Rough sample size computations for *in vivo* experiments
Osborne - January, 2005

Sample size requirements require

1. An assessment of minimum detectable effect of the scaffold on a given response. ($\mu_1, \mu_2, \mu_3, \mu_4$ below)
2. An assessment of the variability of the response across rats for a given treatment. (σ^2 below)
3. Distributional assumptions about the response.
4. Specification of an acceptable level of type II (false negative) error rates.

We are not really in a position to provide much to requirements 1 and 2. However, I provide a simple table below which expresses things in terms of relative errors and *effect sizes*. Perhaps the experts can pow-wow and decide on reasonable expectations for these effect sizes and arrive at acceptable sample sizes.

Consider a one-factor analysis of variance to test for an effect of the scaffold construction factor with four levels (1=no scaffold, 2=scaffold alone, 3=scaffold w/ cells, 4=loaded scaffold). Suppose the mean of some response, say energy absorbed to failure, among many rats receiving the same scaffold treatment are given below:

	i: Treatment			
	1	2	3	4
mean, μ_i	$\mu_1 = 90$	$\mu_2 = 90$	$\mu_3 = 110$	$\mu_4 = 110$

Suppose further that the standard deviation among many rats receiving a fixed scaffold treatment is $\sigma = 10$ units of energy. Then the size effect, Δ, is defined as

$$\Delta = \sum_{1}^{4} \frac{(\mu_i - \bar{\mu})^2}{\sigma^2} = \frac{400}{100} = 4$$

The power of the F-test for a treatment effect using level of significance $\alpha = 0.05$ in a one-factor analysis of variance with n rats per scaffold treatment depends on the sample size through the *non-centrality parameter* λ, defined by $\lambda = n\Delta$ and the following equation:

$$\beta = \Pr(F > F_{\alpha/2,3,4(n-1)}, \lambda).$$

Here, F denotes a random variable with an F-distribution with $df = 3, 4(n-1)$ and non-centrality parameter λ. F_{α,ν_1,ν_2} denotes a critical value from the (central) F distribution.

The code and output on the next pages give the minimum number of rats, n, per treatment, to detect a treatment effect of size Δ in a balanced one-factor experiment to attain a power of $1 - \beta = 0.8$. It can be used to decide upon the smallest number of rats needed to detect effects of the scaffold treatment on a given mechanical testing response such as energy, provided meaningful effect sizes on this response can be specified. Note that similar computations can be carried out for other reasonable effect sizes if these are inappropriate for the experiment under consideration.

Note that these sample sizes must be doubled (half for histology) and multiplied by four (# of treatments) for TOTAL sample sizes.

```
/* begin SAS code */
data one;
   do delta=0.1 to 4 by 0.1;
      do n=2 to 10;
         ncp=n*delta;
         fstar=finv(0.95,3,4*(n-1));
         beta=probf(fstar,3,4*(n-1),ncp);
   if beta<0.2 then output;
      end;
   end;
run;
proc print;run;
proc sort;  by delta; run;
proc means noprint;
   by delta;
   var n;
   output out=two min=nmin;
run;

data two;
   set two;
   ncp = delta*nmin;
   label nmin="minimum sample size for beta<0.2";
run;
proc print labels;
   var delta ncp nmin;
run;
/* end SAS code */
```

Obs	delta	n	ncp	fstar	beta
1	1.3	10	13.0	2.86627	0.17160
2	1.4	9	12.6	2.90112	0.19050
3	1.4	10	14.0	2.86627	0.14231
4	1.5	9	13.5	2.90112	0.16186
5	1.5	10	15.0	2.86627	0.11739
6	1.6	8	12.8	2.94669	0.19131
7	1.6	9	14.4	2.90112	0.13692
	(output abbreviated)				
129	3.9	10	39.0	2.86627	0.00040
130	4.0	4	16.0	3.49029	0.17757
131	4.0	5	20.0	3.23887	0.07297
132	4.0	6	24.0	3.09839	0.02747

Obs	delta	ncp	minimum sample size for beta<0.2
1	1.3	13.0	10
2	1.4	12.6	9
3	1.5	13.5	9
4	1.6	12.8	8
5	1.7	13.6	8
6	1.8	14.4	8
7	1.9	13.3	7
8	2.0	14.0	7
9	2.1	14.7	7
10	2.2	15.4	7
11	2.3	13.8	6
12	2.4	14.4	6
13	2.5	15.0	6
14	2.6	15.6	6
15	2.7	16.2	6
16	2.8	14.0	5
17	2.9	14.5	5
18	3.0	15.0	5
19	3.1	15.5	5
20	3.2	16.0	5
21	3.3	16.5	5
22	3.4	17.0	5
23	3.5	17.5	5
24	3.6	18.0	5
25	3.7	18.5	5
26	3.8	15.2	4
27	3.9	15.6	4
28	4.0	16.0	4