REPORT

(D)

FINAL REPORT

Lethality Rate Estimation and Testing Procedures

To

U.S. Army Medical Research

and Development Command

Institute of Chemical Defense

SEPTEMBER 11, 1989

Battelle

AD_____

MULTIPLE ANIMAL STUDIES FOR MEDICAL CHEMICAL DEFENSE PROGRAM IN SOLDIER/PATIENT DECONTAMINATION AND DRUG DEVELOPMENT

Subtitle: Lethality Rate Estimation and Testing Procedures

Final Report

Dr. Ronald L. Joiner
Mr. Frederick R. Todt
Dr. Paul I. Feder
Dr. Steven W. Rust

September 11, 1989

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-83-C-3129

Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201-2693

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Distribution unlimited; authorized for public release.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Battelle Memorial Institute; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: U.S. Army Medical Research Institute of Chemical Defense
6c. ADDRESS (City, State, and ZIP Code): 505 King Avenue, Columbus, OH 43201-2693	7b. ADDRESS (City, State, and ZIP Code): Aberdeen Proving Ground, MD 21010-5425
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: Contract DAMD17-83-C-3129
8c. ADDRESS (City, State, and ZIP Code): Fort Detrick, Frederick, Maryland 21701-5012	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 63751A; PROJECT NO. 3M4-63751D993; TASK NO. BB; WORK UNIT ACCESSION NO. 023

11. TITLE (Include Security Classification)
Multiple Animal Studies for Medical Chemical Defense Program in Soldier/Patient Decontamination and Drug Development

12. PERSONAL AUTHOR(S)
Steven W. Rust, Paul I. Feder, Frederick R. Todt, Ronald L. Joiner

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Draft Final	FROM 850101 TO 890101	890911	76

16. SUPPLEMENTARY NOTATION
Lethality Rate Estimation and Testing Procedures

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Statistical methods; historical comparison, model fitting, nonlinear regression analysis; RA V (A W)
06	15		
06	20		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
Statistical methods were developed for the incorporation of the historical data accumulated on animal studies to enhance the statistical sensitivity of individual comparisons between standards and candidates. The basic idea is that the historical levels and variability of the standard can be used to predict a likely range for the concurrent standard in testing.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Mrs. Virginia Miller	22b. TELEPHONE (Include Area Code): 301-663-7325; 22c. OFFICE SYMBOL: SGRD-RMI-S

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

EXECUTIVE SUMMARY

In 1985, the Medical Research and Evaluation Facility (MREF) developed a standardized first-stage screen (MREF Protocol 21, May 1985) to compare liquid or powder experimental decontaminants against the dual-component M258A1 skin decontamination kit for their effectiveness in mitigating the toxic effects of percutaneous exposure to organophosphate chemical surety materiels (CSM). The testing protocol calls for strict standardization of methods, materials, and agent doses across the individual screening tests that are spread out over a multi-year time frame. A concurrent standard decontaminant group is included in each individual test of the experimental decontaminants. Thus, a considerable data base is amassed over time pertaining to the standard decontamination procedure results.

A principal objective of this report is to describe and illustrate statistical methods for the incorporation of the historical data accumulated on the standard decontaminant results to enhance the statistical sensitivity of individual comparisons between the standard and experimental decontaminants. The basic idea is that the historical levels and variability of the standard decontaminant test results can be used to predict a likely range for the concurrent standard decontaminant test results. This information can be incorporated into the concurrent test procedures. The tradeoff in using this historical information is that the test procedure will be more sensitive if concurrent results fall within the range of the past results, but may perform worse in terms of the Type 1 error being too large or too small than a test that ignores the historical information if the concurrent standard decontaminant response level is substantially discrepant from the distribution of historical response levels. For this reason, the test procedure recommended in this report compares the experimental decontaminant response rate to a weighted average of the concurrent standard decontaminant response rate and the historical rate. The weight associated with the historical rate increases as the observed time to time variability in the historical data decreases and as the agreement between the concurrent rate and the historical average rate increases.

Several additional statistical aspects of the screen are also discussed. Control chart procedures are suggested to detect drifts over time or sudden jumps in the standard decontaminant responses rates. In the event the control charts indicate that the LD_{50} for the standard decontaminant has shifted, procedures are presented for carrying out studies to update the LD_{50}. These procedures are designed to conserve experimental resources. The dose allocation for the LD_{50} studies is carried out in a stagewise, adaptive fashion. The dose selection for each stage of the design is based on the test results from all previous stages. It is designed to accommodate unanticipated aspects of the dose-response relation. The stagewise, adaptive dose-allocation strategy introduces a number of nonstandard considerations that necessitate the use of specialized dose-response model fitting procedures. Specialized probit analysis model fitting procedures, based on nonlinear regression analysis, are discussed and illustrated by example.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

LETHALITY RATE ESTIMATION AND TESTING PROCEDURES

1.0 INTRODUCTION

In 1985 the Medical Research and Evaluation Facility (MREF) developed a standardized first stage screening test entitled "Assessment of Liquid or Powder Decontaminants Against GD, Thickened GD, and VX Administered Topically to Rabbits" (MREF Protocol 21, May 1985) to compare liquid or powder experimental decontaminants against the dual-component M258A1 skin decontamination kit for their effectiveness in mitigating the toxic effects of percutaneous exposure to organophosphate (OP) chemical surety materials (CSM). The standardized screen is based on a lethality endpoint in laboratory albino rabbits. An essential aspect of this testing protocol is the strict standardization of methods, materials, and agent dose levels (LD_{50} doses associated with standard decontamination procedure) that are used to screen numerous experimental decontaminants throughout a time period extending over multiple years.

Each test decontaminant is necessarily evaluated in the first-stage screen using just a limited number of animals (n = 24); a similarly small group of standard decontaminant animals (n = 24) is tested concurrently with the experimental decontaminant animals. It was recognized by MREF personnel in 1985 that a considerable data base would be amassed over time pertaining to the lethality rates associated with the standard decontamination procedure. This historical information can be incorporated into the efficacy comparisons with the experimental decontaminants to considerably increase the statistical sensitivity of these comparisons.

Initial statistical methods were adopted in 1985 to make use of the historical information. The agent dose for the screening test was set on the basis of an extensive LD_{50} study with standard decontaminant animals; the nominal standard lethality rate is thus 50 percent. The lethality rates observed in the concurrent standard decontaminant animals were compared to this nominal 50 percent value. If the concurrent lethality rate differs from the nominal by more than three standard deviations, then the concurrent test is considered suspect and both the standard and experimental decontaminant tests are repeated. If the concurrent lethality rate is within three standard

deviations of 50 percent, the experimental decontaminant lethality rates are compared to a fixed 50 percent standard. If the concurrent standard decontaminant lethality rate is between two and three standard deviations above the nominal, the experimental decontaminant lethality rate is also compared to the concurrent standard decontaminant lethality rate. The experimental decontamination procedure is declared to be inferior to the standard procedure if the observed lethality rate among the experimental decontaminant animals is statistically significantly greater ($P = 0.05$) than 0.5 (and the concurrent standard decontaminant lethality rate in those instances when it is compared to the concurrent standard also).

The present statistical procedures thus compare the experimental decontaminant lethality rate either to the concurrent standard decontaminant lethality rate or to a fixed 50 percent standard. This is an all or nothing procedure, with a discontinuity in the decision point. The current work was undertaken to build on these ideas and to refine them to arrive at a procedure which allows for compromises between the all or nothing use of the historical data. Under the updated procedure, the experimental decontaminant lethality rate is compared to a weighted average of the concurrent standard decontaminant lethality rate and the historical lethality rate. The weight associated with the concurrent standard lethality rate increases as the variability in the concurrent rate decreases (i.e., as more concurrent animals are tested), as the observed time to time variability in the historical standard lethality rates increases, and as the agreement between the concurrent rate and the historical rate decreases.

The updated procedure is based on the assumption that the historical standard lethality rates for individual tests are completely randomly distributed about an overall lethality rate, with no shifts or systematic drifts in the rates. Control chart methods are recommended as the surveillance procedure to monitor the validity of this assumption. If a shift or a systematic drift over time is detected in the historical standard decontaminant lethality rates, the data from the far past should be excluded and only the more recent past data should be used to form a historical average standard lethality rate, to be averaged with the concurrent standard lethality results. While specific procedures for the elimination of data from the

historical data set are not provided, the control chart procedures described in this report can be employed to determine when a change in the overall lethality rate has occurred, indicating the need for elimination of data from the historical data set.

It is also necessary to establish appropriate methods for redetermining an agent LD_{50} dose for standard decontaminant protected animals for the case when the ongoing surveillance procedure detects a significant shift in the standard decontaminant lethality rate, signaling a corresponding significant shift in the LD_{50} level. The LD_{50} determination procedure in the current version of MREF Protocol 21 (May 1985) calls for using a minimum of three replicates with 40 animals per replicate to determine the LD_{50}. If insufficient numbers of groups are obtained with observed lethality rates strictly between 0 and 1, then additional replicates may be required. Thus, 200 or more animals might be used to determine an LD_{50} value and associated confidence limits.

Procedures have been developed in conjunction with work carried out for MREF Tasks 85-18 and 87-34 to estimate the LD_{50} with acceptable precision based on many fewer animals. To accomplish this, it is necessary that the test animals be distributed among appropriate percentiles of the dose-response distributions for that agent and treatment regimen. These dose-response distribution percentiles should be centered around the true LD_{50} dose, with sufficient spread that the dose-response distribution slope may be determined. Since the dose-response relationships are either a priori unknown or just partially known based on historical data, the allocation of animals to agent doses is made in a stagewise, adaptive fashion as more and more information about the current dose-response relation becomes available.

The methods developed here should be considered as modifications, refinements, and improvements to the methods that have been used in conjunction with the previous MREF Protocol 21 screening program. The basic methodological concepts and approaches have remained unchanged.

Statistical problems occurring in the first-stage screening test are addressed in this report. The first problem, addressed in Section 2.0, is the development of a test procedure for comparing each experimental decontaminant

with the standard decontaminant. Computer programs to implement the recommended methods have been developed and are documented in Appendices A and B.

The second problem, addressed in Section 3.0, is the development of control chart procedures for monitoring the standard decontaminant lethality rates over time. Section 4.0 addresses the problem of comparing the standard decontaminant lethality rates observed in replicate subsets of the concurrent test. The results of the comparison determine whether these replicates can be pooled to arrive at an overall concurrent lethality rate, and if not which replicates differ from the others.

Section 5.0 addresses the problem of redetermining an LD_{50} dose level. It discusses procedures and associated computer programs to carry out stagewise, adaptive dose allocation designs when redetermining the LD_{50}. It also discusses specialized probit analysis model fitting procedures, based on nonlinear regression analysis, for updating the estimated standard decontaminant dose-response distribution following each stage in the LD_{50} study.

2.0 THE TESTING PROBLEM

Each time a first-stage screening test is performed for a set of experimental decontaminants, observed lethality rates at the standard decontaminant LD_{50} dose are obtained for the standard decontaminant and for each of the experimental decontaminants, based on a limited number of laboratory albino rabbits (nominally 24 animals for each decontaminant). A statistical model for the lethality data associated with the standard and experimental decontaminants is described in Section 2.1 and the testing problem is stated in terms of the parameters of this model. Recommended testing procedures are developed in Section 2.2 and are characterized in Section 2.3.

2.1 Statement of the Testing Problem

Each time a first-stage screen is performed for a set of experimental decontaminants, a limited number (n_c) of animals receive the standard decontaminant treatment and a nominal LD_{50} dose of agent. Also, limited numbers of animals receive each of the experimental decontaminant treatments and the same nominal LD_{50} dose of agent determined for the standard decontaminant. Each experimental decontaminant is compared to the standard decontaminant in a separate statistical test. Thus, without loss of generality, assume that there is only one experimental decontaminant being tested; denote the number of animals receiving this experimental decontaminant by n_t.

The number of lethalities obtained with the standard decontaminant (x_c) is assumed to have a binomial distribution with n_c trials and success probability p_c, where

$$\arcsin(\sqrt{p_c}) = \arcsin(\sqrt{\mu_c}) + \delta. \qquad (1)$$

μ_c is the long-term lethality rate for the standard decontaminant and δ is a random (block) effect associated with this particular first-stage screening test.

The number of lethalities for the experimental decontaminant (x_t) is assumed to have a binomial distribution with n_t trials and success probability p_t, where

$$\arcsin(\sqrt{p_t}) = \arcsin(\sqrt{\mu_t}) + \delta. \qquad (2)$$

μ_t is the long-term lethality rate for the experimental decontaminant and δ is the same random (block) effect as for the concurrent lethality rate, p_c.

The random effect term δ is included in the model to account for those sources of experimental variation that simultaneously affect the true lethality rates for all the standard and experimental decontaminants that are tested at the same time. It is assumed that δ is distributed as a mixture of two normal distributions: a normal$(0,\sigma_\delta^2)$ distribution with probability $1-\epsilon$

and a normal$(0,\sigma_{\delta_1}^2)$ distribution with probability ϵ. Selecting small values of $\sigma_{\delta_0}^2$ and ϵ and a larger value of $\sigma_{\delta_1}^2$ provides a model that results in small values of δ the majority of the time, but also allows for an occasionally large random effect. This reflects the situation where, on most occasions, the true lethality rates are close to their long run average values but on infrequent occasions may differ considerably from those long run values.

The arcsin-square root transformation is utilized in the models for p_c and p_t in anticipation of applying the same variance-stabilizing transformation to the observed lethality rates. Let $r_c = x_c/n_c$ and $r_t = x_t/n_t$ denote the observed lethality rates for the standard and experimental decontaminants, respectively. Also, let σ_c^2 denote $0.25/n_c$ and σ_t^2 denote $0.25/n_t$. For the purpose of deriving a test statistic for comparing the standard and experimental decontaminants, it will be assumed that the conditional distribution of $\arcsin(\sqrt{r_c})$ given δ is approximately normal$(\arcsin(\sqrt{p_c}),0.25/n_c)$ or normal$(\arcsin(\sqrt{\mu_c})+\delta,\sigma_c^2)$. Similarly, it will be assumed that the conditional distribution of $\arcsin(\sqrt{r_t})$ given δ is approximately normal$(\arcsin(\sqrt{p_t}),0.25/n_t)$ or normal$(\arcsin(\sqrt{\mu_t})+\delta,\sigma_t^2)$.

The purpose of the first-stage screen is to eliminate those and only those experimental decontaminants from consideration that are obviously inferior to the standard decontaminant. The problem is to test the null hypothesis H_0: $\mu_t \leq \mu_c$ versus the alternative H_1: $\mu_t > \mu_c$ and to fail the experimental decontaminant if the null hypothesis can be rejected in favor of the alternative that the experimental decontaminant is inferior.

2.2 Recommended Testing Procedures

Before beginning the development of a test procedure, we must first determine the criteria that the test must satisfy. Consider the criterion that "the test must have a (conditional) significance level of $\alpha = 0.05$, conditioning on the value of the random effect δ at the time the screening test is performed". If the significance level $\alpha = 0.05$ is to be attained for every individual realization of δ, then δ is being treated as a fixed nuisance parameter and so the information concerning the random behavior of δ cannot be used. Lehmann (1959) indicates that the standard two-sample test is best in

this situation. However, the two-sample binomial test does not make use of the historical information available concerning the fluctuations of δ, and therefore, it is not fully efficient. For this reason, the above criterion was not pursued.

Consider the alternative criterion that "the test must have a (unconditional) significance level of $\alpha = 0.05$ where the random nature of δ is factored into the significance level" and assume that the null hypothesis will be rejected if $\arcsin(\sqrt{r_t}) > k$, where k is a critical value to be determined.

