2010

An Empirical Approach to Evaluating Sufficient Similarity: Utilization of Euclidean Distance As A Similarity Measure

Scott Marshall
Virginia Commonwealth University

AN EMPIRICAL APPROACH TO EVALUATING SUFFICIENT SIMILARITY IN DOSE-RESPONSE IN COMPLEX CHEMICAL MIXTURES: UTILIZATION OF EUCLIDEAN DISTANCE AS A SIMILARITY MEASURE

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics at Virginia Commonwealth University.

by

Scott Marshall

M.S. Biostatistics, University of Alabama at Birmingham, 2006
B.S. Applied Mathematics, University of Tennessee at Chattanooga, 2003

Director: Chris Gennings, Ph.D.
Professor, Department of Biostatistics

Virginia Commonwealth University
Richmond, Virginia
September, 2009

Acknowledgement

Throughout the process of pursuing my Ph.D. I encountered many challenges and to my own surprise I conquered them all, however, I had very little to do with these successes. I have been extremely blessed to be surrounded by people who provided constant support, motivation, and that can do attitude that is so desperately important in graduate school and life. So it is now that I would like to say thank you to the many people who have so graciously and selflessly lent me their support. Thank you to my parents, Tim and Jo Ann Marshall, who gave me the foundation of discipline and hard work and unconditional love and support; thank you to my sister, Dr. Julie Marshall for all the dissertation complaint sessions that no one else could have understood; thank you to my grandmother, Gram, who without fail called every week to check on me; thank you to Mandi and Jordan for their unconditional love and support through both easy and hard times; thank you to all my committee members and faculty members that put up with me and taught me so much more than statistics; thank you to Yvonne Hargrove who without her constant reminders I would have never registered for a class; last but certainly not least, thank you to Dr. Chris Gennings for being the best advisor, mentor, and friend that a student could ever ask for. It can definitely be stated that the pursuit of and accomplishment of obtaining my Ph.D. was truly a team effort. Thank you team.

Table of Contents

List of Tables

List of Figures

Abstract

AN EMPIRICAL APPROACH TO EVALUATING SUFFICIENT SIMILARITY IN DOSE-RESPONSE IN COMPLEX CHEMICAL MIXTURES: UTILIZATION OF EUCLIDEAN DISTANCE AS A SIMILARITY MEASURE

By Scott Leo Marshall

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics at Virginia Commonwealth University.

Virginia Commonwealth University, 2009.

Research Director:

Dr. Chris Gennings
Professor
Department of Biostatistics

Individuals are exposed to chemical mixtures while carrying out everyday tasks, with unknown risk associated with exposure. Given the number of resulting mixtures it is not economically feasible to identify or characterize all possible mixtures. When complete dose-response data are not available on a (candidate) mixture of concern, EPA guidelines define a similar mixture based on chemical composition, component proportions and expert biological judgment (EPA, 1986, 2000).

Current work in this literature is by Feder et al. (2009), evaluating sufficient similarity in exposure to disinfection by-products of water purification using multivariate statistical

techniques and traditional hypothesis testing. The work of Stork et al. (2008) introduced the idea of sufficient similarity in dose-response (making a connection between exposure and effect). They developed methods to evaluate sufficient similarity of a fully characterized reference mixture, with dose-response data available, and a candidate mixture with only mixing proportions available. A limitation of the approach is that the two mixtures must contain the same components.

It is of interest to determine whether a fully characterized reference mixture (representative of the random process) is sufficiently similar in dose-response to a candidate mixture resulting from a random process. Four similarity measures based on Euclidean distance are developed to aid in the evaluation of sufficient similarity in dose-response, allowing for mixtures to be subsets of each other. If a reference and candidate mixture are concluded to be sufficiently similar in dose-response, inference about the candidate mixture can be based on the reference mixture. An example is presented demonstrating that the benchmark dose (BMD) of the reference mixture can be used as a surrogate measure of BMD for the candidate mixture when the two mixtures are determined to be sufficiently similar in dose-response. Guidelines are developed that enable the researcher to evaluate the performance of the proposed similarity measures.

Chapter 1: Introduction and Prospectus

Section 1.1 Introduction

Individuals are exposed to chemical mixtures throughout the processes of carrying out everyday tasks. Whether an individual takes a drink of water from the faucet or eats a piece of fruit or a vegetable from the grocery store they are exposed to some type of chemical mixture. The particular chemical mixtures to which individuals are exposed can be thought of as the result of randomly occurring processes. This process could be the application of pesticides in different locations, such as child care centers (Tulve et al., 2006); the disinfection of drinking water at different purification stations with different source waters (Bull et al., 2009); the breakdown of environmental estrogens and the subsequent leaching into the soil and drinking water at different waste sites (Vom Saal and Hughes, 2005). Each of these random processes creates different chemical mixtures, which implies that individuals are exposed to different levels and different chemicals as a result of the different mixtures.

Characterizing the toxicologic effect of these mixtures, particularly the effect of exposure to these mixtures, is of great importance with respect to risk assessment. However, it is important to note that the toxicological effect of a chemical mixture depends on the toxicity of the components and how the components interact with each other in a dose-dependent way

(Gennings et al., 2004). Given the number of mixtures generated as the result of these random processes, it is not economically feasible to identify or characterize all possible mixtures.

Consider the case of mixtures of polycyclic aromatic hydrocarbons (PAHs) that can occur in the environment through residential heating and cooking, incinerators, and by many other means. One of the suggested available approaches to PAH risk assessment is the surrogate approach, described in an EPA guidance document as:

> "a whole mixture approach based on the assumption that any mixture of PAHs in the atmosphere (or mixture of concern) is merely a dilution of a 'surrogate' mixture of PAHs, the 'surrogate' being a potent PAH-containing mixture that has been well characterized both chemically and toxicologically. Under this assumption, the risk from any PAH mixture of concern is directly related to the extent of this dilution. The extent of dilution is based on examining the ratios of several PAHs common to both the mixture of concern and the surrogate mixture. The surrogate approach is based on the Agency's mixtures' guideline that recognizes and endorses whole mixture approaches. Fundamental difficulties of this approach include the appropriate choice of a 'surrogate' whole mixture and evaluation of 'sufficient similarity' to the mixture of concern, based on EPA's mixtures' guidelines. Major advantages to the surrogate approach include: (1) a mixture (as compared to single components) is used as the reference compound, and (2) the composition and toxicity of the surrogate mixture as a whole is known." (EPA, 2001)

Ideally, environmental health risk assessments are conducted using dose-response data from the mixture of concern (Feder et al., 2009). However, when complete dose-response data are not available on a (candidate) mixture of concern, EPA guidelines define a similar mixture based on similar chemical composition and similar component proportions, based on expert biological judgment (EPA, 1986, 2000). The guidance does not provide specific direction on methods to approach this type of problem (Feder et al., 2009) and as Stork et al. (2008) point out, one of the inherent difficulties with this approach to using the concept of sufficient similarity in a whole

mixture approach to risk assessment is making an appropriate choice of the surrogate (reference) mixture.

To this point there have been relatively few attempts at developing empirical approaches to evaluating sufficient similarity. The most current work in the literature is by Feder et al. (2009), Bull et al. (2009) and Rice et al. (2009). In an example evaluating sufficient similarity in disinfection by-products of water purification, Feder et al. (2009) present multivariate statistical techniques for evaluating sufficient similarity, that include utilization of Hotelling's T^2 and principal components analysis. They characterize similarity with respect to process input and output variables. For purposes of demonstrating how to implement their proposed method, we will only consider the case of assessing similarity with respect to process input variables. Consider that there are nine process input variables where two of the input variables are dichotomous, such that i=1,2 and j=1,2. The other p=7 input variables are considered to be continuous response variables so that for each combination (i, j) a p-dimensional joint distribution exists with mean vector μ_{ij} and covariance matrix Σ_{ij}. It is assumed that

$$X_{ij} \sim MVN(\mu_{ij}, \Sigma_{ij}) \quad i, j = 1, 2 . \tag{1.1}$$

They then suggest that with sufficient data, distribution parameters can be estimated with a linear model:

$$\mu_{ij} = \mu_0 + \alpha I_i + \beta J_j + \gamma I_i J_j \quad \text{where } I_i = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{o.w.} \end{cases}$$

$$\text{and} \quad J_j = \begin{cases} 1 & \text{if } j = 1 \\ 0 & \text{o.w.} \end{cases}$$

where I and J are indicator functions for the previously defined dichotomous variables. If the interaction term, γ, is significant this implies that at a given level of i there is a difference in means at different levels of j and vice versa. If the interaction term is not significant the model reduces to a main effects model where when α and β are both significant it implies that there is a difference in means within i and j, respectively. There is one main issue with applying this linear model that Feder et al. (2009) do not address. Concluding that means are different with respect to "some" input variables, i and j, is not of interest in determining sufficient similarity. Further, failing to conclude that means are different (failing to reject the null hypothesis) does not imply that the means are the same (equivalent to accepting the null hypothesis). Also, recall that there exists a p-dimensional joint distribution among the continuous input variables, implying some covariance structure exists.

Feder et al. (2009) also consider comparing a new treatment process to a reference set of similar processes. Assume that a reference set distribution exists across i and j, and that it is modeled as normally distributed with mean and covariance as defined above. Now, $N = \sum_{i,j} N_{ij}$ samples are selected from this single population of similar processes and the p-dimensional response vector for this set of responses is $\underline{X}_{ij}$. Suppose an additional treatment plan has a p-dimensional response vector Y and it is desired to test the null hypothesis that the process that generates Y is the same as the process that generates $\underline{X}_{ij}$ (Feder et al., 2009). Assuming a common covariance matrix among X and Y, Hotelling's T^2 statistic can be used where Hotelling's T^2 is defined as

$T^2 = \left(\bar{Y} - \bar{X}\right)' \left[\left(1 + \frac{1}{N}\right) S\right]^{-1} \left(\bar{Y} - \bar{X}\right)$ where under the null hypothesis T^2 is proportional to that of a central *F* distribution with p and $v - (p-1)$ degrees of freedom

$\frac{v-(p-1)}{vp} T^2 \sim F_{p,v-(p-1)}$ (Feder et al., 2009). Similar to the situation with the linear model, failing to reject the null hypothesis does not imply that one can accept the null and conclude the means are the same. This method does allow for determining if a joint mean vector is different. However, when evaluating sufficient similarity the concept is to be able to conclude that X and Y are the same.

The work of Stork et al. (2008) makes an attempt to evaluate the sufficient similarity of chemical mixtures in an empirical manner, utilizing dose-response data. Stork et al (2008) propose using mixed model theory and the principle of confidence interval/region inclusion to test for sufficient similarity in dose-response. Because these chemical mixtures can be thought of as the result of random processes, the corresponding dose-response curves can be thought of as representing random exposure/dilutions and this needs to be accounted for in the total variability in the dose-response curve. Stork et al. (2008) suggest using a non-linear mixed effects model to account for the random changes in the exposure/dilutions, where the random effect (random coefficient), denoted as *h*, can be thought of as the similarity measure. It is of interest to determine how much variability due to the random process is associated with mixtures that are sufficiently similar in dose-response. It should be noted that any cumulative distribution function can be used, such as the Gompertz, logistic, or exponential (this is not an exhaustive

list). In evaluating risk in mixture studies it is often of interest to be able to detect a threshold. Stork et al. (2008) suggest a non-linear mixed effects model of the following form

$$y = f(t, \underline{\omega}, b) + \varepsilon = \begin{Bmatrix} \alpha + \gamma k & \text{if } \beta th > \delta \\ \alpha + \gamma F(\beta th - \delta) & \text{if } \beta th \le \delta \end{Bmatrix} + \varepsilon \text{ if } \beta < 0 \qquad (1.2)$$

where y is the response variable; $\underline{\omega} = [\alpha, \gamma, \beta, \delta]'$ is a $[p \times 1]$ vector of unknown parameters; β is a parameter associated with the slope of the curve and the random coefficient; t is the total dose of the mixture; δ is a parameter associated with the dose-threshold of the curve; k is a constant satisfying the nonlinear threshold constraint $k = F(\beta th = \delta)$; $h = 1 + b$ is the random effect where $b \sim N(0, \sigma_h^2)$ so that $h \sim N(1, \sigma_h^2)$; ε are iid random errors such that $\underline{\varepsilon} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]'$ is an N-vector assumed to follow $N(0, \underline{R})$ where $\underline{R} = \sigma^2 I_N$; and ε and b are assumed to be distributed independent of each other. Without loss of generality, when increasing curves are considered, the inequalities in eq. (1.2) are switched.

Consider the following notation and general setup presented by Stork et al (2008) to describe the mixtures and dose-response curves that result from some dynamic process. Let j (j=1,2,…) be the number of randomly sampled mixtures where each mixture contains the same c(c=1,2,…) components/chemicals but for each j the mixing ratios are different; the c mixing ratios (proportions) for the j^{th} mixture are $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{cj} \end{bmatrix}$ where $\sum_{i=1}^{c} a_{ij} = 1$; y_{jk} is the response from the

k^{th} observation of the j^{th} mixture where $1 \le k \le n_j$; $x_{ij} = a_{ij} t_j$ is the dose of the i^{th} chemical $(1 \le i \le c)$ of the j^{th} mixture where $t_j = \sum_{i=1}^{c} x_{ij}$ is the total dose of the j^{th} mixture; and $N = \sum_j n_j$ are the total number of observations for a given dose-response curve. Recall it is assumed that complete dose-response data are available on all *j* sampled mixtures (i.e., the data rich case). In order to test if two curves are sufficiently similar in dose-response, Stork et al. (2008) suggest following equivalence testing logic and further presents the idea of reparameterizing the dose-response curve in eq. (1.2) as functions of the model parameters (conditional on the minimum and maximum effect parameters). Consider *D* functions of the model parameters $g(\underline{\omega}) = \begin{bmatrix} g_1(\underline{\omega}) \\ \vdots \\ g_D(\underline{\omega}) \end{bmatrix}$ where the d=1,…,D have an intuitive meaning such as an

$$ED(\mu_0) = \frac{F^{-1}\left(\frac{\mu_0 - \alpha}{\gamma}\right) + \delta}{\beta h} \text{ or} \tag{1.3}$$

$$\delta^* = \frac{\delta}{\beta} \tag{1.4}$$

a dose-threshold. For visualization the dose-response curve is reparameterized in two dimensions (D=2). Using expert judgment, shifts of the curve associated with inappreciable differences representing a biologically significant region can be determined. Let's call these

shifts $\begin{bmatrix} (\Delta_1, \Delta_2) \\ (\Delta_3, \Delta_4) \end{bmatrix}$. Boundary curves defined by these shifts specify a region of similarity. Stork et al. (2008) developed an approach to test for sufficient similarity following the principles of equivalence testing logic and mixed model theory (described in Chapter 2) that, when *D*=2, tests the hypothesis

$$\begin{aligned} H_0 &: g_1(\underline{\omega}) < \Delta_1 \text{ or } g_1(\underline{\omega}) > \Delta_2 \text{ or } g_2(\underline{\omega}) < \Delta_3 \text{ or } g_2(\underline{\omega}) > \Delta_4 \\ H_a &: \Delta_1 \le g_1(\underline{\omega}) \le \Delta_2 \text{ and } \Delta_3 \le g_2(\underline{\omega}) \le \Delta_4 \end{aligned} \tag{1.5}$$

In order to test the hypothesis stated in eq. (1.5) graphical methods were developed that extend the principles of confidence interval/region inclusion (Berger and Hsu, 1996) to confidence region inclusion with multidimensional hypotheses (see chapter 2 appendix). The acceptable shifts define a similarity region that can be plotted and further a conservative *D*-dimensional $100(1-\alpha)\%$ confidence region, for any parametric function $g(\underline{\omega})$, can be plotted (described in chapter 2 appendix). The variance-covariance matrix of this confidence region is a function of the variance of the random exposure/dilution factor, σ_h^2. If the ellipse is contained within in the region then the null hypothesis in eq. (1.5) can be rejected. Berger and Hsu (1996, Theorem 4) argue that this is a valid $\alpha-$level test. When *D*=2 the resulting confidence region is an ellipse. This method is generalizable to the case when $D > 2$, however it becomes harder to visualize.

The situation described above is the "ideal" data rich situation. In the case when data are only available on a surrogate/reference mixture (i.e., the data poor case) Stork et al. (2008) describe a method that utilizes the concepts of mixed models and confidence region inclusion as previously described. Consider the mixed model in eq. (1.2). When this mixed model is fit to a

single dose-response curve, it is equivalent to fitting a mixed effects model with $\sigma_h^2 = 0$ and *h*=1. This is the same as fitting a fixed effects model to the single dose-response curve. Again, think of reparameterizing the dose-response curve in terms of some parametric functions, $g(\underline{\omega})$, that have an intuitive or toxicologically relevant meaning. Once again, shifts of the dose-response curve (in terms of the specified functions) associated with inappreciable differences can be specified through expert judgment as $\begin{bmatrix} (\Delta_1, \Delta_2) \\ (\Delta_3, \Delta_4) \end{bmatrix}$. These shifts create the similarity region. Once the similarity region is determined, the associated *D*-dimensional $100(1-\alpha)\%$ confidence region is formed for $g(\underline{\omega})$. Recall that the variance-covariance matrix of this confidence region is a function of σ_h^2. Given that there is only one dose-response curve, $\sigma_h^2 = 0$. The additional variability that can be added to this ellipse such that it is still contained in the similarity region can be found by incrementing through values of σ_h^2 while holding fixed the mean parameters and variance estimates. The form of this variance-covariance matrix is obtained through the zero order Taylor series expansion (Wolfinger and Lin, 1997) as described in the appendix of chapter 3. Based on the maximum size of σ_h^2, the acceptable interval (h_L, h_U) for the random exposure/dilution factor (the similarity measure),*h*, is obtained based on the similarity region such that $h_L = 1 - z_{\max}\sigma_h^2$ and $h_U = 1 + z_{\max}\sigma_h^2$ where $z_{\max}$ is chosen such that at least one of the following holds for some acceptably small value of $\varepsilon_j^* > 0,\ j = 1,2,3,4$

$$
\begin{aligned}
g_1(\underline{\omega})h_L - \Delta_1 &< \varepsilon_1^* \\
g_1(\underline{\omega})h_U - \Delta_2 &< \varepsilon_2^* \\
g_2(\underline{\omega})h_L - \Delta_3 &< \varepsilon_3^* \\
g_2(\underline{\omega})h_U - \Delta_4 &< \varepsilon_4^*
\end{aligned}.
$$

The acceptable interval of the similarity measure is decomposed to form acceptable intervals for the mixture components (a_{iL}, a_{iU}) for $i = 1,...,c$ such that

$$a_{iL} = \frac{a_i h_L}{a_i h_L + (1-a_i)h_U} \text{ and } a_{iU} = \frac{a_i h_U}{a_i h_U + (1-a_i)h_L}.$$

Given a randomly sampled candidate mixture, if each $a_i \in (a_{iL}, a_{iU})$ for $i = 1,...,c$ then it is concluded that the reference and candidate mixtures are sufficiently similar in dose-response.

The methods of Stork et al. (2008) require calculation of σ_h^2 and that all chemical mixtures of concern contain the same *c* components. The methods proposed in the following chapters allow for mixtures of concern to be subsets of each other and a relationship is developed that does not require the calculation of σ_h^2. Guidelines are developed that enable the researcher to evaluate the performance of the proposed methods.

Section 1.2 Prospectus

In Chapter 2 the concept of adjusting dose scale due to chemicals that are in 'active' and 'inactive' dose ranges is explored in the data rich case when it is possible to use the "gold standard" test for sufficient similarity. It is suggested to use the equivalence test, which we coin

the "gold standard" test in Chapter 2, as proposed by Stork et al. (2008) when complete dose-response data are available on both the reference and candidate mixtures.

While the multivariate statistical methods, such as principal components analysis presented by Bull et al. (2009) present viable techniques for the data poor case we will focus on advancing and extending the method proposed by Stork et al. (2008). This method is most suitable in the case when there are complete data on the reference mixture and only mixing ratios for the candidate mixture. This approach made advances in providing an empirical approach to evaluating sufficient similarity in dose-response. However, the one inherent restriction in this method is that it requires the reference mixture and candidate mixture to have the same c chemicals in both mixtures. The research in Chapter 3 proposes an extension to this method and a working example that utilizes the concept of Euclidean distance to provide a similarity measure, h, that can ultimately be used for the purposes of risk assessment. Utilizing the concept of Euclidean distance provides for either the reference or candidate mixture being a subset of the other. There are four different similarity measures presented in this research which allows for the measure to account for additional chemicals/components (in either mixture) that are either sub-threshold (in an inactive range) or at the threshold or beyond (in an active range). This method provides additional flexibility by allowing for a weight matrix, W, to be used to up weight more potent components or down weight less potent components, for example.

In Chapter 4, simulation studies are conducted for two cases that evaluate the properties of the four proposed similarity measures. This research addresses how to characterize the properties of the methods described by Stork et al. (2008) as well as the proposed method in

Chapter 3. Measures such as sensitivity and specificity are evaluated for the two cases that are presented.

Chapter 5 addresses general technical issues that can arise throughout the process of evaluating sufficient similarity. How to address issues such as additivity when conducting simulations, power of the "gold standard" method, study design, general simulation issues, and issues regarding model parameterization are outlined in this chapter.

Chapter 6 includes a discussion and summary of Chapters 2, 3, 4, and 5. Also included are possible extensions to the methods presented in Chapter 3. The Appendices include some additional figures and tables and the associated SAS programs. While reading Chapters 2 and 3 be aware that there will be some redundancy as these chapters are essentially written as stand alone papers.

This research was partially supported by the National Institute of Environmental Health Sciences, National Institute of Health (NIEHS, NIH) training grant #T32ES007334.

Chapter 2 : Testing for Sufficient Similarity: Do Interaction and Dose Scale Matter?

Section 2.1 Introduction

The protection of human health from adverse effects of environmental exposure to chemical mixtures is an important issue (Gennings et al., 2004). Recently, interest in understanding the potential human health effects of exposure to chemical mixtures has increased due to congressional acts, such as, the Food Quality Protection Act of 1996 and the Safe Drinking Water Act Amendments of 1996 (Food Quality Protection Act, 1996; Safe Drinking Water Act Amendments, 1996). These congressional acts direct that assessments of pesticide safety include consideration of the risk(s) associated with the cumulative effects of chemicals that have a common mechanism of toxicity and request the development of new approaches for studying complex chemical mixtures (Gennings et al., 2004). The U.S. EPA has developed guidance and suggests procedures for conducting health risk assessment for complex chemical mixtures (Feder et al., 2009; U.S. EPA, 1986, 2000) and ideally these assessments are conducted using dose-response data from the chemical mixture of concern (Feder et al., 2009). When toxicity data are not available (data poor situation) for a chemical mixture of concern, U.S. EPA guidelines allow risk assessment to be based on data for a surrogate mixture considered "sufficiently similar" in terms of chemical composition and component proportions. As a supplementary approach, using statistical equivalence testing logic Stork et al. (2008) developed

methodology to define sufficient similarity in dose-response for mixtures of many chemicals containing the same components with different ratios when complete dose-response data are available (i.e., the data rich situation) on all mixtures of concern. Using the method of confidence region inclusion as described by Stork et al. (2008), the appropriate equivalence type test for sufficient similarity can be conducted (Berger and Hsu, 1996). The equivalence test conducted in the data rich situation is coined the "gold standard" test for sufficient similarity.

Motivating Example

To illustrate, consider (Gennings et al., 2004) 5 OP pesticides (acephate, chlorpyrifos, diazinon, dimethoate, and malathion) in a mixing ratio based on an average exposure level as specified by the Dietary Evaluation Exposure Model (DEEM) conducted by the U.S. EPA (Table 2.1). It is known that there is an interaction between malathion and the other OP pesticides (Gennings et al., 2004). To evaluate potential neurotoxicity, the endpoint for this analysis was a dichotomized gait score: normal gait (Y=0) vs abnormal gait (Y=1).

Consider Mixture 1 (reference mixture) which contains all five OP pesticides and Mixture 2 (candidate mixture) which excludes malathion. For ease of notation in later sections Mixture 1 is referred to as the 'full' mixture and Mixture 2 as the 'reduced' mixture. The full mixture consists of 82.5% of malathion, which is inactive alone given the dose range of the study (0-450 mg/kg), while the remaining 17.5% of the mixture is a mixture of four dose-responsive chemicals; and the reduced mixture contains only the four dose-responsive chemicals (0-78.8 mg/kg). This is to say that given a large enough dose, malathion could elicit an effect, however, given the dose range of the study it is not dose-responsive. In fact, doses of malathion alone (0-

500 mg/kg) yielded no effect on gait score. Essentially malathion acts in a way that artificially inflates total dose as mass is added to the total dose that is not expected to elicit any response. However, given these two mixtures and the knowledge that malathion interacts with the other four chemicals, it is of interest to determine what effect, if any, malathion has on the dose-response relationship of the other four pesticides with respect to the concept of sufficient similarity.

