CASWELL FILE
2-13-80

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

DATE: February 13, 1980

SUBJECT: EPA Reg. No. 707-75; Stampede 3E on Wheat; PP#8F2106; Addendum (Accession No. 098949) Caswell #325

FROM: Bertram D. Litt
Statistician, Toxicology Branch/HED (TS-769)

TO: Laurence D. Chitlik
Toxicologist, Toxicology Branch/HED (TS-769)

Action Requested: Review of Critical Statistical Issues in Teratology Study.

1. This study to evaluate fetal toxicity and teratogenicity of STAM Technical in rabbits was designed to administer corn oil vehicle control or 4mg/kg/day or 20 mg/kg/day or 100 mg/kg/day of STAM from day 6 through 20 during gestation of 20 female rabbits per group. The evaluation of maternal factors was based on survivors at day 29 who had been pregnant: 11, 19, 16 and 12 animals respectively. The evaluation of teratogenicity findings were based on the pups examined at day 29 from 10, 11, 13 and 5 does per group respectively.

2. To have maximum power the number of experimental subjects evaluated should include approximately 1.75 times as many control animals as there are in each of the 3 groups administered STAM (for example 18/controls and 10/4 mg/kg). The power of the experiment is the probability (expressed in percent) that an important difference (for example 10% resorptions among controls vs 25% resorptions in the mid-dose group) will be detected at a predetermined level of significance (for example $P \leq .05$). If number of animals per group is small there is only a small chance of observing a statistically significant result. This means that if a difference is found to be statistically significant we can be sure of being right in rejecting the hypothesis of no difference 19 times out of 20. But if we do not find a statistically significant difference the result does not mean that we accept the null hypothesis as proven unless the power is high (95% or higher). Thus differences which are statistically significant are important and when there is a consistant pattern the results become even more persuasive. When one of several types of findings is significant and $P \leq .05$ we say there is a 5% chance that we declare a difference to be present based on the study sample when it really is not true for the entire population. However, if 2 events are found to be statistically significant (for example statistically higher resorption rates and lower fetal viability in the low dose animals compared to controls), $P \leq (.05)^2$ or .0025; thus the chance of improperly rejecting these two hypothesis is 1 in 400.

3. Additional problems which could be mentioned include:

   a) It would be desirable to utilize implantation, resorption, born dead and born live data wherever possible for all animals up to and including the day of death or sacrifice.

   b) When a study, as here, begins study treatment for half the study of animals on one day (10/31/78) and the balance on another (11/2/78) it should be shown that this factor does not influences the outcomes of the study.

   c) Differences due to other factors (implantation rates) or inbalances between groups resulting from randomization (doe weight dt day 0) should be taken into account in performing statistical analyses.

4. From the preceeding it follows that there is reasonable evidence from this study that:

   a) A dose of 4 mg/kg/day of STAM during days 6 through 18 of pregnancy has a true treatment effect.

   b) The lack of treatment effect in the 20 mg/kg/day animals during the same treatment period may be an experimental artifact.

LITT:sp:2/13/80