This formulation adopts a random effects viewpoint. Namely, hypothesis tests to compare concurrent standard and test decontamination procedures will be carried out numerous times throughout the course of the screening program. The random (block) effect δ will vary across these tests according to a probability distribution, which can be estimated based on historical data and which is discussed below. If we select a test at random from the "population" of such tests, we wish that it has specified Type 1 error level α (e.g., $\alpha = 0.05$). If δ varies regularly and randomly across tests, then this formulation permits us to utilize the information about the standard decontamination procedure lethality rates observed in the previous tests to obtain a more precise estimate of the current standard decontamination procedure lethality rate. This, in turn, results in increased sensitivity of the current test of hypothesis to compare the response rates of the current standard and experimental decontamination procedures, relative to what would be obtained if the historical data were ignored.

The price for this is that this test procedure may perform more poorly (e.g., in terms of significance levels that are too large or too small) for values of δ that are substantially discrepant from past values. The variable weighting scheme proposed in this report accounts for the possibility of occasional values of δ that are somewhat discrepant from historically observed values.

Since the distribution of r_c does not depend on μ_t (the parameter about which we are making an inference), the principal of conditionality implies that k should be determined from the conditional null distribution of $\arcsin(\sqrt{r_t})$, given r_c. Denote this conditional null distribution by $F(\arcsin(\sqrt{r_t})|r_c)$. Then k should be set equal to $F_{1-\alpha}(\arcsin(\sqrt{r_t})|r_c)$, the

100(1-α)th percentile of the conditional null distribution. Since the conditional significance level is equal to α for each value of r_c, the overall marginal significance level is exactly α. However, the conditional (on δ) significance level is a function of δ, which can be significantly greater than α for values of δ beyond the primary support of the assumed probability distribution of δ, although such values of δ should occur only very rarely based on the previous history of the screen.

Consider first the special case when $\epsilon = 0$, so that δ is normal$(0,\sigma_{\delta_0}^2)$. Then, under the null hypothesis that $\mu_t = \mu_c$, $F(\arcsin(\sqrt{r_t})|r_c)$ is approximately normal$(\mu_\theta,\sigma_\theta^2)$, where

$$\mu_\theta = \arcsin(\sqrt{\mu_c}) + w\,(\arcsin(\sqrt{r_c}) - \arcsin(\sqrt{\mu_c})) \qquad (3)$$
$$= (1 - w)\arcsin(\sqrt{\mu_c}) + w \arcsin(\sqrt{r_c}),$$

$$\sigma_\theta^2 = \sigma_t^2 + w\,\sigma_c^2, \qquad (4)$$

and

$$w = \sigma_{\delta_0}^2 / (\sigma_{\delta_0}^2 + \sigma_c^2). \qquad (5)$$

Equation (3) demonstrates that the mean μ_θ is a weighted average of the transformed long-term lethality rate for the standard decontaminant ($\arcsin(\sqrt{\mu_c})$) and the transformed observed lethality rate for the standard decontaminant based on the concurrent control animals. Note that the weight given to the currently observed lethality rate increases as the variability of the random effect δ ($\sigma_{\delta_0}^2$) increases and as the variability of the transformed observed lethality rate for the standard decontaminant (σ_c^2) decreases. However, the weight does not depend on r_c.

Thus for $\epsilon = 0$, the critical value is $k = \mu_\theta + 1.645\,\sigma_\theta$. The critical region $\arcsin(\sqrt{r_t}) > k$ employing this critical value satisfies the criterion stated at the beginning of this section. However, for large values of $|\delta|/\sigma_{\delta_0}$, the conditional (on δ) significance level can be as large as 1. Such values of the ratio are highly unlikely under the assumption that $\epsilon = 0$, yet they are of considerable concern in the development of the hypothesis testing procedure. This suggests that a single normality assumption for δ may not reflect the true state of prior feelings about the performance of the test

system. It is for this reason that we have chosen to assume that δ is distributed as a mixture of two normal distributions with $\varepsilon > 0$. Such a model allows for the occurrence, on infrequent occasions, of more extreme values of δ than would be predicted by a simple normal distribution model.

Under the mixture model, the random variable δ can be written as

$$\delta = (1-I)\ Y_0 + I\ Y_1, \tag{6}$$

where I has a binomial distribution with 1 trial and success probability ϵ and Y_0 and Y_1 have independent normal$(0,\sigma_{\delta_0}^2)$ and normal$(0,\sigma_{\delta_1}^2)$ distributions. Then the conditional distribution of I given r_c is binomial with 1 trial and success probability

$$\epsilon^* = \epsilon R\ /\ (1 - \epsilon + \epsilon R), \tag{7}$$

where

$$R = \frac{\phi(\arcsin(\sqrt{r_c});\arcsin(\sqrt{\mu_c}),\sigma_{\delta_1}^2+\sigma_c^2)}{\phi(\arcsin(\sqrt{r_c});\arcsin(\sqrt{\mu_c}),\sigma_{\delta_0}^2+\sigma_c^2)}\ . \tag{8}$$

$\phi(x;\mu,\sigma^2)$ is the normal density function with mean μ and variance σ^2 evaluated at x. Further, $F(\arcsin(\sqrt{r_t})|r_c)$ is a mixture of two normal distributions: a normal(μ_0,σ_0^2) distribution with probability $1 - \epsilon^*$ and a normal(μ_1,σ_1^2) distribution with probability ϵ^*, where

$$\mu_i = \arcsin(\sqrt{\mu_c}) + w_i\ (\arcsin(\sqrt{r_c}) - \arcsin(\sqrt{\mu_c})) \tag{9}$$

$$= (1 - w_i)\ \arcsin(\sqrt{\mu_c}) + w_i\ \arcsin(\sqrt{r_c}),$$
$$\sigma_i^2 = \sigma_t^2 + w_i\ \sigma_c^2, \tag{10}$$

and

$$w_i = \sigma_{\delta_i}^2\ /\ (\sigma_{\delta_i}^2 + \sigma_c^2). \tag{11}$$

The critical value k is then the 95th percentile of this mixture distribution and the recommended test procedure is to reject the null hypothesis H_0: $\mu_t \leq \mu_c$ in favor of the alternative H_1: $\mu_t > \mu_c$, if $\arcsin(\sqrt{r_c}) > k$.

The critical value k can be determined in an iterative fashion, starting with $k_0 = \mu + 1.645\sigma$, where $\mu = (1-\epsilon^*)\mu_0 + \epsilon^*\mu_1$ and $\sigma = (1-\epsilon^*)\sigma_0 + \epsilon^*\sigma_1$. The ith iterative solution (k_i) can be expressed as a function of the previous solution (k_{i-1}) as follows:

$$k_i = k_{i-1} + (0.95 - F(k_{i-1}|r_c)) \;/\; f(k_{i-1}|r_c), \tag{12}$$

where

$$F(k_{i-1}|r_c) = (1-\epsilon^*)\,\Phi(k_{i-1};\mu_0,\sigma_0^2) + \epsilon^*\,\Phi(k_{i-1};\mu_1,\sigma_1^2), \tag{13}$$

$$f(k_{i-1}|r_c) = (1-\epsilon^*)\,\phi(k_{i-1};\mu_0,\sigma_0^2) + \epsilon^*\,\phi(k_{i-1};\mu_1,\sigma_1^2), \tag{14}$$

$\Phi(x;\mu,\sigma^2)$ is the cumulative normal distribution function with mean μ and variance σ^2 evaluated at x, and $\phi(x;\mu,\sigma^2)$ is the normal density function with mean μ and variance σ^2 evaluated at x. The iterative process should be continued until $F(k_i|r_c)$ is sufficiently close to 0.95.

It should be noted that, under the mixture model, the relative weighting of the concurrent and historical standard decontaminant responses depends on the value of r_c. The farther r_c is from μ_c, the larger is R. This implies that increasingly more weight is given to the normal(μ_1,σ_1^2) distribution, which in turn implies that increasingly more weight is given to the current r_c in the determination of the concurrent standard decontaminant response rate.

Based on a limited examination of the information available to support a selection of the parameters ϵ, σ_{δ_0}, and σ_{δ_1}, it is recommended that the values $\epsilon = 0.1$, $\sigma_{\delta_0} = 0.1$, and $\sigma_{\delta_1} = 0.4$ be used initially. It should be noted that the primary reasons for the selection of these values are that they are consistent with the historical database and they appear to provide a test procedure with desirable overall properties as illustrated in Section 2.3. Further work must be performed to develop a procedure for the selection of test procedure parameters, allowing these parameters to vary with the agent and test system. The following example illustrates the use of the procedure for $n_t = n_c = 24$ and $\mu_c = 0.5$.

Suppose we have $n_t = n_c = 24$, $\mu_c = 0.5$, $\epsilon = 0.1$, $\sigma_{\delta_0} = 0.1$, and $\sigma_{\delta_1} = 0.4$. Then:

$$\arcsin(\sqrt{\mu_c}) = \pi/4 = 0.7854$$

$$R = \frac{\phi(\arcsin(\sqrt{r_c});0.7854,0.1704)}{\phi(\arcsin(\sqrt{r_c});0.7854,0.0204)}$$

$$\epsilon^* = 0.1\ R/(0.9 + 0.1\ R)$$
$$= R/(9 + R)$$

$$w_0 = 0.01/0.0204 = 0.4898$$

$$w_1 = 0.16/0.1704 = 0.9389$$

$$\mu_0 = 0.7854 + 0.4898\ (\arcsin(\sqrt{r_c}) - 0.7854)$$

$$\mu_1 = 0.7854 + 0.9389\ (\arcsin(\sqrt{r_c}) - 0.7854)$$

$$\sigma_0^2 = 0.0104 + 0.4898\ (0.0104) = 0.0155$$

$$\sigma_0 = 0.1246$$

$$\sigma_1^2 = 0.0104 + 0.9389\ (0.0104) = 0.020197$$

$$\sigma_1 = 0.1421$$

Values of R, ϵ^*, $\mu 0$, $\mu 1$, k, and $\sin^2(k)$ are listed in Table 2.2.1 for various values of r_c. Either of the last two columns of Table 2.2.1 can be used to easily carry out the procedure by rejecting the null hypothesis H_0: $\mu_t \leq \mu_c$ in favor of the alternative H_1: $\mu_t > \mu_c$ if $\arcsin(\sqrt{r_t}) > k$ or if $r_t > \sin^2(k)$.

2.3 Characterization of the Testing Procedures

In this section, we characterize and compare the performance of the recommended test procedure with that of the current test procedure and the standard two-sample binomial test procedure. The current test procedure involves the following steps:

TABLE 2.2.1. EXAMPLE CALCULATIONS FOR THE PROPOSED TEST PROCEDURE WITH $n_t = n_c = 24$, $\mu_c = 0.5$, $\epsilon = 0.1$, $\sigma_{\delta_0} = 0.1$, AND $\sigma_{\delta_1} = 0.4$

r_c	R	ϵ^*	μ_0	μ_1	k	$\sin^2(k)$
0.00	206,026.54	1.00	0.40	0.05	0.28	0.08
0.10	35.62	0.80	0.56	0.35	0.67	0.38
0.20	3.22	0.26	0.63	0.48	0.82	0.53
0.30	0.86	0.09	0.68	0.59	0.89	0.60
0.40	0.43	0.05	0.74	0.69	0.94	0.65
0.50	0.35	0.04	0.79	0.79	0.99	0.70
0.60	0.43	0.05	0.83	0.88	1.04	0.75
0.70	0.86	0.09	0.89	0.98	1.11	0.80
0.80	3.22	0.26	0.94	1.09	1.23	0.89
0.90	35.62	0.80	1.01	1.22	1.44	0.98
1.00	206,026.54	1.00	1.17	1.52	1.57	1.00

(1) Calculate the trigger statistic

$$Z = (r_c - \mu_c)/(\mu_c(1-\mu_c)/n_c)^{1/2}. \tag{15}$$

(2) If $Z \leq 2$, calculate the test statistic

$$T_1 = (r_t - \mu_c)/(\mu_c(1-\mu_c)/n_t)^{1/2} \tag{16}$$

and reject the null hypothesis H_0: $\mu_t \leq \mu_c$ in favor of the alternative H_1: $\mu_t > \mu_c$ if $T_1 > 1.645$.

(3) If $Z > 2$, calculate the test statistic

$$T_2 = \frac{r_t - r_c}{\left[\frac{r_t(1-r_t)}{n_t} + \frac{r_c(1-r_c)}{n_c}\right]^{1/2}} \tag{17}$$

and reject the null hypothesis H_0: $\mu_t \leq \mu_c$ in favor of the alternative H_1: $\mu_t > \mu_c$ if $T_2 > 1.645$.

The standard two-sample binomial test procedure is based on the test statistic T_2 defined above regardless of the value of the trigger statistic and the null hypothesis H_0: $\mu_t \leq \mu_c$ is rejected in favor of the alternative H_1: $\mu_t > \mu_c$ if $T_2 > 1.645$.

In Tables 2.3.1 and 2.3.2, the three test procedures are characterized in terms of the probability of rejecting the experimental decontaminant under the assumption that $\mu_c = 0.5$. In Table 2.3.1, the tabled values are the exact conditional probabilities (x 1,000) of rejecting the experimental decontaminant conditioned on the value of the lethality rate (p_c) in effect at the time of the test. The input parameters used for the recommended procedure are $\sigma_{\delta_0} = 0.1$, $\epsilon = 0.1$, $\sigma_{\delta_1} = 0.4$. It is assumed that $n_t = n_c = 24$ and $\mu_c = 0.5$.

TABLE 2.3.1 CONDITIONAL PROBABILITIES (x 1,000) OF REJECTING EXPERIMENTAL DECONTAMINANT FOR SPECIFIC VALUES OF μ_t AND p_c ASSUMING THAT $n_t = n_c = 24$ and $\mu_c = 0.5$

μ_t	p_c								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	062[a]	018	011	018	042	091	158	178	156
	000[b]	000	000	002	031	155	249	106	056
	055[c]	056	057	054	056	054	057	056	055
0.6	161	069	065	104	194	327	452	489	
	000	000	002	032	186	458	426	276	
	250	219	189	184	184	189	219	250	
0.7	340	215	235	344	515	696	803		
	000	002	032	192	547	750	623		
	569	475	427	423	427	475	569		
0.8	579	481	542	695	852	947			
	002	032	192	564	884	886			
	826	733	702	702	733	826			
0.9	810	770	842	939	989				
	032	192	565	909	981				
	951	909	900	909	951				

[a] Top value is for recommended test procedure ($\sigma_{\delta_0} = 0.1$, $\epsilon = 0.1$, $\sigma_{\delta_1} = 0.4$)
[b] Middle value is for current test procedure
[c] Bottom value is for two-sample binomial test procedure

TABLE 2.3.2. UNCONDITIONAL PROBABILITIES (x 1,000) OF REJECTING EXPERIMENTAL DECONTAMINANT FOR SPECIFIC VALUES OF μ_t AND VARIOUS DISTRIBUTIONS FOR THE RANDOM EFFECT δ WITH $\sigma_{\delta_1} = 0.4$ ASSUMING THAT $n_t = n_c = 24$ and $\mu_c = 0.5$

	$\epsilon = 0$	$\epsilon = 0.05$			$\epsilon = 0.10$			$\epsilon = 0.20$		
μ_t	σ_{δ_0}	σ_{δ_0}			σ_{δ_0}			σ_{δ_0}		
	0	0.05	0.10	0.20	0.05	0.10	0.20	0.05	0.10	0.20
0.5	042[a]	051	057	074	055	059	077	066	068	082
	031[b]	043	064	074	047	067	072	049	063	068
	056[c]	055	053	053	055	051	054	054	054	053
0.6	194	195	207	222	202	212	223	208	224	233
	186	197	210	192	192	207	190	189	193	182
	184	179	188	182	176	181	179	168	183	181
0.7	515	515	504	483	511	505	492	508	508	495
	547	510	468	385	499	451	379	471	436	359
	427	424	429	425	419	427	425	413	418	418
0.8	852	835	823	779	835	815	787	829	816	784
	884	821	752	612	801	732	608	765	711	593
	733	725	732	720	725	722	719	712	718	709
0.9	989	984	980	966	982	978	965	979	972	965
	981	955	936	837	940	918	825	896	880	809
	951	944	942	924	938	937	917	921	918	911

[a] Top value is for recommended test procedure ($\sigma_{\delta_0} = 0.1$, $\epsilon = 0.1$ $\sigma_{\delta_1} = 0.4$)
[b] Middle value is for current test procedure
[c] Bottom value is for two-sample binomial test procedure

The values in Table 2.3.1 for $\mu_t = 0.5$ are the conditional probabilities that an experimental decontaminant with the same lethality rate as the standard decontaminant would be rejected (false rejection rate) in a particular screening test for various values of p_c. Note that the two-sample binomial test procedure has a relatively constant false rejection rate for all values of p_c. The false rejection rate of the current test procedure decreases rapidly as p_c becomes smaller than 0.5 and increases rapidly as p_c goes from 0.5 to approximately 0.7 and then decreases as p_c goes from approximately 0.7 to 0.9. The false rejection rate for the recommended test procedure follows the same pattern as that for the current test procedure but with less rapid increases and decreases. The remainder of Table 2.3.1 illustrates that the current and recommended test procedures have better conditional power relative to the two-sample binomial test procedure for values of p_c greater than 0.5 and worse power for values of p_c less than 0.5.