While it is preferable for risk assessments of chemical mixtures to be based on toxicity and exposure (dose-response) data on the chemical mixtures of concern (Rice et al., 2009), it is not the actual risk assessment that is of concern in this example but how an additional chemical affects the dose-response curve with respect to sufficient similarity. In practice, it is often the situation that one chemical mixture is a subset of the other and it is of particular interest to evaluate the effect the subset (or the additional chemical(s)) has on the dose-response relationship with respect to determining sufficient similarity.

An important extension to the work of Stork et al. (2008) is to develop methodology such that sufficient similarity can be determined for a full and a reduced mixture which contains only a subset of the chemicals in the full mixture. Consider the example where the full mixture has an additional chemical (as compared to the reduced mixture), malathion, that is not dose-responsive given the dose range of the study. It may be reasonable to assume (under the assumption of no interaction) that a mixture of chemicals in an active dose range, such as the reduced mixture, should be sufficiently similar to the same mixture with the addition of a chemical that is not dose-responsive or is in an inactive dose range. The following methodology follows the work of

Stork et al. (2008) and presents a dose adjustment factor to be utilized in testing for sufficient similarity in dose-response.

Section 2.2 Methods

It is of interest to test whether two mixtures, a full mixture and a reduced mixture, are sufficiently similar in dose-response; one mixture contains an additional chemical in an inactive dose range in a substantial quantity (i.e. one mixture is a subset of the other) that causes the dose scales to be significantly different. It is assumed that this additional chemical in the mixture adds additional mass in terms of total dose that will shift the dose-response curve to the right, solely as a function of the increasing total dose. That is, due to dose scale there might not be sufficient evidence to conclude that the two curves are sufficiently similar and this conclusion is an artifact of the differing dose scales. This suggests that it is necessary to make an adjustment to the dose scales before performing an equivalence test (utilizing graphical methods and the concept of confidence region inclusion) to test whether the two dose-response curves are sufficiently similar in dose-response. Following the methods suggested by Stork et al. (2008) (steps 2-5), the following steps will form the hypothesis test for sufficient similarity

1. Rescale the total dose for the mixture with chemicals that are inactive (in an inactive range/subthreshold).

2. Fit the appropriate mixed effects/random coefficients model.

3. Fully characterize the curve as functions of the model parameters and determine the similarity bounds (region) based on allowable shifts of the functions of model parameters.

4. Construct a $100(1-\alpha)\%$ confidence region for the defined functions of model parameters.

5. By inspection determine if the confidence region is contained within the similarity region to determine if sufficient similarity can be concluded.

Step 1

Consider the case where there are two mixtures of interest; a full mixture and a reduced mixture. The following notation is established for the reduced (*red)* and full (*full*) mixtures, respectively, $\underline{a}_{red} = \{a_{i,red}\}$ where $i = 1,...,k$, $\underline{a}_{full} = \{a_{i,full}\}$ where $i = 1,...,c$ and $a_{i,full}$ and $a_{i,red}$ are the individual chemical proportions of each mixture. Define the two mixtures as:

$$\underline{a}_{red} = \begin{bmatrix} a_{1,red} \\ a_{2,red} \\ \vdots \\ a_{k,red} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ and } \underline{a}_{full} = \begin{bmatrix} a_{1,full} \\ a_{2,full} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ a_{c,full} \end{bmatrix}$$ where there are $c-k=s$ "placeholder" zeros in $\underline{a}_{red}$. Also,

$\underline{a}_{red}t = \underline{x}_{red}$ and $\underline{a}_{full}t = \underline{x}_{full}$. The total dose groups for the reduced mixture are

$\begin{bmatrix} t_{0,red} & t_{1,red} & \cdots & t_{d,red} \end{bmatrix}$ where $j = 0,1,\ldots,d$, $\{a_{i,red}t_{j,red}\} = x_{ij,red}$ and $x_{ij,red}$ is the dose of the i^{th} component in the mixture at the j^{th} total dose, $j = 0,1,\ldots,d$, such that, $\sum_{i=1}^{k} x_{ij,red} = t_{j,red}$. The total dose groups for the full mixture are

$\left[t_{0,full} \;\; t_{1,full} \;\; \cdots \;\; t_{J,full} \right]$ where $j = 0,1,\ldots,J$, $\left\{ a_{i,full} t_{j,full} \right\} = x_{ij,full}$ and $x_{ij,full}$ is the dose of the i^{th} component in the mixture at total dose, t_j , such that, $\sum_{i=1}^{c} x_{ij,full} = t_{j,full}$. The total dose for the full mixture can be rescaled by adjusting the value based on the percentage of active (i.e., dose-responsive) chemicals in the mixture. We will define the dose adjustment factor as the proportion of chemicals in the full mixture that are also in the reduced mixture, $D_{AF} = \sum_{i=1}^{k} a_{i,full}$ such that $t_{j,full\_adj} = t_{j,full} \sum_{i=1}^{k} a_{i,full}$ is the rescaled total dose for the full mixture. This adjustment attempts to bring the dose scales closer together. The reasoning for adjusting the total dose scale follows the logic of Casey et al. (2004) using data described in Moser et al. (2005).

Step 2

Following the logic of Stork et al. (2008) and without loss of generality, consider the following mixed-effects model

$$E[y] = f(t, \underline{\omega}, h) = F\left(\beta_0 + \beta_1 h t\right) \qquad (2.1)$$

where *y* is the response variable; $\underline{\omega} = \left[\beta_0, \beta_1\right]$ is a $\left(p \times 1\right)$ vector of unknown model parameters (p=2); β_0 is a parameter associated with the intercept of the curve; *t* is the total dose of the mixture; β_1 is a parameter associated with the slope and the random coefficient; *h* is the random effect where $h \sim N(1, \sigma_h^2)$.

Step 3

Again following Stork et al. (2008), for the purpose of determining shifts in the dose-response curve that are associated with inappreciable differences, it is useful to reparameterize eq. (2.1) as functions of the model parameters that have an intuitive meaning. Consider D functions of the model parameters $g(\underline{\omega}) = \begin{bmatrix} g_1(\underline{\omega}) \\ \vdots \\ g_D(\underline{\omega}) \end{bmatrix}$ where the d=1,…,D functions have an intuitive meaning such as an effective dose (*ED*) of interest or a dose-threshold depending on the specified model. For example, when D=2, the similarity region is defined in terms of allowable percentage shifts of $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$ that represent a biologically relevant similarity region. Without loss of generality, for this example the allowable shifts are as follows

$$\begin{aligned} \Delta_1 &= g_1(\underline{\omega}) - \pi_1 \times g_1(\underline{\omega}),\ \Delta_2 = g_1(\underline{\omega}) + \pi_2 \times g_1(\underline{\omega}) \\ \Delta_3 &= g_2(\underline{\omega}) - \pi_3 \times g_2(\underline{\omega}),\ \Delta_4 = g_2(\underline{\omega}) + \pi_4 \times g_2(\underline{\omega}) \end{aligned} \tag{2.2}$$

so that the similarity region is defined as $\begin{bmatrix} (\Delta_1, \Delta_2) \\ (\Delta_3, \Delta_4) \end{bmatrix}$. The π_i are the allowable percentage shifts in $g_d(\underline{\omega})$, d=1,2.

Step 4

Following Stork et al. (2008), in order to estimate/construct the $100(1-\alpha)\%$ confidence region in terms of $g(\underline{\omega})$, the asymptotic variance-covariance matrix for the model parameters is

estimated. Let Ω be the inverse of the Hessian matrix evaluated at the parameter estimates denoted as, $\hat{\Omega}$. Consider the parameter vector

$\underline{\omega} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ where $\hat{\underline{\omega}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$. A conservative *D*-dimensional 100(1-α)% confidence region (Spjotvoll, 1972; Rao, 1973) for ω is given by

$$P\left[\inf_{\omega \in C} \omega \le \omega \le \sup_{\omega \in C} \omega\right] \ge 1-\alpha \tag{2.3}$$

where *C* is the 100(1-α)% region for ω. The region, *C*, is defined by the ellipsoid centered at $\hat{\omega}$ (Seber and Wild,1989) and is given by the Wald-type confidence region (quadratic form):

$$C = \left\{\omega : \left(\hat{\omega}-\omega\right)' \hat{\Omega}^{-1} \left(\hat{\omega}-\omega\right) \le \chi^2_{1-\alpha,D}\right\} \tag{2.4}$$

The quadratic form in eq. (2.4) has an approximate $\chi^2_{1-\alpha,D}$ distribution.

The asymptotic variance-covariance matrix is for functions of the model parameters and is calculated by using the Multivariate Delta Method.

Recall that $\underline{\omega} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ and the asymptotic variance covariance matrix of $\underline{\omega}$ is $\underline{\Omega}$. Consider a function of $\underline{\omega}$, say $g(\underline{\omega})$, and

$$G = \begin{pmatrix} \frac{dg_1(\underline{\omega})}{d\omega_1} & \cdots & \frac{dg_1(\underline{\omega})}{d\omega_P} \\ \frac{dg_2(\underline{\omega})}{d\omega_1} & \cdots & \frac{dg_2(\underline{\omega})}{d\omega_P} \\ \vdots & & \vdots \\ \frac{dg_D(\underline{\omega})}{d\omega_1} & & \frac{dg_D(\underline{\omega})}{d\omega_P} \end{pmatrix} \tag{2.5}$$

The asymptotic variance-covariance matrix of $g(\hat{\underline{\omega}})$, $\hat{\underline{\Gamma}}$,is $G\Omega G'|_{\hat{\omega}}$.

A conservative *D*-dimensional (Spjotvoll, 1972; Rao 1973) 100(1-α)% confidence region for $g(\omega)$ is given by

$$P\left[\inf_{\omega\in C} g(\underline{\omega}) \le g(\omega) \le \sup_{\omega\in C} g(\underline{\omega})\right] \ge 1-\alpha \tag{2.6}$$

where *C* is the 100(1-α)% region for $g(\underline{\omega})$. The region, *C*, is defined by the ellipsoid centered at $\hat{\omega}$ and is given by the Wald-type confidence region (quadratic form):

$$C = \left\{\omega : \left(g(\hat{\underline{\omega}}) - g(\underline{\omega})\right)' \hat{\Gamma}^{-1} \left(g(\hat{\underline{\omega}}) - g(\underline{\omega})\right) \le \chi^2_{1-\alpha,D}\right\} \tag{2.7}$$

In order to plot the confidence region, the bounds of the confidence region need to be obtained first. Following the logic of Carter (1986) to calculate the bounds of the confidence region, it is first necessary to identify the points on the boundary *C*. Anderson (1958) gives a transformation from rectangular coordinates to polar coordinates that permits identification of points on the boundary of *C* which expedites this search in *D* dimensions (Appendix A.2). For the case when *D*=2 this reduces to the appropriate confidence ellipse.

Step 5

Evaluating the confidence region to determine if all boundary points of *C* are contained within the similarity region performs the equivalence test for sufficient similarity in dose-response. For the case when *D*=2, plotting the confidence region as described in Step 3 simultaneously with the specified similarity region allows for the following hypothesis test of equivalence to be conducted

$$\begin{aligned} H_0 &: g_1(\underline{\omega}) < \Delta_1 \text{ or } g_1(\underline{\omega}) > \Delta_2 \text{ or } g_2(\underline{\omega}) < \Delta_3 \text{ or } g_2(\underline{\omega}) > \Delta_4 \\ H_a &: \Delta_1 \le g_1(\underline{\omega}) \le \Delta_2 \text{ and } \Delta_3 \le g_2(\underline{\omega}) \le \Delta_4 \end{aligned} \tag{2.8}$$

Section 2.3 Example/Results

Example

To illustrate the methodology described in Section 2.2, we return to the example described by Gennings et al. (2004). The dose-response data (Appendix A;Tables A.2.1 and A.2.2) are binomial data. Assume the data are distributed $bin(\pi_i, n_i)$ $i = 1,...,14$ where *i* denotes the 14 dose groups between the two mixtures. Recall that Mixture 1 is the full mixture and Mixture 2 is the reduced mixture. Utilizing the dose adjustment factor, D_{AF} =0.175, the total dose scale for the full mixture was adjusted (scaled), resulting in the following total dose scales in Table 2.1. That is, a total dose of 100 mg/kg for the full mixture is comprised of 17.5% of chemicals in an active range, so the adjusted corresponding total dose is 17.5 mg/kg. With this adjustment, the adjusted total doses for the two curves are nearly identical. A nonlinear mixed effects model (eq. (2.1)) for these data is based on the logistic function

$$\pi = \frac{e^{\beta_0 + \beta_1 ht}}{1 + e^{\beta_0 + \beta_1 ht}}$$

where $h \sim N(1, \sigma_h^2)$, β_0 is the intercept parameter, β_1 is the slope parameter, and π is the probability of abnormal gait. This model was fit using Adaptive Gaussian Quadrature to evaluate the likelihood (PROC NLMIXED in SAS v. 9.1) which results in the corresponding maximum likelihood parameter estimates and asymptotic variance-covariance matrix (Table 2.2; Figure 2.1). This procedure produces empirical Bayes estimates for the random effect, *h*. Because the random effect has two levels (i.e., two dose-response curves exist) two estimates are calculated for *h* resulting in predicted dose-response curves for both the full and reduced mixtures.

Table 2.1. Adjusted total dose for the full mixture (Mixture 2) resulting from using the D_{AF} to adjust for percentage of chemicals in an active dose range.

Mixture	Total Dose (mg/kg)	Adjusted Total Dose (mg/kg)
1	0	0
1	10	1.750
1	55	9.625
1	100	17.50
1	200	35.00
1	300	52.50
1	450	78.750
2	0	0
2	1.75	1.75
2	9.6	9.6
2	17.5	17.5
2	35	35
2	52.5	52.5
2	78.8	78.8

Table 2.2. Parameter estimates from the fitted model with adjusted total dose.

Parameter	Estimate	Standard Error	P-value
β_0	-2.33	0.36	0.1
β_1	0.095	0.019	0.13

Note: Estimate of σ_h^2=0.035

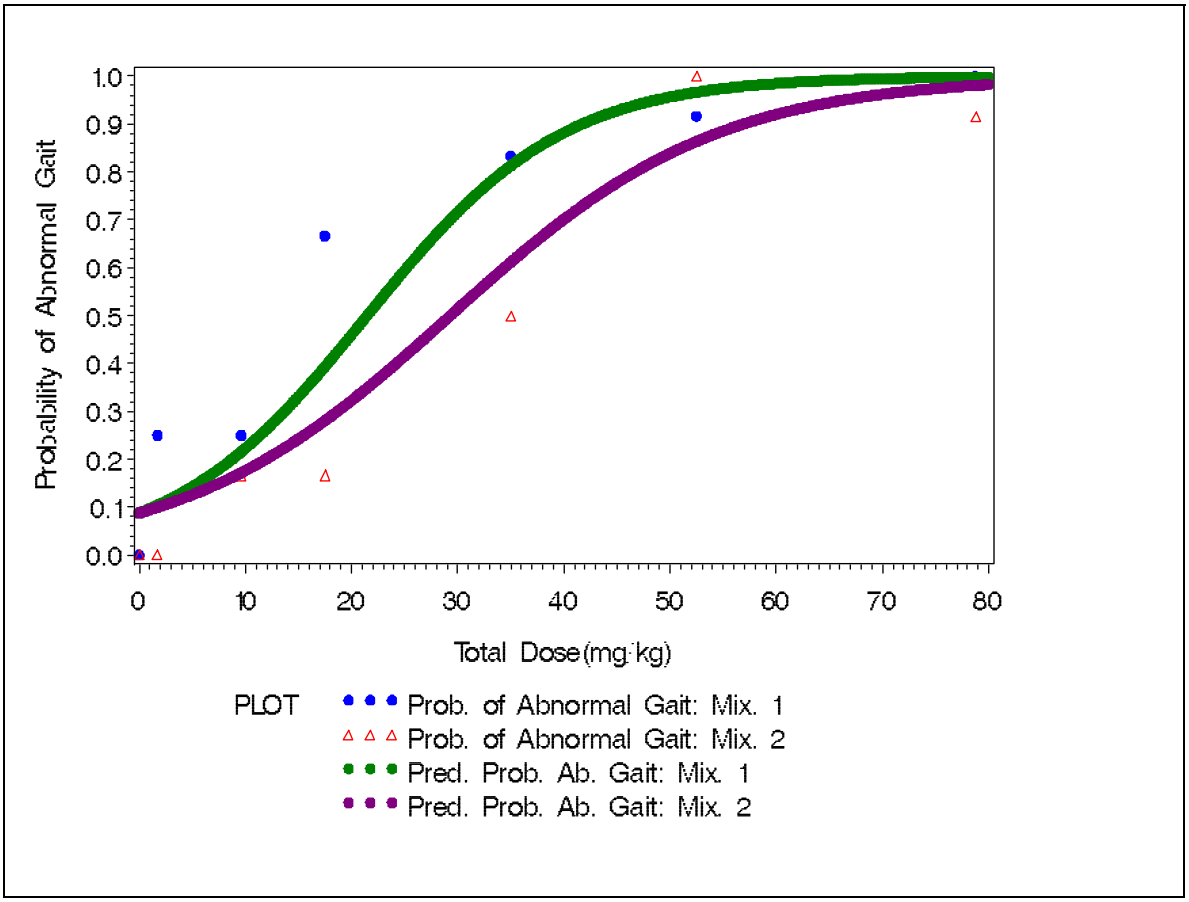


Figure 2.1. Plot of predicted probability of abnormal gait for the full mixture (mixture 1) with adjusted total dose and for the reduced mixture (mixture 2) with original total dose and raw dose-response data (σ_h^2=0.035).

For ease of defining the similarity region, the dose-response curve was fully characterized in terms of two ED's of interest, the ED(20) and ED(50). (estimates for the ED(20) and ED(50) are in Table 2.3) For this logistic model, the general form of the ED(μ_0) is

$$ED(\mu_0) = \frac{\log(\frac{\pi_{\mu_0}}{1-\pi_{\mu_0}}) - \beta_0}{\beta_1 h} \qquad (2.9)$$

In terms of the ED(20) and ED(50)

$$\beta_0 = \frac{1}{ED(20) - ED(50)}\left[ED(20)\log(\frac{.50}{1-.50}) - ED(50)\log(\frac{.2}{1-.2}) \right]$$

and (2.10)

$$\beta_1 = \frac{1}{\{ED(20) - ED(50)\}h}\left[\log(\frac{.2}{1-.2}) - \log(\frac{.50}{1-.50}) \right].$$

The reparameterized logistic model is

$$\pi = \exp\left\{ \begin{array}{l} \frac{1}{ED(20) - ED(50)}\left[ED(20)\log(\frac{.50}{1-.50}) - ED(50)\log(\frac{.2}{1-.2}) \right] + \\ \left(\frac{1}{\{ED(20) - ED(50)\}h}\left[\log(\frac{.2}{1-.2}) - \log(\frac{.50}{1-.50}) \right] t \right) \end{array} \right\}. \quad (2.11)$$

This is to say that $g(\underline{\omega}) = \begin{bmatrix} g_1(\underline{\omega}) \\ g_2(\underline{\omega}) \end{bmatrix} = \begin{bmatrix} ED(20) \\ ED(50) \end{bmatrix}$ and the similarity region is defined in terms of allowable shifts in the ED(20) and ED(50) (Figure 2.2). Now, suppose expert judgment suggests that a 90% shift in the ED(20) and a 50% shift in the ED(50) is associated with an inappreciable shift in the dose-response curve and represents a biologically relevant/significant region (Table 2.4) so that

$$\Delta_1 = ED(20) - 0.9 \times ED(20),\ \Delta_2 = ED(20) \times 1.9$$
$$\Delta_3 = ED(50) - .50 \times ED(50),\ \Delta_4 = ED(50) \times 1.50 .$$

These $\begin{bmatrix} (\Delta_1, \Delta_2) \\ (\Delta_3, \Delta_4) \end{bmatrix}$ make up the similarity bounds /similarity region (Figure 2.3; Table 2.4).

Table 2.3. Estimates of the ED(20) and ED(50) from the fitted model with adjusted total dose.

Parameter	Estimate	Standard Error	P-value
ED(20)	9.98	3.10	0.19
ED(50)	24.56	4.16	0.11

Table 2.4. Estimates of the ED(20) and ED(50) and their respective similarity bounds from the fitted model with adjusted total dose.

Adjusted Total Dose: Similarity Region

	Estimate	Lower Bound	Upper Bound
ED(20)	9.98	$1.00(\Delta_1)$	$18.96(\Delta_2)$
ED(50)	24.56	$12.28(\Delta_3)$	$36.85(\Delta_4)$

For the purposes of this example, we constructed and plotted the 95% confidence region (Figure 2.2).

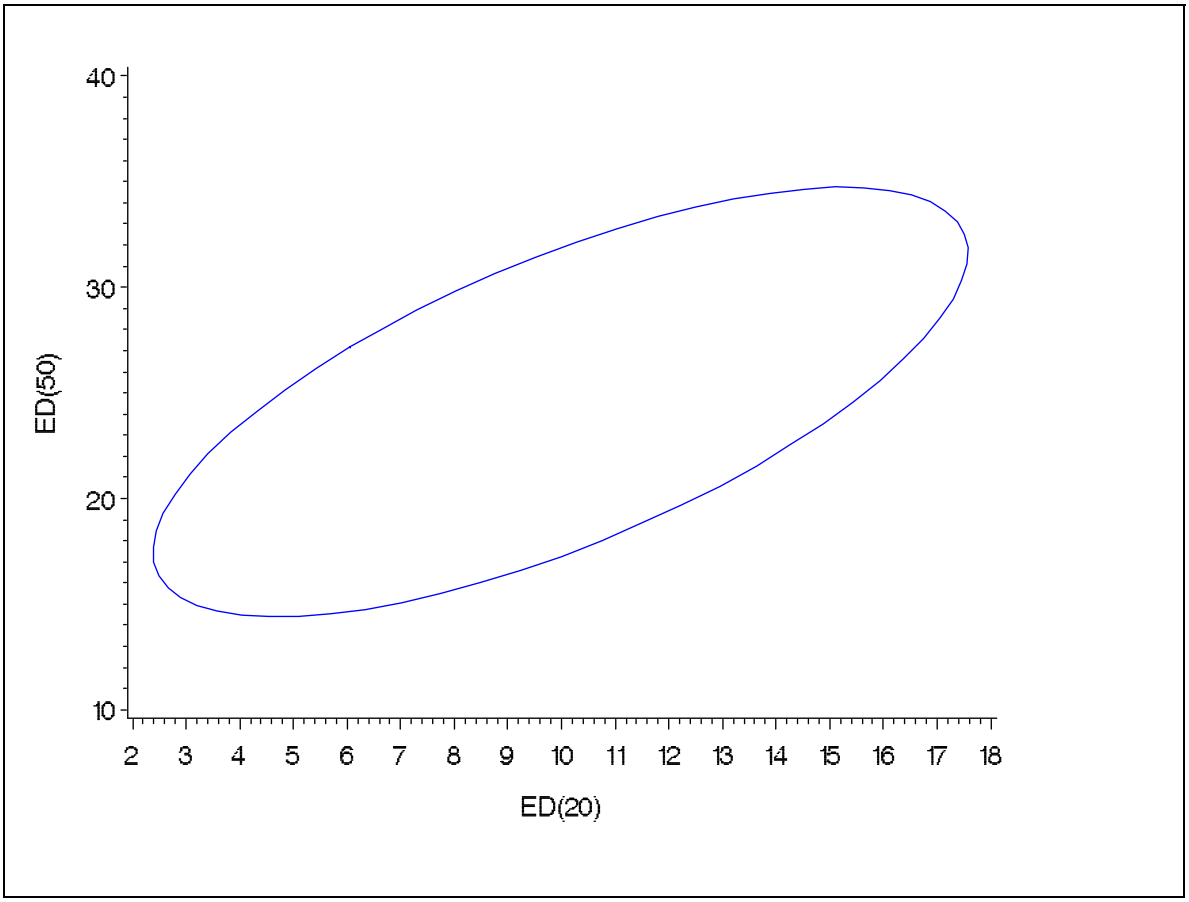


Figure 2.2. Plot of 95% confidence region in terms of the ED(20) and ED(50) on the adjusted total dose scale.

The box in Figure 2.3 is determined by the similarity bounds and constitutes the similarity region.

