U.S. Consumer Product Safety Commission
Office of Hazard Identification and Reduction

Accuracy of In-vivo Limit Dose Tests

Prepared for the Acute Toxicity Working Group
Interagency Committee on Validation of Alternative Methods

Michael A. Greene
Mathematical Statistician
Division of Hazard Analysis

March 2000

Accuracy of In-vivo Limit Dose Tests

Michael A. Greene, Ph.D.
Mathematical Statistician

Division of Hazard Analysis
Directorate for Epidemiology
U. S. Consumer Product Safety Commission

The analysis in this paper is intended to determine the accuracy of various limit dose tests. A limit dose test involves dosing a number of animals with a chemical at a single dose, the limit dose. All animals may be dosed at once or animals may be dosed one or two at a time. The test outcome is a series of deaths and survivals. A set of rules associates a test outcome with a decision as to whether the median lethal dose or LD50 is above or below the limit dose. An example of a decision rule would be to classify the LD50 as over the limit dose when more than half the animals die.

The analysis in this paper uses a computer model to evaluate the accuracy of these decision rules. A decision rule is defined to be correct when the LD50 is correctly classified as above or below the limit dose. This classification is probabilistic because it depends on the deaths and survivals observed in the limit dose test. In assessing the test accuracy, the model begins by assuming the existence of a probit dose-response curve with a known LD50 and slope. This curve is used to estimate the probability that an animal will die or survive at a given dose. The computer model then extends this result to the number of animals tested by calculating the probability of each possible sequence of deaths and survivals for all these animals. The computer model then adds up the probability that the correct outcomes occur. This would be

- the probabilities associated with outcomes that classify the LD50 below the limit dose if the true LD50 is below the limit dose, or
- the probabilities associated with outcomes that classify the LD50 above the limit dose if the true LD50 is above the limit dose.

The test accuracy is defined as the probability that the test result is correct. This is the probability that the correct outcomes occur.

The accuracy of different plans is compared in this paper. Plans differ by the number of animals involved and whether a fixed or sequential sample design is used. Accuracy is evaluated at a wide range of hypothetical LD50's and slopes of the dose-response curve. For sequential testing plans, the model also estimates the expected number of animals that would be required.

The limit dose test provides a gross classification of the toxicity of a chemical. Using a limit dose test, it is possible to determine if a chemical has an LD50 above the limit dose by using a small number of animals. A precise estimate of the LD50 may not

be required for such low toxicity chemicals. For chemicals where the test classifies the LD50 below the limit dose, an estimate of the LD50 can be obtained from an up and down test (Dixon 1991). A more general discussion of limit dose tests is in Springer et al (1993).

The limit dose test is part of the draft OECD Guideline for the Testing of Chemicals (OECD 425). It is under review by the Acute Toxicity Working Group of the Interagency Committee on Validation of Alternative Methods (ICCVAM). This committee represents a number of government agencies including the Environmental Protection Administration, the Department of Transportation, the Consumer Product Safety Commission, and the Food and Drug Administration. The guideline specifies a limit dose test at 5000 mg/kg body weight. This is in accordance with the Federal Hazardous Substances Act Regulation for acute oral toxicity in section 1500.3 (1997, page 377). Limit dose tests at 2000 mg/kg body weight are in use in Europe.

The next section describes the methods. It is followed by results and the discussion. Only limit dose tests at 5000 mg/kg are discussed in the paper. Tests at 2000 mg/kg are presented in Appendix 1.

Methods

This section describes the procedure for computing the accuracy of a limit dose test.

It is assumed that animal mortality at a given dose follows a probit dose-response curve. Let p be the probability that an individual animal dies following a dose at a given level . Then, with hypothesized values for the LD50 and σ, p is computed from the dose response curve using the following equation:

$$p = p(death; dose, LD_{50}, s) = \Phi\left(\frac{\log_{10}(LimitDose) - \log_{10}(LD_{50})}{s}\right) \qquad (1)$$

where Φ is the standard normal cumulative distribution.

The probabilities associated with individual outcomes are then aggregrated to possible sequences of test outcomes. Each animal represents an independent trial, i.e. an identical, independent (i.i.d.) realization of equation (1). The probability distribution of any given outcome involving m deaths and n animals is given by the binomial distribution as

$$P(m; n, p) = \binom{n}{m} p^m (1-p)^{n-m} \qquad (2)$$

where p is from equation (1).