In Table 2.3.2, the tabled values are estimates of the unconditional probabilities (x 1,000) of rejecting the experimental decontaminant where the random effect δ (and therefore p_c) is allowed to vary according to an assumed probability distribution. Again, the input parameters used for the recommended procedure are $\sigma_{\delta_0} = 0.1$, $\epsilon = 0.1$, $\sigma_{\delta_1} = 0.4$. It is assumed that $n_t = n_c = 24$, $\mu_c = 0.5$ and that $\sigma_{\delta_1} = 0.4$ for all the assumed distributions for δ.

Each value in Table 2.3.2 is the result of 10,000 replications of the following process. Generate a δ value from the mixture of two normal distributions defined by $\sigma_{\delta_1} = 0.4$ and the values of σ_{δ_0} and ϵ at the top of the column. Generate independent binomial test results for the standard and experimental decontaminant using the values of p_c and p_t defined by the $\mu_c = 0.5$, μ_t, and δ. Record the result of each of the three test procedures based on this simulated test data.

The values in Table 2.3.2 for $\mu_t = 0.5$ are the probability that an experimental decontaminant with the same lethality rate as the standard decontaminant would be rejected (false rejection rate) for various distributions of the random effect δ. These values illustrate that the false rejection rate is reasonably controlled by all three test procedures in the neighborhood of the assumed distribution for δ ($\sigma_{\delta_0} = 0.1$, $\epsilon = 0.1$, $\sigma_{\delta_1} = 0.4$).

The remainder of Table 2.3.2 illustrates that the recommended test procedures have better power relative to the current test procedure and the two-sample binomial test procedure in the neighborhood of the assumed distribution for δ ($\sigma_{\delta_0} = 0.1$, $\epsilon = 0.1$, $\sigma_{\delta_1} = 0.4$). This increased power is the motivation for the development of the recommended test procedure.

3.0 THE CONTROL CHART PROBLEM

Over time, a data base is accumulated for the standard decontaminant, consisting of the observed lethality rates from the individual screening tests. A statistical model for the lethality data associated with the standard decontaminant is described in Section 3.1 and the control chart problem is stated in terms of the parameters of this model. The problem of estimating the model parameters from the historical database is discussed in Section 3.2. Recommended control chart procedures are developed in Section 3.3 and characterized in Section 3.4.

3.1 Statement of the Control Chart Problem

Each time a first-stage screening test is performed for a set of experimental decontaminants, a limited number of animals receive the standard decontaminant treatment and a nominal LD_{50} dose of agent. Let

- k = the number of first stage screening tests in the historical database,
- n_i = the number of animals receiving the standard decontaminant during the ith screening test, and
- x_i = the number of lethalities observed with the standard decontaminant during the ith screening test.

x_i is assumed to have a binomial distribution with n_i trials and success probability p_i, where

$$\arcsin(\sqrt{p_i}) = \arcsin(\sqrt{\mu}) + \delta_i. \qquad (18)$$

μ is the long-term lethality rate for the standard decontaminant, and δ_i is a random effect associated with the ith screening test.

The random effect term δ_i is included in the model to account for all factors that randomly affect the true lethality rate for the standard decontaminant from test to test. It is assumed that each δ_i has a Normal$(0,\sigma_{\delta 2})$ distribution and that the δ_i's associated with separate tests are statistically independent of one another.

The arcsin-square root transformation is utilized in the model for p_i in anticipation of applying the same variance-stabilizing transformation to the observed lethality rates. Let $r_i = x_i/n_i$ denote the observed lethality rate for the ith screening test. Also, let σ_i^2 denote $0.25/n_i$. For the purpose of deriving control charting procedures for monitoring the standard decontaminant lethality rate over time, it will be assumed that the conditional distribution of $\arcsin(\sqrt{r_i})$, given δ, is approximately normal$(\arcsin(\sqrt{p_i}),0.25/n_i)$ or normal$(\arcsin(\sqrt{\mu})+\delta_i,\sigma_i)$.

The purpose of the control chart procedures is to monitor the standard decontaminant lethality rate over time to detect any shifts or trends that may occur as a result of random and inadvertent variations in methods, materials, or agent doses employed. The problem is thus to plot (a standardized version of) $r_1, r_2, \ldots, r_k$ versus time along with upper and lower control limits that characterize the expected extreme variations according to the statistical model and associated parameter estimates. Values beyond the control limits are evidence that either the statistical model or the parameter estimates being employed may no longer be valid. Aspects of the tests may then need to be adjusted, for example, by adjusting the agent dose.

3.2 Estimating Model Parameters

The first step in forming a control chart for the standard decontaminant lethality rate is to determine the values of μ and σ_δ^2 that will be assumed in the statistical model. The following procedure may be used to estimate these parameters from the historical database. Let

$$\hat{\mu} = \sin^2(M), \qquad M = \frac{\Sigma\, w_i \arcsin(\sqrt{r_i})}{\Sigma\, w_i} \tag{19}$$

and

$$w_i = \frac{1}{\sigma_i^2 + \theta_\delta^2}. \tag{20}$$

w_i is an approximation to the inverse of the variance of arcsin($|r_i|$). θ_δ^2 is defined below in equation (22). Then $\hat{\mu}$ is an unbiased estimate of μ with approximate standard deviation

$$SE(\hat{\mu}) = 2 \left| \sin(M) \cos(M) \right| (1/\Sigma w_i)^{1/2}. \tag{21}$$

Let

$$\theta_\delta^2 = \frac{\Sigma z_i \theta_i^2}{\Sigma z_i}, \tag{22}$$

where

$$z_i = \frac{1}{2\left[\sigma_i^2 + \theta_\delta^2\right]^2} \tag{23}$$

and

$$\theta_i^2 = (\arcsin(|r_i|) - M)^2 - \sigma_i. \tag{24}$$

z_i is an approximation to the inverse of the variance of θ_i^2. Then θ_δ^2 is an approximately unbiased estimate of σ_δ^2 with approximate standard deviation

$$SE(\theta_\delta^2) = (1/\Sigma z_i)^{1/2}. \tag{25}$$

All summations above are over $i = 1,\ldots,k$.

Because equations (20), (23), and (24) involve the parameter estimates, equations (19) and (22) must be solved in an iterative fashion. The following procedure may be employed. Begin with initial values, say

$M = \pi/4 = 0.7854$ ($\beta = 0.5$) and $\sigma_\delta^2 = 0.01$, and solve equations (19) and (22) using these values in equations (20), (23), and (24). Repeatedly solve equations (19) and (22), substituting the estimates from the previous iteration into equations (20), (23), and (24) until the estimates do not change appreciably from one iteration to the next.

3.3 Recommended Control Chart Procedures

As stated previously, the purpose of the control chart procedures is to monitor the standard decontaminant lethality rates over time to provide a timely signal in the event that the statistical model or the parameter estimates being employed for the standard decontaminant are no longer valid. Since the number of animals used may vary from test to test, it is most convenient to standardize the individual transformed lethality rates to achieve approximate uniform variance over time. Let

$$Z_i = \frac{\arcsin(\sqrt{r_i}) - \arcsin(\sqrt{\mu_{c0}})}{\left[\frac{0.25}{n_c} + \sigma_{\delta 0}^2\right]^{1/2}}, \qquad (26)$$

where μ_{c0} and $\sigma_{\delta 0}$ are the assumed values of the parameters μ_c and σ_δ. Then $Z_1, \ldots, Z_k$ are distributed approximately as independent standard normal random variables. It is recommended that $Z_1, \ldots, Z_k$ be plotted (as the vertical variable) versus time (as the horizontal variable) along with horizontal lines across the entire plot at -3.00, 0, and 3.00. Three types of control chart procedures are considered:

(A) If the current Z-value falls below -3.00 or above 3.00, the test involving this observation should be repeated. If the repeated test also exceeds these limits, the cause of this exceedance should be investigated and corrected.

(B) If the three previous Z-values fall either all below -1.22 or all above 1.22, this exceedance is strong evidence that a shift in the process has occurred. The cause of this shift should be investigated and corrected.

(C) If the seven previous Z-values fall either all below -0.28 or all above 0.28, this exceedance is strong evidence that a shift in the process has occurred. The cause of this shift should be investigated and corrected.

These three procedures provide short-term, intermediate, and long-term tests respectively for a shift in process behavior. As with the testing procedure, it is recommended that the value σ_{δ_0} be set to 0.1 initially.

If any of the three control chart procedures exceed their critical values, the historical database should be scrutinized. If some of the older data are no longer pertinent to current tests, they might be eliminated from calculations of model parameters. Consideration might be given to carrying out a new LD_{50} study and adjusting the agent dose.

The above procedures are designed to be easy to carry out with a calculator and a simple plot of the data. Similar tests based on the median can be employed as follows. Let

M_1 = the Z-value (Equation 26) for the most previous test

M_3 = the median of the Z-values (Equation 26) for the three previous tests, and

M_7 = the median of the Z-values (Equation 26) for the seven previous tests.

The following procedures are analogous to procedures A, B, and C defined above.

(A2) If M_1 falls below -3.00 or above 3.00, the test involving this observation should be repeated. If the repeated test also exceeds these limits, the cause of this exceedance should be investigated and corrected.

(B2) If M_3 falls below -2.17 or above 2.17, this exceedance is strong evidence that a shift in the process has occurred. The cause of this shift should be investigated and corrected.

(C2) If M_7 falls below -1.42 or above 1.42, this exceedance is strong evidence that a shift in the process has occurred. The cause of this shift should be investigated and corrected.

Similar tests based on the mean can be employed as follows. Let

T_1 = the Z-value (Equation 26) for the most previous test

T_3 = the average of the Z-values (Equation 26) for the three previous tests, and

T_7 = the average of the Z-values (Equation 26) for the seven previous tests.

The following procedures are analogous to procedures A, B, and C and A2, B2, and C2 defined above.

(A3) If T_1 falls below -3.00 or above 3.00, the test involving this observation should be repeated. If the repeated test also exceeds these limits, the cause of this exceedance should be investigated and corrected.

(B3) If T_3 falls below -1.73 or above 1.73, this exceedance is strong evidence that a shift in the process has occurred. The cause of this shift should be investigated and corrected.

(C3) If T_7 falls below -1.13 or above 1.13, this exceedance is strong evidence that a shift in the process has occurred. The cause of this shift should be investigated and corrected.

3.4 Characterization of the Control Chart Procedures

In Table 3.4.1, the three sets of control chart procedures [(A,B,C), (A2,B2,C2), and (A3,B3,C3)] are characterized in terms of the probability of signaling a process shift. The model parameters employed are $\mu_{c\theta} = 0.5$ and $\sigma_{\delta\theta} = 0.1$ and it is assumed that $n_i = 24$.

The values in Table 3.4.1 are the probability that a process shift will be detected by the various procedures. The results are based on the assumption that the Z values are independent standard normal random variables. The results for the B2 and C2 procedures are based on a normal approximation to the distribution of median of independent standard normal random variables. While the procedures based on counts and medians are simpler to carry out, it is recommended that the procedures based on means (A3, B3, and C3) be employed due to the significant power advantage for detecting moderate shifts in the standard decontaminant lethality rate.

4.0 COMPARISONS OF THE STANDARD DECONTAMINANT LETHALITY RATES AMONG REPLICATES WITHIN TESTS

Each screening test involves the simultaneous testing of n_c animals with the standard decontaminant and n_t animals with each of the experimental decontaminants. Usually $n_c = n_t = 24$. For logistical reasons, particularly if a number of test decontaminants are to be evaluated at the same time, the test is divided into K replicate portions and each portion is carried out on separate days. Usually $K = 3$, and $n = 8$ animals are tested per group per day.

Comparisons between the standard decontaminant results and the test decontaminant results usually incorporate the assumption that the individual replicate results within tests can be pooled to arrive at overall lethality rates. Preliminary comparisons among the standard decontaminant lethality rates observed in each replicate are carried out to examine the reasonableness of this assumption. If there is no evidence of heterogeneity among the standard decontaminant replicates, then it is presumed that the replicates were carried out under homogeneous conditions and the test decontaminant results, as well as the standard decontaminant results, are pooled across

TABLE 3.4.1. PROBABILITIES (x 1,000) OF SIGNALING A PROCESS SHIFT FOR SPECIFIC VALUES OF μ_c ASSUMING THAT $n_i = 24$, $\mu_{c0} = 0.5$, AND $\sigma_{\delta_0} = 0.1$

μ_c	Control Chart Procedure		
	A	B	C
0.5	003[a]	003	003
	003[b]	003	003
	003[c]	003	003
0.6	011	028	057
	011	021	066
	011	038	128
0.7	059	202	399
	059	157	517
	059	306	791
0.8	227	612	842
	227	545	960
	227	816	998
0.9	597	937	989
	597	931	1,000
	597	996	1,000

[a]Top value is for procedures A, B, and C (Counts)
[b]Middle value is for procedures A2, B2, and C2 (Medians)
[c]Bottom value is for procedures A3, B3, and C3 (Means)

replicates. The pooled standard decontaminant results are then compared with the historical standard decontaminant results and with the control chart limits, in the manner discussed in Sections 2.0 and 3.0. The resulting estimate of current standard decontaminant lethality rate is compared with each of the test decontaminant rates, in the manner discussed in Section 2.2.

If there is statistically significant heterogeneity among the standard decontaminant replicates, further tests are carried out to determine which replicates differ from the others and/or from the long run historical results. These outlying replicates are then considered for deletion (both the standard decontaminant and the test decontaminant results) and additional replicates are carried out to replace them.

4.1 Comparison of Individual Replicate Standard Decontaminant Results with the Overall Average

Suppose that the current test is divided into K replicates (days), that the ith replicate includes n_i animals in the standard decontaminant group, and that x_i responses (deaths) are observed among these animals. Let $r_i = x_i/n_i$ denote the observed response rate in the ith replicate, $x_c = \sum_{i=1}^{K} x_i$, $n_c = \sum_{i=1}^{K} n_i$, $\bar{n} = n_c/K$, and $r_c = x_c/n_c$. Let $p_i = E(r_i)$ denote the population average response rate in the ith replicate. The analysis of means (Ott, 1975) is used to compare each replicate response rate, r_i, to the average rate, r_c. The analysis of means test is designed to be sensitive to the presence of an extreme replicate that differs from the others, much like a control chart inference. The hypothesis

$$H_o: \quad p_1 = p_2 = \ldots = p_K$$

is tested by the analysis of means procedure. Let

$$Z = \max_{i = 1,\ldots,K} \left[|r_i - r_c| / (r_c(1 - r_c)/\bar{n})^{\frac{1}{2}} \right]$$

The hypothesis H_o is rejected at significance level α if $Z > H_\alpha$, where H_α is tabulated by Ott (1975), Schilling (1973), and others. If $K = 3$ and $\alpha = 0.05$, then $H_{0.95} = 1.93$.