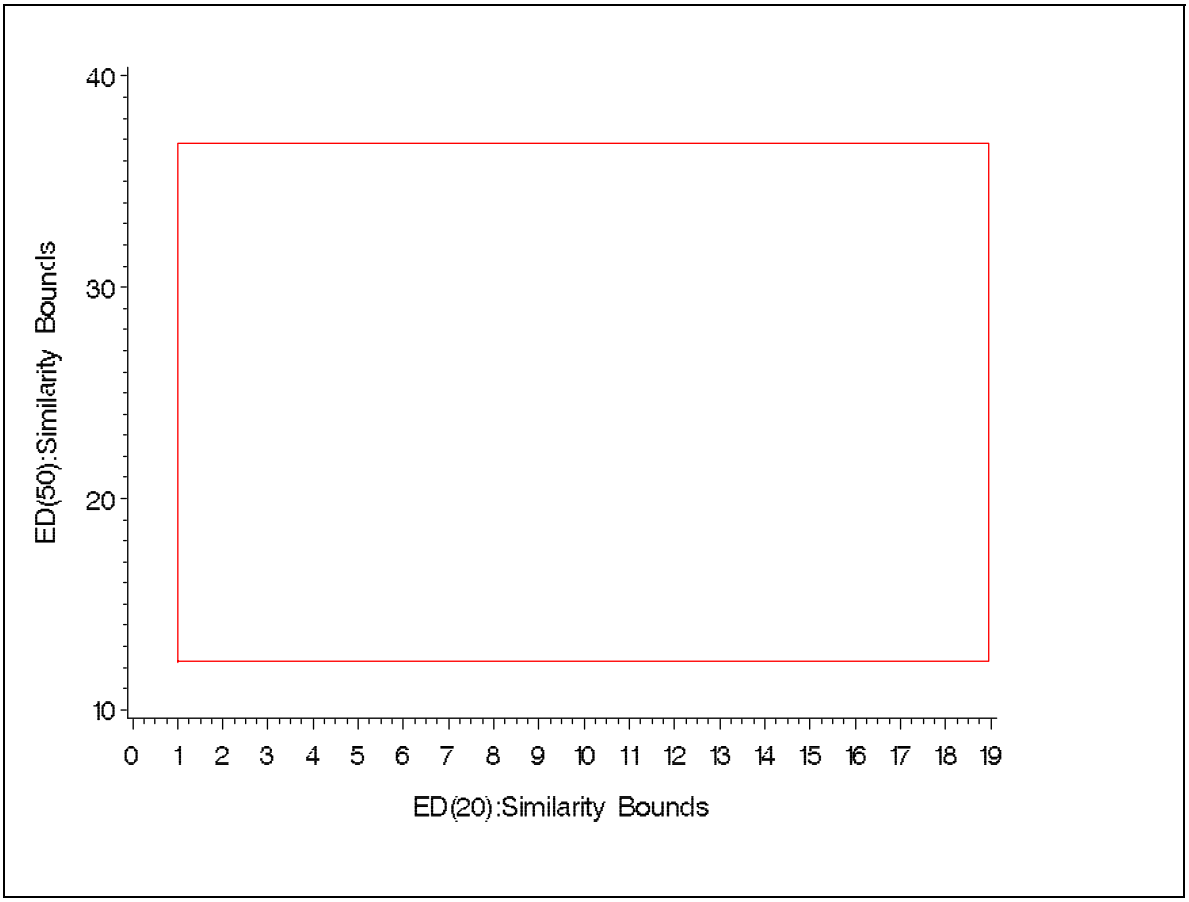


Figure 2.3. Plot of the similarity region resulting from a 90% shift in the ED(20) and a 50% shift in the ED(50) on the adjusted total dose scale.

Following the logic of Stork et al. (2008), to test for sufficient similarity in dose-response, the approach determines if the variability associated with the random exposure factor, *h*, in model (2.1) is "too" large based on the pre-specified similarity region. Plotting Figure 2.2 with Figure 2.3 (Figure 2.4) results in the following hypothesis test of sufficient similarity from eq. (2.8)

$$H_0 : ED(20) < \Delta_1 \text{ or } ED(20) > \Delta_2 \text{ or } ED(50) < \Delta_3 \text{ or } ED(50) > \Delta_4$$
$$H_1 : \Delta_1 \le ED(20) \le \Delta_2 \text{ and } \Delta_3 \le ED(50) \le \Delta_4 .$$

If any part of the confidence region (in terms of the ED(20) and ED(50)) crosses any of the similarity bounds, then sufficient similarity cannot be concluded. Because the confidence ellipse is completely contained within the similarity region, the null hypothesis is rejected and it is concluded that Mixture 1 and Mixture 2 are sufficiently similar in dose-response (see Figure 2.4).

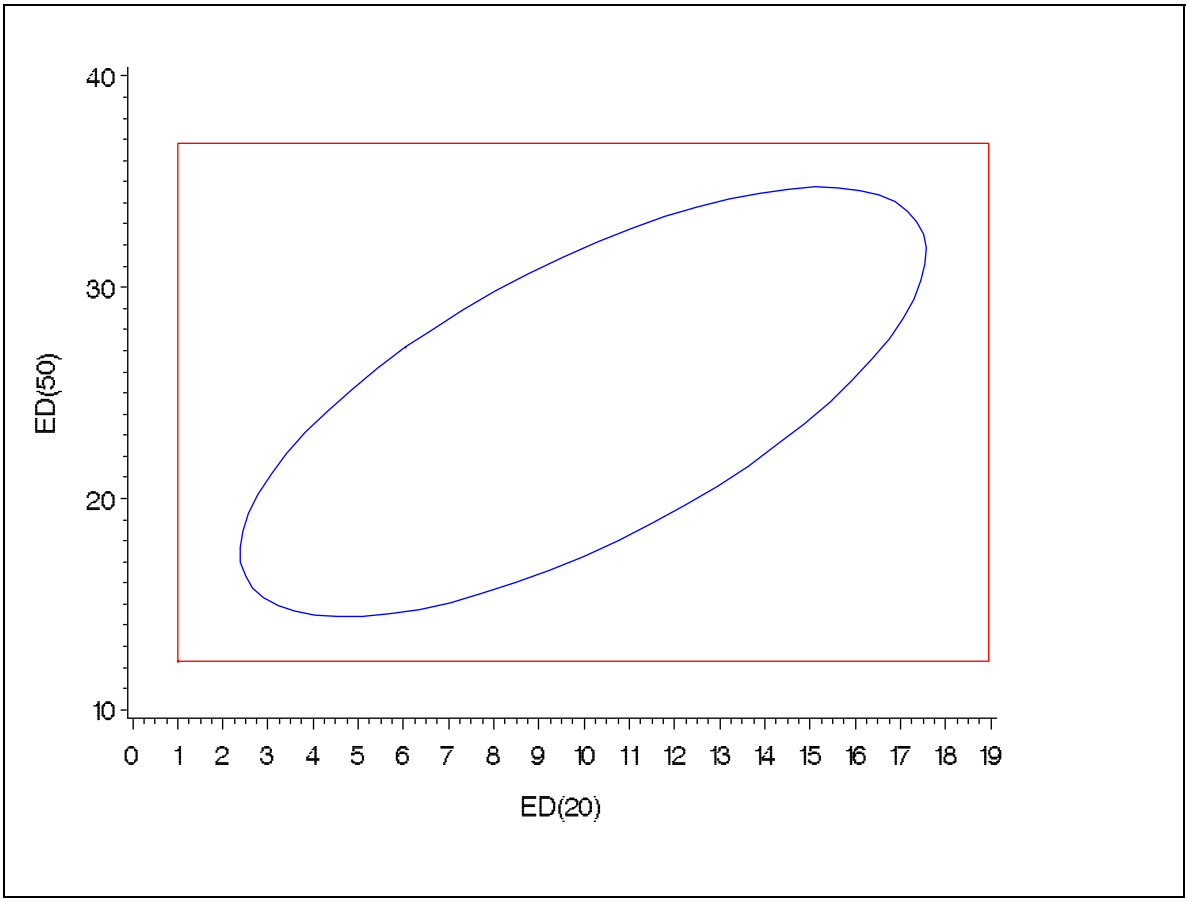


Figure 2.4. Plot of the 95% confidence region for the ED(20) and ED(50) overlayed with the similarity region on the adjusted total dose scale.

The analysis was repeated without adjusting the total dose scale of Mixture 2. Estimates of the parameters and ED's are in tables 2.5 and 2.6. Plots of the 95% confidence ellipse (Figure 2.6) and the similarity region (figure 2.7) are below. Because the confidence region (Figure 2.8) is not contained in the similarity region, it cannot be concluded that the two mixtures are sufficiently similar in dose-responsiveness. Many factors can contribute as to why the two mixtures are not sufficiently similar, including differing dose scales.

Table 2.5 . Parameter estimates from the fitted model with original dose scales.

Parameter	Estimate	Standard Error	P-value
β_0	-2.33	0.37	0.1
β_1	0.047	0.02	0.26

Note: Estimate of σ_h^2=0.31

Table 2.6. Estimates of the ED(20) and ED(50) with original dose scales.

Parameter	Estimate	Standard Error	P-value
ED(20)	19.97	9.71	0.29
ED(50)	49.31	20.16	0.25

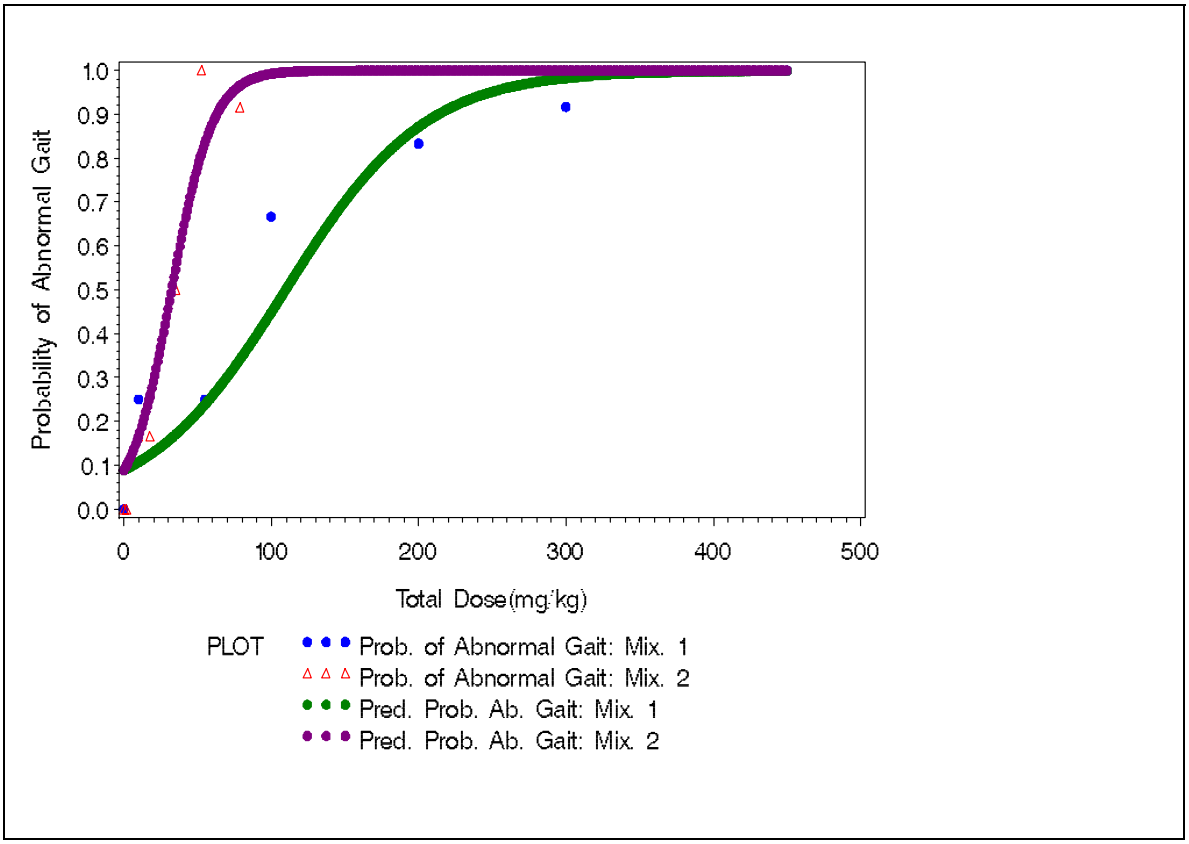


Figure 2.5. Plot of predicted probability of abnormal gait for full (mixture 1) and reduced (mixture 2) mixtures vs. original total dose (σ_h^2=0.31).

Table 2.7. Estimates of the ED(20) and ED(50) and their respective similarity bounds on the original dose scales.

		Similarity Region	
	Estimate	**Lower Bound**	**Upper Bound**
ED(20)	19.97	2.00(Δ_1)	37.94(Δ_2)
ED(50)	49.31	24.66(Δ_3)	73.97(Δ_4)

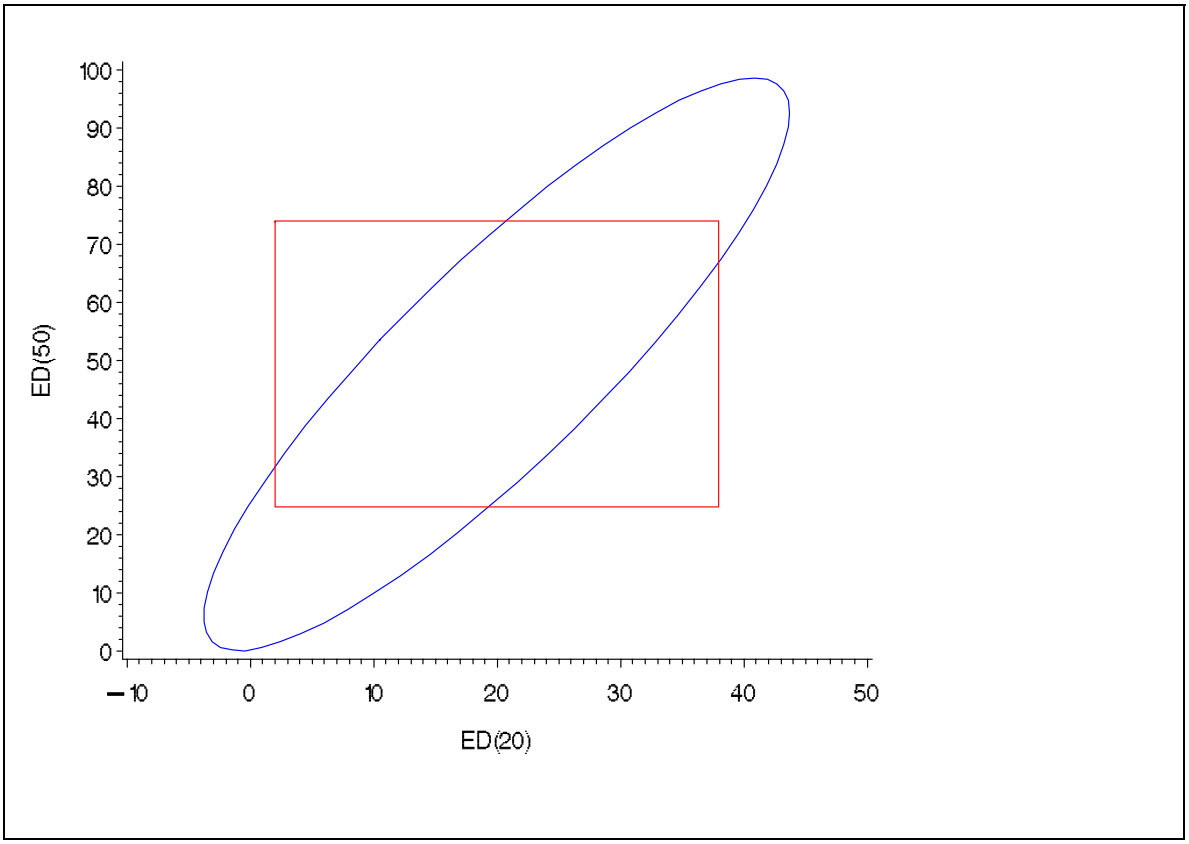


Figure 2.6. Plot of the 95% confidence region for the ED(20) and ED(50) overlayed with the similarity region on the original dose scales.

Section 2.4 Discussion

There are two main points of interest to be addressed:

1. If sufficient similarity is to be defined based on dose-responsiveness of a mixture, is the scale of dose important to consider?

2. Gennings et al. (2004) determined that, although malathion is not dose-responsive alone, it is associated with an interaction among the five chemicals in the mixture. What are the implications of the presence of interactions in chemical mixtures in an equivalence testing framework?

For the example presented in this chapter, there were two dose-response curves of interest, a full mixture and a reduced mixture. The full mixture contains one additional chemical that is not in a dose-responsive range as compared to the reduced mixture that has all chemicals in a dose-responsive range. Due to the differing dose scales, the dose adjustment factor, D_{AF} was used to scale the total dose scale of the full mixture (Table 2.1) resulting in the same dose scale. Figure 2.4 contains the plot of the 95% confidence ellipse with the specified similarity bounds. Because the ellipse is completely contained in the similarity region, it is concluded that the full and reduced mixtures are sufficiently similar in dose-response. In order to investigate the effect of utilizing the D_{AF} a sensitivity analysis was performed where the D_{AF} was not used. Figure 2.8 contains the plot of the 95% confidence ellipse that is not contained in the specified similarity bounds (not enough evidence to conclude sufficient similarity) demonstrating that adjusting the dose scale is an important factor when testing for sufficient similarity. It can also be observed that after adjusting total dose for percent of chemicals in an active dose-range, the

observed dose-response data of Mixtures 1 and 2 appear more similar (Appendix A; Figure A.2.1). When two mixtures are measured on different dose scales, this adds variability to the analysis (in terms of the random effect). This point is observed in the example as the variance of the random effect decreases by an order of magnitude (0.31 compared to 0.035) when total dose is rescaled in the full mixture. These points demonstrate that the relationship between scale of analysis and the percent of chemicals in a mixture in an active dose range needs to be accounted for in the analysis.

The second point is more complicated. Malathion is present in the full mixture but not in the reduced mixture, and Gennings et al. (2004) determined that malathion is not dose-responsive alone (given the dose range of the study). Instead, it is associated with a statistically significant interaction among the other four chemicals in the mixture. Although interactions may be present and statistically significant, it may not be of consequence when evaluating sufficient similarity based on the size of the similarity region. This is to say that an interaction does not necessarily imply the curves are "different", as what is not different is defined by the similarity region. The hypothesis test in eq. (2.8) is not a test of interaction, but a test for sufficient similarity. If the confidence region of interest is contained in the similarity region, the curves are concluded to be sufficiently similar in dose-response at the stated alpha level. This example demonstrates that in the presence of a statistically significant interaction, two curves can be sufficiently similar given the acceptable shifts (Figure 2.4) when dose scale is adjusted for percentage of active chemicals. The knowledge of the presence of an interaction or the assumption of interaction when testing for sufficient similarity is not necessary as the similarity region should be determined from expert judgment based on relevant biological information.

This is to say that in the presence of an interaction the similarity region should not be chosen to account for the interaction but to signify a biologically relevant region. For this example the similarity region was defined based on shifts that were assumed to be biologically relevant.

Based on the current analyses it can be concluded that it is necessary and important to evaluate the dose scales prior to performing a test for sufficient similarity as it was demonstrated that given the circumstances this can affect the outcome of the test. It can also be concluded that a chemical or subset of chemical(s) can affect the dose-response relationship with respect to sufficient similarity, however, this is dependent on the size of the similarity region. This is to say that although malathion was known to have a statistically significant interaction with the other four chemicals in the full mixture, we still claimed that the two curves were sufficiently similar in dose-response given the similarity region. These analyses demonstrate the importance of adjusting dose scale when the additional chemical(s) are not in a dose-responsive range and that size of the similarity region is important whether or not an interaction is present. These points are important to consider as determination of sufficient similarity is a function of dose scale and similarity region, as demonstrated in the example.

Chapter 3: An Empirical Approach to Sufficient Similarity in Dose-Responsiveness: Utilization of Euclidean Distance as a Similarity Measure

Section 3.1 Introduction

Each year, roughly 13 million U.S. children are in day care during some or all of the work day and in some instances can be in supervised day care for as many as 10 hours per day (Tulve et al., 2006). Though such a large number of children are in these facilities, little research has been done to characterize the exposure of the children to the environmental chemicals that may be present. The Food Quality Protection Act of 1996 requires that the U.S. Environmental Protection Agency upgrade the risk assessment process for setting pesticide residue tolerances in food by assessing the potential susceptibility of children to both aggregate and cumulative exposure to pesticides (Food Quality Protection Act, 1996). Environmental health issues, such as asthma and pesticide exposure, have been assessed in child care centers; however, these assessments are limited by geographic area, location samples within the site, diversity of analytes and accompanying survey/questionnaire data (Nafstad et al., 2005; Wilson et al., 2005; Fritz and Herbarth, 2005; Tulve et al., 2006). In order to understand aggregate exposure to pesticides, environmental exposure data are needed. The concept of collecting exposure data, referred to as

exposure assessment, is one step in the risk assessment process (EPA, 1992), and is defined by the following four steps:

1. Identify pollutants that may be released.
2. Estimate the amount of pollutants released from all sources.
3. Estimate concentration of pollutants.
4. Estimate number of people exposed to different concentrations.

Depending on the intended use of the exposure assessment, the numerical output may be an estimate of either exposure or dose (EPA 1992). The exposure and dose data are often combined with exposure-response or dose-response data to estimate risk, the probability of an adverse event occurring (EPA 1992). Exposure data can be in the form of external exposure data, such as individual loadings of particular chemicals (measured in ng/cm^2 for example), as internal concentrations (e.g., μg/kg in blood), or it could be presented as mixing ratios (proportions) and respective total doses/concentrations

Ideally the risk assessments would be conducted using dose-response data for the mixture(s) of concern (Feder et al., 2009). When complete dose-response data are available (i.e., the data rich case) for the mixture(s) of concern, these data can be used to characterize each of the individual mixtures for the purpose of risk assessment. For example, for each observed chemical mixture a benchmark dose (BMD) could be calculated and compared back to a reference BMD. While it is preferable to conduct risk assessments of complex chemical mixtures based on toxicity and exposure data for the complete mixture (Rice et al., 2009), in most practical applications data are not available on all possible mixtures of concern.

When toxicity data are not available on a mixture of concern, the U.S. EPA (1986, 2000) has developed a framework for conducting risk assessments of chemical mixtures based on observed data from a "sufficiently similar" mixture as a surrogate for the mixture of concern. The EPA guidelines define a similar mixture based on similar chemical composition and similar component proportions, based on expert biological judgment. However, these guidelines do not provide specific information on how to approach this type of problem (Feder et al., 2009) and, as noted by Seed et al. (1995), a challenge with the use of this method is associated with determining how similar an observed mixture is to an unobserved mixture of interest. The approach described by the EPA for determining sufficient similarity is limited in that it uses no empirical data (dose-response) or statistical modeling approach to account for uncertainty in the process, as this is an important issue (Feder et al., 2009).

Motivating Example

Characterizing the exposure of children to pesticides in child care centers is an important issue with such a large number of children in child care centers each day. To begin to understand the aggregate exposure of children to pesticides a collaborative project of the U.S. Department of Housing and Urban Development, the U.S. Consumer Product Safety Commission and the U.S. Environmental Protection Agency, called The First National Environmental Health Survey of Child Care Centers was created (Tulve et al., 2006). This collaborative project was the first probability-based (including multi-stage sampling and clustering) national study of child care centers. This study included 168 child care centers from 30 sampling units and measured lead, allergen and pesticide exposure. Children in the child care

centers studied were less than 6 years of age, although the children were not directly studied. Samples were collected from July through October 2001 at multiple locations within a center by means of floor, tabletop, and desk swipes and soil samples (Tulve et al., 2006). For purposes of this example, we only consider data obtained from floor swipes. Of the pesticides measured, five (permethrin, cypermethrin, beta-cyfluthrin, deltamethrin, esfenvalerate) in a class of chemicals called pyrethroids, were targeted as the chemicals of interest in a dose-response study.

Each of the 168 exposure profiles to the five pyrethroids observed in the child care centers is different. That is, there is a unique mixture (candidate mixture) associated with each child care center based on floor swipe data. If it were feasible, it would be ideal to perform a dose-response study for each of the 168 unique mixtures and analyze the distribution of one or more summary measures for each dose-response curve, such as the benchmark dose (BMD). It is important to note that exposure does not necessarily imply risk, as risk is dependent on the dose or level of exposure. This is to say that by utilizing the distribution of BMD's we could determine if the observed mixtures were associated with BMD's below the BMD associated with a benchmark response (BMR) of interest, thus linking exposure and dose. In this instance, where there are not data available (i.e., the data poor case) on the candidate mixtures of concern, the EPA allows for a surrogate/reference mixture that has been concluded to be sufficiently similar, to be used for the purpose of inference and risk assessment; however, the existing guidelines do not contain procedures for determining sufficient similarity (Feder et al., 2009) or for estimating summary measures for the candidate mixture, such as the benchmark dose for a particular response of interest. Our objective is to compute estimates for the benchmark doses for the candidate mixtures that are sufficiently similar to the surrogate/reference mixture, as well as for

the individual components of the candidate mixtures. In order to provide these resulting estimates and inference, a more flexible and adaptable procedure for determining sufficient similarity needs to be developed that provides an overall summary statistic to be used in risk assessment.

