightforward interpretations. In particular, the error in this model represents all other day-to-day variation within subjects. Any treatment effects would be part of this.

| Source | SS | Percentage | |
|---|---|---|---|
| Dose | 930 | 45.6 | Initial Dose Received |
| Subject(Dose) | 228 | 11.2 | Subject Genetics; Initial Dose Absorbed |
| Time<br>Subject*Time | 522 | 25.6 | Subject Genetics; Resorption Rates? |
| Error | 357 | 17.5 | Treatments? Diet? Experimental Error? Other Variables? |
| Total | 2037 | | |

Table 2. Breakdown of Sums of Squares

## 4. *Heteroscedasticity: The Natural Log Transformation*

In assessment of assumptions for the model described in Section 2, it was found that the homogeneity of variance assumption was not satisfied. In fact the subjects may be divided into two groups, and that the divisive factor is initial dose size. The subjects having the larger initial dose display about five times the variation among their residuals when compared to subjects having the smaller dose.

Measurement error is the likely cause of this phenomenon. The accelerator mass spectrometer obtains measurements by counting atoms [3]. Thus a Poisson distribution would apply and we would expect the variability to increase as there are more atoms to count. In particular,

because the higher dose is ten times the lower one, the standard deviations would be expected to differ by a factor of $\sqrt{10}$. The variation in the residuals actually reflects a slightly larger difference – probably because of differences in formation rates.

One simple way by which to circumvent the issue and obtain a situation in which the homogeneity assumption is satisfied is to consider the log of the response variable. In general, we will let

$$L(t) = \log(R(t))$$

where the logarithm is taken on the natural scale. Our new model becomes

$$L_k(t_j) = \log(R_k(t_j)) = b_{0,k} + b_{1,k} t_j + e_k(t_j)$$

for a given subject $k$ at time $t_j$. Note that the parameters are not the same parameters as in the previous model, but are related to the previous model by the transformation. A simple check of diagnostic plots showed that the problems with homogeneity of variance were substantially reduced by this transformation.

## 5. *Treatments*

The final piece to our model is the inclusion of treatments. Note that the "treatment" variable here includes both pre-treatment and treatment periods. To determine whether or not a treatment is effective, subjects are used as their own controls. This is accomplished by the inclusion of subject described in Section 2. The $^{41}Ca/Ca$ ratio is also adjusted for time as previously described. It is then reasonable to compare measurements taken during treatment with measurements taken prior to treatment and to do so across different subjects.

One may assume that time-post-dose has nothing to do with the effect of a treatment on bone resorption. Hence the treatment-by-time interaction term is excluded from the model. The treatment-by-subject term is included since it is conceivable that subjects might react differently to a treatment.

Repeated measures also come into play because approximately five measurements are taken over the course of each fifty-day treatment (or pretreatment) phase of the study. Because of the repeated measures, the mean square error represents within subject variation *over a fifty day interval*, which is not the correct denominator for testing treatment effects. In fact, using the MSE as the error term for this test would result test statistics that are at least one order of magnitude too large!

The correct 'error term' (for a balanced design) is instead the treatment-by-subject interaction as derived from the following expected mean squares:

$$E(MS_{treatment}) = s^2 + f(treatment) + n s^2_{subj,trt}$$

$$E(MS_{subj,trt}) = s^2 + n s^2_{subj,trt}$$

Note that $f(treatment)$ is the fixed effect term associated to the treatments, and $n = 5$ is the number of observations for each treatment/subject combination. In our experiment, we do not always have exactly five observations per treatment period. Thus SAS software is used to obtain the expected mean squares and approximate tests for treatment effects.

## 6. *Testing for Treatment Effects*

As this study is designed in such a way that the treatment variable consists of both treatment and pre-treatment phases, contrasts are necessary in order to compare each treatment period to the pre-treatment phase directly preceding it. The null hypothesis for each contrast is that, for a specified treatment, the log urinary $^{41}Ca/Ca$ ratio adjusted for subject and time is not different from the immediately preceding period during which there was no actual treatment. Using the appropriate combination of least-squares means, it is also possible to obtain confidence intervals for the difference between a treatment period and its corresponding pre-treatment period – again adjusted for subject and time. Any confidence interval that did not contain 0 would indicate a positive treatment effect. While we are working in the natural log scale, the CI's for

$$L_{Diff,TRT} = \log\left(Ratio_{PRE-TRT|SUBJ,TIME}\right) - \log\left(Ratio_{TRT|SUBJ,TIME}\right)$$

are perfectly functional for answering the question of whether a treatment is or is not effective. They do not, however, carry any particular biological meaning. Thus we consider the *relative resorption* defined as follows:

$$RR = Exp\left\{-L_{Diff,TRT}\right\} = \frac{Ratio_{TRT|SUBJ,TIME}}{Ratio_{PRE-TRT|SUBJ,TIME}}$$

From a biological standpoint, the resorption reduction is the factor by which the resorption rate has changed between the pre-treatment and treatment phases in question. A resorption ratio of 0.85 would indicate that resorption rate during the treatment was 85% of resorption rate before the treatment. An even more concise statement would be that the resorption rate was reduced by 15%. The interpretation of these numbers as percentages is quite important as it allows us to consider the size of each treatment effect. Approximate confidence limits for the RR may be obtained by simply exponentiating the confidence limits for $-L_{Diff,TRT}$.

## 7. *References*


[1] D.C. Montgomery. *Design and Analysis of Experiments, 5th Ed.* New York: John Wiley & Sons, 2001.

[2] J. Neter, M.H. Kutner, C.J. Nachtsheim, W. Wasserman. *Applied Linear Statistical Models, 4th Ed.* Boston: McGraw-Hill, 1996.

[3] D. Elmore, N. Conard, and P.W. Kubik. Computer controlled isotope ratio measurements and data analysis. *Nuclear Instruments and Methods in Physics Research B*, 5:233-237, 1984.

## 8. *Appendix*

The following SAS code was used to execute the analysis.

```
PROC GLM data=Ca41;
  class Subject Treatment;
  model LNratio = Subject|Time Treatment Subject*Treatment;
  test H=Treatment E=Subject*Treatment;
  lsmeans Treatment /adjust=Tukey pdiff tdiff cl E=Subject*Treatment;
  contrast 'Treatment 0' trt 1 0 0 -1 0 0 /E=Subject*Treatment;
  contrast 'Treatment 97' trt 0 1 0 0 -1 0 /E=Subject*Treatment;
  contrast 'Treatment 135' trt 0 0 1 0 0 -1 /E=Subject*Treatment;
RUN; QUIT;
```