If H_o is rejected, each r_i is compared to the others and to the historical control rate (nominally 0.50).

4.2. Comparisons of Replicate Responses with the Historical Control Rate and Among Each Other

If the analysis of means test rejects the above H_o, then each r_i is compared to the long run historical standard decontaminant response to determine which differ. Assume for purposes of this discussion that the historical response rate is close to the nominal, 0.50. The hypotheses

$$H_o: \ p_i = 0.50 \text{ and } i = 1, 2, \ldots, K$$

are tested using individual one-sample, two-sided tests. The significance levels are adjusted by Bonferroni's method so that the overall Type 1 error across the K tests does not exceed α. Let

$$Z_i = |r_i - 0.5| / [(0.5)(0.5)/n_i]^{\frac{1}{2}}$$

The hypothesis

$$H_o: \ p_i = 0.5$$

is rejected if $Z_i > Z_\alpha$. Values of Z_α for $\alpha = 0.10$, 0.05, and 0.01 and $K = 1(1)$ 10 and above are tabulated by Ott (1975), Miller (1966), and others. If $K = 3$, then $Z_{0.95} = 2.39$ and $Z_{0.10} = 2.11$.

It should be noted that if $K = 3$ and $n_i = 8$, only $r_i = 0$ or $r_i = 1$ would cause H_o to be rejected at $\alpha = 0.05$; $r_i \geq 0.875$ or $r_i \leq 0.125$ would cause H_o to be rejected at $\alpha = 0.10$. When $n_i = 8$, these tests are rather insensitive; they will detect only very large departures from consistency across replicates or from the historical average rate.

Pairwise comparisons among replicates are carried out using Fisher's exact test (two-sided). With just $n_i = 8$ animals per group, these tests are also insensitive. Critical values are tabulated by Pearson and Hartley (1958) When K = 3, the single-tailed 0.01 level critical values correspond to an $\alpha = 0.06$ (0.01 x 2 x 3) two-tailed simultaneous significance level for all (three) pairwise comparisons among replicates. The Pearson-Hartley table demonstrates that 8 of 8 responses can be distinguished from 2 of 8, 7 of 8 from 1 of 8, and 6 of 8 from 0 of 8 at this significance level. The single-tailed 0.025 level critical values correspond to an $\alpha = 0.12$ (0.020 x 2 x 3) two-tailed simultaneous significance level. At this significance level, 8 of 8 can be distinguished from 3 of 8, 7 of 8 from 2 of 8, 6 of 8 from 1 of 8, and 5 of 8 from 0 of 8. Thus, when $n_i = 8$, these pairwise comparisons will detect only substantial departures from consistency across replicates.

The discussion in this section demonstrates that only sizeable inconsistencies among replicate response levels will be flagged by these procedures. In all other instances, the responses will be pooled across replicates and the principal comparisons will proceed.

5.0 REDETERMINATION OF LD_{50} DOSES

An important aspect of the screening program is to initially establish and then periodically update LD_{50} doses for standard treatment. If the control chart inferences discussed in Section 3.0 detect drift in the standard decontaminant response rates or repeated exceedences of the control chart limits, then the LD_{50} dose needs to be redetermined. A new LD_{50} study must be carried out to determine the new agent dose. This section discusses experimental design and data analysis methods and associated computer programs that have been developed to determine LD_{50} doses in an efficient manner; fewer animals are needed to attain the desired levels of estimation precision, relative to a classical LD_{50} design.

Section 5.1 discusses a stagewise dose allocation experimental design strategy that has been developed to accomplish this aim. Such stagewise designs lead to nontraditional dose allocations that utilize relatively large numbers of doses with relatively small numbers of animals per

dose. It is possible that each animal will be tested at a different dose. Standard probit analysis computer programs, therefore, cannot be used to fit dose-response models to the lethality data.

Specialized procedures, based on nonlinear regression analysis, have been developed to fit dose-response models to these data. These procedures have been developed in a series of computer programs based on the general purpose nonlinear regression procedure, PROC NLIN, in the SAS statistical computing system (SAS, 1985). Section 5.2 discusses the procedures and programs.

5.1 Stagewise, Adaptive Dose Allocation Procedures

The LD_{50} is estimated based on a small to moderate number of animals. The precision of estimation of the LD_{50} and slope depends on the numbers of animals tested as well as on the allocation of animals to appropriate portions of the (unknown) dose-response distribution. To obtain relatively precise estimates of the LD_{50} and slope with the numbers of animals available, the test doses should be centered around the LD_{50} with enough spread to permit good estimation of the slope. The test doses should not, however, extend too far beyond the central portion of the dose-response region (e.g., they should lie between the 10th and 90th percentiles). The desired dose allocation heavily depends on the underlying dose-response distribution. It is assumed that the dose-response relation for the standard decontaminant animals can be described by a two-parameter probit model without background, at least in the central portion of the dose-response region.

The relative sensitivities of alternative dose allocations can be evaluated before any data have been collected. This permits "target designs" to be selected before the start of the experiment and to be updated as the experiment proceeds.

Since the underlying dose-response distributions are not known prior to the start of the test, the LD_{50} test is carried out in a stagewise fashion. The dose allocation for the first stage is based on historical results. Previous LD_{50} study results, augmented by observed response rates in more recent standard decontaminant tests, can be used to obtain initial estimates

of the LD_{50} and slope and the associated first-stage dose allocation. Following each stage, the information concerning the underlying dose-response distribution is updated based on fitting dose-response models to the results obtained in the current and previous stages. Doses are selected for the next stage to best approximate the target design, over and above the previous allocations, based on the updated dose-response distribution. This process is iterated until the completion of all the stages or until the LD_{50} and/or slope are estimated with the required level of precision.

This approach is in the spirit of, but is more flexible and adaptive than, the formal up-down method (Dixon and Mood, 1948). It attempts to incorporate relatively large numbers of test doses within each stage and uses information from all previous stages to make decisions about the doses to be selected in subsequent stages.

The application of the stagewise dose allocation approach to determining the standard decontaminant LD_{50} is illustrated by an example pertaining to percutaneous application of GD in albino rabbits and treatment with both components of the M258A1 standard decontamination kit. An updated program to screen new candidate decontaminants is to be implemented, utilizing whatever information can be obtained from previously completed screening programs.

The a priori assumptions for the dose-response relation applicable to the forthcoming screening program are based on a probit model fit to the results from a previous LD_{50} study with this same agent, animal model, and decontamination regimen that was carried out in May-June 1985 in and MREF Final Report entitled "Task 85-10: Validation of a Protocol to Compare the Effectiveness of Experimental Decontaminants With Both Components of the M258A1 Kit Against Percutaneous Application of Undiluted Organophosphate Chemical Surety Materiels to the Laboratory Albino Rabbit," (December 1987, Table 3.1.4). The LD_{50} was estimated to be 13.0 μg/kg and the slope 3.732, based on n = 360 animals. The parameters of this distribution are displayed in Table 5.1.1.

Ten alternative "target designs" were considered. These are numbered GD1 to GD10 and are shown in Table 5.1.2. Each target design consists of n = 100 animals, allocated equally or unequally to various

TABLE 5.1.1. PARAMETERS SPECIFYING THE A PRIORI PROBIT DOSE-RESPONSE DISTRIBUTION. THESE PARAMETERS DETERMINE THE CENTRAL DISTRIBUTION AND PERTURBATIONS INCORPORATED IN SUBSEQUENT SENSITIVITY ANALYSES

FITID	B0	B1	V0	C01	V1
GD/DECON	0.843	3.732	0.336	-0.299	0.274

FITID = Fit identification
B0,B1 = Slope and intercept of the a priori dose-response distribution
V0,V1,C01 = Variance of the intercept, variance of slope, and covariance between the intercept and slope of the a priori dose-response distribution and quantify the uncertainty in these parameters.

TABLE 5.1.2. ALTERNATIVE TARGET DESIGNS FOR THE FORTHCOMING LD_{50} STUDY. THE DOSES, BASED ON SPECIFIED PERCENTILES ON THE A PRIORI DOSE-RESPONSE DISTRIBUTION, AND THE NUMBER OF ANIMALS PER DOSE ARE SPECIFIED FOR EACH DESIGN

	Candidate designs for Stage 4 - Proposed doses and numbers of subjects																	
Design ID	Dose 1[a]	N Subs /Dose 1[b]	Dose 2	N Subs /Dose 2	Dose 3	N Subs /Dose 3	Dose 4	N Subs /Dose 4	Dose 5	N Subs /Dose 5	Dose 6	N Subs /Dose 6	Dose 7	N Subs /Dose 7	Dose 8	N Subs /Dose 8	Dose 9	N Subs /Dose 9
GD1	7.7	25.00	11.1	25.00	15.2	25.00	21.9	25.00	.	.	.	.	.	.	.	.	.	.
GD2	5.9	11.11	7.7	11.11	9.4	11.11	11.1	11.11	13.0	11.12	15.2	11.11	18	11.11	21.9	11.11	28.7	11.11
GD3	5.9	20.00	9.4	20.00	13.0	20.00	18.0	20 00	28.7	20.00	.	.	.	.	.	.	.	.
GD4	7.7	33.33	13.0	33.34	21.9	33.32	.	.	.	.	.	.	.	.	.	.	.	.
GD5	7.7	30.00	13.0	40.00	21.9	30.00	.	.	.	.	.	.	.	.	.	.	.	.
GD6	9.4	33.33	13.0	33.34	18.0	33.33	.	.	.	.	.	.	.	.	.	.	.	.
GD7	5.9	25.00	11.1	25.00	15.2	25.00	20.7	25.00	.	.	.	.	.	.	.	.	.	.
GD0	5.9	10.00	7.7	20.00	11.1	20.00	15.2	20.00	21.9	20.00	28.7	10.00	.	.	.	.	.	.
GD9	5.9	33.33	13.0	33.34	28.7	33.33	.	.	.	.	.	.	.	.	.	.	.	.
GD10	5.9	10.00	7.7	25.00	13.0	30.00	21.9	25.00	28.7	10.00	.	.	.	.	.	.	.	.

[a]Dose level 1 in the target design.
[b]Number of subjects tested at dose level 1.

combinations of the 10, 20,...,90 percentiles of the assumed prior dose-response distribution. For example, target design GD1 allocates 25 animals to each of the 20, 40, 60, and 80 percentiles. The absolute sensitivities calculated for these designs pertain to 100 animals. However, these sensitivities are each scaled up or down by the factor $(100/n)^{\frac{1}{2}}$ if n animals are used instead. The relative sensitivities for these designs are thus invariant to sample size.

Although each target design is presented for initial planning purposes as a single-stage design, the LD_{50} study is in fact carried out in stages, with doses adjusted from stage to stage. The target allocations are updated, over and above the doses previously tested, in light of the most current estimate of the dose-response relation. The stagewise, adaptive dose allocation helps assure conformance to the target design even if the estimate of the underlying dose-response distribution shifts from stage to stage as additional results are obtained. Furthermore, if the attained sensitivity to estimate dose-response distribution parameters or to compare dose-response distributions exceeds that predicted at the outset of the experiment, the stagewise design strategy can lead to early stopping.

The predicted sensitivities for each design are calculated from the information obtained in the previous stages, combined with the expected information to be obtained in the current and future stages. The information associated with each design is evaluated for the distribution specified in Table 5.1.1 (the "central" distribution), as well as for distributions that are perturbations about the central distribution.

Table 5.1.3 displays the "results" from the "previous" stages. For each stage and dose, the logarithm of dose (X), the number of animals on test (NN), and the number of responses (Y) are given. In this example, the designs are being evaluated prior to the first stage. There is no previous data and so NN = 0. For evaluations following later stages, the observed results at all the previous stages and doses would be used.

Detailed sensitivity analyses are carried out for each target design to assess its performance under a variety of distributions that might be likely to occur (i.e., perturbations about the central distribution).

TABLE 5.1.3. EXPERIMENTAL "RESULTS" FROM "PREVIOUS" STAGES[a]

OBS	GROUP	STAGE	DOSE	X	NN	Y
1	DECON	1	20	1.30103	0	0

[a]Prior to the first stage, the numbers of animals (NN) and the numbers of responses (Y) are each 0.
OBS = Record number
GROUP = Identification variable
STAGE = Stage at which dose was administered
DOSE = Dose administered
X = Common logarithm of dose

Table 5.1.4 displays the detailed sensitivity results for design GD1. The middle line in the table (underlined) corresponds to the central distribution. The remaining 48 lines correspond to perturbations about this central distribution. Standard errors of the estimates of the specified logarithmic percentiles (50, 80, 90 in this example) and of the slope are calculated for each of these distributions. The results of the sensitivity analyses for each target design are summarized in Tables 5.1.5 and 5.1.6. Table 5.1.5 displays weighted averages of the standard errors over all of the 49 distributions in the sensitivity analysis; the distributions closer to the central distribution receive the greater weight. Table 5.1.6 displays the minima and the maxima of these standard errors over these same distributions. The maxima can be regarded as "worst cases", over the range of distributions considered plausible based on the current information.

Table 5.1.5 shows that the weighted averages of the standard errors of the $\log_{10}$ (LD_{50}) are similar across all of the target allocations considered. This is not surprising, since they were all selected to be symmetric about the a priori 50th percentile. Design GD6 has the smallest and design GD9 has the largest. By contrast the standard errors of the $\log_{10}$ (LD_{90}) and the slope vary to a greater extent across the target allocations considered. Design GD9 has the smallest and design GD6 is the largest.

Similar considerations hold for the maxima of the standard errors. Those for the $\log_{10}$ (LD_{50}) are similar across all the target allocations considered. Those for the $\log_{10}$ (LD_{90}) and the slope vary more across the target allocations considered; design GD9 has the smallest and design GD6 has the largest.

Design GD6 allocates animals evenly to the 30, 50, and 70 percentiles; design GD9 allocates animals evenly to 10, 50 and 90 percentiles. Designs GD3 and GD10 are compromises between the two extremes. They allocate animals equally among the 10, 30, 50, 70 and 90 percentiles and among the 10, 20, 50, 80, and 90 percentiles, respectively.