The decision rules involve specifying the outcomes that classify the chemical's LD50 under the limit dose and the outcomes that involve classifying the LD50 as over the limit dose. Outcomes with more deaths tend to be associated with decision rules that classify the LD50 as under the limit dose. Suppose that n animals are to be dosed all at once with a decision rule that m or more deaths are required to classify the LD50 as under the limit dose. Then the probability that m or more deaths occur is given in equation (3) as

$$P(LD50 \le LimitDose) = \sum_{j=m}^{n} P(j; n, p) \qquad (3)$$

where $P(j;n,p)$ is given in the binomial distribution found in equation (2).

If the hypothetical LD50 is under the limit dose, then the accuracy of a test is measured by adding all the probabilities for the outcomes that lead to classifying the LD50 as under the limit dose. This requires equation (3). On the other hand, if the LD50 for the chemical is above the limit dose, the accuracy is measured by adding all the probabilities associated with the outcomes that classify as over the limit dose. This can be computed as $1\text{-}P(LimitDose \le LD50)$.

So far, the discussion has assumed that there will be a fixed sample size. In such a plan, all animals are dosed at one time. For fixed sample size plans with n animals tested, the LD50 is considered to be below the limit dose when $n/2$ or more animals die (n even) or $(n+1)/2$ or more die (n odd). For example, three or more deaths out of five animals, or five or more deaths with ten animals would be classification rules for establishing the LD50 dose below the limit dose.

Sequential sampling plans are defined to have a nominal size of n animals, indicating that no more than n animals can be dosed. Animals are dosed one or two at a time, depending on the outcomes from earlier animals in the same study. Sequential sampling plans can follow almost the same decision rules for classifying outcomes, with the exception that once enough animals survive or die to reach a conclusion, it becomes unnecessary to test more animals. When sequential sampling plans have the same decision rules as fixed sampling plans, they have the same accuracy. However, sequential plans do not have to follow the same rules and can take advantage of the order of survivals or deaths. A sequential plan can have a rule like "if the first or second animal dies then …"

The sequential plans that are considered in this paper depart from the "majority rule" classifications. They have the following general characteristics:

1. If the first animal dies, the chemical is suspected as having an LD50 below the limit dose. Limit testing is then discontinued and an up and down test conducted.
2. Otherwise animals are dosed one or two at a time. Testing is discontinued when $(n+1)/2$ die or survive (n odd).
3. If there were $(n+1)/2$ deaths, then the chemical is classified as having the LD50 below the limit dose. If the testing is discontinued when $(n+1)/2$ animals survive, the chemical is classified as having an LD50 above the limit dose. For example, in a five animal test plan with the first animal surviving, the LD50 would be classified as under the limit dose as soon as three die. It would be classified as over the LD50 if three (i.e. two more after the first) survive.

The first characteristic takes advantage of the order of deaths or survivals. This can only be done with sequential designs.

The equations presented above have only addressed the accuracy of a plan with a fixed sample size. When fixed and sequential plans have the same classification rules, such as "majority rules," the procedures for calculating accuracy are identical, because the outcome probabilities are identical. However, equations (2) and (3) can be used with sequential testing plans even when there is no fixed plan equivalent. A mathematically correct, but tedious approach is to write all the fixed sample outcomes that would correspond to a sequential plan outcome and then sum all the probabilities. There are more clever approaches that take into account the independence of the events.

The last issue for this analysis is the computation of the expected or average number of animals used in a sequential sample plan. Recall that an animal used in the trial counts toward the expected value whether the animal survives or dies, because a surviving animal cannot be used for other tests. However, animals do not count if the test is discontinued before the animal is (scheduled to be) used. The various outcomes with different numbers of animals need to be identified and the probability of the simple events needs to be calculated. For example, here are the outcomes for a five sequential sample plan:

- one animal (the first animal dies)
- three animals all survivors (S SS),
- four animals (S DD D or S SD S or S DS S) or
- five animals (all other sequences)

Let j denote the number of animals used in a test plan. Then the expected number of animals used is given in equation (4)

$$ExpectedAnimalsUsed = \sum_{j=1}^{n} j \sum_{k \subset J} p^k (1-p)^{n-k} \qquad (4)$$

where p is given in equation (1) and J is the set of sequences that use j animals.

These equations are implemented in the SAS program in Appendix 2. Equation (1) is in the linked routine *getprob*, called in *data test*. Equation (2) computes the binomial distribution in the linked routine *fillprob*, also in *data test*. This step uses either the built-in binomial cumulative distribution function in the SAS function *probbnml* or the binomial density function in (*%macro pbinom*) or some combination of the two. The rules, which are specific to each test plan, are found in an external routine called by *fillprob*. An example is on the last page of the appendix shown as *rule5f.sas*. This produces the components of equation (3), with the summation completed by *proc summary* following the data step. The calculation for the expected value in equation (4) uses similar logic. This requires a separate run of the program with a different external routine to be linked in by fillprob. See *rule5x.sas* at the end of the appendix.