Due to the lack of data, a supplementary approach, suggested by Stork et al. (2008), using statistical equivalence testing logic and mixed model theory, has been developed to define sufficient similarity in dose-response for chemical mixtures containing the same chemicals with different ratios, where complete dose-response data are available on one mixture and only mixing ratios are available on the other mixture (data poor case) (Stork et al., 2008). The current methodology (Stork et al., 2008) requires that both the reference and candidate mixtures contain the same c components and this presents a limitation, as this is rarely the case. Feder et al. (2009) suggest using classical multivariate statistical techniques, such as the multivariate form of the traditional t-test, Hotelling's T^2, and principal components analysis. While the method presented by Feder et al. (2009) is formulated in the traditional hypothesis testing framework and attempts to evaluate whether a process is 'out of control' their method can easily be adapted to perform tests for similarity using traditional equivalence testing methods.

An extension to the method of Stork et al. (2008) is developed that allows for the absence of chemicals in either the reference or candidate mixture, i.e., one mixture is a subset of the other. The proposed test for sufficient similarity follows the same statistical mechanics described by Stork et al. (2008): equivalence testing with mixed-effects models in terms of total dose. However, the proposed test defines a similarity measure, h, that is a function of the mixture

proportions and based on the concept of Euclidean distance, where the distance between the corresponding mixing ratios in the reference and candidate mixtures is of interest. It is of interest to determine if the BMD's for different chemical mixtures are similar in terms of total dose and if the dose-response data from a reference mixture can be used as a surrogate to predict BMD's for other mixtures determined to be sufficiently similar in dose-response. An example is provided that illustrates how to use the calculated similarity measure and the benchmark dose as an overall summary statistic for use in risk assessment.

Section 3.2 Statistical Methods

Recall that it is of interest to test whether the reference mixture and a candidate mixture of concern are sufficiently similar in dose-response. In this section, as in chapter 2, the mixtures will be referred to as 'reduced' (*red*) and 'full' (*full*). The terms reduced and full are used mainly for the development of notation for the different distance measures and the proofs of some of the associated properties. Depending on the example either the reference or candidate mixture can assume the label as the reduced or full mixture. The application of the label is dependent on which mixture (reference or candidate) contains the subset of the other.

To evaluate if the full and reduced mixtures are sufficiently similar the following steps may be taken

1. Determine the appropriate similarity measure
2. Fit the appropriate non-linear model; characterize the curve as a function of model parameters; determine the similarity bounds (following Stork et al. (2008))

3. Compute the appropriate similarity measure between the reference and candidate mixing ratios and test for sufficient similarity

Once the appropriate similarity measure (distance measure) is chosen the step of fitting the appropriate non-linear model and determining the similarity bounds (region) is virtually the same as the method of Stork et al. (2008) described in Chapter 1. However, instead of computing bounds on the mixing ratios, bounds are computed for the similarity measure *h*, denoted as (h_L, h_U). If the similarity measure is within the similarity bounds, i.e., $h \in (h_L, h_U)$ then it is concluded that the reduced and full or reference and candidate mixtures are sufficiently similar in dose-response. It is important to note that this procedure is not a true statistical hypothesis test of equivalence, as we cannot claim a significance (α) level and we are attempting to evaluate a mixture of concern that has no associated dose-response data. The proposed methods represent an evaluation of sufficient similarity, rather than a true test of sufficient similarity as described in Chapter 2.

Notation/Definitions

Consider the following notation for the reduced (*red)* and full (*full*) mixtures respectively, $\underline{a}_{red} = \{a_{i,red}\}$ where $i = 1,...,k$, $\underline{a}_{full} = \{a_{i,full}\}$ where $i = 1,...,c$ and $a_{i,full}$ and $a_{i,red}$ are the proportions of the individual chemical components of each mixture. Let

$\underline{a}_{red} = \begin{bmatrix} a_{1,red} \\ a_{2,red} \\ \vdots \\ a_{k,red} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ and $\underline{a}_{full} = \begin{bmatrix} a_{1,full} \\ a_{2,full} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ a_{c,full} \end{bmatrix}$ where there are $c-k=s$ "placeholder" zeros in $\underline{a}_{red}$ be the two mixtures of interest.

Also, $\underline{a}_{red}t = \underline{x}_{red}$ and $\underline{a}_{full}t = \underline{x}_{full}$ for $t = 1,...,J$, where t are the total dose groups and $x_{()}$ are the doses of the individual chemicals at each dose group.

The Euclidean distance between vectors $\underline{x}$ and $\underline{y}$ of the same dimension is defined as (Johnson and Wichern, 1998)

$$d(\underline{x},\underline{y}) = \sqrt{(\underline{x}-\underline{y})'(\underline{x}-\underline{y})} \tag{3.1}$$

The statistical distance between the same two vectors is (Johnson and Wichern, 1998)

$$d(\underline{x},\underline{y}) = \sqrt{(\underline{x}-\underline{y})'\underline{A}(\underline{x}-\underline{y})} \tag{3.2}$$

where $\underline{A}$ is a symmetric matrix. In most cases, such as in discrimination and classification, $\underline{A}$ is the sample variance covariance matrix, which makes this measure the statistical distance. There is no implicit constraint on the $\underline{A}$ matrix other than it must be a square matrix. We define $\underline{A} = \underline{W}$ to be a diagonal weight matrix to define a weighted Euclidean distance. It should be noted that there are other measures of statistical distance, such as Minkowski's measure

$$d\left(\underline{x},\underline{y}\right)=\left[\sum_{i=1}^{p}\left|x_i-y_i\right|^{m}\right]^{\frac{1}{m}} \tag{3.3}$$

When m=1, $d\left(\underline{x},\underline{y}\right)$ is the city block distance and when m=2 it is Euclidean distance. Varying m changes the weight given to larger or smaller distances.

In the following suggested distance measures, d is distance and t is total dose so that $\frac{d}{t}$ is interpreted as distance relative to total dose and the similarity measures are $1+\frac{d}{t}$. Following are four proposed similarity measures based on Euclidean distance between $\underline{a}_{red}$ and $\underline{a}_{full}$.

Step 1 In Determining Sufficient Similarity: Determine the Appropriate Similarity Measure

Unadjusted Unweighted Distance

In this case, the distance between the chemical proportions, a_i, of two mixtures is directly compared.

$$d=\sqrt{\left(\underline{x}_{red}-\underline{x}_{full}\right)'\left(\underline{x}_{red}-\underline{x}_{full}\right)}$$
$$d=\sqrt{\left(\underline{a}_{red}t-\underline{a}_{full}t\right)'\left(\underline{a}_{red}t-\underline{a}_{full}t\right)}$$
$$d=\sqrt{t\left(\underline{a}_{red}-\underline{a}_{full}\right)'t\left(\underline{a}_{red}-\underline{a}_{full}\right)}$$
$$d=\sqrt{t^2\left(\underline{a}_{red}-\underline{a}_{full}\right)'\left(\underline{a}_{red}-\underline{a}_{full}\right)}$$
$$d=t\sqrt{\left(\underline{a}_{red}-\underline{a}_{full}\right)'\left(\underline{a}_{red}-\underline{a}_{full}\right)}$$

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red} - \underline{a}_{full}\right)' \left(\underline{a}_{red} - \underline{a}_{full}\right)} \qquad (3.4)$$

No adjustments are made for potency, absence/presence of chemical components, or for any other environmentally or toxicologically relevant characteristics. This approach might be taken when the two mixtures have the same composition and dimension and the proportions are similar in both mixtures. This measure also might be used in the case when there are additional chemical(s) in the full mixture that are present in an inactive range but constitute a small proportion of the mixture.

Adjusted Unweighted Distance

It may be the case that one mixture is a subset of the other mixture. For example, the reference mixture could be a subset of the candidate mixture. The reference mixture is then referred to as the reduced mixture and the candidate mixture is referred to as the full mixture. When the *s* additional components of the full mixture are in a subthreshold range (inactive), adjustments can be made to total dose so that the dose scales (metameters) are comparable (Casey et al., 2004). Consider the proposed dose adjustment factor, $D_{AF} = \sum_{i=1}^{k} a_{i,full}$ which is the proportion of the *k* chemicals that are in both the full and reduced mixtures.

$$d = \sqrt{\left(\underline{a}_{red} t \sum_{i=1}^{k} a_{i,full} - \underline{a}_{full} t\right)' \left(\underline{a}_{red} t \sum_{i=1}^{k} a_{i,full} - \underline{a}_{full} t\right)}$$

$$d = t\sqrt{\left(\underline{a}_{red} D_{AF} - \underline{a}_{full}\right)' \left(\underline{a}_{red} D_{AF} - \underline{a}_{full}\right)}$$

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red} D_{AF} - \underline{a}_{full}\right)' \left(\underline{a}_{red} D_{AF} - \underline{a}_{full}\right)} \qquad (3.5)$$

where $D_{AF} = \sum_{i=1}^{k} a_{i,full}$. The reduced mixture is adjusted for total dose in relation to the full mixture.

Unadjusted Weighted Distance

When the $s = c - k$ additional components of the full mixture are at or beyond threshold levels and it is expected that the additional components may contribute to the effect of total dose, the proportions, a_i, may be weighted to emphasize or deemphasize certain properties. For example, the additional components may be present in small quantities but have a large effect. In this case, it may be expected that the curves are not sufficiently similar, however, the distance measure might be near zero, indicating the mixtures are 'close' to one another with respect to Euclidean Distance. In this case, weights may be added to magnify the effect of particular proportions. These weights may be constructed to control for relative potency, presence/absence of chemicals, or for other environmentally or toxicologically relevant characteristics. Recall that the only restriction on the weight matrix, $\underline{W}$ is that it is a square matrix. For convenience we restrict $\underline{W}$ to be a diagonal matrix with the respective weights, w_i, on the diagonal, subject to the constraint, $\sum_{i=1}^{c} w_i = c$. When all chemical components are weighted equally, each chemical receives a weight of one and the weights sum to *c*. One possible weighting scheme is to construct the weights, w_i, as functions of relative potencies of the individual chemicals, where

$\underline{W} = diag(w_i);\ w_i = \frac{RP_i}{TRP}c;$ where $TRP = \sum_{i=1}^{c} RP_i$. The weights are applied to the proportions by inserting the weight matrix, W , into the quadratic form. The unadjusted weighted distance is constructed in the following manner

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red} - \underline{a}_{full}\right)' \underline{W}\underline{W} \left(\underline{a}_{red} - \underline{a}_{full}\right)}$$
$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red}\underline{W} - \underline{a}_{full}\underline{W}\right)' \left(\underline{a}_{red}\underline{W} - \underline{a}_{full}\underline{W}\right)}$$

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red,w} - \underline{a}_{full,w}\right)' \left(\underline{a}_{red,w} - \underline{a}_{full,w}\right)} \qquad (3.6)$$

where $\underline{a}_{\bullet,w} = \underline{a}_{\bullet}\underline{W}$.

Adjusted Weighted Distance

The adjusted weighted distance allows for adjusting for total dose and for weighting the individual proportions. This similarity measure may be used when the additional components in the full mixture are present in an inactive/subthreshold range but constitute such a significant proportion of the mixture that adjusting for total dose is not a strong enough adjustment (described later in Theorem 1) when it is expected that the two mixtures are sufficiently similar. The adjusted weighted distance measure is constructed in the following manner

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red} D_{AF} - \underline{a}_{full}\right)' \underline{W}\underline{W} \left(\underline{a}_{red} D_{AF} - \underline{a}_{full}\right)}$$

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red} D_{AF} \underline{W} - \underline{a}_{full} \underline{W}\right)' \left(\underline{a}_{red} D_{AF} \underline{W} - \underline{a}_{full} \underline{W}\right)}$$

$$\frac{d}{t} = \sqrt{\left(\underline{a}_{red,w} D_{AF} - \underline{a}_{full,w}\right)' \left(\underline{a}_{red,w} D_{AF} - \underline{a}_{full,w}\right)} \qquad (3.7)$$

where $\underline{a}_{red,w} = \underline{a}_{red} \underline{W}$. While the weights, w_i are subjective in nature, this subjectivity can be minimized through practice and the consensus of experts.

Step 2 In Determining Sufficient Similarity: Fit the Appropriate Non-linear Model/ Characterize the Curve as a Function of Model Parameters/Determine the Similarity Bounds

Once the similarity measure is chosen, an appropriate non-linear model is fit and the similarity bounds are determined. Stork et al. (2008) suggest using a non-linear mixed effects threshold model, however, any appropriate non-linear model/CDF (i.e. fits the shape of the data) can be used to model the dose-response curve. Consider the general form of the non-linear mixed effects model for the mean

$$E\left[y \mid h\right] = \mu_h = \alpha + \gamma f(\boldsymbol{\omega}, t, h) \qquad (3.8)$$

where f is some nonlinear function; $\boldsymbol{\omega}$ is a vector of model parameters; h is the similarity measure and is modeled as a random effect with assumed distribution $N(1, \sigma_h^2)$; t is the total dose; α and γ are minimum and maximum effect parameters. For the purpose of determining shifts in the dose-response curve that are associated with inappreciable differences it is useful to reparameterize the model (eq. (3.8)) as functions of the model parameters, $g(\underline{\omega})$ (Section 2.2, step 3) that have a meaning in application. Without loss of generality, reparameterize eq. (3.8) as

two functions of model parameters, say $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$ that have an intuitive meaning such as two ED's of interest or an ED and a threshold parameter. Following Stork et al. (2008), using the Principle of Confidence Interval Inclusion, inappreciable shifts in $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$ are determined (as percentage shifts) through expert opinion that define a biological region of significance. The similarity region is specified (see Figure 3.1) as $\begin{bmatrix}(\Delta_1, \Delta_2)\\(\Delta_3, \Delta_4)\end{bmatrix}$ such that Δ_i, $i = 1,2,3,4$ are defined as

$$\begin{aligned}\Delta_1 &= g_1(\underline{\omega}) - \pi_1 \times g_1(\underline{\omega}),\ \Delta_2 = g_1(\underline{\omega}) + \pi_2 \times g_1(\underline{\omega})\\ \Delta_3 &= g_2(\underline{\omega}) - \pi_3 \times g_2(\underline{\omega}), \Delta_4 = g_2(\underline{\omega}) + \pi_4 \times g_2(\underline{\omega})\end{aligned} \tag{3.9}$$

where π_i are the inappreciable percentage shifts that result in the similarity region of biologic significance. The 95% confidence ellipse for the $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$, where σ_h^2 =0, is plotted (Figure 3.1). For details on how the confidence ellipse is computed see Chapter 2 and Appendix A.2.

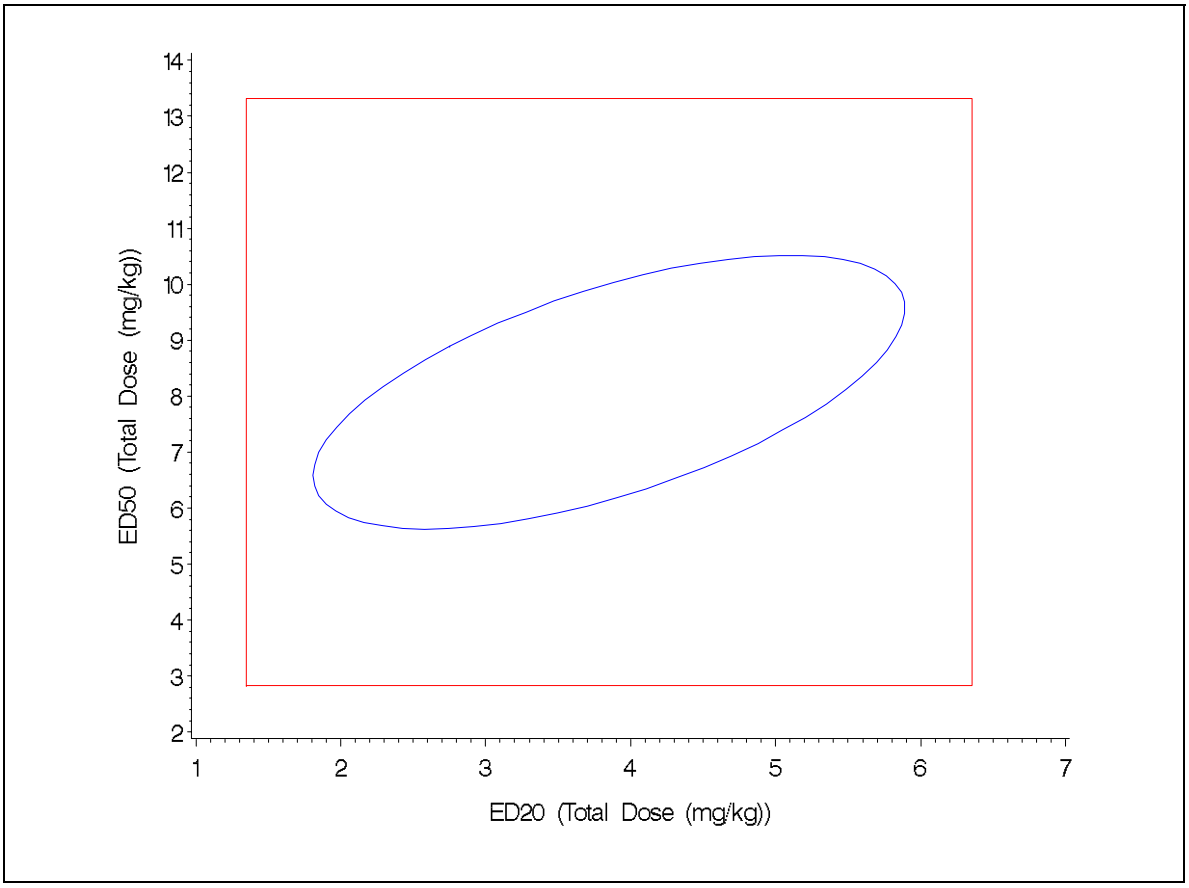


Figure 3.1. Plot of 95 confidence ellipse for the ED(20) and ED(50) with the plotted similarity bounds (box).

Following Stork et al. (2008), the maximum level of additional variability (i.e., $\sigma_h^2 > 0$) is determined such that the confidence ellipse is contained in the similarity region (see Appendix A.3 for calculation of σ_h^2; see box in Figure 3.2). The similarity bounds, (h_L, h_U), are found using the following procedure (Stork et al., 2008):

Determine

$$h_L = 1 - z_{\max}\sigma_h$$
$$h_U = 1 + z_{\max}\sigma_h \qquad (3.10)$$

such that the value of $z_{\max}$ is chosen as the maximum value such that at least one of the following holds for some acceptably small value of $\varepsilon_j > 0, j = 1,2,3,4$:

$$\begin{aligned} g_1(\underline{\omega}) \times h_L - \Delta_1 &< \varepsilon_1 \\ \Delta_2 - g_1(\underline{\omega}) \times h_U &< \varepsilon_2 \\ g_2(\underline{\omega}) \times h_L - \Delta_3 &< \varepsilon_3 \\ \Delta_4 - g_2(\underline{\omega}) \times h_U &< \varepsilon_4 \end{aligned} \tag{3.11}$$

where (Δ_1, Δ_2) and (Δ_3, Δ_4) are the inappreciable shifts in $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$.

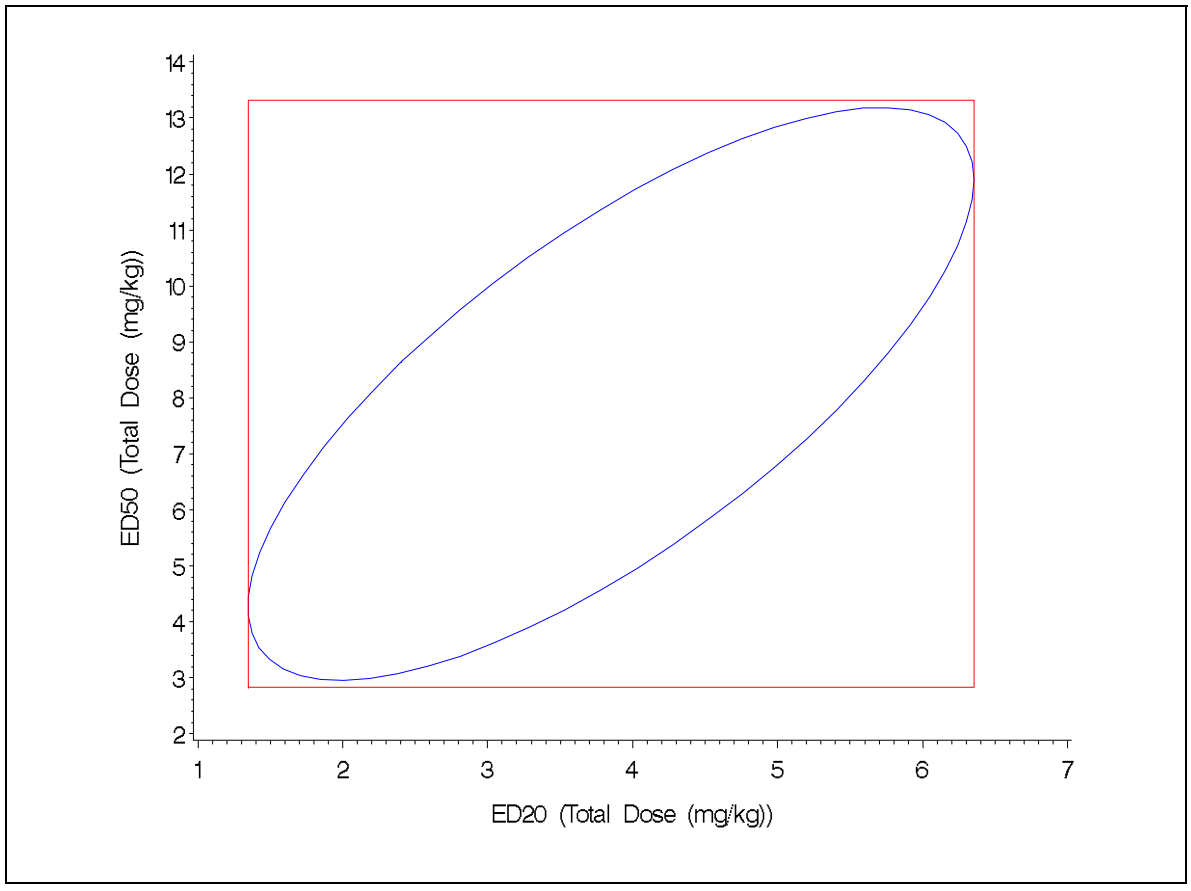


Figure 3.2 Plot of 95% confidence ellipse with $\sigma_h^2 > 0$ and the associated similarity region (box).

Step 3 In Determining Sufficient Similarity: Compute the Similarity Measure

In order to compute the similarity measure, *h*, define the similarity measure as

$$h = 1 + \frac{d}{t} \tag{3.12}$$

where $\frac{d}{t}$ is one of the suggested distance measures (eqs. 3.4-3.7). If the mixtures are exactly the same (have all the same chemicals in the same proportions) then $\frac{d}{t} = 0$. The more 'similar' the mixtures are the closer to zero $\frac{d}{t}$ becomes and subsequently the closer to one *h* becomes. If $h \in (h_L, h_U)$, the reference and candidate mixtures are concluded to be sufficiently similar in dose-response. Following these three steps results in the heuristic evaluation of sufficient similarity that is related to the "gold standard" test of sufficient similarity

$$\begin{aligned} H_0 &: g_1(\underline{\omega}) < \Delta_1 \text{ or } g_1(\underline{\omega}) > \Delta_2 \text{ or } g_2(\underline{\omega}) < \Delta_3 \text{ or } g_2(\underline{\omega}) > \Delta_4 \\ H_1 &: \Delta_1 \le g_1(\underline{\omega}) \le \Delta_2 \text{ and } \Delta_3 \le g_2(\underline{\omega}) \le \Delta_4 \end{aligned} \tag{3.13}$$

which is the formal equivalence test ("gold standard") for sufficient similarity.

Properties and Simplifications of the Adjusted Unweighted Distance

In this section we establish some properties in the form of a theorem and a postulate that act as a guide in utilizing this method in practice. In order to establish Theorem 1 and the postulate, it is necessary to make these simplifying assumptions

A1. The study is designed such that the first $c - s = k$ (where s is the number of chemicals that are absent from the reduced mixture, but are in the full mixture) chemicals in the full mixture are in the same relative proportions as the *k* chemicals in the reduced mixture. That is $\frac{a_{i,full}}{\sum_{i=1}^{k} a_{i,full}} = a_{i,red}$ which implies that

$a_{i,full} = a_{i,red} \sum_{i=1}^{k} a_{i,full}$.

A2. The additional $c - k = s$ chemicals in the full mixture are sub-threshold or in an inactive range; this implies that these *s* chemicals will not change the dose response curve other than to shift it to the right in terms of increasing the total dose unless there is an interaction.