To utilize this information for a stagewise dose allocation, the numbers of stages and the numbers of animals per stage would be decided upon and design GD3, for example, appropriately scaled down, might be run for the first stage. Following the first-stage, the dose-response distribution

TABLE 5.1.4. DETAILED SENSITIVITY ANALYSIS RESULTS FOR DESIGN GD1. UNDERLINED DISTRIBUTION IS THE CENTRAL DISTRIBUTION

A and B and their standard errors for each design in the sensitivity analyses

OBS	DSGNID	B0	B1	PCT1	PCT2	PCT3	SEPCT1	SEPCT2	SEPCT3	SEB1	FACTOR
1	GD1	1.580460	3.81872	50	90	99	0.0478863	0.105432	0.144400	0.111388	0.007225
2	GD1	1.28344	3.24360	50	90	99	0.0428211	0.087662	0.110306	0.088200	0.012325
3	GD1	0.97630	2.67824	50	90	99	0.0388816	0.077833	0.106347	0.086861	0.014450
4	GD1	0.67048	2.11200	50	90	99	0.0388650	0.080306	0.083270	0.068787	0.017000
5	GD1	0.37462	1.60312	50	90	99	0.0348612	0.081843	0.062831	0.068230	0.014450
6	GD1	0.08727	1.28364	50	90	99	0.0324384	0.064565	0.073817	0.083448	0.012325
7	GD1	0.66281	1.77028	50	90	99	0.0260068	0.047653	0.082673	0.017048	0.007225
8	GD1	1.03040	2.81872	50	90	99	0.0458488	0.088725	0.130232	0.107648	0.012325
9	GD1	1.37730	3.24360	50	90	99	0.0413682	0.082288	0.114488	0.021642	0.021025
10	GD1	1.02624	3.67824	50	90	99	0.0367146	0.074148	0.101283	0.032858	0.014450
11	GD1	0.74443	3.11200	50	90	99	0.0382852	0.088245	0.088681	0.048208	0.017000
12	GD1	0.44458	4.60312	50	90	99	0.0340484	0.058413	0.078887	0.081788	0.014450
13	GD1	0.13121	4.28364	50	90	99	0.0320774	0.053511	0.071482	0.080048	0.021025
14	GD1	0.38487	4.77028	50	90	99	0.0283662	0.048141	0.060828	0.013644	0.012325
15	GD1	1.66010	3.81872	50	90	99	0.0451834	0.096067	0.134308	0.108163	0.014450
16	GD1	1.32867	3.24360	50	90	99	0.0400188	0.080488	0.111440	0.020228	0.014450
17	GD1	1.06853	3.67824	50	90	99	0.0383606	0.071824	0.088721	0.031888	0.028900
18	GD1	0.78371	3.11200	50	90	99	0.0360628	0.064317	0.087718	0.044907	0.034000
19	GD1	0.48788	4.60312	50	90	99	0.0338181	0.057828	0.078208	0.060628	0.028900
20	GD1	0.18046	4.28364	50	90	99	0.0318483	0.052218	0.068882	0.078888	0.014450
21	GD1	0.33838	4.77028	50	90	99	0.0283183	0.048202	0.060743	0.012648	0.014450
22	GD1	1.72638	3.81872	50	90	99	0.0448234	0.092603	0.130648	0.106488	0.017000
23	GD1	1.37886	3.24360	50	90	99	0.0407232	0.077828	0.108518	0.019817	0.028900
24	GD1	1.01882	3.67824	50	90	99	0.0382803	0.088888	0.088278	0.031068	0.034000
25	GD1	0.64308	3.11200	50	90	99	0.0366880	0.082478	0.085872	0.044153	0.028900
26	GD1	0.84714	4.60312	50	90	99	0.0336884	0.066314	0.078611	0.060186	0.034000
27	GD1	0.22878	4.28364	50	90	99	0.0318308	0.060688	0.068868	0.078444	0.028900
28	GD1	0.28868	4.77028	50	90	99	0.0283488	0.044318	0.058722	0.012783	0.017000
29	GD1	1.72648	3.81872	50	90	99	0.0448628	0.088234	0.128813	0.008638	0.014450
30	GD1	1.42525	3.24360	50	90	99	0.0407888	0.078248	0.106714	0.018612	0.024650
31	GD1	1.06611	3.67824	50	90	99	0.0383237	0.087481	0.083622	0.031347	0.028900
32	GD1	0.69228	3.11200	50	90	99	0.0360447	0.080718	0.083718	0.044844	0.034000
33	GD1	0.89643	4.60312	50	90	99	0.0338404	0.064874	0.074388	0.060873	0.028900
34	GD1	0.27808	4.28364	50	90	99	0.0320002	0.060618	0.067274	0.078282	0.024650
35	GD1	0.24060	4.77028	50	90	99	0.0283782	0.043476	0.057758	0.012648	0.014450
36	GD1	1.82787	3.81872	50	90	99	0.0482716	0.085868	0.123417	0.008382	0.012325
37	GD1	1.47488	3.24360	50	90	99	0.0411137	0.072778	0.103018	0.020813	0.021025
38	GD1	1.01741	3.67824	50	90	99	0.0366008	0.068618	0.081678	0.032468	0.024650
39	GD1	0.84158	3.11200	50	90	99	0.0362783	0.068034	0.081848	0.046062	0.028900
40	GD1	0.84573	4.60312	50	90	99	0.0341378	0.053602	0.073388	0.062068	0.024650
41	GD1	0.32838	4.28364	50	90	99	0.0321853	0.048707	0.066017	0.080614	0.021025
42	GD1	0.18140	4.77028	50	90	99	0.0285001	0.042884	0.058858	0.018442	0.012325
43	GD1	1.86181	3.81872	50	95	99	0.0448672	0.081338	0.118388	0.008148	0.007225
44	GD1	1.54648	3.24360	50	90	99	0.0410408	0.068258	0.088184	0.020834	0.012325
45	GD1	1.28136	3.67824	50	90	99	0.0383838	0.082568	0.088481	0.038884	0.014450
46	GD1	1.01882	3.11200	50	90	99	0.0360281	0.060652	0.078188	0.048534	0.017000
47	GD1	0.71867	4.60312	50	90	99	0.0340660	0.061588	0.073178	0.068788	0.014450
48	GD1	0.60231	4.28364	50	90	99	0.0325828	0.047147	0.064248	0.084888	0.012325
49	GD1	0.11746	4.77028	50	90	99	0.0288071	0.041578	0.055583	0.018888	0.007225

OBS = Observation number
DSGNID = Identifier of target design
B0,B1 = Intercept and slope of probit distribution
PCT1,PCT2,PCT3 = Percentiles that are being evaluated
SEPCT1,SEPCT2,SEPCT3 = Standard errors of the estimates of the logarithm percentiles
SEB1 = Predicted standard error of the slope following ...
FACTOR = Weights associated with each distribution

TABLE 5.1.5. WEIGHTED AVERAGES OF THE STANDARD ERRORS OF THE LOGARITHMIC 50, 80, AND 90 PERCENTILES AND THE SLOPE OVER ALL THE DISTRIBUTIONS IN THE SENSITIVITY ANALYSIS

Weighted averages of standard errors of points varied around A and B over the designs in the sensitivity analysis

APERCENT	IDPCT	NPCT	SQRTSUM
	Design I.D. = GD1		
50	SEFVX50	49	0.036863
80	SEFVX80	49	0.065796
90	SEFVX90	49	0.090660
B1	SEFVXB1	49	0.850717
	Design I.D. = GD10		
50	SEFVX50	49	0.038451
80	SEFVX80	49	0.058843
90	SEFVX90	49	0.077879
B1	SEFVXB1	49	0.704621
	Design I.D. = GD2		
50	SEFVX50	49	0.037861
80	SEFVX80	49	0.061135
90	SEFVX90	49	0.082221
B1	SEFVXB1	49	0.759807
	Design I.D. = GD3		
50	SEFVX50	49	0.038780
80	SEFVX80	49	0.058818
90	SEFVX90	49	0.077645
B1	SEFVXB1	49	0.705694
	Design I.D. = GD4		
50	SEFVX50	49	0.037343
80	SEFVX80	49	0.062219
90	SEFVX90	49	0.084367
B1	SEFVXB1	49	0.778833

TABLE 5.1.5.
(Continued)

APERCENT	IDPCT	NPCT	SQRTSUM
	Design I.D. = GD5		
50	SEFVX50	49	0.037063
80	SEFVX80	49	0.064222
90	SEFVX90	49	0.087923
B1	SEFVXB1	49	0.820809
	Design I.D. = GD6		
50	SEFVX50	49	0.03609
80	SEFVX80	49	0.08256
90	SEFVX90	49	0.11852
B1	SEFVXB1	49	1.14917
	Design I.D. = GD7		
50	SEFVX50	49	0.039233
80	SEFVX80	49	0.058390
90	SEFVX90	49	0.076798
B1	SEFVXB1	49	0.693325
	Design I.D. = GD8		
50	SEFVX50	49	0.038148
80	SEFVX80	49	0.059842
90	SEFVX90	49	0.079798
B1	SEFVXB1	49	0.728632
	Design I.D. = GD9		
50	SEFVX50	49	0.040943
80	SEFVX80	49	0.056355
90	SEFVX90	49	0.071733
B1	SEFVXB1	49	0.623474

APERCENT = Quantity being estimated (percentile or slope)
IDPCT = Identifier of quantity being estimated
NPCT = Number of distributions entering into the weighted average
SQRTSUM = Weighted averages of the standard errors

TABLE 5.1.6. MINIMA AND MAXIMA OF THE STANDARD ERRORS OF THE LOGARITHMIC 50, 80, AND 90 PERCENTILES AND THE SLOPE OVER ALL THE DISTRIBUTIONS IN THE SENSITIVITY ANALYSIS

Minimum and maximum of unweighted standard errors of points varied around A and B over the designs in the sensitivity analysis

VARNAME	N	MINIMUM	MAXIMUM
	Design I.D. = GD1		
SEFVX50	49	0.029316	0.047555
SEFVX80	49	0.041578	0.105432
SEFVX90	49	0.055582	0.144400
SEFVXB1	49	0.805459	0.919895
	Design I.D. = GD10		
SEFVX50	49	0.031554	0.047595
SEFVX80	49	0.041651	0.087573
SEFVX90	49	0.053014	0.117208
SEFVXB1	49	0.641773	0.799141
	Design I.D. = GD2		
SEFVX50	49	0.030646	0.047480
SEFVX80	49	0.041790	0.092691
SEFVX90	49	0.054465	0.125116
SEFVXB1	49	0.690566	0.861523
	Design I.D. = GD3		
SEFVX50	49	0.031852	0.047782
SEFVX80	49	0.042300	0.086001
SEFVX90	49	0.054073	0.114733
SEFVXB1	49	0.627240	0.822425
	Design I.D. = GD4		
SEFVX50	49	0.030076	0.047356
SEFVX80	49	0.041038	0.097405
SEFVX90	49	0.053542	0.132328
SEFVXB1	49	0.733219	0.849374

TABLE 5.1.6.
(Continued)

VARNAME	N	MINIMUM	MAXIMUM
	Design I.D. = GD5		
SEFVX50	49	0.029622	0.047440
SEFVX80	49	0.041387	0.101758
SEFVX90	49	0.054838	0.138902
SEFVXB1	49	0.772841	0.894420
	Design I.D. = GD6		
SEFVX50	49	0.02775	0.04980
SEFVX80	49	0.04441	0.14178
SEFVX90	49	0.06528	0.19805
SEFVXB1	49	1.12243	1.19243
	Design I.D. = GD7		
SEFVX50	49	0.032382	0.048034
SEFVX80	49	0.042927	0.083840
SEFVX90	49	0.054734	0.111264
SEFVXB1	49	0.605445	0.826605
	Design I.D. = GD8		
SEFVX50	49	0.031101	0.047525
SEFVX80	49	0.041628	0.089953
SEFVX90	49	0.053529	0.120888
SEFVXB1	49	0.0664329	0.824324
	Design I.D. = GD9		
SEFVX50	49	0.034806	0.048910
SEFVX80	49	0.044264	0.077431
SEFVX90	49	0.054499	0.100846
SEFVXB1	49	0.539006	0.755271

VARNAME = Identifier of quantity being estimated
N = Number of distributions entering into the estimate
MINIMUM = Minimum standard error of the logarithm
MAXIMUM = Maximum standard error of the logarithm

estimate would be updated and the sensitivity analysis would be carried out, as above, to determine which second-stage dose allocations best augment the first-stage doses, in light of what has been learned about the dose-response distribution from the first stage. This procedure would be iterated following each stage of experimentation.

5.2 Probit Dose Response Estimation Based on Nonlinear Regression Analysis

Following each stage of experimentation, the estimates of the underlying dose-response distributions are updated. Probit dose-response models in logarithmic dose (Finney 1977) are fitted to the data for each treatment regimen to quantify the relationships. Distribution percentiles are estimated based on these models. Background response is not incorporated into the models, due to the relatively short durations of the tests (hours, days, or at most one or two weeks).

Standard probit analysis computer programs cannot be used to fit these models to the dose-response data due to the nonstandard dose-allocation strategy and due to a number of nonstandard aspects of the model specifications. These nonstandard aspects include individual animal responses rather than pooled group lethality rates, common probit slopes shared by several treatment regimens, the possible presence of stage effects and the capability to adjust for such effects, and the incorporation of covariates, such as body weight, into the models. These model aspects are discussed in greater detail in this section.

Specialized procedures, based on nonlinear regression analysis, have been developed to fit dose-response models to such data. These procedures are described and illustrated.

5.2.1 Individual Animal Responses

The dose-allocation strategy discussed in the previous section results in many different doses with few animals tested per dose, possibly just one. The model fitting methods thus need to accommodate the possibility

of dose-response data with each animal tested at a unique dose. All the responses would then be 0's and 1's. This is in contrast to the usual probit analysis situation where multiple animals are tested at a relatively small number of repetitive, discrete doses and dose-response models are fitted to the observed response proportions at each dose.

5.2.2 Separate Slopes and Common Slopes Models

The model fitting procedures are sometimes used to compare dose-response distributions corresponding to several treatment regimens. For example, no treatment, standard treatment, and one or more candidate treatments may be compared simultaneously. A fully general model fits separate dose-response distributions to each regimen. Submodels incorporating the assumption of common slopes among various subsets of the treatment afford the possibility of substantially greater estimation precision and interpretation simplicity. Provisions have been incorporated into the model fitting procedures to fit common slopes to various subsets (of size 2 to 5) of the treatments and to test the adequacy of fit of the submodels relative to the separate slopes model or to less restrictive common slopes models. Common slopes models are also sometimes used to augment information about the dose response for a current treatment with that based on historical data. Although the dose-response distributions may be shifted relative to one another, the slopes may have remained the same.

Tests of adequacy of the submodels are carried out by comparing the values of the log likelihoods under the more restrictive and the less restrictive models. The log likelihood ratio is referred to the upper percentiles of a chi-square distribution with an appropriate number of degrees of freedom.

5.2.3 Stage to Stage Variation

The basic design strategy calls for carrying out the dose-response experiment in stages, utilizing the results from all previous stages to design the following stage. A test for the presence of stage to stage variation is

incorporated into the model fitting procedures. The test is carried out by fitting probit models to the combined data across all stages. Residuals from these fits are standardized by dividing them by estimates of their standard deviations. In the absence of stage to stage variation, these standardized residuals would be expected to have mean approximately 0 and standard deviation approximately 1. If systematic stage to stage variation exists, then the residuals from some stages would have positive means while those from other stages would have negative means.

A one-way analysis of variance is carried out on the standardized residuals, incorporating stage as a grouping variable and the logarithm of dose as a covariate. Statistical significance of the stage factor ($\alpha = 0.05$) provides evidence of stage to stage variation. Possible causes of such stage effects might be drift across stages, isolated outlying responses, or variation of some of the experimental conditions across stages.

The nature of the stage to stage variation would need to be studied by more in-depth examination of the data, such as diagnostic plots, multiple comparison procedures, or the incorporation of additional explanatory variables into the models. The nature and extent of such additional analyses, and possible actions taken as a result, would necessarily be decided upon on a case by case basis. They are not incorporated into the more general model fitting procedures discussed here.

5.2.4 Covariates

Body weight (kg) at the time of dosing is incorporated into the models as a covariate. Models incorporating separate covariate effects for different treatment regimens and models incorporating common covariate effects are fitted to the data. Likelihood ratio tests for common covariate effects are carried out in the same manner as likelihood ratio tests for common dose-response slopes.

It should be noted that including covariates such as body weight in the dose-response models necessitates fitting to the individual animal 0-1 responses rather than pooling across animals that were tested at the same dose. This is because, in general, each animal has a different body weight and so presents a different set of explanatory variables.

5.2.5 Dose-Response Model Fitting Procedures

A series of computer programs, based on PROC NLIN in the SAS statistical computing system, have been developed to fit the dose-response models to the experimental results. These procedures utilize as input the individual animal 0-1 responses, as well as the treatment dose and any covariates, such as body weight. Programs are available to fit separate probit models to each individual treatment (separate slopes model) and to fit joint probit models having a common slope to several treatments (common slopes model). Covariates can be included in or excluded from the models.

Table 5.2.1 displays the output from a common slopes probit model fit to the results from two treatment regimens. The parameter B1 represents the common slope and the parameters B01, B02 represent the intercepts for the treatments. No covariate is included in this model. If the model fits the data, then the expected value of the residual mean square is asymptotically 1.0. The attained residual mean square of 0.85 indicates no evidence of lack of fit of the model. The "sum of loss" is proportional to -2 times the (natural) logarithm of the likelihood function; it is used to compare the adequacy of alternative models.