The question addressed in this paper is how these limit dose test plans work over a wide variety of chemicals. We used LD50 values of 1.5, 50, 250, 1500, 2000, 3000, 5000, and 6000 mg/kg body weight. Values for σ (the inverse of the slope of the dose response curve) were 0.12, 0.25, 0.5, 1.25, and 2.00. Each pair of LD50 and σ values were modeled, i.e. 1.5 and 0.12, 1.5 and 0.25, etc, resulting in a total of 40 values for each test plan.

Both fixed and sequential test plans were modeled. Fixed sample size plans of five, seven and ten animals and sequential plans using up to five and seven animals were modeled. Limit doses were evaluated at 2000 mg/kg and 5000 mg/kg. Tables for 2000 mg/kg are in Appendix 1.

Results

This section contains results for fixed and sequential test plans at 5000 mg/kg. First, the ten animal fixed sample test plan is presented. This is the present standard procedure for limit dose tests. Next, seven animal and five animal sequential test plans are shown. The purpose of these comparisons is to determine how much (or how little) is lost when using sequential test plans that economize on the number of animals.

In the third part of the results section, fixed sample size plans with seven and five animals are presented. The purpose is to examine the difference between fixed and sequential using the same nominal number of animals. The next part of the section compares results between fixed and sequential sampling plans. The last part of the section presents the expected number of animals used in five and seven animal test plans.

The appendix contains tables in the same sequence for the 2000 mg/kg results.

The results show for each combination of LD50 and σ, the probability that the limit dose plan classifies correctly.

Ten Animal Fixed Sample (5000 mg/kg limit dose)

Table 1 shows the probability of correct classifications using the ten animal fixed sample test plan for the 5000 mg/kg limit dose.

Table 1

Probability of Correct Classification for Ten Animal Fixed Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	1.00
250	1.00	1.00	1.00	1.00	0.98
1500	1.00	1.00	1.00	0.92	0.84
2000	1.00	1.00	0.99	0.87	0.80
3000	1.00	1.00	0.93	0.78	0.73
5000	0.62	0.62	0.62	0.62	0.62
6000	0.92	0.69	0.54	0.44	0.42

Rule: five or more deaths classifies as under the limit dose. A classification is correct if the LD50 is 5000 or below, and the outcome leads to a classification of 5000 or below. It is also correct if the LD50 is 6000 and the outcome leads to a classification of over 5000.

Each entry in the table represents the probability that the correct classification would occur given the values of the LD50, σ and the classification rule of five or more deaths classifies the LD50 below the limit dose. Table 1 shows that the plan is very accurate for chemicals with low LD50s. For example, the ten animal test plan is perfect (to 2 decimal places) with LD50s between 1.5 and 3000 mg/kg for $\sigma = 0.12$ and 0.25. When $\sigma = 0.5$, there is a 93% correct classification rate at 3000 mg/kg. With σ at 2.0, there is a 98% correct classification rate at 250 mg/kg, 84% correct at 1500 mg/kg, 80% correct at 2000 and 73% correct at 3000.

To summarize the results from table 1, both low and high values of the LD50 produce the most accuracy.[1] Values close to the LD50 produce the least accuracy in fact, just above the limit dose of 5000 mg/kg, the accuracy is only (100%-62%=) 38%. The decision is correct at 5000 mg/kg if the outcome is consistent with under 5000 mg/kg. So at 5000 the probability of an incorrect decision is 38%. Just above 5000 mg/kg a decision is correct when the outcome is consistent with over 5000 mg/kg. For a dosage

[1] This finding is even more apparent in Appendix 1, which uses a limit dose of 2000 mg/kg., In the tables in the Appendix, 3000, 5000 and 6000 mg/kg are above the limit dose. The accuracy can be seen to increase as the LD50 becomes much greater than the limit dose.

infinitesmally greater than 5000, the outcomes would be just about the same as at 5000. So then the probability of a correct decision (over 5000) would be 38% and the probability of an incorrect decision (under 5000) would be 62%.

In a similar manner, increases in σ result in decreases in accuracy. Equation (1) shows that as σ increases, the term inside the parentheses approaches zero and the normal cumulative distribution function approaches 0.5. Consequently, when the LD50 is below the limit dose, increases in σ cause the accuracy to approach 62% asymptotically. When the LD50 is above the limit dose, increases in σ, would have the accuracy approaching 38%.