A3. The variance-covariance matrix expands proportionately. Consider the following form of the variance-covariance matrix

$\Omega^*\left(\sigma_h^2\right) \approx (X_0^{*T} V_0^{-1} X_0^*)^{-1}$

where:

$X_0^* = \frac{\partial f}{\partial \omega}$ is an $N \times p$ matrix of derivatives,

$f_0(t, \omega) = f(t, \omega, b)|_{b=0}$,

$V_0^{-1} = R^{-1} - R^{-1} Z_0^* D(1 + Z_0^{*T} R^{-1} Z_0^* D)^{-1} Z_0^{*T} R^{-1}$, where in this case $D = \sigma_h^2$, and

$Z_0^* = \frac{\partial f}{\partial b}|_{b=0}$.

This is to say that as *D* increases the other components of the matrix all change proportionally by some factor of *D*. The other components are only changing as a result of *D* increasing.

Theorem 1

Without loss of generality, when using the adjusted unweighted similarity measure, as long as

$$1 - D_{AF} = \sum_{i=k+1}^{c} a_{i,full} \leq \pi_0 \qquad \textbf{(3.14)}$$

where π_0 is the smallest percentage shift allowed in $g_1(\underline{\omega})$ or $g_2(\underline{\omega})$ and $D_{AF} = \sum_{i=1}^{k} a_{i,full}$ then the conclusion will be that the two mixtures are sufficiently similar in dose-response under assumptions A1 and A2.

Proof

The quadratic form in eq. (3.5) can be broken into two terms, where one term is the sum of squared differences of the first *k* common chemicals (eq. 3.15) and one term is the sum of squared differences of the *s* chemicals that are not in the reduced mixture (eq. 3.16).

$$\sum_{i=1}^{k} (a_{i,red} D_{AF} - a_{i,full})^2 = 0 \text{ and} \qquad (3.15)$$

$$\sum_{i=k+1}^{c} (a_{i,red} D_{AF} - a_{i,full})^2 = \left(-a^2{}_{k+1,full} - a^2{}_{k+2,full} - \cdots - a^2{}_{c,full}\right) = \sum_{i=k+1}^{c} a^2{}_{i,full}. \qquad (3.16)$$

By design eq. 3.15 is equal to zero and $\left\{a_{i,red} D_{AF}\right\} = 0$ for $i = k+1,...,c$. Now, consider sufficient similarity to be defined by $h \in \left(h_L, h_U\right)$ where $h_L = 1 - \pi_0$ and $h_U = 1 + \pi_0$ and h is calculated as $1 + \frac{d}{t}$ in step 3 (eq. 3.12). In the case when *s*=1

$\frac{d}{t} = \sqrt{\sum_{i=k+1}^{c} a^2{}_{i,full}} = \sum_{i=k+1}^{c} a_{i,full} = a_{c,full}$. The larger that *s* becomes, the smaller $\frac{d}{t} = \sqrt{\sum_{i=k+1}^{c} a^2{}_{i,full}}$ becomes. As the number of additional components, *s*, increases but the additional proportion of the mixture that the *s* components constitute stays fixed, the additional proportions become smaller (Table 3.1). The allocation among the *s* chemicals at subthreshold levels can be in any manner, with the only constraint being that $\sum_{i=k+1}^{c} a_{i,full} \leq \pi_{()}$. In the case when there is one additional chemical the distance, $\left(\frac{d}{t}\right)$, is equal to $1 - D_{AF} = \sum_{i=k+1}^{c} a_{i,full} = \pi_{()}$. This creates an upper bound on the proportion of the mixture that can be at subthreshold levels while sufficient similarity can still be concluded without utilizing the weight matrix. Therefore, as long as $1 - D_{AF} = \sum_{i=k+1}^{c} a_{i,full} \leq \pi_{()}$ then the conclusion will be that the two mixtures are sufficiently similar in dose-response using the adjusted unweighted similarity measure. *QED.*

Consider the example in Table 3.1 where in one case there are 7 additional chemicals and in the other case there is 1 additional chemical.

Table 3.1. Calculated adjusted unweighted distances when *s*=1 and 7.

	***s* Additional Chemicals**	
	7	*1*
	0.1	0.65
	0.1	
	0.1	
	0.1	
	0.1	
	0.1	
	0.05	
$\pi_0 = \sum_s a_i$	0.65	0.65
Sum Squared Distances	0.0625	0.4225
$\frac{d}{t}$	0.25	0.65

Summing and squaring these small distances when there are 7 additional chemicals results in a smaller distance than when there is 1 additional chemical.

We can then think of $1 - D_{AF} = \sum_{i=k+1}^{c} a_{i,full} = \pi_0$ as being the proportion of chemicals (components) at subthreshold (inactive) levels that can be in the full mixture and sufficient

similarity is still concluded. However, if the mixtures are not designed so that they are proportional to each other, as is the case in Theorem 1, the proportion of chemicals at subthreshold levels that can be present without weighting necessarily decreases. Furthermore, if the *s* additional chemicals are mixed with some at subthreshold levels and others at the threshold and beyond, these properties and relationships begin to break down. Theorem 1 establishes a "rule of thumb" for using this method. In light of these properties, any additional chemical(s) at subthreshold levels that are greater than π_0 indicates that further steps may need to be taken in determining the appropriate similarity measure to use, if it is assumed the full and reduced mixtures are sufficiently similar.

Relationship between $z_{\max}$, σ_h^2, and π_0

Stork et al. (2008) suggest using the algorithm in eqs. 3.10 and 3.11, along with the Taylor-series expansion about *h*=1 in eq. (3.8) to estimate $\underline{\Omega}$ and to calculate $z_{\max}$, σ_h^2, and ultimately h_L and h_U. However, it can be argued mathematically and shown through examples that the quantity $z_{\max}\sigma_h = \min(\pi_1, \pi_2, \pi_3, \pi_4) = \pi_0$.

Postulate/Conjecture 1

Without loss of generality, given the algorithm described in eqs. 3.10 and 3.11, the upper and lower bounds (h_L, h_U) for the similarity measure *h* can be calculated as $1 \pm z_{\max}\sigma_h$ where $z_{\max}\sigma_h = \min(\pi_1, \pi_2, \pi_3, \pi_4) = \pi_0$ **under the assumption A3.**

Proof

Recall eq. 3.11

$$
\begin{aligned}
g_1(\underline{\omega}) \times h_L - \Delta_1 &< \varepsilon_1 \\
\Delta_2 - g_1(\underline{\omega}) \times h_U &< \varepsilon_2 \\
g_2(\underline{\omega}) \times h_L - \Delta_3 &< \varepsilon_3 \\
\Delta_4 - g_2(\underline{\omega}) \times h_U &< \varepsilon_4
\end{aligned}
$$

Define $\min(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) < \varepsilon^*$; $\Delta_1 = g_1(\underline{\omega})[1-\pi_1]$; $\Delta_2 = g_1(\underline{\omega})[1+\pi_2]$; $\Delta_3 = g_2(\underline{\omega})[1-\pi_3]$; $\Delta_4 = g_2(\underline{\omega})[1+\pi_4]$.

Putting these together eq. 3.11 becomes

$$
\begin{aligned}
& g_1(\underline{\omega})h_L - g_1(\underline{\omega})[1-\pi_1] < \varepsilon_1 \\
& = (h_L - [1-\pi_1]) < \varepsilon_1^* \\
& \\
& g_1(\underline{\omega})[1+\pi_2] - g_1(\underline{\omega})h_U < \varepsilon_2 \\
& = ([1+\pi_2] - h_U) < \varepsilon_2^* \\
& \\
& g_2(\underline{\omega})h_L - g_2(\underline{\omega})[1-\pi_3] < \varepsilon_3 \\
& = (h_L - [1-\pi_3]) < \varepsilon_3^* \\
& \\
& g_2(\underline{\omega})[1+\pi_4] - g_2(\underline{\omega})h_U < \varepsilon_4 \\
& = ([1+\pi_4] - h_U) < \varepsilon_4^*
\end{aligned}
\qquad (3.20)
$$

Choosing $\min(\varepsilon_1^*, \varepsilon_2^*, \varepsilon_3^*, \varepsilon_4^*) < \varepsilon^*$ the algorithm will choose the smallest value of π_i under the assumption that the variance-covariance matrix expands proportionally (computing the additional variability of σ_h^2). *QED.*

This is a useful property because this implies that the bounds for the similarity measure, *h*, can be calculated by finding the minimum allowable percentage shift in $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$ and it is not necessary to calculate the additional variability (σ_h^2) that can be added to the variance-covariance matrix of the fixed effects model. This adds additional flexibility to the method of Stork et al. (2008), which requires the calculation of σ_h^2.

Section 3.3 Example

Example: Part I

Consider the child care center study (Tulve et al., 2006) described in the motivating example. Potential exposure data (measured in ng/cm^2) were collected by means of floor swipes for 15 pyrethroids in 168 child care centers that constitute a nationally representative sample (Tulve et al., 2006). Of the 168 child care centers, all chemicals were below the limit of detection in 24.4% (41) of the centers. For the 15 pyrethroids studied, the percent of observations that were below the limit of detection ranged from 66% to 99%. When the observed exposure of a chemical was below the limit of detection it was given a standard value. The assumption is that observations that have exposure levels below the limit of detection are not of concern in the context of exposure to humans. To avoid the issue of observed exposures being below the limit of detection, for each of the 168 centers the total loading was calculated as, $total\ loading = \sum_{i=1}^{15} l_i$, where l_i is the loading or exposure for each of the 15 pyrethroids in ng/cm^2. The centers were then sorted in descending order by total loading so as to rank the centers by greatest total exposure. Percentiles were assigned to each

center based on total loading and centers that were in the top (10%) with respect to total loading, or the 16 centers with the highest total loading were used to calculate average mixing ratios for 15 pyrethroids. The individual chemical proportions, a_i, of the 15 pyrethorids are average proportions calculated from the observed exposure data (data not shown). The 5 pyrethroids with the lowest percent of observations below the limit of detection (with the exception of deltamethrin) were selected by EPA scientists (Tornero et al.) (permethrin, cypermethrin, beta-cyfluthrin, deltamethrin, esfenvalerate; Table 3.2) to use to conduct a dose-response study (Figure 3.3). The mixing ratios for these 5 pyrethroids constitute 96% of the 15 pyrethroids on average and the mixing ratios for the 5 selected pyrethroids (Table 3.2) were determined by dividing the calculated average proportions by 0.96. This step insures that the mixing ratios for the 5 pyrethroids sum to one.

Table 3.2. Mixing ratios and % below the limit of detection of 5 pyrethroids (reference mixture) obtained from the Tulve et al (2006) child care center study.

Chemical	Mixing Ratios Reference Mixture(Red)	% Below the Limit of Detection
Permethrin	0.522	66%
Cypermethrin	0.288	79%
B-Cyfluthrin	0.129	93%
Deltamethrin	0.034	98%
Esfenvalerate	0.027	93%

Table 3.3. Dose levels (mg/kg) for the five pyrethroid mixture study (Table 3.2).

Dose Levels
0
0.274
1.096
2.74
9.042
13.7
18.084
27.40

A dose-response study was conducted following the methods of Wolansky et al. (2005) with the proposed mixture (Tables 3.2 and 3.3) to characterize the neurotoxicological effects of the mixture (Figure 3.3). One question that arises is whether it makes sense to conduct a dose-response study for the purposes of studying neurotoxicological effects on the mixture of 5 pyrethroids when exposure was measured on 15 pyrethroids. This reduction to 5 pyrethroids makes sense if one can claim that the mixture of 5 pyrethroids is sufficiently similar in dose-response to the mixture of 15 pyrethroids. It is assumed that 65% shifts in the ED(20) and ED(50) are associated with inappreciable shifts in the dose-response curve. The D_{AF}=0.65 which implies $\pi = 1 - D_{AF} = 0.35$. Now, utilizing Theorem1 and making the assumption that the additional 9 chemicals are present at subthreshold levels, it can be concluded that the two mixtures are sufficiently similar in dose-response, as the additional 10 chemicals only constitute

4% of the mixture (which is less than 35%). That is, the mixture of 5 pyrethroids is representative of the mixture of 15 pyrethroids under the stated assumption. Verification of this assumption can be evaluated with additional (external) dose-response data for these 10 chemicals.

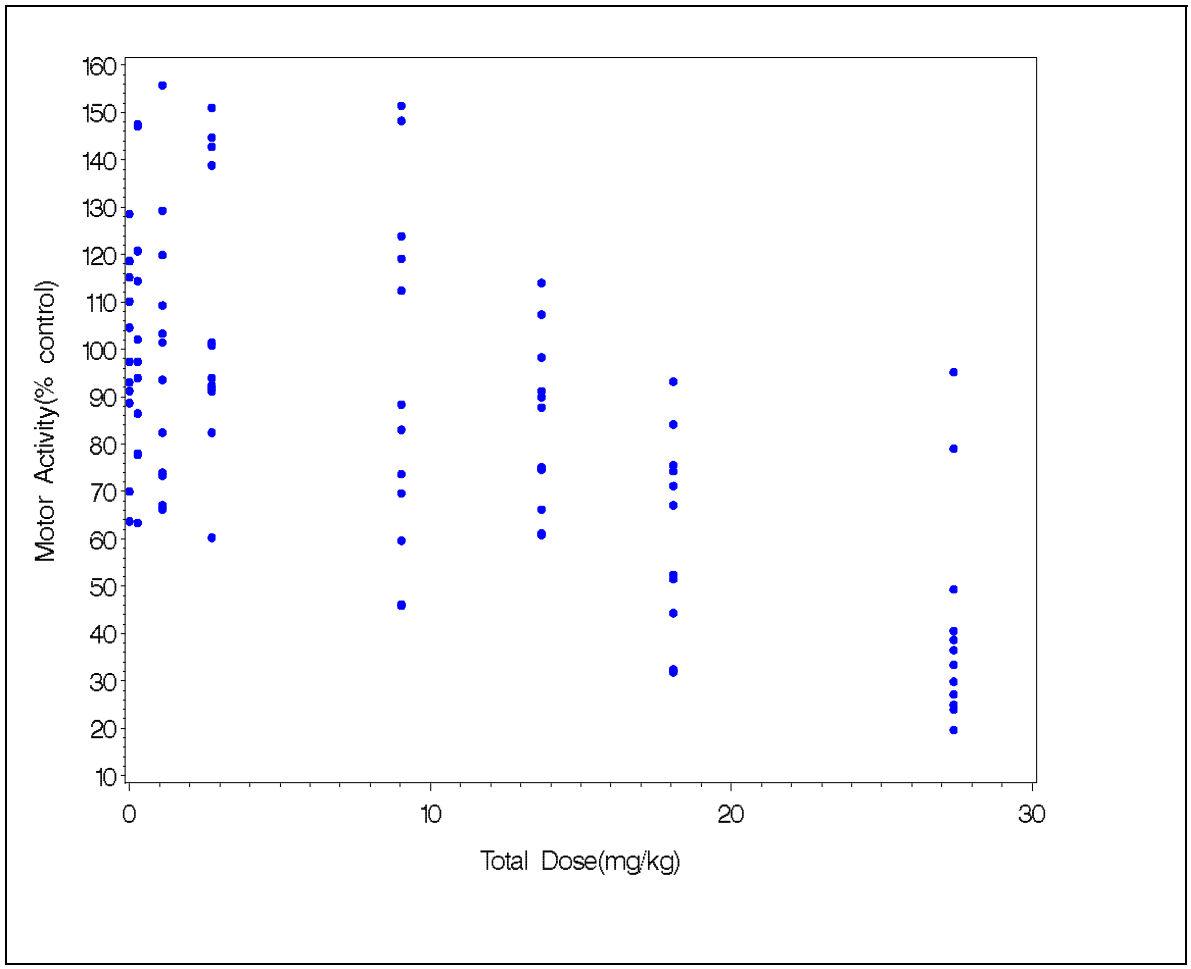


Figure 3.3. Observed (percent control) dose-response data (Tornero et al.) for the study conducted on the mixture in Table 3.2, where decrease in motor activity is the measure of neurotoxicity in adult male Long Evans rats.

Now, recall that it is of particular interest to use the mixture from the Tornero et al. dose-response study (Figure 3.3) as the surrogate/reference mixture for testing for sufficient similarity among the 20 child care centers with the greatest exposure levels (by total loading). For each

center, the proportions, a_i, of the observed mixture were calculated as $a_i = \frac{l_i}{\sum_{i=1}^{5} l_i}$. For each of the 20 selected centers, these proportions constitute a candidate mixture. It is desired to evaluate the 20 candidate mixtures, where there are no dose-response data available, to determine if the individual candidate mixtures are sufficiently similar in dose-response to the fully characterized reference mixture. For the candidate mixtures that are concluded to be sufficiently similar in dose-response to the reference mixture, it is of interest to investigate whether inferential statements relating to the reference mixture can be applied to the respective candidate mixtures. This is to say that risk (hazard) can be estimated for sufficiently similar mixtures without any additional test data. For example, the benchmark dose can be calculated for the reference mixture dose-response study and then used to provide pseudo-estimates of benchmark doses for the candidate mixtures in the child care center study that are concluded to be sufficiently similar in dose-response to the reference mixture.

Step 1

To be able to provide the pseudo-estimates of benchmark dose, the appropriate similarity measure must be selected. Recall the four proposed distance measures from Section 3.2. These measures provide flexibility so that the reference/candidate mixtures can be subsets of one another; adjustments can be made to account for differing total dose scales/additional chemical components; and weights can be assigned to the individual chemical components to account for aspects such as relative potency. If the candidate mixtures were subsets of the reference mixture, then one might choose to use the adjusted unweighted similarity measure. However, because the

candidate mixtures have the same components as the reference mixture, the unadjusted unweighted distance, eq. (3.4), is chosen to compute the similarity measure.

Step 2

Once the appropriate similarity measure is chosen an appropriate model needs to be fit to the dose-response data for the reference mixture. For the data depicted in Figure 3.3 a non-linear mixed effects exponential threshold model is selected for eq. (3.8), i.e.,

$$\mu = \begin{cases} \alpha + \gamma & t \le \delta \\ \alpha + \gamma \exp\left(\beta h(t-\delta)\right) & t > \delta \end{cases}$$

where $h \sim N(1, \sigma_h^2)$ is the random effect; β is the slope parameter; δ is the dose threshold; and t is total dose. The corresponding fixed effects model (for h=1 with $\sigma_h^2 = 0$) is fit to the reference mixture data set (Figure 3.4). For purposes of this example, it is assumed that through expert judgment 65% shifts in the ED(20) and ED(50) are associated with inappreciable shifts in the dose-response curve. These shifts create the similarity region of biologic significance (dashed and dotted curves in figure 3.4; box in figures 3.5a and b). Thus,

$g_1(\underline{\omega}) = ED(20)$ and $g_2(\underline{\omega}) = ED(50)$ where

$$ED(\mu_{20}) = \frac{\log\left(\frac{\mu_{20} - \alpha}{\gamma}\right)}{\beta} + \delta$$

and . (3.21)

$$ED(\mu_{50}) = \frac{\log\left(\frac{\mu_{50} - \alpha}{\gamma}\right)}{\beta} + \delta$$

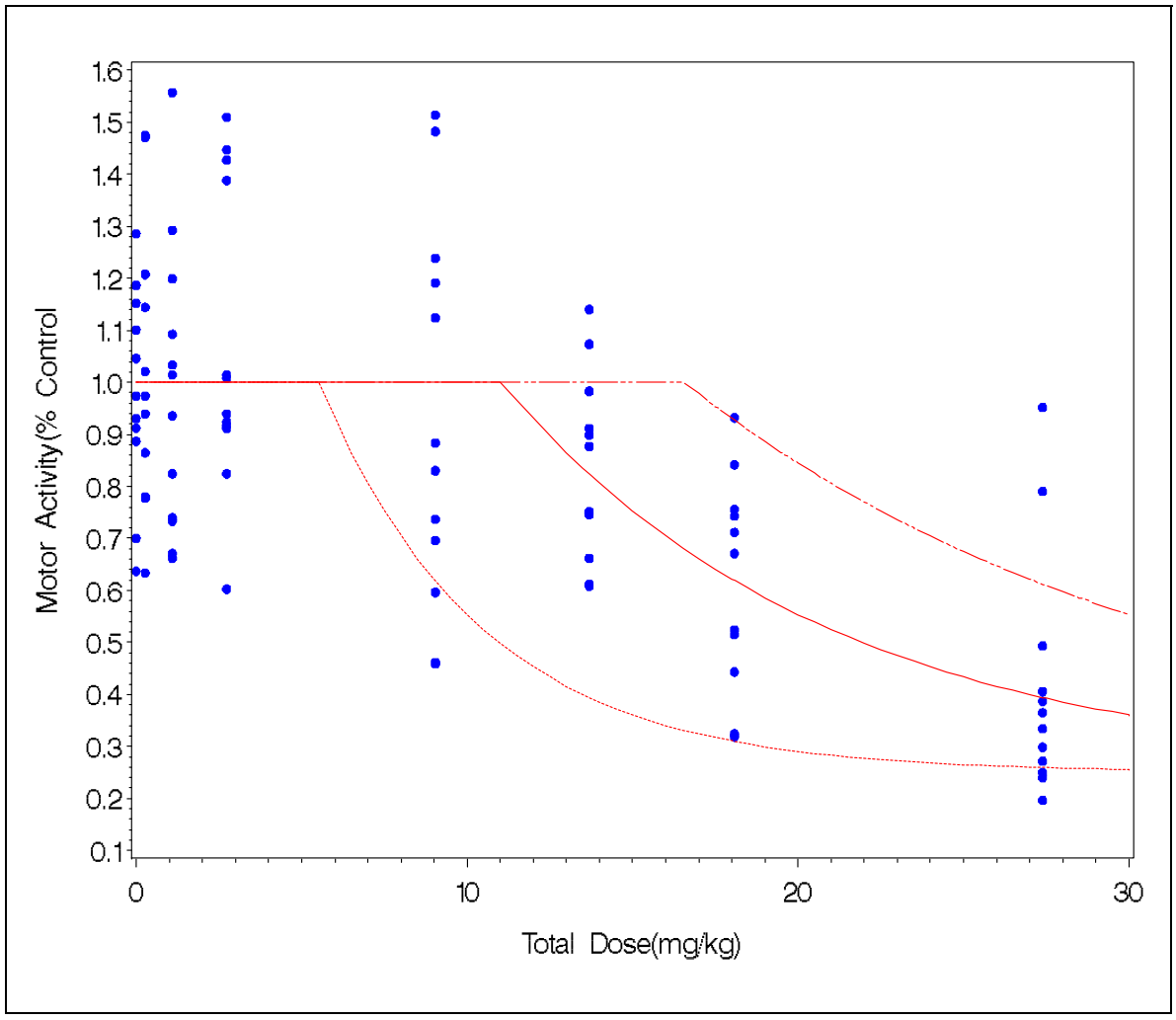


Figure 3.4. Plot of raw dose response data (dots) overlayed with the predicted curve from the model in eq. (3.21) (solid line) and the upper (broken/dashed line) and lower (dotted line) similarity bounds (resulting from 65% shifts in the ED(20) and ED(50)).

The resulting parameter estimates for the fitted model, the ED(20), ED(50) and benchmark dose associated with the benchmark response (resulting from a two standard deviation shift below the mean of the control group) are in the Table 3.4.

Table 3.4. Parameter estimates of the fitted non-linear exponential threshold model with the associated estimates of the ED(20) and ED(50).