Tables 5.2.2 and 5.2.3 display estimates of the dose-response distribution percentiles, associated standard errors, and upper and lower 95 percent confidence bounds. The confidence bounds in Table 5.2.2 are based onpropagation of errors, while those in Table 5.2.3 are based on Fieller's method. If the estimated slope is somewhat more than two standard errors from 0, as in this example, then both confidence intervals are similar, particularly for doses in the central portion of the design. If the estimated slope is less than two standard errors from 0, then the Fieller's method confidence intervals will be substantially wider than the propagation of

TABLE 5.2.1. COMMON SLOPES PROBIT DOSE-RESPONSE MODEL FIT TO THE RESULTS FROM TWO TREATMENT REGIMENS

AGENT=GD

NON-LINEAR LEAST SQUARES SUMMARY STATISTICS DEPENDENT VARIABLE NDEAD

SOURCE	DF	WEIGHTED SS	WEIGHTED MS
REGRESSION	3	30178.105509	10059.368503
RESIDUAL	41	34.866132	0.850393
UNCORRECTED TOTAL	44	30212.971641	
(CORRECTED TOTAL)	43	27602.015889	
SUM OF LOSS		34.330548	

PARAMETER	ESTIMATE	ASYMPTOTIC STD. ERROR	ASYMPTOTIC 95 % CONFIDENCE INTERVAL LOWER	UPPER
B1	12.84032759	4.0026351428	4.557078575	20.723576612
B01	-5.08247685	3.3026552693	-11.752285854	1.587342158
B02	-10.00687366	4.7811388706	-19.682733385	-0.331013942

NOTE: STANDARD ERRORS COMPUTED USING SIGSQ= 1

ASYMPTOTIC CORRELATION MATRIX OF THE PARAMETERS

CORR	B1	B01	B02
B1	1.0000	-0.9875	-0.9983
B01	-0.9875	1.0000	0.9858
B02	-0.9983	0.9858	1.0000

TABLE 5.2.2. ESTIMATED DOSE-RESPONSE DISTRIBUTION PERCENTILES AND ASSOCIATED CONFIDENCE BOUNDS BASED ON PROPAGATION OF ERRORS

Agent=GD TRTSYM=1

Agent	Percentile	Probit of Percentile	Log(Leth Dose) for Percentile	Standard Error for Log(L.D.)	Lethal Dose for Percentile	Lower Confidence Bound	Upper Confidence Bound
GD	10	-1.2816	0.698258	0.0557182	4.98887	3.88412	6.38948
GD	13	-1.1264	0.708533	0.0531765	5.11132	4.02075	6.48769
GD	16	-0.9945	0.718970	0.0511485	5.23585	4.15842	6.59510
GD	20	-0.8418	0.731061	0.0489733	5.38348	4.31592	6.71505
GD	25	-0.6745	0.744283	0.0468380	5.54988	4.48243	6.85624
GD	30	-0.5244	0.756157	0.0451658	5.70371	4.65192	6.99330
GD	35	-0.3853	0.767160	0.0438477	5.85006	4.78875	7.13020
GD	40	-0.2533	0.777601	0.0428221	5.99240	4.93935	7.26998
GD	45	-0.1257	0.787702	0.0420540	6.13341	5.07314	7.41528
GD	50	-0.0000	0.797644	0.0415254	6.27543	5.20300	7.56891
GD	55	0.1257	0.807585	0.0412311	6.42074	5.33055	7.73388
GD	60	0.2533	0.817686	0.0411770	6.57183	5.45732	7.91395
GD	65	0.3853	0.828127	0.0413817	6.73174	5.58494	8.11400
GD	70	0.5244	0.839130	0.0418798	6.90446	5.71638	8.34093
GD	75	0.6745	0.851004	0.0427308	7.09584	5.85128	8.60510
GD	80	0.8416	0.864226	0.0440378	7.31519	5.99369	8.92380
GD	84	0.9945	0.876317	0.0455384	7.52172	6.12437	9.23788
GD	87	1.1264	0.886755	0.0470483	7.70468	6.23080	9.52720
GD	90	1.2816	0.899030	0.0490459	7.92555	6.35184	9.88918

Agent=GD TRTSYM=2

Agent	Percentile	Probit of Percentile	Log(Leth Dose) for Percentile	Standard Error for Log(L.D.)	Lethal Dose for Percentile	Lower Confidence Bound	Upper Confidence Bound
GD	10	-1.2816	1.08584	0.0409865	12.1853	10.1270	14.6619
GD	13	-1.1264	1.09811	0.0377770	12.5348	10.5698	14.8648
GD	16	-0.9945	1.10855	0.0351459	12.8395	10.9583	15.0485
GD	20	-0.8418	1.12084	0.0322533	13.2020	11.4136	15.2706
GD	25	-0.6745	1.13388	0.0293367	13.6101	11.9223	15.5388
GD	30	-0.5244	1.14574	0.0270037	13.9874	12.3825	15.8002
GD	35	-0.3853	1.15674	0.0251523	14.3463	12.8068	16.0708
GD	40	-0.2533	1.16718	0.0237373	14.6953	13.2024	16.3570
GD	45	-0.1257	1.17728	0.0227450	15.0411	13.5738	16.6871
GD	50	-0.0000	1.18722	0.0221783	15.3894	13.9238	17.0098
GD	55	0.1257	1.19718	0.0220527	15.7457	14.2541	17.3935
GD	60	0.2533	1.20728	0.0223816	16.1163	14.5678	17.8282
GD	65	0.3853	1.21771	0.0231850	16.5084	14.8884	18.3284
GD	70	0.5244	1.22871	0.0244804	16.9326	15.1803	18.9108
GD	75	0.6745	1.24058	0.0263472	17.4013	15.4505	19.5985
GD	80	0.8416	1.25380	0.0288542	17.9382	15.7489	20.4342
GD	84	0.9945	1.26590	0.0314611	18.4457	16.0041	21.2598
GD	87	1.1264	1.27633	0.0338985	18.8944	16.2141	22.0178
GD	90	1.2816	1.28861	0.0369381	19.4380	16.4516	22.9818

TABLE 5.2.3. ESTIMATED DOSE-RESPONSE DISTRIBUTION PERCENTILES AND ASSOCIATED CONFIDENCE BOUNDS BASED ON FIELLER'S METHOD

Agent=GD TRTSYM=1

Agent	Percentile	Probit of Percentile	Log(Leth Dose) for Percentile	Lethal Dose for Percentile	A = b1.b1 - Z.Z.VAR(b1)	B.B - A.C	Lower Confidence Bound	Upper Confidence Bound
GD	10	-1.2816	0.686258	4.86887	88.2343	240.447	2.81152	8.0231
GD	13	-1.1264	0.708533	5.11132	88.2343	213.303	3.11358	8.1746
GD	16	-0.9845	0.728870	5.33585	88.2343	192.554	3.28347	8.3120
GD	20	-0.8416	0.739081	5.38348	88.2343	171.186	3.51082	8.4828
GD	25	-0.6745	0.744283	5.54988	88.2343	151.131	3.75784	8.6870
GD	30	-0.5244	0.756157	5.70371	88.2343	136.042	3.88715	8.8886
GD	35	-0.3853	0.767160	5.85008	88.2343	124.536	4.20427	7.0840
GD	40	-0.2533	0.777601	5.99240	88.2343	115.818	4.41287	7.3085
GD	45	-0.1257	0.787702	6.13341	88.2343	108.425	4.61548	7.5368
GD	50	-0.0000	0.797644	6.27543	88.2343	105.092	4.81412	7.7845
GD	55	0.1257	0.807585	6.42074	88.2343	102.703	5.01048	8.0578
GD	60	0.2533	0.817688	6.57183	88.2343	102.268	5.20628	8.3641
GD	65	0.3853	0.828127	6.73174	88.2343	103.824	5.40355	8.7143
GD	70	0.5244	0.838130	6.90448	88.2343	107.881	5.60488	9.1230
GD	75	0.6745	0.851004	7.08584	88.2343	115.050	5.81388	9.6128
GD	80	0.8416	0.864226	7.31518	88.2343	126.175	6.03702	10.2215
GD	84	0.9845	0.876317	7.52172	88.2343	139.357	6.23263	10.8385
GD	87	1.1264	0.886755	7.70468	88.2343	153.048	6.39552	11.4222
GD	90	1.2816	0.899030	7.92558	88.2343	171.893	6.58132	12.1688

Agent=GD TRTSYM=2

Agent	Perc entile	Probit of Percentile	Log(Leth Dose) for Percentile	Lethal Dose for Percentile	A = b1.b1 - Z.Z.VAR(b1)	B.B - A.C	Lower Confidence Bound	Upper Confidence Bound
GD	10	-1.2816	1.08584	12.1853	88.2343	146.477	7.8387	13.8245
GD	13	-1.1264	1.08811	12.2348	88.2343	121.808	8.4188	14.1148
GD	16	-0.9845	1.10855	12.8485	88.2343	102.788	8.8368	14.3745
GD	20	-0.8416	1.12064	13.2020	88.2343	83.668	9.5708	14.8848
GD	25	-0.6745	1.13386	13.6101	88.2343	66.051	10.2887	15.0782
GD	30	-0.5244	1.14574	13.9874	88.2343	53.161	10.8812	15.4558
GD	35	-0.3853	1.15674	14.3483	88.2343	43.681	11.8273	15.8508
GD	40	-0.2533	1.16718	14.6853	88.2343	36.906	12.2436	16.2777
GD	45	-0.1257	1.17728	15.0411	88.2343	32.383	12.8307	16.7536
GD	50	-0.0000	1.18722	15.3884	88.2343	29.890	13.3873	17.2883
GD	55	0.1257	1.19718	15.7457	88.2343	28.341	13.8118	17.8333
GD	60	0.2533	1.20726	16.1163	88.2343	30.774	14.4038	18.6818
GD	65	0.3853	1.21771	16.5084	88.2343	34.385	14.8881	19.5706
GD	70	0.5244	1.22871	16.9320	88.2343	40.488	15.3133	20.8340
GD	75	0.6745	1.24058	17.4013	88.2343	48.726	15.7527	21.9242
GD	80	0.8416	1.25380	17.9382	88.2343	63.288	16.2052	23.5307
GD	84	0.9845	1.26590	18.4457	88.2343	78.718	16.5858	25.1553
GD	87	1.1264	1.27633	18.8844	88.2343	94.342	16.8204	26.6788
GD	90	1.2816	1.28861	19.4360	88.2343	115.458	17.2830	28.6178

errors intervals; the propagation of errors intervals are too narrow, while the Fieller's method intervals are too wide. A compromise interval cannot be obtained analytically; it likely requires a resampling method, such as bootstrapping, to account for the inherent nonlinearities.

Tables 5.2.4 and 5.2.5 display summaries, by stage, of the standardized residuals from the probit model fit and a one-way analysis of variance to test for the presence of stage to stage variation in these residuals. There is no evidence of a significant stage effect in this example. If the stage effect in Table 5.2.5 was significant, then the stagewise means, standard deviations, minima, and maxima in Table 5.2.4 would be studied to determine the nature of the variation, and which stage or stages differ from the remainder.

Tables 5.2.6 and 5.2.7 display the outputs from two probit model fits to the results from a different dose-response experiment, with two treatments. Body weight (kg) is included as a covariate in these models. Table 5.2.6 displays the results of a four-parameter common slopes model, with a common covariate effect, fitted to the two treatments. The parameters B1 and B2 represent the common slope and the common body weight effects, respectively; B01 and B02 represent the intercepts corresponding to treatments 1 and 2, respectively. Table 5.2.7 displays the results of a single three-parameter probit model fitted to the combined results from both treatments. B1 and B2 represent the slope and the body weight effect, respectively; B0 represents the intercept. Based on the residual mean square, both models appear to fit the data.

A log likelihood ratio test for differences between the dose-response distributions is carried out by comparing the difference between the "sum of loss" values for the two models to a chi-square distribution with 1 degree of freedom (4 parameters minus 3 parameters). Namely, 60.989 - 60.538 = 0.451 is significant at the $\alpha = 0.50$ level, based on the chi-square distribution with 1 degree of freedom. Thus, there is no evidence of differences between the dose-response distributions associated with each of the treatments and so the single model is accepted.

TABLE 5.2.4. SUMMARIES, BY STAGE, OF THE STANDARDIZED RESIDUALS FROM THE PROBIT MODEL FIT

[illegible]	LABEL	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM	VARIANCE
				AGENT=GD STAGE=1					
STUDRES	STUDENTIZED RESIDUAL	6	-0.04620300	0.88888431	-0.23877878	-0.00000000	0.03872877	-0.28881863	0.00800446
				AGENT=GD STAGE=2					
STUDRES	STUDENTIZED RESIDUAL	8	-0.33138185	0.84447330	-1.30823012	0.20826452	0.24432870	-1.86780868	0.38817620
				AGENT=GD STAGE=3					
STUDRES	STUDENTIZED RESIDUAL	8	0.01107827	0.66821268	-0.81821438	0.48282888	0.18831240	0.08848884	0.23835280
				AGENT=GD STAGE=4					
STUDRES	STUDENTIZED RESIDUAL	8	0.82731873	1.84864473	-2.84888818	1.80833888	0.78222487	6.18381288	3.78871388
				AGENT=GD STAGE=5					
STUDRES	STUDENTIZED RESIDUAL	3	-0.88887818	0.84888848	-1.30823012	0.20826452	0.48883888	-1.88783788	0.71888323
				AGENT=GD STAGE=6					
STUDRES	STUDENTIZED RESIDUAL	3	0.18883118	0.88782481	-0.81821438	0.78824872	0.51841713	0.31888338	0.88838888
				AGENT=GD STAGE=7					
STUDRES	STUDENTIZED RESIDUAL	3	1.14824884	0.88178178	0.43283888	1.80833888	0.38288838	3.43874881	0.43782887
				AGENT=GD STAGE=8					
STUDRES	STUDENTIZED RESIDUAL	3	-0.25384283	0.84837731	-0.81821438	0.78824872	0.81284718	-0.78882788	0.78883888
				AGENT=GD STAGE=9					
STUDRES	STUDENTIZED RESIDUAL	8	0.01286822	0.82283488	-1.11878278	1.88218185	0.28884838	0.18824878	0.87872817

TABLE 5.2.5. ONE-WAY ANALYSIS OF VARIANCE OF THE STANDARDIZED RESIDUALS FROM THE PROBIT MODEL FIT

AGENT=GD

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: STUDRES STUDENTIZED RESIDUAL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR > F	R-SQUARE	C.V.
MODEL	8	5.84425246	0.66047250	0.70	0.7017	0.156887	6684.7328
ERROR	34	31.84437775	0.93954052		ROOT MSE		STUDRES MEAN
CORRECTED TOTAL	43	37.68863021			0.96929388		-0.01450019

SOURCE	DF	TYPE I SS	F VALUE	PR > F	DF	TYPE III SS	F VALUE	PR > F
STRAIN	8	5.78962207	0.77	0.6288	8	5.84098471	0.78	0.6145
LOGDOSE	1	0.14463040	0.15	0.6972	1	0.14463040	0.15	0.6972

TABLE 5.2.6. COMMON SLOPES PROBIT DOSE-RESPONSE FIT TO THE RESULTS FROM THE TWO TREATMENT REGIMENS. BODY WEIGHT IS INCLUDED AS A COVARIATE.