Also, increases in σ result in decreases in accuracy. However, the tests perform well in the upper part of the table, where the LD50 is low, representing the most toxic chemicals.

In the 10 animal fixed plan, the probability of a correct result when the LD50 is just below the limit dose is much greater than the probability of a correct result when the LD50 is slightly above the limit dose. This is a characteristic of a biased plan. Biased tests are discussed later in this paper.

Seven and Five Animal Sequential Test Plans

Tables 2 and 3 show seven and five animal sequential test plans.

Table 2

Probability of Correct Classification for Seven Animal Sequential Test Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.99
250	1.00	1.00	1.00	0.99	0.95
1500	1.00	1.00	0.99	0.89	0.82
2000	1.00	1.00	0.98	0.85	0.79
3000	1.00	0.98	0.90	0.78	0.74
5000	0.67	0.67	0.67	0.67	0.67
6000	0.72	0.53	0.43	0.37	0.35

Rule: LD50 is under limit dose if first animal dies, or 4 animals die. LD50 is over 5000 mg/kg if 4 animals survive.

Table 2 shows the same pattern as table 1. In comparing the probabilities between this plan and the 10 animal fixed plan of table 1, the results appear to be fairly close. The difference between correct classification probabilities for the two plans for LD50s at 3000 mg/kg and under is never more than 0.03. The difference of 0.03 is reached when σ is 0.5 at 3000 mg/kg, where table 1 shows 93% correct classification, while table 2 shows 90%. Also at σ = 1.25 and the LD50 of 1500, table 1 shows 92% correct classifications while table 2 shows 89%.

When the LD50 is equal to the limit dose, the seven animal sequential test plan has a correct classification probability of 67%, somewhat higher than the 62% in table 1. This means that for values slightly above the limit dose, the seven animal plan will be correct 33% of the time, while the 10 animal plan will be correct 38% of the time. For example as shown in table 1, 92% of the time chemicals with LD50s of 6000 mg/kg will be classified as above the limit dose at σ=0.12, while 72% of the time this will occur with the seven animal test plan.

Table 3 shows the correct classification probability from a five animal sequential test plan. The purpose of this table is to determine how much is lost by using a plan that would nominally have fewer animals.

Table 3

Probability of Correct Classification for Five Animal Sequential Test Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.98
250	1.00	1.00	1.00	0.98	0.93
1500	1.00	1.00	0.98	0.86	0.79
2000	1.00	1.00	0.96	0.82	0.76
3000	1.00	0.97	0.87	0.75	0.72
5000	0.66	0.66	0.66	0.66	0.66
6000	0.71	0.53	0.44	0.38	0.37

Rule: LD50 is under limit dose if first animal dies, or three animals die. LD50 is over if three animals survive.

As would be expected from a plan with fewer animals, the correct classification probabilities decrease somewhat from the seven animal plan in table 2. For LD50 values of 3000 mg/kg or lower, the largest difference between a five animal and ten animal plan is 6%. The largest differences occur in the same place as the seven animal plan compared with ten animals. These are at σ = 1.25 and LD50 = 1500 mg/kg (92% vs.

86%) and $\sigma = 0.5$ and LD50 = 3000 (93% vs. 87%). At an LD50 of 6000 mg/kg, the five animal test plan has almost the same results as the seven animal test plan, differing by less than 1% in probability of correct classification.

To summarize, five and seven animal sequential test plans produce very similar results to the ten animal fixed test plan. For low values of the LD50 the results are very close among all three plans. For values of the LD50s over the limit dose, the sequential plans tend to classify correctly less frequently than the ten animal fixed dose plan. This means that more chemicals would be erroneously considered to have the LD50 below the limit dose. This type of misclassification is probably better than erroneously classifying the LD50 above the limit dose.

Before comparing the five and seven animal sequential plans with fixed sample size plans, it is important to address bias in test plans.

Bias

Some definitions are necessary. An unbiased test plan classifies the LD50 as under the limit dose with exactly the same probability that a single animal would die when administered the limit dose. That means $p = P(LD50 \leq \text{Limit Dose})$, where p is the probability of death and the probability $P(LD50 \leq \text{Limit Dose})$ can be found in equation (3). In general most plans will be somewhat biased, because the two probabilities will not be exactly equal. This is really a small sample problem.[2]

However, many but not all limit dose tests will be unbiased when $p = 0.5$. Since the value of p in equation (1) is 0.5 when the limit dose is equal to the LD50, a biased plan occurs when there are more outcomes resulting in a classification of under (over) 5000 than over (under) 5000. This means that all fixed sample size plans with an even number of animals and a majority rule classification scheme are biased. For example, with a two animal plan, no deaths would classify the LD50 as over the limit dose, while two deaths would classify it as under the limit dose. The way that one death would be classified would determine the direction of the bias.