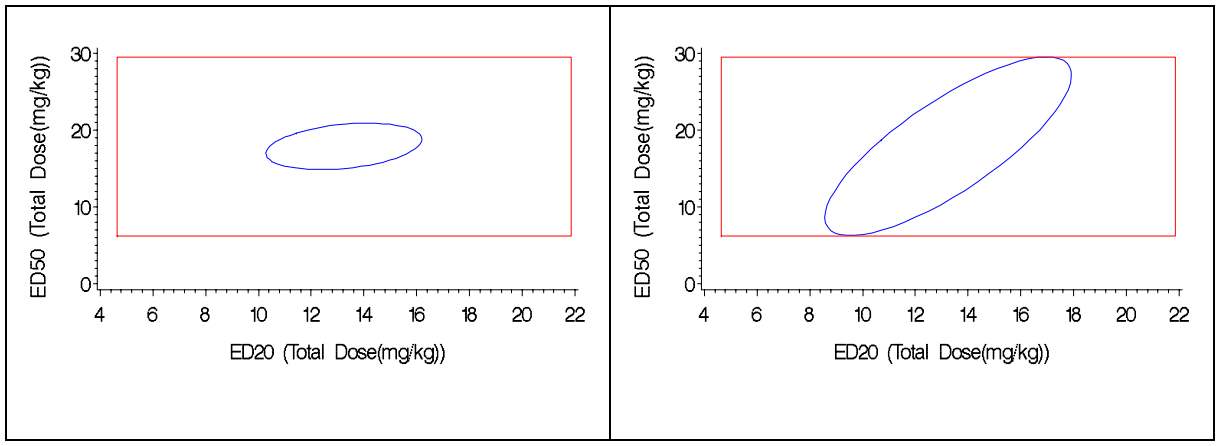
Parameter	Estimate	Standard Error	P-value
β	-0.10	0.03	0.004
δ	11.03	1.84	<0.0001
Additional Estimates			
ED(20)	13.24	1.29	<0.0001
ED(50)	17.91	1.32	<0.0001
Benchmark Dose*	16.10	1.05	<0.0001

* lower one sided 95% confidence interval (14.35)

Following Stork et al. (2008) the 95% confidence region is plotted with the similarity bounds for the ED(20) and ED(50) and the additional variability that can be added to the dose-response study is then calculated ($\sigma_h^2 = 0.43$) and the 95% confidence interval is plotted with the similarity bounds. It was determined through expert judgment that 65% shifts in either direction of both the ED(20) and ED(50) resulted in inappreciable shifts in the dose-response curve. Because the percentage shifts are the same and symmetric, the bounds, (h_L, h_U) on the similarity measure, *h*, were determined to be (0.35, 1.65) (Postulate 1). It should be noted that if the shifts were not symmetric, according to Postulate 1 the smallest percentage shift in any direction of

$g(\underline{\omega})$ would determine the bounds. The resulting similarity bounds (region) are

$$\begin{bmatrix} (4.64, 21.86) \\ (6.27, 29.54) \end{bmatrix} \text{ for the } \begin{bmatrix} ED(20) \\ ED(50) \end{bmatrix}.$$



Figures 3.5a and 3.5b. Plot of 95% confidence region and associated similarity region and plot of 95% confidence region as σ_h^2 increases.

Step 3

The similarity measures, *h*, for each of the 20 centers in the child care study were then calculated and it was concluded that 16 of the 20 (80%) centers were sufficiently similar in dose-response to the reference mixture (The distribution of the similarity measures with respect to total loading (ng/cm^2) are in Figure 3.6). For the 16 mixtures considered to be sufficiently similar it is of interest to investigate whether inferential statements relating to risk can be made based on the reference mixture. The two observed mixtures with the largest total loading and the two mixtures with relatively low total loadings were not sufficiently similar. If additivity is a viable assumption, an additivity model can be used to calculate individual measures of risk for the four observations not concluded to be sufficiently similar.

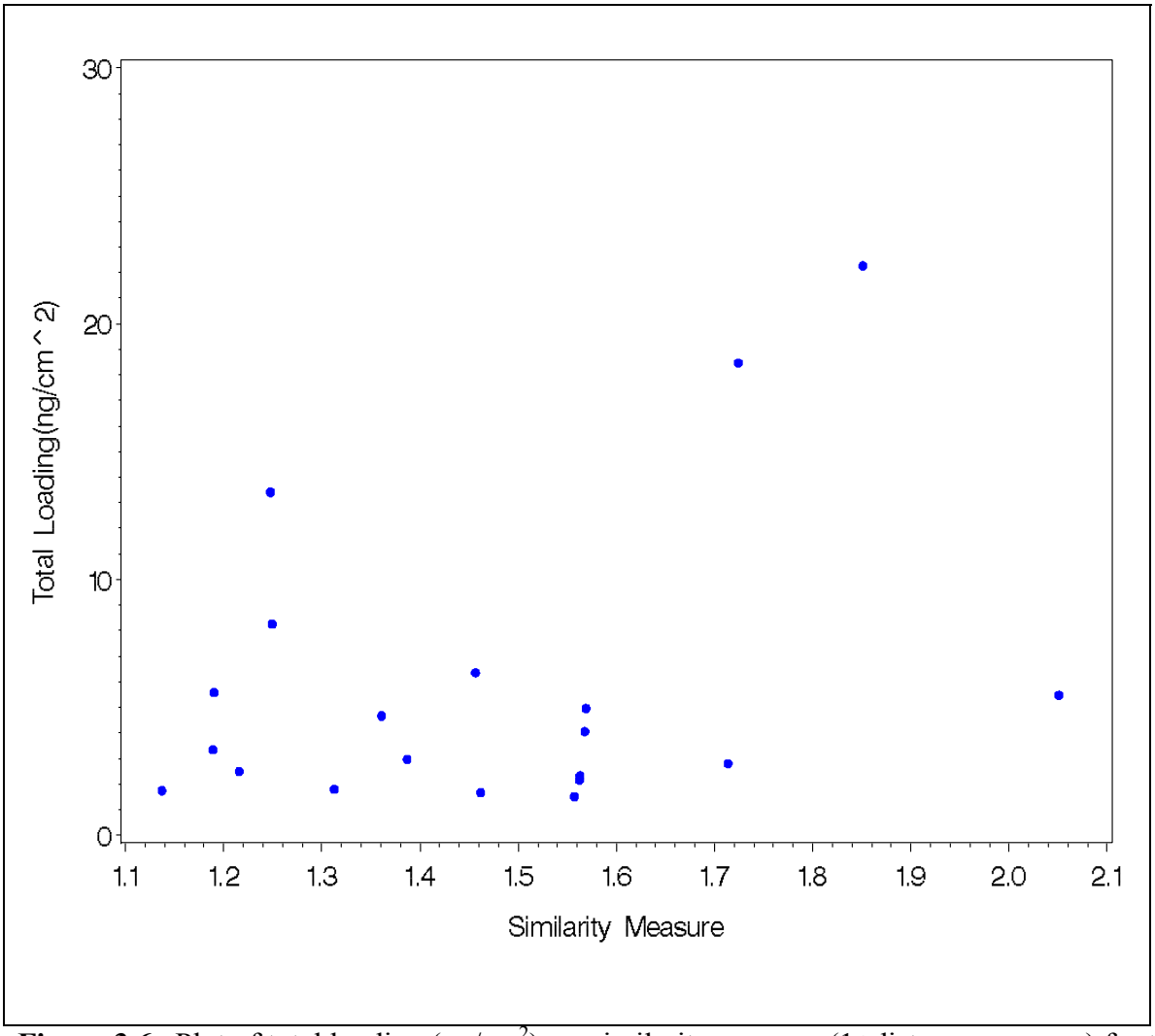


Figure 3.6. Plot of total loading (ng/cm^2) vs. similarity measure (1+distance measure) for the 20 centers with highest total loading.

To calculate the benchmark dose for the reference mixture (BMD_{ref}) dose-response study, the benchmark response was first calculated by finding the standard deviation in the control group (0.199) and subtracting 2 times this quantity from the mean of the control group (1) resulting in a benchmark response of 0.60, equivalent to calculating an ED(40). The associated benchmark dose was calculated using this value (eq. (3.21)) and is in Table 3.4. The associated lower one sided 95% confidence interval for the BMD_{ref} was found and is denoted as

the lower benchmark dose level, $BMD_{L,ref}$ (14.35, Table 3.4). Using the $BMD_{L,ref}$ the individual $BMD_{L,ref}$'s ($IBMD_{L,ref}$'s) for each chemical were calculated as $IBMD_{L,ref} = \underline{a}_{ref} BMD_{L,ref}$ (the reference mixture times $BMD_{L,ref}$). The similarity bounds for the $IBMD_{L,ref}$'s were calculated as 65% shifts in either direction of the $IBMD_{L,ref}$'s (Table 3.5). The lower $IBMD_{L,ref}$'s ($LIBMD_{L,ref}$) were calculated as $LIBMD_{L,ref} = IBMD_{L,ref} - 0.65 * IBMD_{L,ref}$ while the upper $IBMD_{L,ref}$'s ($UIBMD_{L,ref}$) were calculated as $UIBMD_{L,ref} = IBMD_{L,ref} + 0.65 * IBMD_{L,ref}$.

Thus, we use the BMD_{ref} for the reference mixture to generate confidence intervals on components of the mixture. Given the appropriate conversion factor(s) to relate exposure/total loading (ng/cm^2) to total dose would allow for statements to be made about the benchmark dose (BMD_{cand}'s) of candidate mixtures considered to be sufficiently similar in terms of actual exposure. In the next section we investigate if the confidence interval on the $IBMD_{L,ref}$'s for the reference mixture can be used to evaluate sufficient similarity in terms of individual benchmark doses for possible candidate mixtures.

Table 3.5. Individual chemical benchmark doses and their associated lower and upper similarity bounds from the observed mixture data.

Chemical	Mixing Ratios	Individual Benchmark Doses	Lower Benchmark Dose*	Upper Benchmark Dose*
Permethrin	0.522	7.491	2.622	12.360
Cypermethrin	0.288	4.133	1.446	6.819
B-Cyfluthrin	0.129	1.851	0.648	3.054
Deltamethrin	0.034	0.488	0.171	0.805
Esfenvalerate	0.027	0.387	0.136	0.639

*results from 65% shifts in either direction of the lower individual benchmark dose

Example: Part II

It is of interest to determine which of the 20 candidate mixtures (where no dose-response data are available) are sufficiently similar in dose-response to the reference mixture and then use the reference benchmark dose (BMD_{ref}) and the lower reference benchmark dose($BMD_{L,ref}$) calculated from the reference mixture dose-response study to make inferential statements about the candidate benchmark doses (BMD_{cand}'s), lower candidate benchmark doses ($BMD_{L,cand}$'s), and individual lower candidate benchmark doses ($IBMD_{L,cand}$'s) for the candidate mixtures that are concluded to be sufficiently similar in dose-response. Recall that if it were possible to generate data for the candidate mixtures of concern, it would be of interest to analyze the distribution of the BMD_{cand}'s and $BMD_{L,cand}$'s. In order to analyze these distributions, data need to be generated. It is possible to generate data through simulation studies; however, it is necessary to make an assumption about additivity. It has been shown in a previous analysis (data

not shown) that there is evidence of departure from additivity in the reference mixture dose-response data set. For the purpose of being able to evaluate sufficient similarity between the candidate mixtures and the reference mixture we simulated a reference dose-response data set under the assumption of additivity. The additivity model used to generate the dose-response data for the reference mixture and for the simulations to characterize the candidate mixtures is given by

$$\mu = \begin{cases} \alpha + \gamma & t\sum_{i=1}^{c} a_i\beta_i \geq \delta \\ \alpha + \gamma \exp\left\{ t\sum_{i=1}^{c} a_i\beta_i - \delta \right\} & t\sum_{i=1}^{c} a_i\beta_i < \delta \end{cases} \qquad (3.22)$$

A normal random deviate , (r), was added to the mixture such that $r \sim N(0, \sigma^2)$ where $\sigma^2 = 0.0648$ is the mean square error of the original reference mixture data set. The slopes for the β_i and δ estimates were obtained from Wolansky et al. (2005) and for convenience α was fixed at 0.25 where $\gamma = 1 - \alpha$.

Table 3.6. Slopes for the individual chemicals from the model in eq. (3.22) and the common threshold used in the simulation studies from Wolansky et al (2005).

Chemical	Slope
Permethrin	-0.0139
Cypermethrin	-0.0554
β -Cyfluthrin	-0.2686
Deltamethrin	-0.2364
Esfenvalerate	-0.4959
Resmethrin	-0.002
λ -cyhalothrin	-0.4505
δ	-0.2359

Using the parameter values in Table 3.6 and the model in eq. 3.22, the analysis in *Part I* was repeated for the simulated dose-response data set. The same non-linear mixed effects exponential threshold model was fit to the data and was reparamterized in terms of the ED(20) and ED(50). It was assumed that through expert judgment 65% shifts in the ED(20) and ED(50) are associated with inappreciable shifts in the dose-response curve. Once again, the likelihood was evaluated using Adaptive Gaussian Quadrature (PROC NLMIXED SAS V. 9.1) (Figure 3.7) and the resulting estimates for the parameters and BMD_{ref} and $BMD_{L,ref}$ are in Table 3.7.

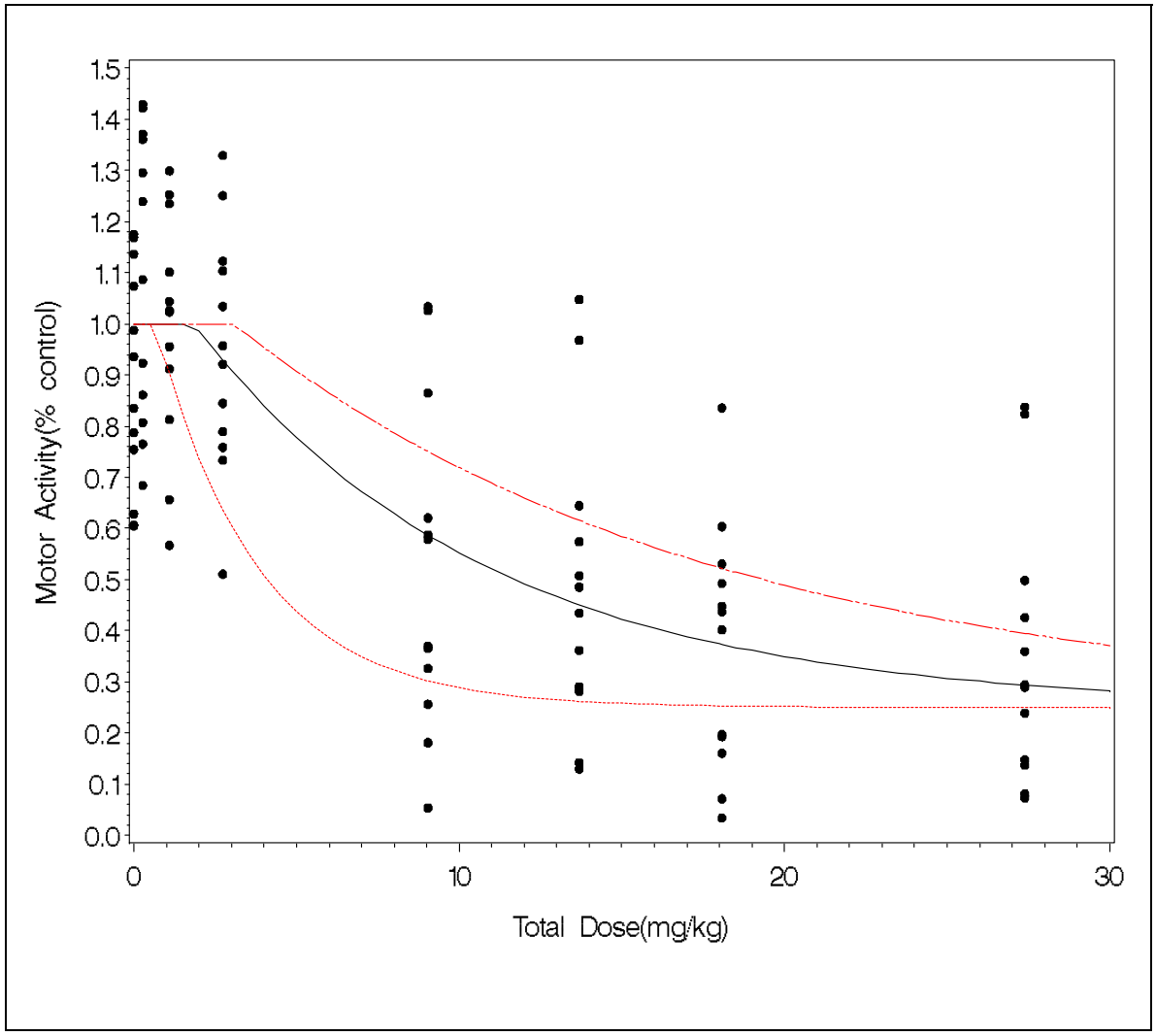


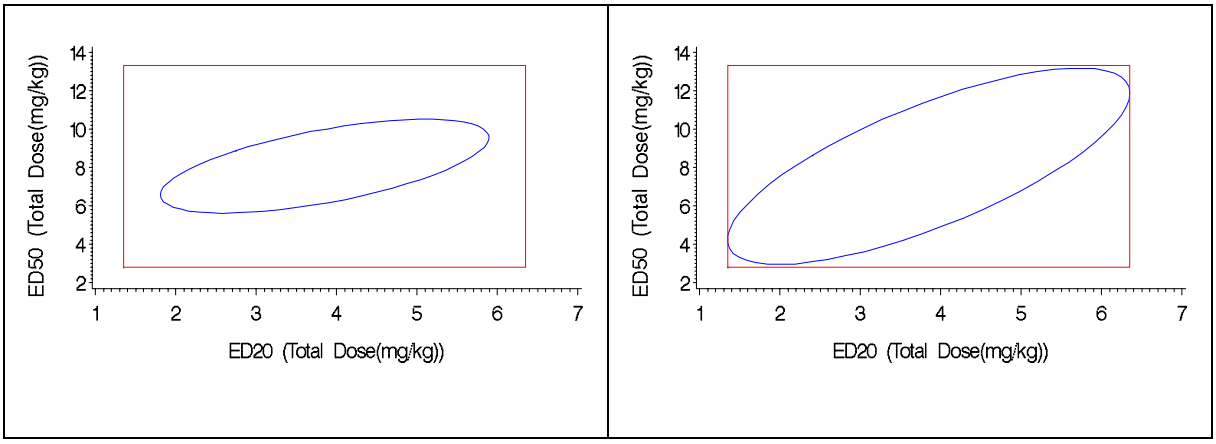
Figure 3.7. Plot of observed simulated dose response data (dots) overlayed with the predicted curve (solid line) and the upper (broken/dashed line) and lower (dotted line) similarity bounds (resulting from 65% shifts in the ED(20) and ED(50).

Table 3.7. Model parameter estimates and estimates of functions of model parameters of interest for simulated data.

Parameter	Estimate	Standard Error	P-value
β	-0.11	0.02	<0.0001
δ	1.85	1.09	0.09
Additional Estimates			
ED(20)	3.85	1.29	<0.0001
ED(50)	8.07	1.32	<0.0001
Benchmark Dose*	7.05	0.95	<0.0001

* lower one sided 95% confidence interval (5.47)

The 95% confidence ellipse was plotted with the similarity region and the additional variability was determined (Stork et al., 2008) such that the ellipse was still contained within the similarity region (Figures 3.8a and b). The resulting similarity bounds (region) are $\begin{bmatrix} (1.35, 6.35) \\ (2.82, 13.32) \end{bmatrix}$ for the $\begin{bmatrix} ED(20) \\ ED(50) \end{bmatrix}$. The bounds, (h_L, h_U) on the similarity measure, *h*, were determined to be (0.35, 1.65) using both Stork et al., (2008) and Postulate 1.



Figures 3.8a and 3.8b. Plot of 95% confidence region and associated similarity region and plot of 95% confidence region as σ_h^2 increases.

The similarity measures, *h*, for each of the 20 centers in the child care study were then calculated and it was concluded that 16 of the 20 centers (80%) were sufficiently similar in dose-response to the reference mixture and the $BMD_{L,ref}$'s, $IBMD_{L,ref}$'s and the associated similarity bounds of the $IBMD_{L,ref}$'s were calculated in the same manner as in *part I*.

Table 3.8. Reference mixture mixing ratios and lower individual benchmark doses and bounds.

Chemical	Mixing Ratios	Individual Benchmark Doses	Lower Benchmark Dose*	Upper Benchmark Dose*
Permethrin	0.522	2.855	0.999	4.710
Cypermethrin	0.288	1.575	0.551	2.599
B-Cyfluthrin	0.129	0.705	0.247	1.164
Deltamethrin	0.034	0.186	0.065	0.307
Esfenvalerate	0.027	0.148	0.052	0.244

The 16 centers (candidate mixtures) that were concluded to be sufficiently similar in dose-response were sorted by the magnitude of their respective similarity measure values and for illustration, three candidate mixtures were chosen based on their strength of similarity. Similarity measures closer to 1 are more similar. The three selected candidate mixtures range from being near 1 to being near the similarity bounds (Table 3.9).

Table 3.9. Three candidate mixtures from the observed simulated data, their associated similarity measures and how often they actually were sufficiently similar.

	Candidate Mixture		
Chemical	1	2	3
Permethrin	0.486	0.25	0.995
Cypermethrin	0.215	0.445	0.0009
B-Cyfluthrin	0.229	0.302	0.0009
Deltamethrin	0.0064	0.002	0.002
Esfenvalerate	0.0636	0.001	0.0012
Similarity Measure	1.14	1.36	1.57
% Sufficiently Similar	99.12	98.7	0.1

For each of the three candidate mixtures in Table 3.9, 1,000 dose-response data sets were simulated in the same fashion as the reference mixture data set (eq. 3.18 and Table 3.6). Using the similarity region specified previously, the "gold standard" test for sufficient similarity between the simulated reference mixture data set and the respective candidate mixtures was

conducted (see Chapter 2 for details) to determine how often the selected candidate mixtures were actually sufficiently similar in dose-response to the reference mixture; and to yield empirical distributions of BMD_{cand}'s, $BMD_{L,cand}$'s, and $IBMD_{L,cand}$'s. We observed that as the similarity measure increased the percent of reference and candidate mixtures concluded to be sufficiently similar decreased (Table 3.9). The percent of the time that the estimated $IBMD_{L,cand}$'s fell within the similarity bounds was calculated (Tables 3.10-3.12) as well as pseudo-estimates for the $IBMD_{L,cand}$'s (*pe*- $IBMD_{L,cand}$). The $BMD_{L,cand}$'s were estimated from the simulated candidate data sets (as described previously) and the $IBMD_{L,cand}$'s are calculated as $BMD_{L,cand}\underline{a}_{cand}$. The *pe*- $IBMD_{L,cand}$'s are calculated as $BMD_{L,ref}\underline{a}_{cand}$. Where the *pe*- $IBMD_{L,cand}$'s are located in relation to the empirical distribution of the 1,000 $IBMD_{L,cand}$ estimates can be seen in the histograms in Figures 3.9-3.11.

Table 3.10. Components of candidate mixture 1 with their estimated $IBMD_L$ and the percent contained in the bounds.

Chemical	Candidate Mixing Ratios	Individual Lower Benchmark Doses*	% Contained in BMD Bounds
Permethrin	0.486	2.6577882	86.77
Cypermethrin	0.215	1.1757705	96.14
B-Cyfluthrin	0.229	1.2523323	17.86
Deltamethrin	0.0064	0.0349997	10.03
Esfenvalerate	0.0636	0.3478093	3.86

*"estimate" obtained from using original lower benchmark dose

Table 3.11. Components of candidate mixture 2 with their estimated $IBMD_L$ and the percent contained in the bounds.

Chemical	Candidate Mixing Ratios	Individual Lower Benchmark Doses*	% Contained in BMD Bounds
Permethrin	0.25	1.367175	95.67
Cypermethrin	0.445	2.4335715	29.63
B-Cyfluthrin	0.302	1.6515474	3.03
Deltamethrin	0.002	0.0109374	0
Esfenvalerate	0.001	0.0054687	0

*"estimate" obtained from using original lower benchmark dose

Table 3.12. Components of candidate mixture 3 with their estimated $IBMD_L$ and the percent contained in the bounds.