NON-LINEAR LEAST SQUARES SUMMARY STATISTICS DEPENDENT VARIABLE NDEAD

SOURCE	DF	WEIGHTED SS	WEIGHTED MS
REGRESSION	4	425303.04381	106325.76088
RESIDUAL	68	57.25548	0.84188
UNCORRECTED TOTAL	72	425360.28840	
(CORRECTED TOTAL)	71	533.35580	
SUM OF LOSS		60.53762	

PARAMETER	ESTIMATE	ASYMPTOTIC STD. ERROR	ASYMPTOTIC 95 % CONFIDENCE INTERVAL LOWER	UPPER
B1	12.63071801	3.3932670606	5.858547034	19.401888981
B01	-5.42784638	2.8545267431	-11.123975881	0.268283122
B02	-5.17115781	2.7536868482	-10.666064285	0.323748665
B2	0.13842270	0.0582372860	0.022211779	0.254633620

NOTE: STANDARD ERRORS COMPUTED USING S(GSQ= 1

ASYMPTOTIC CORRELATION MATRIX OF THE PARAMETERS

CORR	B1	B01	B02	B2
B1	1.0000	-0.9858	-0.9853	0.0107
B01	-0.9858	1.0000	0.9812	-0.0047
B02	-0.9853	0.9812	1.0000	-0.0103
B2	0.0107	-0.0047	-0.0103	1.0000

B1 = Slope
B01 = Intercept from the first treatment regimen
B02 = Intercept from the second treatment regimen
B2 = Body weight covariate

TABLE 5.2.7. SINGLE PROBIT DOSE-RESPONSE FIT TO THE RESULTS FROM TWO TREATMENT REGIMENS. BODY WEIGHT IS INCLUDED AS A COVARIATE.

NON-LINEAR LEAST SQUARES SUMMARY STATISTICS DEPENDENT VARIABLE NDEAD

SOURCE	DF	WEIGHTED SS	WEIGHTED MS
REGRESSION	3	387827.83243	129275.87748
RESIDUAL	68	56.66170	0.82118
UNCORRECTED TOTAL	72	387884.28413	
(CORRECTED TOTAL)	71	518.17751	
SUM OF LOSS		60.88858	

PARAMETER	ESTIMATE	ASYMPTOTIC STD. ERROR	ASYMPTOTIC 95 % CONFIDENCE INTERVAL LOWER	UPPER
B0	-4.86860080	2.6587824168	-10.170736880	0.437535281
B1	12.10482134	3.2223485072	5.676287483	18.533245198
B2	0.14117782	0.0580788088	0.025313543	0.257042103

NOTE: STANDARD ERRORS COMPUTED USING SIGSQ= 1

ASYMPTOTIC CORRELATION MATRIX OF THE PARAMETERS

CORR	B0	B1	B2
B0	1.0000	-0.9978	-0.0114
B1	-0.9978	1.0000	0.0148
B2	-0.0114	0.0148	1.0000

B0 = Intercept
B1 = Slope
B2 = Body weight covariate

The body weight parameter can be interpreted as follows. The logarithm of the LD_{50} dose for a W kg animal, based on the model in Table 5.2.7, can be calculated by solving the equation

$$B0 - 5 + B1 * \hat{x}_{50} + B2(W - \bar{W}) = 0,$$

where $\hat{x}_{50}$ represents the estimated common logarithm of the LD_{50} dose and $\bar{W}$ is the average body weight. Thus,

$$\hat{x}_{50} = - \frac{B0-5}{B1} - \frac{B2}{B1} (W - \bar{W}).$$

If B2 is positive, then this relation can be interpreted as a decrease in the LD_{50} of $100\ (1 - 10^{-B2/B1})$ percent for each 1-kg increase in body weight. For the present example, this is an estimated 2.6 percent decrease in the LD_{50} for each kg increase in body weight, at least for body weights around the average.

6.0 POSSIBLE DIRECTIONS FOR FURTHER EXTENSIONS

This section considers several possible directions for further development, extension, or modification of the methods and procedures that were discussed in the previous sections.

6.1 Parameter Selection

The primary reasons for the selection of the input parameter values of the recommended test procedure are that these values are consistent with the historical database and they appear to provide a test procedure with desirable overall properties as illustrated in Section 2.3. Further work must be performed to develop a procedure for the selection of test procedure parameters, allowing these parameters to vary with the agent and test system. The two major activities that would be required are more extensive simulation studies to characterize the behavior of the recommended test procedure for

various input parameter values and a statistical analysis of the historical database, using the methods of Section 3.2, to determine plausible values of the input parameters.

6.2 Historical Data

The methods discussed in Sections 2.0 and 3.0 assume that an extensive amount of historical data is available. The parameters μ_c, σ_{δ_0}, σ_{δ_1}, and ϵ that characterize the historical response distribution for the standard decontaminant are assumed known. In some applications, however, there might be just a small or moderate amount of historical data available. The historical distribution parameters would then have uncertainty associated with their estimates. This uncertainty would inflate the variability of the historical estimates relative to the expressions presented in Section 2.0 and would thereby result in greater weight being given to the current estimate. The methods discussed in this report can be extended to account for this additional source of variability and its influence on the recommended weighting procedure.

6.3 Deleting Far Past Historical Data

Procedures might be developed for determining when and to what extent to delete the far past standard decontaminant results when they are no longer compatible with the current and more recent past standard decontaminant results. Such decisions would be based on exceedences observed with the control chart procedures discussed in Section 3.0.

6.4 Discounting Historical Data Based on Its Age

Current procedures utilize all the historical data as equivalent as long as they remain within the control limits and as long as weighted averages of various durations remain within the control limits. This is the case whether the historical values were obtained a day, a week, or a year ago. An alternative procedure is to routinely discount the historical data based on

their age, irrespective of whether or not they lie within the control limits. For example, if exponential discounting were used with a discount parameter of $-\lambda$ per month, then one current observation would be discounted to $e^{-\lambda}$ observations in a month, $e^{-12\lambda}$ observations in a year, etc. Thus, if the screen were not used on a regular basis, the extent of the historical information would gradually diminish over time. If the amount of discounted historical information about the standard decontaminant response rate drops below a specified level, then additional tests would be carried out with the standard decontaminant to increase the amount of historical information up to a specified minimum threshold. These additional tests would be carried out at the same agent dose as that used in past tests. If the standard decontaminant response rates drift (or jump) out of control, as determined by the control chart procedures, then a new LD_{50} study would be carried out to determine how to modify the agent dose for future tests.

6.5 Beta Binomial Distribution

The observed response rates, for both the standard and test decontaminants, are currently modelled as being approximately normally distributed. Following an arc sin transformation, the variances of these response rates are assumed to be independent of the mean. The normal approximation to the binomial distribution is reasonable for response rates near 50 percent, as is the case with the current application. For other applications, with response rates closer to 0 or 100 percent, the normality assumption may not be as appropriate. An alternative formulation for such problems would be to model the responses within each individual test as binomially distributed with response probability varying among tests according to a beta distribution. The resulting marginal distribution of the observed standard decontaminant response rates across tests can be described by the beta binomial distribution. This distribution is bounded between 0 and 1 and incorporates skewness in the appropriate direction when the true response probabilities are near 0 or 1.

There are no conceptual differences in a model formulation based on the beta binomial distribution from one based on the normal distribution. There are, however, a number of technical differences; the expressions for the weights would need to be modified.

6.6 Determination of Control Chart Boundaries

In Section 2.0, the test to test variation of standard decontaminant response rates is modelled as a mixture of normal distributions. The control chart limits in Section 3.0, however, are based on a single normal distribution. The distributional assumptions made in Section 3.0, and control limits based on them, might be modified to be brought into conformance with the assumptions made in Section 2.0. Namely, the control limits might be based on the upper percentiles of the mixture distribution, using the iterative calculation recommended in Section 2.2 to determine the critical value. This would probably result in wider control limits than those based on the normal approximation to this distribution.

6.7 Determination of Control Chart Statistics

The discussion in Section 3.3 refers to three alternative statistics to indicate when the standard decontaminant response rates are drifting away from historical levels. An individual, standardized transformed response rate is associated with each test. One statistic is based on the numbers of consecutive individual values that exceed control limits. A second statistic is based on comparing the medians of consecutive individual values to control limits. A third statistic is based on comparing the means of consecutive individual values to control limits.

Table 3.4.1 in Section 3.4 shows that the statistic based on the means is more powerful for detecting small to moderate departures than those based on counts or medians. However, the statistic based on the means is more sensitive to the effects of a small number of outlying values. A compromise between the means-based statistic and the medians-based statistic might be found that simultaneously provides much of the improved sensitivity to detect

systematic departures, yet resists much of the insensitivity and is not as influenced by isolated outlying values. Such a compromise procedure might be based on trimmed means of consecutive values.

6.8 Generalization of the Dose-Response Models

The discussion in Sections 4.0 and 5.0 pertains to experimental design and data analysis considerations in determining the LD_{50} associated with the standard decontaminant. It is assumed there that the dose-response relation can be described by a probit model, without background. This model is adequate for many applications to which the screening methodology has been applied. Other applications, however, might necessitate the use of more general models.

A test period of relatively long duration might result in a nonzero background lethality response rate. Treatment with a specified drug regimen might not be efficacious for all the animals, no matter how much the drug dose is increased. Morbidity responses, such as deterioration of neurological function, may be exhibited by some animals, no matter how high the drug dose, and may not be exhibited by some animals, even in the absence of drug treatment.

The sensitivity analysis procedures and the dose-response model fitting procedures discussed in Sections 4.0 and 5.0 might be extended to accommodate minimum and maximum response rates strictly between 0 and 1. These rates would be additional model parameters, to be estimated from the data.

7.0 REFERENCES

1. Dixon, W.J. and A.M. Mood, "A Method for Obtaining and Analyzing Sensitive Data," Journ. of Amer. Stat. Assoc., 43: 109-126, 1948.

2. Finney, D.J., Probit Analysis, Third Edition, Cambridge University Press, 1971.

3. Joiner, R.L., Dill, G.S., Hobson, D.W., Hassler, C.R., and Feder, P.I., "Task 87-34 (Report 1 of 3): The Effect of Treatment Regimens of Variable Concentrations of Atropine Sulfate in Combination with Pralidoxime Chloride on the Survival of Soman-Challenged Rhesus Monkeys Pretreated with Pyridostigmine Bromide," Appendix C, Stagewise Dose Allocation Procedures, Draft Final Report from Battelle Memorial Institute to U.S. Army Medical Research Institute of Chemical Defense (DAMD17-83-C-3129), 1989.

4. Joiner, R. L., Harroff, H. H., Jr., Snider, T. H., Kiser, R. C., and Feder, P. I., Final Report, "Task 85-10: Validation of a Protocol to Compare the Effectiveness of Experimental Decontaminants With Both Components of the M258A1 Kit Against Percutaneous Application of Undiluted Organophosphate Chemical Surety Materiels to the Laboratory Albino Rabbit," Final Report from Battelle Memorial Institute to U.S. Army Medical Research Institute of Chemical Defense (DAMD17-83-C-3129), 1987.

5. Joiner, R.L., and Kluwe, W.M., "Task 85-18: Conduct of Pralidoxime Chloride, Atropine in Citrate Buffer and Pyridostigmine Bromide Pharmacokinetic Studies, and Comparative Evaluation of the Efficacy of Pyridostigmine Plus Atropine and Pralidoxime Versus Atropine and Pralidoxime Alone Against Acute Soman Poisoning in Male Rhesus Monkeys," Final Report from Battelle Memorial Institute to U.S. Army Medical Research Institute of Chemical Defense (DAMD17-83-C-3129), 1988.

6. Lehmann, E.L., Testing Statistical Hypotheses, John Wiley and Sons, 1959.

7. Miller, R.G., Simultaneous Statistical Inference, McGraw-Hill Book Company, 1966.

8. Ott, E.R., Process Quality Control. Troubleshooting and Interpretation of Data, McGraw-Hill Book Company, 1975.

9. Pearson, E.S. and H.O. Hartley, Biometrika Tables for Statisticians Volume I, University Press, 1958.

10. SAS User's Guide: Statistics, Version 5, SAS Institute, 1985.

11. Schilling, E.G., "A Systematic Approach to the Analysis of Means, Part I," Journ. of Qual. Tech., 5: 93-108, 1973.

APPENDIX A

DOCUMENTATION FOR NEW SAS PROGRAMS

DOCUMENTATION FOR NEW SAS PROGRAMS

A.1 DATA STRUCTURES

A.1.1 Historical Data File

The historical data file is an ASCII file containing standard decontaminant results from first stage screens. There is a separate file for each agent. Each record contains information about a screen, including the starting date, the number of animals dosed, and the number of lethalities. The first field in each record is a "USEFLAG" which can be set to either use or ignore that record when computing historical estimates and creating control charts. As new standard decontaminant data becomes available, it can be appended to the end of the existing historical data file.

File name: <agent>.HIS, where agent is the agent code.

Field Name	Columns	Description
USEFLAG	1	Flag which is set to 0 or 1 to ignore or use data when calculating the historical estimate.
TSEQ	3 - 4	First stage screen identifier (A, B, C...)
AGTCD	6 - 8	Agent code (GD, TGD, VX)
DCNCD	10-12	Decontamination code (STD = standard, A,B,C...=test)
STRTDATE	14-21	First date of testing for the screen (mm/dd/yy format)
DURATION	24-30	Duration of screen in days: (last date) - (first date) + 1
NDOSED	32-38	Number of animals which were dosed
NDEAD	40-46	Number of animals which died

A.1.2 Current Data Files

The current data file is an ASCII file containing results of a single first-stage screen using the standard system and the test decontaminants for a particular agent. This file is in the same format as the historical data file described above. Each record corresponds to a single standard or test decontaminant.

File name: <agent>.CUR where <agent> is the agent code.

A.1.3 Nominal Parameter Values Files

One file per agent, containing the nominal values for the lethality rate and the screen-to-screen variability.

File name: <agent>.NOM and <agent>.EST where agent is the agent code.

Field Name	Columns	Description
MUC	1-20	Nominal lethality rate, μ_c.
SIGD2	21-40	Nominal screen-to-screen variance, σ_δ^2.

A.2 DOCUMENTATION FOR SCREEN PROGRAMS

The following programs perform the analyses in the attached Report on Lethality Rate Estimation and Testing Procedures. The programs are written using in the Statistical Analysis System (SAS) and are designed to run on Battelle's VAX system. It is assumed that the program and data files reside in the default directory of the analyst.

To run the programs, log on to the Battelle computing network requesting a destination of VMSF, and at the VMS prompt (F$), type:

SAS <program name>

followed by a carriage return. Each program produces a file named <program name>.LIS which contains the analysis results, and a file named <program name>.LOG which contains the SAS log.

A.2.1 Documentation for "HISTLETH"

Program file: HISTLETH.SAS

This program uses a supplementary file named HISLETHS.SAS

This program computes estimates of lethality rate and screen-to-screen variance using historical data from first-stage screens. The required input to the program is a file called <agent>.HIS which contains historical data on the standard system for a particular agent. Output of the program consists of a printed report and a data file. The printed report includes the input data, estimates of lethality rate and screen-to-screen variability, and standard errors for these estimates. The data file is called <agent>.EST and includes the estimates of lethality rate and screen-to-screen variance. The format of this file is given in A.1.3.

Prior to running the program, edit HISTLETH.SAS to include the names of the .HIS historical data file to be used and the .EST file to be created.

A listing of the HISTLETH program is provided in Section B-1 of Appendix B. A listing of the HISLETHS file is provided in Section B-2.

A.2.2 Documentation for "COMPARE"

Program file: COMPARE.SAS

This program uses supplementary files named CONTAMC.SAS and CONTAMS.SAS.

This program carries out the statistical test procedures for comparing the current behavior of the standard decontaminant with historical behavior, and for comparing each experimental decontaminant with the standard decontaminant. Two data files are required as input to the program. The first is a file called <agent>.CUR which contains current test system data

including standard system and test decontaminant results. The second is called <agent>.NOM and contains estimates of the nominal historical lethality rate and screen-to-screen variance. Output of the program consists of a report and a data file. The printed report contains listings of the raw data and computed values and significance levels from the test procedures. The data file is named <agent>.NEWHIS and is an ASCII file containing the current standard decontamination data. This file can be APPENDed to the historical data file by the analyst.