Plans can be arbitrarily made to be biased as well. A fixed or sequential sample plan with an odd number of animals could be almost unbiased. However, a sequential plan could stop after the first death (as shown in this paper) classifying the outcome as under the limit dose. This plan would then be biased.

---

[2]For a very simple example, consider a fixed test plan with 3 animals. Outcomes associated with classification of a chemical's LD50 above the limit dose would be 0 or 1 death, while 2 or 3 deaths would lead to classification below the limit dose. An unbiased plan would put the probability of classification below the limit dose at p. It can be shown that the probability of 2 or 3 deaths is $p^2(3-2p)$ where p is the probability that an animal dies. The probability the chemical is classified below the limit dose is can be shown to be below p for $p \leq 0.5$ and above p for $p > 0.5$. Some values for this probability of 2 or 3 deaths, i.e. the probability that the chemical is classified below the limit dose are 0.03 (p=0.1), 0.16 (p=0.25), 0.5 (p=0.5), 0.84 (p=0.75), and 0.97 (p = 0.9).

Comparison Between Five and Seven Animal Fixed Sample Size Plans

Table 4 shows the probability of correct classifications for seven animal fixed test plans. Recall that a fixed test plan involves dosing all the animals at once.

Table 4

Probability of Correct Classification for Seven Animal Fixed Test Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.99
250	1.00	1.00	1.00	0.99	0.92
1500	1.00	1.00	0.99	0.82	0.72
2000	1.00	1.00	0.96	0.76	0.67
3000	1.00	0.97	0.83	0.65	0.60
5000	0.50	0.50	0.50	0.50	0.50
6000	0.93	0.76	0.64	0.56	0.53

Rule: Classify as LD50 under the limit dose if four or more animals die, as over if four or more animals survive.

The differences between the seven animal plan and the ten animal plan are considerably greater than with the sequential plans considered in earlier tables. The reason is that the five and seven animal fixed plans are unbiased, in contrast to the sequential plans that are biased. For example, with an LD50 at 3000 mg/kg and σ =1.25, the ten animal plan had a 78% chance of a correct classification, while the seven animal plan in table 4 had a 65% probability Values of σ of 1.25 and 2.0 and LD50s between 1500 and 3000 generally had differences this large between the two plans. However, the seven animal fixed test plan classifies correctly more often than the ten animal plan for values of 6000 mg/kg. The seven animal plan is 76% correct at σ = 0.25 as compared with 69% for the ten animal plan. It is 53% correct, as compared with 42% correct at σ = 2.

For comparison, the five animal fixed sample test plan is shown below in table 5. The results are about the same as the seven animal plan with some small decreases in the percent correctly classified.

Table 5
Probability of Correct Classification for Five Animal Fixed Test Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.97
250	1.00	1.00	1.00	0.97	0.89
1500	1.00	1.00	0.97	0.78	0.69
2000	1.00	1.00	0.93	0.72	0.65
3000	1.00	0.95	0.80	0.63	0.58
5000	0.50	0.50	0.50	0.50	0.50
6000	0.89	0.72	0.62	0.55	0.53

Rule: Classify as LD50 under the limit dose if three or more animals die, as over if three or more animals survive.

Comparison between fixed and sequential sampling plans

Fixed and sequential sampling plans that have the same decision rules will have the same accuracy. This does not require empirical estimates, instead just the understanding that the sequential plan would be identical to the fixed sample plan if the sequential plan is required (unnecessarily) to be carried out even after enough animals have been tested to reach a decision.

But the five and seven animal sequential plans have different rules than the fixed plans. Recall that the sequential plans in this paper stop the test with the death of the first animal. This cannot be done with the fixed plans. The result is that the sequential plans in this paper are more accurate than fixed when the test uses chemicals that have LD50s below the limit dose. The fixed plans are more accurate with chemicals that have an LD50 above the limit dose. When the LD50 is very low or very high and σ is low, both types of tests perform accurately.

Expected Number of Animals Used in Sequential Tests

The benefit of the sequential sample size plans over fixed sample size plans is a decrease in the number of animals used in the test. The expected number of animals used in seven and five animal sequential tests are shown in tables 6 and 7 below.