Chemical	Candidate Mixing Ratios	Individual Lower Benchmark Doses*	% Contained in BMD Bounds
Permethrin	0.995	5.4413565	0
Cypermethrin	0.0009	0.0049218	0
B-Cyfluthrin	0.0009	0.0049218	0
Deltamethrin	0.002	0.0109374	1.25
Esfenvalerate	0.0012	0.0065624	0.1

*"estimate" obtained from using original lower benchmark dose

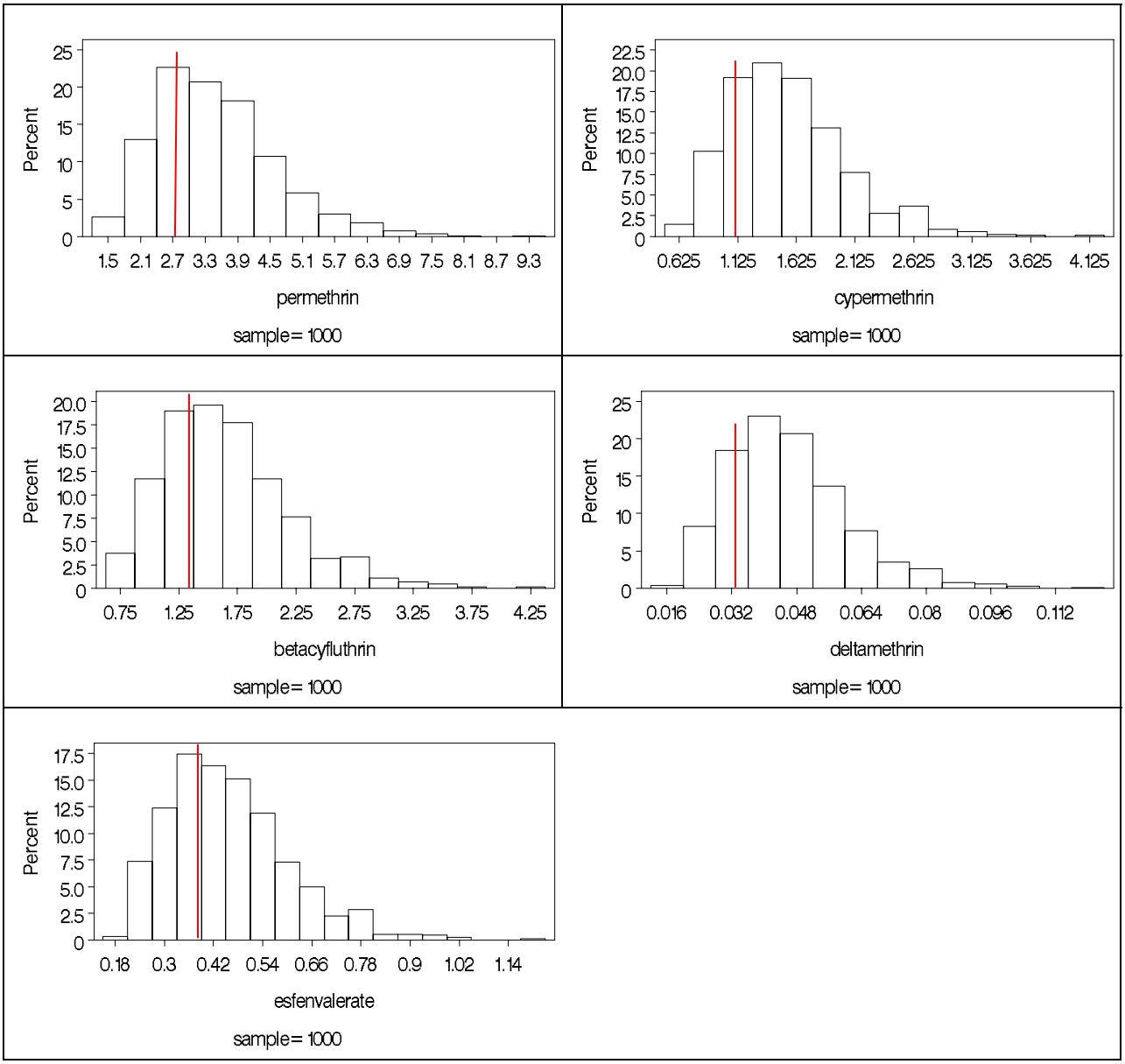


Figure 3.9. Histogram of chemicals from 1000 simulations where the red line is the *pe*- $IBMD_L$ for candidate mixture 1.

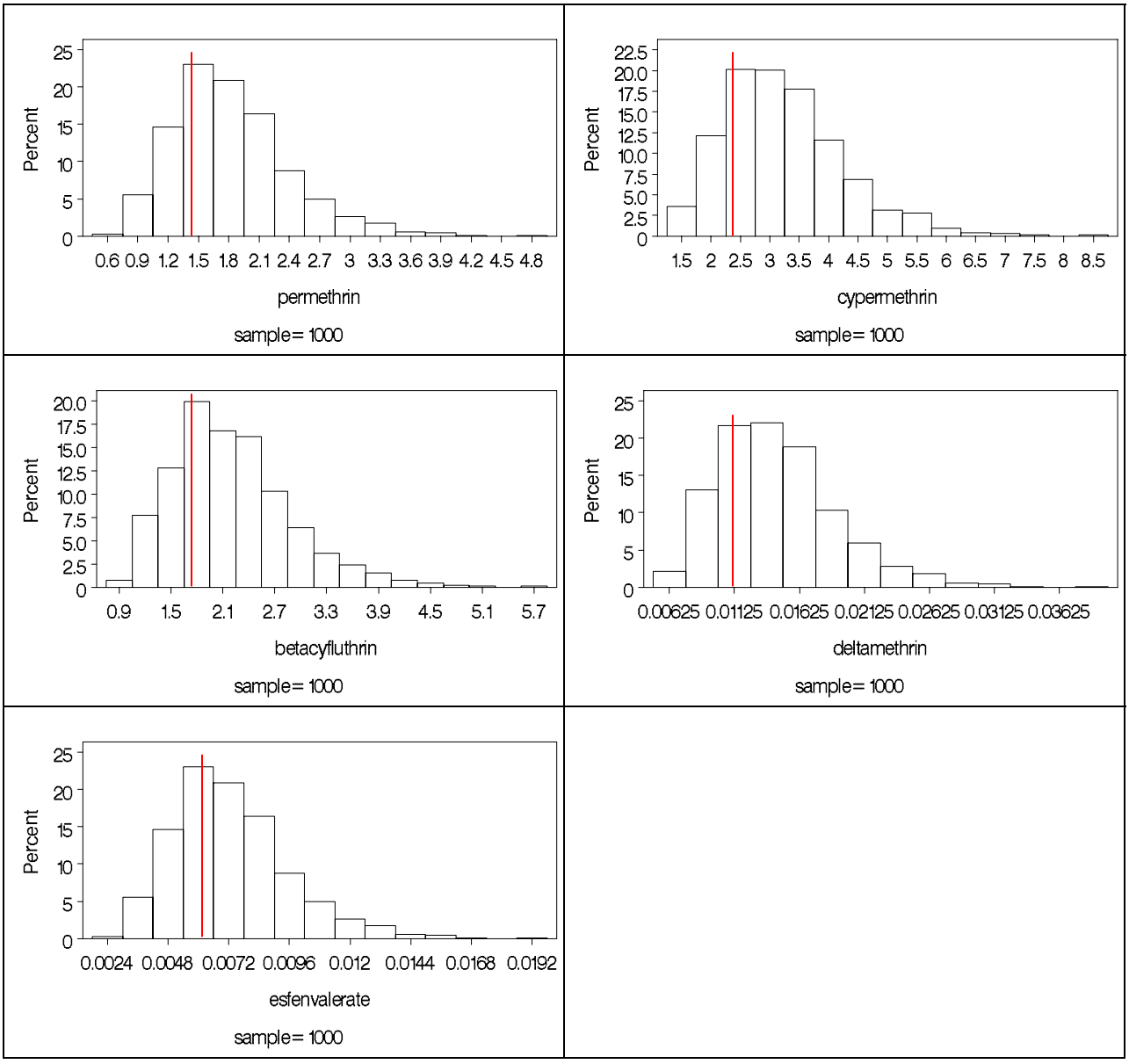


Figure 3.10. Histogram of chemicals from 1000 simulations where the red line is the *pe*-$IBMD_L$ for candidate mixture 2.

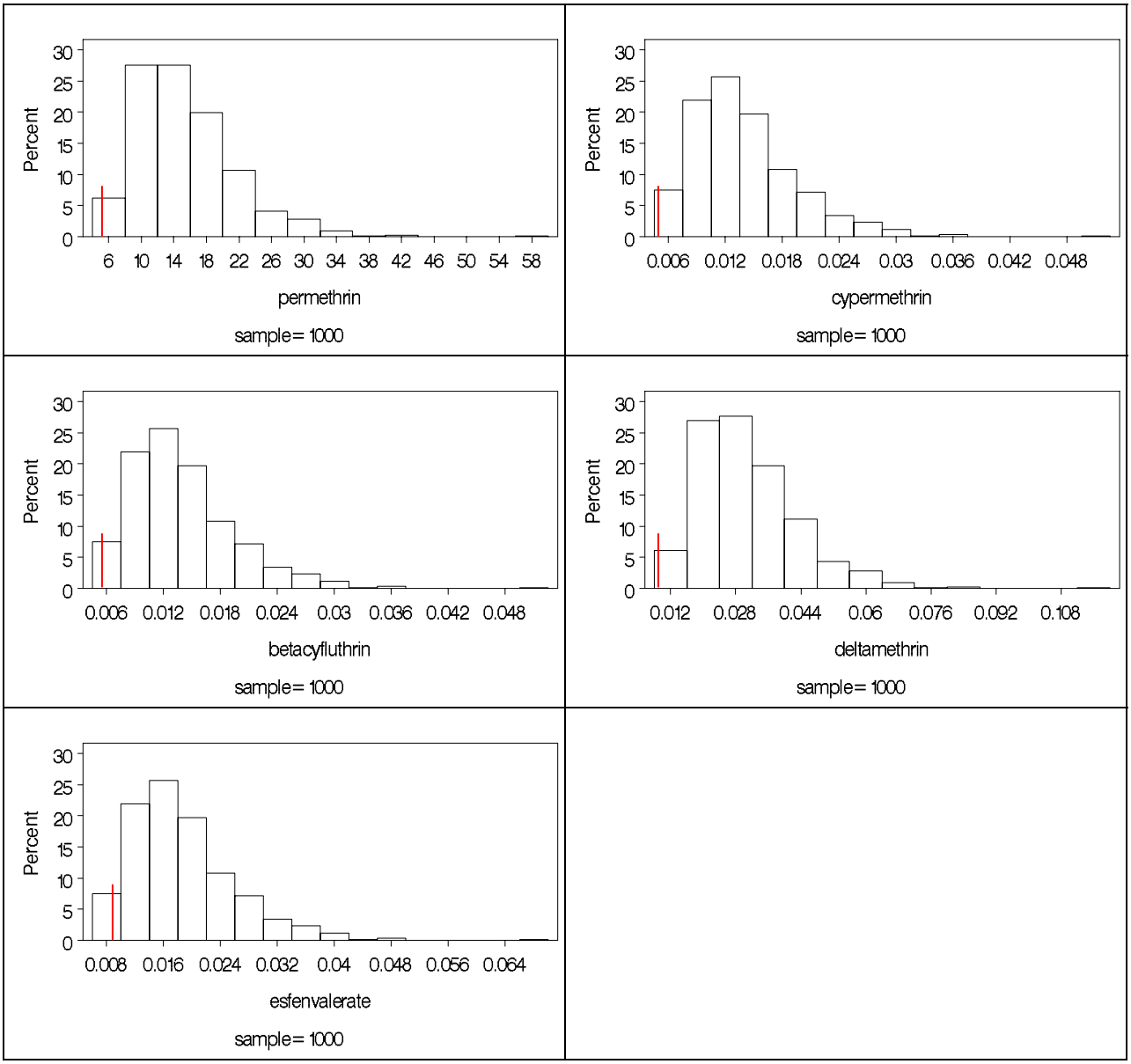


Figure 3.11. Histogram of chemicals from 1000 simulations where the red line is the *pe*-$IBMD_L$ for candidate mixture 3.

It is evident from the histograms in Figures 3.9-3.10 for candidate mixtures 1 and 2 that the *pe*- $IBMD_{L,cand}$ are adequate surrogate estimates for the $IBMD_{L,cand}$ as the estimates fall in the center of the distribution. For candidate mixture three the *pe*- $IBMD_{L,cand}$ does not appear to perform as well in estimating the center of the distribution as they are in the tails of the

distribution (Figure 3.11). These results are not surprising as only 0.10% of the simulated dose-response curves for candidate mixture 3 were found to be sufficiently similar to the simulated reference mixture data set while roughly 99% of the simulated dose response curves for mixtures 1 and 2 were sufficiently similar.

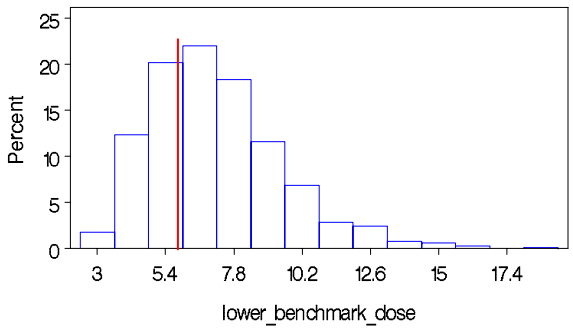


Figure 3.12a Histogram for candidate mixture 1

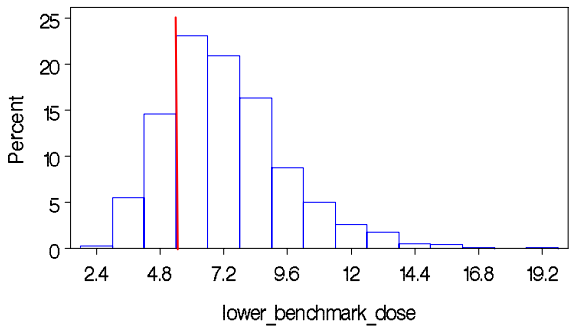


Figure 3.12b Histogram for candidate mixture 2

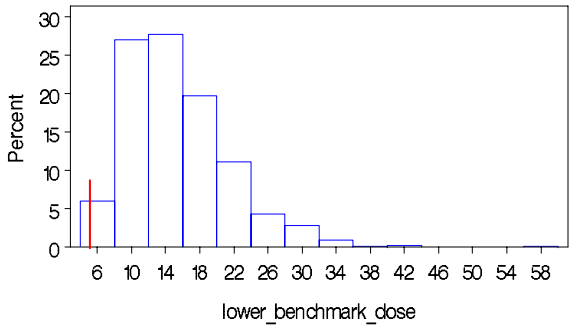


Figure 3.12c Histogram for candidate mixture 3

Figures 3.12a-c. Histograms of the estimates for the BMD_L's from the 100 simulations for each of the three candidate mixtures with the estimate of the BMD_L (5.5 mg/kg; solid red line) from the simulated reference mixture data set.

It is also of interest to view where the estimate of the $BMD_{L,ref}$ (5.5 mg/kg) for the simulated reference data set is located in relation to the distribution of the estimated $BMD_{L,cand}$'s from the 1000 simulations (Figures 3.12a-c). For mixtures 1 and 2, the $BMD_{L,ref}$ is an adequate surrogate for $BMD_{L,cand}$ as it falls in the middle of these two distributions, whereas in mixture 3 it is in the tail. Once again, these results are not surprising as only 0.10% of the simulated dose-response curves for candidate mixture three were found to be sufficiently similar to the simulated reference mixture data set while roughly 99% of the simulated dose response curves for mixtures 1 and 2 were sufficiently similar.

Section 3.4 Discussion

The four similarity measures proposed in the methods section, based on Euclidean distance, demonstrate added flexibility as compared to the method proposed by Stork et al. (2008). The proposed methods have the ability to account for the reference and candidate mixtures having differing dose scales due to one mixture being a subset of the other. These similarity measures also provide the option to include different weights for each of the *c* chemicals. In general, the weights should be decided through expert judgment based on environmentally or toxicologically relevant characteristics. Perhaps one of the most convenient extensions from the work of Stork et al. (2008) is the relationship between (h_L, h_U) and the allowable percentage shifts in $g_1(\underline{\omega})$ and $g_2(\underline{\omega})$ as this does not require a form of the variance-covariance matrix that is a function of the variance parameter for the random effect and allows for use with all types of data.

It appears that for candidate mixtures with similarity measures roughly below 1.40 the *pe*-$IBMD_{L,cand}$ and $BMD_{L,cand}$ perform well in estimating the respective centers of the distributions

of $IBMD_{L,cand}$'s and $BMD_{L,cand}$'s when there are no dose-response data available. Similarly, it would be expected that the BMD_{ref}, $BMD_{L,ref}$, and *pe*- $IBMD_{L,ref}$'s for the original reference mixture in *part I* would provide good surrogate estimates for the BMD_{cand}, $BMD_{L,cand}$, and $IBMD_{L,cand}$'s when there are no dose-response data available (Table 3.13). As the similarity measure approaches the boundary (1.65) the simulation studies suggest that in fact these candidate mixtures are not sufficiently similar in dose-response. One might argue that this suggests the similarity bounds should possibly be tighter (i.e., the percentage shifts should not be so great). However, recall that the similarity bounds are selected with biologic significance in mind. If it is the case that smaller bounds are also biologically significant then it would first need to be determined if this similarity region contains the 95% confidence ellipse. If the ellipse is not contained in the similarity region, this presents a design issue that is addressed in Chapter 6.

Table 3.13. *pe*- $IBMD_L$'s for the 3 candidate mixtures as obtained from the original reference mixture dose-response study.

	Candidate Mixture		
Chemical	1	2	3
Permethrin	6.97	3.59	14.28
Cypermethrin	3.09	6.39	0.01
B-Cyfluthrin	3.29	4.33	0.01
Deltamethrin	0.09	0.03	0.03
Esfenvalerate	0.91	0.01	0.02

The proposed methods demonstrated their use as far as application to risk assessment with use of benchmark doses. When a candidate mixture of concern is determined to be sufficiently similar in dose-response to reference mixture, the benchmark dose and individual benchmark doses from the reference mixture can be used as surrogates for the candidate mixtures of concern. It is safe to assume that there exists a relationship between the observed exposure of children in these child care centers to these pesticides. Given the appropriate conversion factors the exposure data could be compared to the benchmark doses for the purpose of determining risk.

Perhaps one of the most important features of the proposed similarity measures with respect to use in other risk assessment applications is the ability to reduce multi-dimensional data into a single summary measure, h, that describes where a single observation is located with respect to the rest of the distribution. Furthermore, the methods described have a direct link to toxicity through dose-response data and calculation of the benchmark dose, which is a nice feature as compared to the methods of Feder et al. (2009) that refer to similarity in terms of characteristics of a mixture, yet with no link to toxicity.

One issue not addressed in this chapter is the performance of the different similarity measures in different settings. The next chapter (Chapter 4) provides simulation studies to help characterize the properties and performance of each of the similarity measures.

Chapter 4: Properties of the Similarity Measures: Studies Through Simulation

Section 4.1 Introduction

In Chapter 3 we established four different similarity measures and presented an example demonstrating the use of sufficient similarity in risk assessment. However, the methods in Chapter 3 do not represent a true statistical test of sufficient similarity in dose-response, such as the gold standard test in Chapter 2, as there is no α-level established and there are no data available for the candidate mixtures of concern. Rather the methods in Chapter 3 represent a heuristic evaluation of sufficient similarity in dose-response. Even though the methods in Chapter 3 represent a statistical evaluation and not a true statistical test, it is still of interest to evaluate the properties of the methods. For example, when a new type of statistical hypothesis test or estimator is proposed it is often of interest to determine the properties of the test or estimator given different constraints, such as differing variances or distributional assumptions. Following the same logic, we evaluate the proposed similarity measures described in Chapter 3 in different scenarios with different constraints imposed.

Example

Consider a complex chemical mixture of five pyrethroids (permethrin, cypermethrin, β-cyfluthrin, deltamethrin, esfenvelerate) with complete dose-response data and the dose groups

given in Table 4.1. This mixture is completely characterized in dose-response and for the purpose of risk assessment, as described in Chapter 3, can be viewed as the reference mixture (eventually referred to as the reduced mixture) Now, consider a candidate mixture (eventually referred to as the full mixture), with only mixing ratios available, that is determined to be sufficiently similar in dose-response. Following the methods developed in Chapter 3, inferential statements about the benchmark dose (BMD) for the candidate mixture can be made based on this reference mixture. That is, given the BMD for the reference mixture statements about risk for the candidate mixture with respect to the BMD can be based on this value.

Table 4.1. Mixing ratios with their respective upper and lower bounds (Stork et al., 2008) and dose levels of the study in (mg/kg).

Chemical	a_i	$a_{i,lower}$	$a_{i,upper}$	**Total Dose Levels (mg/kg)**
				0
Permethrin	0.522	0.19	0.84	0.274
Cypermethrin	0.288	0.08	0.66	1.096
B-cyfluthrin	0.129	0.03	0.41	2.74
Deltamethrin	0.034	0.007	0.14	9.042
Esfenvelerate	0.027	0.006	0.12	13.7
				18.084
				27.400

Now, assume that this reference mixture represents average mixing proportions of some random process, such as applying pesticides at child care centers. It is logical to assume that the application process varies across child care centers with respect to the mixing ratios and the pesticides used. This is to say that the resulting chemical mixtures might differ slightly in component proportions and possibly contain additional pesticides. For the purposes of this example assume it is known that either resmethrin or λ-cyhalothrin is present in addition to the original five pyrethroids (Table 4.2) in other child care centers in the surrounding areas. (In this example the candidate/full mixtures will contain either resmethrin or λ-cyhalothrin but not both.) Assume that resmethrin is present in the mixture in significant proportions (0.20, 0.50, 0.65, 0.785) such that at the highest dose level it is still subthreshold (resmethrin threshold 117, Table 4.2; Wolansky et al., 2005) while λ-cyhalothrin is present in the mixture in negligible proportions (0.02, 0.07) but at the highest dose level it is at the threshold or beyond (λ-cyhalothrin threshold 1, Table 4.2; Wolansky et al., 2005). It is of interest to determine how well the proposed similarity measures perform with respect to determining sufficient similarity in dose-response in these two cases:

- Case 1: one (or more) chemical(s) in an inactive dose range comprises a significant proportion of the mixture, and
- Case 2: one (or more) chemical(s) in an active dose range comprise a negligible proportion of the mixture.

Performance is defined by sensitivity and specificity of the proposed methods.

Table 4.2. Proportion of the chemicals resmethrin and λ-cyhalothrin added to the simulated mixtures and their respective masses at the highest total dose group.

Chemical	
Resmethrin	**Mass in Mixture at Highest Dose Group (mg/kg)**
0.20	6.85*
0.50	27.40*
0.65	50.89*
0.785	100.04*
λ-cyhalothrin	
0.02	0.56
0.07	2.06

*Indicates that the mass at the given proportion is subthreshold.

Recall the four distance measures ($\frac{d}{t}$) from Chapter 3:

- Unadjusted Unweighted Distance (UUD; eq. (3.4))
- Adjusted Unweighted Distance (AUD; eq. (3.5))
- Unadjusted Weighted Distance (UWD; eq. (3.6))
- Adjusted Weighted Distance (AWD; eq. (3.7))

The similarity measure is one plus the distance measure, $h = 1 + \frac{d}{t}$, and is chosen based on the structure of the reference and candidate mixtures (full and reduced mixtures).

The following summarizes when it is appropriate to utilize each measure. The UUD might be used when the two mixtures have the same composition and dimension or when the additional chemical(s) in the full mixture are present in an inactive range but constitute a negligible proportion of the mixture. The AUD is used when the additional chemical(s) in the full mixture are present in an inactive dose range but constitute a significant proportion of the mixture. The UWD is used when the additional chemical(s) in the full mixture are present at threshold levels or beyond. The AWD is used when the additional chemical(s) in the full mixture are present in an inactive range and weighting is needed in addition to adjusting for total dose.

Section 4.2 Simulation Methods

To be able to evaluate how well each of the proposed similarity measures perform in the two cases it is necessary to define the "truth" with respect to sufficient similarity in dose-response. This is to say it is important to know whether the curves are sufficiently similar. One way to statistically determine if two mixtures are sufficiently similar in dose-response is by being able to perform the "gold standard" test which requires that complete dose-response data are available on both the reference and candidate mixtures. However, we only have complete dose-response data on the reference mixture (five pyrethroids data) and observed exposure data on the five pyrethroids of concern. We propose to generate possible candidate mixtures by utilizing the observed exposure data and then use these mixtures to simulate dose-response data. This allows for the gold standard test to be implemented (described in Chapter 2) as well as the evaluation of sufficient similarity in dose-response described in Chapter 3. The performance of the similarity measures is evaluated by computing the sensitivity and specificity (defined later in the section).

Section 4.2.1 Generating Possible Candidate Mixtures

To ultimately simulate dose-response data, it is first necessary to generate possible candidate mixtures. In order to generate candidate mixtures for the purpose of simulating dose-response data and to finally evaluate the performance of the proposed similarity measures, we use the observed exposure data that are available on the five pyrethroids (ng/cm^2). The top 20 observations with respect to total loading were selected to avoid dealing with observations with an exceedingly high proportion of readings below the limit of detection. For the purposes of simulation it was assumed that the log transformation of the observed exposures were multivariate normal, data $\underline{X}_{n\times k} \sim MVN(\underline{\mu}, \underline{\Sigma})$ where n=20 and k=5. Generating one observation from this distribution results in a vector of the form $\begin{bmatrix} x_1 & x_2 & \cdots & x_k \end{bmatrix}$. The resulting mixture has the following form $\begin{bmatrix} a_{1,cand} & a_{2,cand} & \cdots & a_{k,cand} \end{bmatrix}'$ where $a_{i,cand} = \frac{x_i}{\sum_{i=1}^{k} x_i}$. Simulate N=1000 candidate mixtures in this fashion. Let $\begin{bmatrix} a_{1,cand} \\ a_{2,cand} \\ \vdots \\ a_{k,cand} \end{bmatrix}$ be a $k \times 1$ vector of (chemical) components that constitute a candidate mixture with the constraints that $0 < a_i < 1$ for all i and $\sum_{i=1}^{k} a_i = 1$. At this point we have 1000 candidate mixtures of dimension $k \times 1$ and reference mixture of dimension $k \times 1$. However, it is desired to add an additional component or subset of components of dimension $s \times 1$ to the candidate mixtures to create a new set of full mixtures with dimension $(k+s)\times 1$, where $c = k + s$. Now the candidate mixtures assume the label full mixtures (with dimension

$c \times 1$) and the reference mixture assumes the label reduced mixture. Consider the following,

candidate (full) mixture, $\begin{bmatrix} a_{1,full} \\ a_{2,full} \\ \vdots \\ a_{k,full} \\ a_{k+1,full} \\ \vdots \\ a_{c=k+s,full} \end{bmatrix}$ where $c = k + s$.