Prior to running the program, edit COMPARE.SAS to include the names of the .CUR current system data file and the .NOM file to be used, and the .NEWHIS data file to be created.

A listing of the COMPARE program is provided in Section B-3 of Appendix B. Listings of the CONTAMC and CONTAMS files are provided in Sections B-4 and B-5, respectively.

A.2.3 Documentation for "CRITX"

Program file: CRITX.SAS

This program uses supplementary files named CRITX.DAT (which must be created by the user), CONTAMC.SAS and CONTAMS.SAS.

This program determines the critical values for the test procedure comparing an experimental decontaminant with the standard decontaminant. The analyst specifies values of certain parameters and the program produces a table of critical values for determining when a test decontaminant is no better than the standard decontaminant based on the number of lethalities in the standard and test groups. The six parameters specified by the analyst are:

- n_c number of animals receiving the standard decontaminant
- n_t number of animals receiving the test decontaminant
- μ_c long-term lethality rate for the standard decontaminant
- ϵ mixture probability for the random effect δ

$\sigma_{\delta 0}$ standard deviation of nominal normal distribution

$\sigma_{\delta 1}$ standard deviation of extreme normal distribution

Values for these six parameters should be entered into a one line ASCII file named CRITX.DAT in the order specified above, with at least one blank space separating each value.

The program produces a file named CRITX.LIS displaying the input parameters, and a table of critical values. The critical values are tabulated for values of x_c, the number of lethalities for the standard decontaminant, ranging from 0 to n_c. This information is also printed on the screen as the program runs.

A listing of the CRITX program is provided in Section B-6 of Appendix B. Listings of the CONTAMC and CONTAMS files are provided in Sections B-4 and B-5, respectively.

APPENDIX B

NEW SAS PROGRAM LISTINGS

B-1 PROGRAM LISTING FOR "HISTLETH"

```
DATA HISTORIC;
*******************************************************************
*******************************************************************
*     SUPPLY THE NAME OF THE HISTORICAL DATA FILE IN THE FOLLOWING  *
*                          INFILE STATEMENT                         *
*******************************************************************
*******************************************************************;
  INFILE 'GD.HIS';
  INPUT          USEFLAG    1
                 @3  TSEQ       $2.
                 @6  AGTCD      $3.
                 @10 DCNCD      $3.
                 @14  STRTDATE  MMDDYY8.
                 @24 DURATION
                 @32 NOOSED
                 @40 NOEAD;
  FORMAT STRTDATE MMDDYY8.;
PROC PRINT; TITLE 'HISTORICAL STANDARD DATA';
*
*  GENERATES HISTORICAL ESTIMATES
*;
DATA DUMMY;
  SET HISTORIC;
  IF(USEFLAG EQ 1);
OPTIONS NOSOURCE2;
%INCLUDE HISLETHS;
DATA HISTEST;
  SET CURREST;
*******************************************************************
*******************************************************************
*     SUPPLY THE NAME OF THE NOMINAL VALUE FILE IN THE FOLLOWING  *
*                          FILE STATEMENT                         *
*******************************************************************
*******************************************************************;
  FILE 'GD.EST';
  PUT  @1 PH @21 SIGD2H;
PROC PRINT; TITLE 'HISTORICAL ESTIMATES';
```

B-2 PROGRAM LISTING FOR "HISLETHS"

```
DATA CURREST;
   INPUT PH SIGD2H;
   CARDS;
0.5 0.0
;
* ARCSIN-SQRT TRANSFORMATION IS USED TO TRANSFORM THE BINOMIAL
* MODEL TO CONSTANT VARIANCE
* ;
*
*  ITERATION 1
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NDEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  ppn=pn;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
;
*
*  ITERATION 2
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NDEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  ppn=pn;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
;
*
*  ITERATION 3
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NDEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
```

```
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  pph=ph;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
;
*
*  ITERATION 4
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NOEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  pph=ph;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
;
*
*  ITERATION 5
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NOEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  pph=ph;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
;
*
*  ITERATION 6
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
```

```
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NOEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  ppn=pn;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
;
*
*  ITERATION 7
* ;
DATA ITER;
  SET DUMMY;
  IF _N_ EQ 1 THEN SET CURREST;
  WJ=1/(.25/NOOSED+SIGD2H);
  ZJ=1/(2*(.25/NOOSED+SIGD2H)**2);
  PJ = ARSIN(SQRT(NOEAD/NOOSED)) ;
  SIG2J =(PJ-PH)**2 - .25/NOOSED;
  NUMER1 = WJ*PJ;
  NUMER2 = ZJ*SIG2J;
PROC SUMMARY DATA=ITER;
  VAR NUMER1 NUMER2 WJ ZJ;
  OUTPUT OUT=SUM1 SUM=SUMPJ SUMSIG2J SUMWJ SUMZJ;
DATA CURREST;
  SET CURREST;
  IF (_N_ EQ 1) THEN SET SUM1;
  ppn=pn;
  psigd2h=sigd2h;
  PH = SUMPJ/SUMWJ;
  SEPH = SQRT(1/SUMWJ);
  SIGD2H=SUMSIG2J/SUMZJ;
  SESIGD2H=SQRT(1/SUMZJ);
  PCPH = 100*(PH-PPH)/PPH;
  PCSIGD2H = 100*(SIGD2H-PSIGD2H)/PSIGD2H;
  PUT 'PCT CHANGE IN PH = ' PCPH '.  PCT CHANGE IN SIGD2H = ' PCSIGD2H;
  KEEP PH SEPH SIGD2H SESIGD2H;
*
*  TRANSFORM BACK TO ORIGINAL UNITS
* ;
DATA CURREST;
  SET CURREST;
  SEPH = 2*SIN(PH)*COS(PH)*SEPH;
  PH = SIN(PH)**2;
```

B-3 PROGRAM LISTING FOR "COMPARE"

```
DATA HIST;
*****************************************************************
*****************************************************************
*     SUPPLY THE NAME OF THE NOMINAL VALUE FILE IN THE FOLLOWING  *
*                        INFILE STATEMENT                         *
*****************************************************************
*****************************************************************;
   INFILE 'GD.NOM';
   INPUT   MUC SIGD2;
DATA STDDATA TESTDATA;
*****************************************************************
*****************************************************************
*     SUPPLY THE NAME OF THE CURRENT DATA FILE IN THE FOLLOWING   *
*                        INFILE STATEMENT                         *
*****************************************************************
*****************************************************************;
  INFILE 'GD.CUR';
  INPUT          USEFLAG   1
                 @3  TSEQ      $2.
                 @8  AGTCD     $3.
                 @10 DCNCD     $3.
                 @14  STRTDATE  MMDDYY8.
                 @24 DURATION
                 @32 NOOSED
                 @40 NOEAD;
  FORMAT STRTDATE MMDDYY8.;
  IF (DCNCD EQ 'STD') THEN DO;
    OUTPUT STDDATA;
*****************************************************************
*****************************************************************
*     SUPPLY THE NAME OF THE CURRENT STANDARD DATA FILE IN THE    *
*             FOLLOWING FILE STATEMENT                            *
*****************************************************************
*****************************************************************;
    FILE 'GD.NEWHIS';
    PUT          USEFLAG   1
                 @3  TSEQ      $2.
                 @8  AGTCD     $3.
                 @10 DCNCD     $3.
                 @14  STRTDATE  MMDDYY8.
                 @24 DURATION   7.
                 @32 NOOSED     7.
                 @40 NOEAD      7. ;
    END;
  ELSE OUTPUT TESTDATA;
PROC PRINT DATA=STDDATA; TITLE 'CURRENT STANDARD DATA';
*
* COMPARE HISTORIC & CURRENT LETHALITY RATE ( STANDARD DATA )
*;
DATA COMPARE;
  MERGE HIST STDDATA;
  PC=NOEAD/NOOSED;
  RCSTAR = ARSIN(SQRT(PC));
  MUCSTAR = ARSIN(SQRT(MUC));
  Z1=(RCSTAR-MUCSTAR)/SQRT(.25/NOOSED + SIGD2);
  OSL=2*PROBNORM(-1*ABS(Z1));
  KEEP PC MUC SIGD2 Z1 OSL;
PROC PRINT NOOBS;
  TITLE 'COMPARE CURRENT STANDARD TO HISTORICAL LETHALITY RATE';
  TITLE2 'RATES ARE SIGNIFICANTLY DIFFERENT IF OSL IS LESS THAN 0.05';
DATA CURSTD;
  MERGE STDDATA HIST;
  XC = NOEAD;
  NC = NOOSED;
  SIGD02=SIGD2;
  KEEP XC NC MUC SIGD02;
DATA CURTEST;
  SET TESTDATA;
  XT = NOEAD;
  NT = NOOSED;
  KEEP XT NT DCNCD;
DATA DUMMY;
  SET CURTEST;
  IF _N_ EQ 1 THEN SET CURSTD;
%INCLUDE CONTAMC;
DATA FINAL;
  MERGE TESTDATA CRITICAL; BY DCNCD;
PROC PRINT DATA=FINAL;
  TITLE 'CURRENT TEST DATA, CRITICAL NUMBER AND OSL';
  TITLE2 'TEST DECON FAILS IF NOEAD IS GREATER THAN CRITNUM';
```

B-4 PROGRAM LISTING FOR "CONTAMC"

```
DATA ITER TEMPPRNT;
   SET DUMMY;
*
* INCLUDE NOMINAL VALUES
* ;
   EPS = 0.10;
   SIGDO = SQRT(SIGDO2);
   SIGD1 = 0.4;
*
*  COMPUTE CONSTANTS
* ;
   PI = 2*ARSIN(1.0);
   MUCSTAR = ARSIN(SQRT(MUC));
   SIGC2 = .25/NC;
   SIGT2 = .25/NT;
   SIGDO2 = SIGDO**2;
   SIGD12 = SIGD1**2;
   WO = SIGDO2 / (SIGDO2 + SIGC2);
   W1 = SIGD12 / (SIGD12 + SIGC2);
   RC = XC/24;
   RCSTAR = ARSIN(SQRT(RC));
   R = EXP(-(RCSTAR-MUCSTAR)**2/(2*(SIGC2+SIGD12)))
                  /SQRT(2*PI*(SIGC2+SIGD12)) /
      (EXP(-(RCSTAR-MUCSTAR)**2/(2*(SIGC2+SIGDO2)))
                  /SQRT(2*PI*(SIGC2+SIGDO2)));
   EPSTAR = EPS*R/(1-EPS+EPS*R);
*
* COMPUTE OSL FOR OBSERVED TEST DATA, XT
* ;
   RTSTAR = ARSIN(SQRT(XT/NT));
   ZO = (RTSTAR - (MUCSTAR+WO*(RCSTAR-MUCSTAR)))/SQRT(SIGT2+WO*SIGC2);
   Z1 = (RTSTAR - (MUCSTAR+W1*(RCSTAR-MUCSTAR)))/SQRT(SIGT2+W1*SIGC2);
   FXOSL = (1-EPSTAR) * PROBNORM(ZO)  +  EPSTAR * PROBNORM(Z1);
*
* COMPUTE INITIAL GUESS
* ;
   KNEW = MUCSTAR + (1-EPSTAR)*WO*(RCSTAR-MUCSTAR)
                  + EPSTAR*W1*(RCSTAR-MUCSTAR) +
          1.645*( (1-EPSTAR)*SQRT(SIGT2+WO*SIGC2) +
                     EPSTAR*SQRT(SIGT2+W1*SIGC2) );
IF _N_ EQ 1 THEN OUTPUT TEMPPRNT;
OUTPUT ITER;
PROC PRINT DATA=TEMPPRNT NOOBS;
   TITLE 'PARAMETERS FOR CONTAMINATED NORMAL DISTRIBUTION';
   VAR EPS MUC SIGDO SIGD1 WO W1;
%INCLUDE CONTAMS;
OPTIONS NOSOURCE2;
%INCLUDE CONTAMS;
%INCLUDE CONTAMS;
%INCLUDE CONTAMS;
DATA CRITICAL;
  SET ITER;
  IF(KNEW LT 0) THEN KNEW=0;
  ELSE IF(KNEW GT PI/2) THEN KNEW = PI/2;
  CRITK = SIN(KNEW)**2;
  CRITNUM = CEIL(NT * CRITK);
  OSL=1.0 - FXOSL;
  PUT SIGDO= SIGD1= EPS= NC= NT= XC= CRITNUM= OSL=;
  KEEP DCNCD CRITNUM OSL;
```

B-5 PROGRAM LISTING FOR "CONTAMS"

```
DATA ITER;
   SET ITER;
   KOLD = KNEW;
   Z0 = (KOLD - (MUCSTAR+W0*(RCSTAR-MUCSTAR)))/SQRT(SIGT2+W0*SIGC2);
   Z1 = (KOLD - (MUCSTAR+W1*(RCSTAR-MUCSTAR)))/SQRT(SIGT2+W1*SIGC2);
   FX = (1-EPSTAR) * PROBNORM(Z0)  +  EPSTAR * PROBNORM(Z1) - 0.95;
   FPX = (1-EPSTAR) *
         EXP(-Z0**2/2)/SQRT(2*PI*(SIGT2+W0*SIGC2)) +
        EPSTAR *
         EXP(-Z1**2/2)/SQRT(2*PI*(SIGT2+W1*SIGC2));
   KNEW = KOLD - FX/FPX;
   POIFF = 100*(KNEW-KOLD)/KOLD;
   PUT 'XC = ' XC 'KNEW = ' KNEW '  KOLD = ' KOLD '   POIFF = ' POIFF;
```

B-6 PROGRAM LISTING FOR "CRITX"

```
DATA ITER;
   INFILE 'CRITX.DAT';
   INPUT NC NT MUC EPS SIGDO SIGD1;
   FILE 'SYSSOUTPUT';
   PUT _ALL_=;
*
*  COMPUTE CONSTANTS
* ;
   PI = 2*ARSIN(1.0);
   MUCSTAR = ARSIN(SQRT(MUC));
   SIGC2 = .25/NC;
   SIGT2 = .25/NT;
   SIGDO2 = SIGDO**2;
   SIGD12 = SIGD1**2;
   WO = SIGDO2 / (SIGDO2 + SIGC2);
   W1 = SIGD12 / (SIGD12 + SIGC2);
 DO XC = 0 TO 24;
   RC = XC/NC;
   RCSTAR = ARSIN(SQRT(RC));
   R = EXP(-(RCSTAR-MUCSTAR)**2/(2*(SIGC2+SIGD12)))
                  /SQRT(2*PI*(SIGC2+SIGD12)) /
      (EXP(-(RCSTAR-MUCSTAR)**2/(2*(SIGC2+SIGDO2)))
                  /SQRT(2*PI*(SIGC2+SIGDO2)));
   EPSTAR = EPS*R/(1-EPS+EPS*R);
*
* COMPUTE INITIAL GUESS
* ;
   KNEW = MUCSTAR + (1-EPSTAR)*WO*(RCSTAR-MUCSTAR)
                  + EPSTAR*W1*(RCSTAR-MUCSTAR) +
          1.645*( (1-EPSTAR)*SQRT(SIGT2+WO*SIGC2) +
                     EPSTAR*SQRT(SIGT2+W1*SIGC2) );
   OUTPUT;
   END;
%INCLUDE CONTAMS;
OPTIONS NOSOURCE2;
%INCLUDE CONTAMS;
%INCLUDE CONTAMS;
%INCLUDE CONTAMS;
DATA;
  SET ITER;
  IF(KNEW LT 0) THEN KNEW=0;
  ELSE IF(KNEW GT PI/2) THEN KNEW = PI/2;
  CRITK = SIN(KNEW)**2;
  CRITX = CEIL(NT * CRITK);
  FILE 'SYSSOUTPUT';
  PUT XC= CRITX= ;
PROC PRINT;
  VAR NC NT MUC EPS SIGDO SIGD1 XC CRITX;
  TITLE 'INPUT PARAMETERS AND CRITICAL VALUES FOR XT';
```