Table 6

Expected Number of Animals in Seven Animal Sequential Test Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.01	1.16
50	1.00	1.00	1.00	1.23	1.73
250	1.00	1.00	1.02	1.68	2.26
1500	1.00	1.07	1.68	2.68	2.97
2000	1.00	1.24	2.02	2.87	3.09
3000	1.13	1.89	2.64	3.12	3.24
5000	3.41	3.41	3.41	3.41	3.41
6000	3.94	3.76	3.61	3.49	3.46

Note: for classification rules see table 2.

Table 6 shows that with low values of the LD50, on average slightly more than one animal is used. This is because the test plan calls for classifying LD50 as under the limit dose when the first animal dies. For chemicals with an LD50 of 1.5 or 50 or 250 mg/kg and a limit dose of 5000 mg/kg, survival of the first animal is unlikely.

On the other hand as the LD50 and σ or increases, more animals are required on average, approaching four. Four animals would be the exact number required for a chemical with an infinite LD50, as the most likely outcome to discontinue the test would be four survivals.

Table 7

Expected Number of Animals in Five Animal Sequential Test Plan
(Limit Dose = 5000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.01	1.12
50	1.00	1.00	1.00	1.17	1.53
250	1.00	1.00	1.01	1.49	1.87
1500	1.00	1.06	1.49	2.13	2.30
2000	1.00	1.18	1.71	2.24	2.37
3000	1.10	1.63	2.10	2.39	2.46
5000	2.56	2.56	2.56	2.56	2.56
6000	2.93	2.79	2.69	2.62	2.60

Note: for classification rules see table 3.

Five animal test plans, as shown in Table 7, use fewer animals on average than seven animal sequential test plans. At low LD50's where the most likely outcome is the death of the first animal, the two test plans are not very different in average number of animals. As the LD50 increases, the expected number of animals approaches three, one animal fewer, on average than the seven animal test plan. Three animals would be the exact number required for a chemical with an infinite LD50, because the test termination conditions would be three consecutive survivals.

Appendix 1 shows similar results for the 2000 mg/kg limit dose plan.

Conclusion

From the analysis it appears that sequential testing plans based on five and seven animals classify adequately. This is especially true when the LD50 is either far below or far above the limit dose. The classification deteriorates when the LD50 approaches the limit dose. Classifications are also less accurate when the variance of the dose response curve (symbolized as σ^2) increases.

Theoretically, fixed sample size and sequential plans would have identical accuracy with the same decision rules. However, in contrast to fixed plans, sequential plans can use the order of survivals and deaths as part of the decision rules. The model shows that fixed and sequential plans perform equally well when the LD50 is low relative

to the limit dose and ó is also reasonably low. When the LD50 gets close to the limit dose, the sequential plans tend to perform better than the fixed plans. For values of the LD50 that are above the limit dose, the fixed plans classify more accurately. And finally, as the LD50 continues to increase, the sequential plans start to catch up with the fixed plans in accuracy. The reason for these differences between plans is the use of the bias in the sequential plans. This bias makes it that the more toxic chemicals with low values of the LD50 will be classified correctly.

The other benefit of the sequential plans is that they use fewer animals than the fixed plans. The OECD recommended plan that uses up to five animals sequentially, will average three or fewer animals depending on the LD50 and σ. A seven animal sequential test plan averages up to four animals. The five animal sequential plan produces results that are almost as good as the present ten animal fixed sample plan while averaging one to three animals per test. That is seven to nine fewer animals than the ten animal fixed sample plan.

References

Code of Federal Regulations (1997), "Commerical Practices: Subchapter C-Federal Hazardous Substances Act Regulations, part 1500 to 1512, Revised as of January 1, 1997.

Dixon W J (1991), "Design and Analysis of Quantal Dose-Response Experiments (with Emphasis on Staircase Designs)." Dixon Statistical Associates, Los Angeles, CA.

OECD draft Guideline 425(2000).

Springer JA, Chambers WA, Green S, Gupta KC, Hills RN, Hurley PM, Lambert LA, Lee CC, Lee JK, Liu PT, Lowther DK, Roberts CD, Seabaugh VM and Wilcox NL (1993), "Number of Animals for Sequential Testing," *Food and Chemical Toxicology,* 31,2 pp 105-109.

