

Gebreyes' study to investigate biocide effects on MDR rates
Sample size considerations

Jason A. Osborne

To assess the effect of biocide on the proportion of pigs that are MDR, a completely randomized experimental design will be used in which $4n$ barns will be randomized to four biocide treatments and m pigs will be sampled from each barn. Analysis of data will be based on aggregate measurements at the barn level so that an appropriate model for the proportion MDR for barn j receiving biocide i , Y_{ij} is

$$Y_{ij} = \pi_i + E_{ij}$$

where π_i denotes MDR positive rate among Salmonella positive ($S+$) pigs in a barn receiving biocide i and E_{ij} are mean 0 normal errors with roughly constant variance σ^2 . The null hypothesis to be tested is $H_0 : \pi_i \equiv \pi$. One alternative hypothetical configuration of MDR rates of interest is $H_1 : \pi_1 = 0.7, \pi_2 = \pi_3 = \pi_4 = 0.5$. The “effect size” of the biocide treatments under this configuration may be quantified by the standard deviation of the MDR rates, $SD_\pi = 0.1$. A sample size sufficient to obtain a power near 0.8 to reject H_0 against any alternative with this effect size or larger is needed. The effective sample size, from a given barn will be the number of $S+$ pigs, say M_s . This number is unknown in advance, but some simplifying assumptions may be made to facilitate power calculations. The expected sample size, assuming a 10% rate for $S+$, is $E(M_s) = 10$ (with a standard deviation of 3). Using this mean for the sample size, $m_s \equiv 10$, the standard deviation of a barn average over all $S+$ pigs is at most $\sqrt{0.5(0.5)/m_s} = 0.158$, and not very heterogeneous near $\pi_i = 0.5$. Assuming an error variance of $\sigma^2 \approx 0.158^2 = 0.025$ among the proposed $n = 10$ barn aggregates per biocide treatment, elementary power computations based on the non-central F distribution appropriate for one-way ANOVA lead to an experiment with a power of 0.79. (Power applets available online were used for this computation). If the number of pigs is dropped to $m = 50$, the expectation is $E(M_s) = 5$ and the bound on the error variance estimate jumps to 0.05 and the power decreases to 0.49. To assess the potential impact of stochasticity of the sample sizes, M_s due to uncertainty of $S+$, a simulation of 1000 such datasets was generated using $\Pr(S+) = 0.1$ and the configurations H_0 and H_1 above. The Monte Carlo powers under H_0 and H_1 (with Monte Carlo standard errors in parentheses) were 0.045(0.007) and 0.758(0.014), respectively, suggesting a small impact of the stochasticity of the sample size, M_s of $S+$ pigs on power. A sample of $m = 100$ total pigs per barn should be sufficient to obtain a power of roughly 0.75. (SAS code on the next page.)

```

/* Coded by Jason A. Osborne, Dept. of Statistics, NCSU */
%let pi_s=0.1; %let seed=123; %let npigs=100; %let nsims=1000;

data one;
  array mdr{4} (0.7,0.5,0.5,0.5); /* MDR rates of biocide treatments */
  *array mdr{4} (0.6,0.6,0.6,0.6);
  do sim=1 to &nsims;
    do biocide=1 to 4;
      do barn=1 to 10;
        do pig=1 to &npigs;
          u1=ranuni(&seed);
          u2=ranuni(&seed);
          spos=(u1<&pi_s);
          pimdr=mdr{biocide};
          mdrpos=(u2<pimdr);
          if spos then output;
        end; end; end; end;
  end;
run;

proc means noprint nway;
  by sim;
  class biocide barn;
  var mdrpos;
  output out=two mean=mdrmean;
run;

ods listing close;

proc mixed data=two;
  by sim;
  class biocide;
  model mdrmean=biocide;
  ods output tests3=t3;
run;

ods listing;
/*proc print data=t3; run;*/
data t3;
  set t3;
  power=(ProbF < .05);
run;
proc means data=t3 mean stddev stderr;
  var power; run;

```