Appendix 1
Limit Dose Test Results for 2000 mg/kg

The tables below present the limit test dose results for 2000 mg/kg. The order is the same as in the text. The first five tables present the probability of correct classifications as follows:

Table A1: Ten Animals, Fixed Sample Size
Table A2: Seven Animals Sequential Test
Table A3: Five Animals Sequential Test
Table A4: Seven Animals Fixed Sample Size
Table A5: Five Animals Fixed Sample Size

The last two tables present the expected numbers of animals in the seven and five animal sequential tests.

The results are generally the same as for the 5000 mg/kg dosages. The U-shaped probability function is more apparent in these tables because there are three values of the LD50 above the limit dose (3000, 5000 and 6000 mg/kg). In general the five animal variable sample size plan works adequately.

Table A1

Probability of Correct Classification for 10 Animal Fixed Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.99
250	1.00	1.00	1.00	0.99	0.93
1500	1.00	0.95	0.83	0.72	0.68
2000	0.62	0.62	0.62	0.62	0.62
3000	1.00	0.93	0.72	0.52	0.47
5000	1.00	1.00	0.96	0.69	0.58
6000	1.00	1.00	0.98	0.75	0.62

Majority Rule Classification.

Table A2

Probability of Correct Classification for Seven Animal Sequential Test Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.98
250	1.00	1.00	1.00	0.97	0.90
1500	0.99	0.92	0.82	0.73	0.71
2000	0.67	0.67	0.67	0.67	0.67
3000	0.93	0.73	0.55	0.42	0.39
5000	1.00	0.94	0.77	0.53	0.46
6000	1.00	0.97	0.82	0.57	0.48

Rule: LD50 is under limit dose if first animal dies, or four animals die. LD50 is over 2000 mg/kg if four animals survive.

Table A3

Probability of Correct Classification for Five Animal Sequential Test Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	0.99	0.96
250	1.00	1.00	1.00	0.94	0.87
1500	0.98	0.89	0.79	0.71	0.69
2000	0.66	0.66	0.66	0.66	0.66
3000	0.93	0.72	0.55	0.43	0.40
5000	1.00	0.94	0.76	0.53	0.46
6000	1.00	0.97	0.82	0.57	0.48

Rule: LD50 is under limit dose if first animal dies, or three animals die. LD50 is over 2000 mg/kg if three animals survive.

Table A4

Probability of Correct Classification for Seven Animal Fixed Test Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	1.00	0.96
250	1.00	1.00	1.00	0.94	0.84
1500	0.99	0.86	0.71	0.59	0.55
2000	0.50	0.50	0.50	0.50	0.50
3000	1.00	0.94	0.78	0.62	0.58
5000	1.00	1.00	0.96	0.76	0.67
6000	1.00	1.00	0.98	0.80	0.70

Rule: Classify as LD50 under the limit dose if four or more animals die, as over if four or more animals survive.

Table A5

Probability of Correct Classification for Five Animal Fixed Test Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.00	1.00
50	1.00	1.00	1.00	0.99	0.93
250	1.00	1.00	1.00	0.91	0.80
1500	0.97	0.83	0.68	0.57	0.55
2000	0.50	0.50	0.50	0.50	0.50
3000	1.00	0.91	0.75	0.60	0.57
5000	1.00	1.00	0.93	0.72	0.65
6000	1.00	1.00	0.96	0.76	0.67

Rule: Classify as LD50 under the limit dose if three or more animals die, as over if three or more animals survive.

Table A6

Expected Number of Animals in Seven Animal Sequential Test Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.03	1.25
50	1.00	1.00	1.00	1.44	2.01
250	1.00	1.00	1.15	2.14	2.62
1500	1.68	2.53	3.00	3.25	3.31
2000	3.41	3.41	3.41	3.41	3.41
3000	4.00	3.95	3.79	3.59	3.53
5000	4.00	4.00	3.97	3.76	3.65
6000	4.00	4.00	3.99	3.81	3.69

Note: for classification rules see table A2.

Table A7

Expected Number of Animals in Five Animal Sequential Test Plan
(Limit Dose = 2000 mg/kg)

			σ		
LD50	0.12	0.25	0.5	1.25	2
1.5	1.00	1.00	1.00	1.02	1.19
50	1.00	1.00	1.00	1.32	1.71
250	1.00	1.00	1.11	1.79	2.09
1500	1.49	2.03	2.31	2.47	2.50
2000	2.56	2.56	2.56	2.56	2.56
3000	3.00	2.94	2.81	2.68	2.64
5000	3.00	3.00	2.96	2.80	2.72
6000	3.00	3.00	2.98	2.83	2.75

Note: for classification rules see table A3.

Appendix 2
SAS Program

```
*********************************************
program to compute correct classification property and
  expected values of the number of animals used
  for limit doses

  michael a. greene
  division of hazard analysis
  us consumer product safety commission

  last modified 1/19/2000
*********************************************;

%macro pbinom(n,x,p);
   %* binomial pdf, used in data step;
   ((gamma(&n+1)/ (gamma(&x+1) * gamma(&n-&x+1)))
    * (&p**&x) * (1-&p)**(&n-&x))
%mend;

%macro prt(ds=,title=);
title  &title;
data _null_;  /* pretty printing */
  retain temp1-temp&nsigma;
  array  temp {*} temp1-temp&nsigma;
  file print;
  set &ds;
  by ld_50;
  if first.ld_50 then i=0;
  i+1;
  temp{i}=t_prob;
  if last.ld_50 then put ld_50 6.1 (temp{*})  (8.4);
%mend;


data doseres;                          * read in sigmas and ld50s;
  infile cards missover;
  retain sigma1-sigma99 ld1-ld99;
   input sigma1-sigma99;
   input ld1-ld99;
   call symput("nsigma",trim(left(put(n(of sigma1-sigma99),2.))));
   call symput("nld",   trim(left(put(n(of ld1-ld99)      ,2.))));
cards;
0.12 0.25 .5 1.25 2
1.5 50 250 1500 2000 3000 5000 6000
;

proc print data=doseres;
  var sigma1-sigma&nsigma ld1-ld&nld;
  title1 "dose response assumptions";
run;
```

```
*****************************************
this datastep uses the inputted slopes and ld50s from doseres
to compute the classification probabilities
*****************************************;

data test;
  retain dose 2000.;  *test dosage.  always 5000 micrograms per kg;

  keep   sigma ld_50 rule prob t_prob dose;
  retain sigma1-sigma&nsigma ld1-ld&nld;

  array  sigmaex {*} sigma1-sigma&nsigma;  /* animal char sigma */
  array  ld50x  {*}  ld1-ld&nld;           /* animal ld 50      */

  set doseres;                             /* ld50s and sigmas */

  do i = 1 to &nld;
     ld_50=ld50x{i};        /* get an ld50   */
     do j = 1 to &nsigma;
        sigma = sigmaex{j};/* get a sigma   */
        link getprob;      /* get the one animal death probability */
        link fillprob;     /* get multi animal death probabilites */
     end;
  end;
 return;

getprob:   /* probability of a single animal dying */
   prob = probnorm( (log10(dose) - log10(ld_50))/sigma);
/* probit fn */
return;

fillprob:

%inc "g:\users\epha\mag\pig\425\rule7.sas";  *y=# yx=expectedval;

return;

/* add up the cases by ld50 sigma and rule */
proc summary data=test;
  class ld_50 sigma rule;
  var t_prob;
  output out=new sum=t_prob;

data over under;
  set new;
  if _type_ = 7 & not(rule) then output under;
  else if _type_=7 & rule then output over;

%prt(ds=over,title="Over 5000");
run;
%prt(ds=under,title="Under 5000");
run;
```

```
* rule5.sas 5 animal variable plan;

  rule=0;  /* toxic  */

    t_prob=prob;                    output;
    *1 animal dies;

    t_prob=(1-prob)*(prob**3);      output;
    *S DDD;

    t_prob=(1-prob)*%pbinom(3,2,prob)*prob ;output;
    *S XXX D  XXX=2 of 3 D;

  rule=1; /* over */

    t_prob=(1-prob)**3;               output;
    *3 survivors;

    t_prob=(1-prob)*%pbinom(2,1,prob)*(1-prob); output;
    *S XX  S XX=1 of 2 D;

    t_prob=(1-prob)*%pbinom(3,2,prob)*(1-prob); output;
    *S XXX S XXX=2 of 3 D;




* rule5x.sas   expected value computation 5 animal variable plan;

  rule = 0;  /* toxic  …not used in expected value computations*/

    t_prob = prob;                     output;
    *1 animal dies;

    t_prob = 4* (1-prob)*(prob**3);    output;
    *S DDD;

    t_prob = 5*(1-prob)*%pbinom(3,2,prob)*prob ; output;
    *S XXX D  XXX=2 of 3 D;


    t_prob = 3*(1-prob)**3;             output;
    *3 survivors;

    t_prob = 4*(1-prob)*%pbinom(2,1,prob)*(1-prob);  output;
    *S XX S XX=1 of 2 D;

    t_prob=5*(1-prob)*%pbinom(3,2,prob)*(1-prob);   output;
    *S XXX S        XXX=2 of 3 D;
```