For ease of calculating the mixing ratios of the (candidate) full mixtures and the associated new total dose groups (Section 4.2.2) the first c mixing ratios are constructed so that they are in the same relative proportions as in the (reference) reduced mixture and $\sum_{i=1}^{c} a_{i,full} = 1$. This is to say that $\frac{a_{i,full}}{\sum_{i=1}^{k=c-s} a_{i,full}} = a_{i,red}$; $a_{i,full} = a_{i,red} \sum_{i=1}^{k=c-s} a_{i,full}$ where $\sum_{i=1}^{k=c-s} a_{i,full} = 1 - \sum_{i \in s} a_{i,full}$;

$a_{i,full} = a_{i,red}(1 - \sum_{i \in s} a_{i,full})$. In the case where $s = 1$ this reduces to

$$a_{i,full} = a_{i,red}(1 - a_{k,full}) \tag{4.1}$$

Section 4.2.2 Generating Total Dose Groups for the Full Mixtures

For the purposes of conducting a Monte Carlo simulation and for the example presented here, suppose that $s = 1$ and the total dose groups for the reduced mixture are known to be $\left[t_{0,red} \;\; t_{1,red} \;\; \cdots \;\; t_{d,red} \right]$ where $j = 0,1,\ldots,d$, $\left\{ a_{i,red} t_{j,red} \right\} = x_{ij,red}$ and $x_{ij,red}$ is the dose of the

i^{th} component in the mixture at the j^{th} dose group such that, $\sum_{i=1}^{k} x_{ij,red} = t_{j,red}$. In order to create the dose groups for the full mixture $\{x_{ij,red}\}$ will be held constant for each dose group, $t_{j,red}$. This is to say that we only need to find $x_{c,full}$ such that

$$\frac{x_{cj,full}}{\sum_{i=1}^{c} x_{i,j}} = a_{c,full}$$

$$\frac{x_{cj,full}}{x_{1j,full} + x_{2j,full} + \cdots x_{kj,full}} = a_{c,full}$$

$$x_{cj,full} = a_{c,full}\left(x_{1j,full} + x_{2j,full} + \cdots + x_{cj,full}\right)$$

$$x_{cj,full} = a_{c,full}x_{1j,full} + a_{c,full}x_{2j,full} + \cdots + a_{c,full}x_{cj,full}$$

$$x_{cj,full} = a_{c,full}\sum_{i=1}^{c-1} x_{ij,full} + a_{c,full}x_{cj,full}$$

$$x_{kj,full} - a_{k,full}x_{kj,full} = a_{k,full}\sum_{i=1}^{c-1} x_{i,j}$$

$$x_{cj,full}\left(1 - a_{c,full}\right) = a_{c,full}\sum_{i=1}^{c-1} x_{i,j}$$

$$x_{cj,full} = \frac{a_{c,full}\sum_{i=1}^{c-1} x_{i,j}}{\left(1 - a_{c,full}\right)} \text{ where } \sum_{i=1}^{c-1} x_{ij,full} = t_{j,red} \text{ so that}$$

$$x_{cj,full} = \frac{a_{c,full}t_{j,red}}{\left(1 - a_{c,full}\right)}$$

Now, recall that we want to find $t_{j,full}$.

$$\begin{aligned} t_{j,full} &= x_{1j,full} + x_{2j,full} + \cdots + x_{(c-1)j,full} + x_{cj,full} \\ &= \sum_{i=1}^{c-1} x_{ij,full} + x_{cj,full} \\ &= t_{j,red} + x_{cj,full} \end{aligned}$$

$$= t_{j,red} + \frac{a_{c,full} t_{j,red}}{\left(1 - a_{c,full}\right)}. \quad (4.2)$$

Section 4.2.3 Simulating Dose-Response Data

Now that we have generated a set of possible full mixtures and the associated new total dose groups ($t_{j,full}$), dose-response data can be simulated. Generating dose-response data for the full (candidate) mixtures requires single chemical information such as slopes and a common threshold (Table 4.3 based on published data; Wolansky et al. (2005)).

Table 4.3. Slopes for the individual chemicals and the common threshold parameter from the nonlinear exponential threshold additivity model ($\mu = \alpha + \gamma \exp(t\sum_{i=1}^{c} a_i\beta_i - \delta)$, where *c* is the number of chemicals in the model) used in the simulation studies (Wolansky et al, 2005).

Chemical	Parameter	Estimate
Permethrin	β_1	-0.0139
Cypermethrin	β_2	-0.0554
β-Cyfluthrin	β_3	-0.2686
Deltamethrin	β_4	-0.2364
Esfenvalerate	β_5	-0.4959
Resmethrin*	β_6	-0.002
λ -cyhalothrin**	β_7	-0.4505
Threshold	δ	-0.2359

*Indicates the additional chemical added to the simulated candidate mixtures in the following proportions (0.20,0.50,0.65,0.785)

** Indicates the additional chemical added to the simulated candidate mixtures in the following proportions (0.02, 0.07)

To simulate the data a nonlinear model is utilized and additivity is assumed so that the model has the general form

$$y = \alpha + \gamma f(\sum_{i=1}^{c} a_i\beta_i, \delta, t) + \varepsilon \qquad (4.3)$$

For each of the *N* full mixtures that are generated a corresponding dose-response data set is simulated with fixed $\alpha, \gamma, a_i's, \beta_i's, \delta,$ and $t's$. The ε terms were simulated assuming constant

variance obtained from the mean square error of the Crofton et al. study. For the purposes of this example, each simulated data set is combined with the original 5 pyrethroid data set, resulting in a study that has complete data for two mixtures (reference/reduced and candidate/full). This implies that there are *N* studies created.

Section 4.2.4 Evaluating the Performance of the Chosen Similarity Measure

First, the appropriate percentage shift in functions of model parameters, $g(\underline{\omega})$, is determined. Then, for each of the studies, the gold standard test (Chapter 2) is conducted utilizing the fact that there are complete data on both mixtures of concern. The proposed methods (Chapter 3) for evaluating sufficient similarity in dose-response utilizing the different similarity measures is performed following the methods developed in Chapter 3. Having both results allows the performance of the proposed methods to be evaluated.

To assess how well the methods perform sensitivity and specificity are utilized. Recall that sensitivity is defined as the probability that given the condition exists the test indicates that the condition exists; specificity is the probability that given the condition does not exist the test indicates that the condition does not exist (Agresti, 2002). In our application, sensitivity is the conditional probability that we conclude sufficient similarity when the gold standard test concludes sufficient similarity and specificity is the conditional probability that we fail to conclude sufficient similarity when the gold standard test fails to conclude sufficient similarity. The $100(1-\alpha)\%$ confidence intervals for sensitivity and specificity can be determined in the following manner. Consider the 2x2 table below.

Table 4.4. 2x2 table used to calculated sensitivity and specificity and the associated standard deviations.

	Proposed Test $h \in (h_L, h_U)$			
Gold Standard	yes	No		
Yes	n_{11}	n_{12}	n_{1+}	**Sensitivity =** $\frac{n_{11}}{n_{1+}}$
No	n_{21}	n_{22}	n_{2+}	**Specificity =** $\frac{n_{22}}{n_{2+}}$

Consider each row of the table to be independent binomial distributions. Sensitivity is $\frac{n_{11}}{n_{1+}} = \pi_{1|1}$ and specificity is $\frac{n_{22}}{n_{2+}} = \pi_{2|2}$. Following the logic in Agresti (2002) and Casella and Berger (2002) the general form of the variance for $\pi_{j|i}$ is $\frac{\pi_{j|i}(1-\pi_{j|i})}{n_{i+}}$. The estimate of the variance can then be found by plugging in the maximum likelihood estimates of the respective parameters. The asymptotic $100(1-\alpha)\%$ confidence interval for sensitivity is

$$\frac{n_{11}}{n_{1+}} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\pi}_{1|1}(1-\hat{\pi}_{1|1})}{n_{1+}}}$$ and for specificity is

$$\frac{n_{22}}{n_{2+}} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\pi}_{2|2}(1-\hat{\pi}_{2|2})}{n_{2+}}}$$ where $\hat{\pi}_{j|i}$ are the maximum likelihood estimates. Alternatively, asymmetric confidence intervals could be constructed (Agresti, 2002).

Section 4.3 Results

Before the properties of the similarity measures can be assessed, it is first necessary to fit the appropriate non-linear model to five pyrethroid reference (reduced) dose-response data set, reparameterize the model, determine the similarity bounds (region), and calculate the 95% confidence region. Without loss of generality consider the nonlinear exponential threshold model

$$\mu = \begin{cases} \alpha + \gamma & t \le \delta \\ \alpha + \gamma \exp\left(\beta h(t-\delta)\right) & t > \delta \end{cases} \tag{4.4}$$

where $h \sim N(1, \sigma_h^2)$ is the random effect; β is the slope parameter; δ is the dose threshold; and t is total dose. This proposed fixed effects model is fit to the five pyrethroid (reference/reduced) mixture data set and can be thought of as a mixed effects model where the random effect, *h*, has a mean of 1 and $\sigma_h^2 = 0$. When the variance of the random effect is zero, the random effect model reduces to the fixed effects model. Conditional on α and γ, the model is reparameterized in two dimensions as functions of the model parameters in terms of the ED(20) and ED(50). For example, the ED(20) has the following form

$$ED(\mu_{20}) = \frac{\log\left(\frac{\mu_{20} - \alpha}{\gamma}\right)}{\beta} + \delta. \tag{4.5}$$

The resulting parameter estimates for the fitted model and the ED(20) and ED(50) are in Table 4.5 below.

Table 4.5. Parameter estimates from the fixed effects exponential threshold model for the simulated reference mixture dose-response data.

Parameter	Parameter Estimate	Standard Error	P-value
β	-0.1113	0.02437	<.0001
δ	1.845	1.091	0.0941
Additional Estimates			
ED(20)	3.8494	0.8803	<.0001
ED(50)	8.0714	1.0767	<.0001

For the purpose of this example it was assumed that through expert judgment biologically negligible shifts are 65% shifts in both the ED(20) and ED(50). The resulting similarity bounds (region) are $\begin{bmatrix} (1.35, 6.35) \\ (2.82, 13.32) \end{bmatrix}$ for the $\begin{bmatrix} ED(20) \\ ED(50) \end{bmatrix}$. Utilizing Postulate 1 (Chapter 3), the bounds, (h_L, h_U) on the similarity measure, h, were determined to be (0.35,1.65).

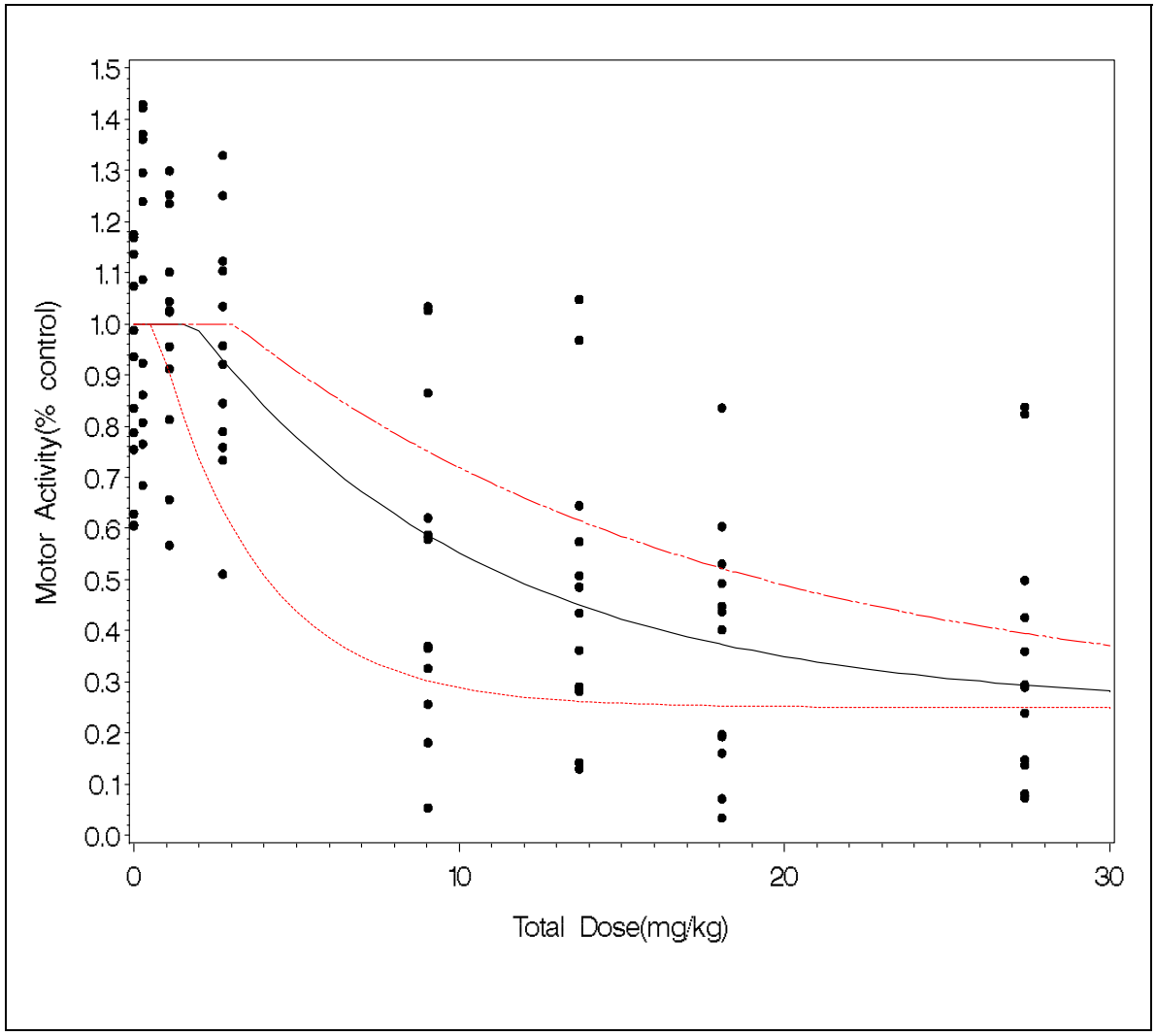


Figure 4.1. Plot of observed simulated dose response data (dots) overlayed with the predicted curve (solid line) and the upper (broken/dashed line) and lower (dotted line) similarity bounds (resulting from 65% shifts in the ED(20) and ED(50).

Recall the example and the two cases of interest, where resmethrin was present (added) at 20, 50, 65, and 78.5 percent of the generated candidate mixtures and λ-cyhalothrin was present (added) at 2 and 7 percent of the generated candidate mixtures. Given the chemicals that are added to the generated candidate mixtures and their characteristics (below threshold or beyond threshold) it is decided that the following similarity measures (Table 4.6) will be used to evaluate sufficient similarity in the different chemical/proportion scenarios for the two cases of interest.

Table 4.6. Proportion of resmethrin added to the candidate mixture and the distance measure used to test for sufficient similarity using the similarity measure of $1+\frac{d}{t}$.

Proportion of Resmethrin	Distance Measure
0, 0.2	Unadjusted Adjusted Unweighted
0.50	Unadjusted Adjsuted Unweighted
0.65	Unadjusted Adjusted Unweighted Adjusted Weighted
0.785	Unadjusted Adjusted Unweighted Adjusted Weighted

Table 4.7. Proportion of λ-cyhalothrin added to the candidate mixture and the distance measure used to test for sufficient similarity using the similarity measure of $1+\frac{d}{t}$.

Proportion of λ-cyhalothrin	Distance Measure
0, 0.02	Unadjusted Unadjusted Weighted
0.07	Unadjusted Unadjsuted Weighted

Following the processes described in Sections 4.2.1, 4.2.2, and 4.2.3, 1000 simulation studies were conducted for the generated candidate mixtures of dimension c=5, where under the assumption of additivity eq. (4.3) becomes

$$\mu = \begin{cases} \alpha + \gamma & , t\sum_{i=1}^{c} a_i\beta_i \geq \delta \\ \alpha + \gamma \exp\left\{ t\sum_{i=1}^{c} a_i\beta_i - \delta \right\} & , t\sum_{i=1}^{c} a_i\beta_i < \delta \end{cases}$$

where c=5 and a_i are the mixing ratios from the 1000 simulated candidate mixtures. Values for the slopes for each proportion (β) and the common threshold (δ) were obtained from Wolansky et al (2005) (Table 4.3). Resmethrin was then added at 20, 50, 65, and 78.5 percent of the simulated candidate mixtures and λ-cyhalothrin was added at 2 and 7 percent of the simulated candidate mixtures which implies that c=6 , t is defined in eq. (4.2), and the candidate mixtures

now assume the label full mixture. For each of the six chemical/proportion combinations, 1000 simulation studies were conducted following the steps outlined in Sections 4.2.1, 4.2.2, and 4.2.3. Including the case when the simulated candidate mixtures have only the original five pyrethroids, seven simulations were conducted with 1000 studies in each simulation.

For each unique study in the seven simulations, the similarity measure, h, was calculated; if the measure was contained in the similarity bounds then the simulated candidate and reference mixtures were concluded to be sufficiently similar in dose-response. The gold standard test was also performed: the associated 95% confidence ellipse was calculated and compared to the similarity region. If the ellipse was completely contained in the similarity region then it was concluded that the candidate and reference or full and reduced mixtures were sufficiently similar in dose-response. The conclusion of the "gold standard" test was viewed as the 'truth'. The results from the gold standard test and the proposed test were compared using sensitivity. The sensitivity and specificity for the proposed similarity measure when the reference and candidate mixtures had the same five chemicals was computed and compared to the method of Stork et al. (2008) (Table 4.8). Recall that the method of Stork constructs intervals around the mixing ratios (Table 4.1) and if each mixing ratio of the candidate mixture is contained in their respective intervals then the reference and candidate mixtures are concluded to be sufficiently similar. The unweighted similarity measure (UUD) was used and the sensitivity was determined to be 0.70 whereas the sensitivity for the method provided by Stork et al. (2008) was 0.10 (Table 4.8). The proposed similarity measure out performs the method of Stork et al. (2008) with respect to sensitivity, illustrating an improvement in the proposed similarity measure.

Table 4.8. Sensitivity and specificity estimates (95% confidence intervals) for the method proposed by Stork et al. (2008) and for the proposed method using UUD with data simulated using parameter values in Table 4.3.

Distance Measure	Sensitivity	Specificity
Unadjusted Unweighted	0.7 (0.64,0.76)	0.34 (0.31,0.38)
Stork et al. (2008)	0.1 (0.06,0.13)	1 (1.00,1.00)

Section 4.3.1 Analysis of Mixture with Resmethrin

The sensitivity and specificity for the proposed similarity measures for each of the six chemical/proportion combinations were calculated for the suggested similarity measures (Tables 4.9 and 4.10). When resmethrin was added to the candidate mixtures at 20 percent, using the unadjusted unweighted similarity measure (UUD) the sensitivity was 0.68. When the dose scale was adjusted (AUD) due to the additional chemical in the reference mixture being in a subthreshold range the sensitivity increased to 0.86 demonstrating that adjusting for differences in total dose scales (or subsets) improves the performance of the test. When adding resmethrin to the candidate mixtures at 50 percent using the unadjusted unweighted similarity measure (UUD), the sensitivity was 0.13. Again, adjusting the total dose scale (AUD) demonstrates an increase in sensitivity from 0.13 to 0.93.

Adding resmethrin to the candidate mixtures at 65 percent and using the unadjusted unweighted similarity measure (UUD) the sensitivity was 0. When the dose scale was adjusted (AUD) because the additional chemical in the reference mixture was in a subthreshold range did

not improve sensitivity, however, utilizing the adjusted weighted similarity measure (AWD) with the specified weight matrix

$$W = \begin{pmatrix} 1.01 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.01 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.01 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.01 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.01 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.95 \end{pmatrix} = \begin{pmatrix} \text{w(permethrin)} & 0 & 0 & 0 & 0 & 0 \\ 0 & \text{w(cypermethrin)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \text{w}(\beta\text{-cyfluthrin)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{w(deltamethrin)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \text{w(esfenvelerate)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{w(resmethrin)} \end{pmatrix}$$ where w( ) are the

weights of the given chemicals, sensitivity improved to 0.61. The weights were chosen with the constraint of summing to six (the number of chemicals in the full mixture) and to down weight resmethrin. While the weights are subjective in nature, through practice and expert judgment identifying weights will become less subjective. Adding resmethrin to the candidate mixtures at 78.5 percent and using the unadjusted unweighted similarity measure the sensitivity was 0. Adjusting the dose scale and utilizing the adjusted weighted similarity measure the weight matrix, *W*, specified did not improve the sensitivity. If the proportion of the chemical becomes "too" large it is possible that a sensible weighting scheme will not exist.

Section 4.3.2 Analysis of Mixture with λ-cyhalothrin

When λ-cyhalothrin was added to the candidate mixture at 2 percent and using the unadjusted unweighted similarity measure (UUD) the sensitivity was 0.74. Utilizing the unadjusted weighted similarity measure (UWD) with the specified weight matrix

$$W = \begin{pmatrix} 0.8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.0 \end{pmatrix} = \begin{pmatrix} \text{w(permethrin)} & 0 & 0 & 0 & 0 & 0 \\ 0 & \text{w(cypermethrin)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \text{w}(\beta\text{-cyfluthrin)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{w(deltamethrin)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \text{w(esfenvelerate)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{w}(\lambda\text{-cyhalothrin)} \end{pmatrix}$$ where w( ) are the

weights of the given chemicals, sensitivity improved to 0.89. The weights were chosen with the constraint of summing to six (the number of chemicals in the full mixture) and to up-weight λ-cyhalothrin. Similar to the case with resmethrin the weights are subjective in nature, however, through practice and expert judgment identifying weights will become less subjective. When λ-cyhalothrin is added to the candidate mixture at 7 percent and using the unadjusted unweighted (UUD) similarity measure the sensitivity was 0.67. Utilizing the unadjusted weighted similarity measure (UWD) with the specified weight matrix the sensitivity improved to 0.87.

Table 4.9. Sensitivity and specificity for the proposed similarity measures when resmethrin is added to the simulated candidate mixtures and the associated 95% CI.

Proportion Resmethrin	Distance Measure	Sensitivity	Specificity
0.2	Unadjusted	0.68 (0.61,0.74)	0.26 (0.23,.30)
	Adjusted	0.86 (0.82,0.91)	0.24 (0.20,0.27)
0.5	Unadjusted	0.13 (0.07,0.19)	0.43 (0.40,0.47)
	Adjusted	0.93 (0.89,0.96)	0.16 (0.13,0.19)
0.65	Unadjusted*	0	1
	Adjusted*	0	1
	Adjusted Weighted	0.61 (0.55,0.67)	0.41 (0.38,0.45)
0.785	Unadjusted*	0	1
	Adjusted*	0	1
	Adjusted Weighted*	0	1

Table 4.10. Sensitivity and specificity for the proposed similarity measures when λ-cyhalothrin is added to the simulated candidate mixtures.

Proportion λ-cyhalothrin	Distance Measure	Sensitivity	Specificity
0.02	Unadjusted	0.74 (0.69,0.79)	0.24 (0.21,0.27)
	Unadjusted Weighted	0.89 (0.86,0.93)	0.14 (0.12,0.17)
0.07	Unadjusted	0.67 (0.63,0.72)	0.16 (0.13,0.19)
	Unadjusted Weighted	0.87 (0.84,0.90)	0.10 (0.07,0.12)

Section 4.4 Discussion

It was demonstrated that adding a chemical or even a subset of chemicals that are at or below the threshold and not making the appropriate adjustment to the dose scale can have a noticeable impact on the performance (sensitivity) of the proposed similarity measure. As the proportion of chemical(s) that are at or below their threshold increases in the mixture it is important to adjust the dose scale and to utilize the adjusted unweighted similarity measure. If the proportion is large enough it will become necessary to utilize the adjusted weighted similarity

measure. In some instances, such as in this example, the selected weighting scheme might not be adequate.

It was expected that when adding a more potent chemical (λ-cyhalothrin) to the candidate mixture, even in small amounts, that the gold standard test would fail to conclude sufficient similarity more often than it would conclude sufficient similarity, which occurred in both cases. It was not expected that as the proportion of λ-cyhalothrin added to the candidate mixture increased the number of curves concluded to be sufficiently similar would increase, however, this is what we observed. We believe that this was due to the similarity region that was chosen and as part of the random variation of the simulation.

The simulation studies reveal important properties of the different proposed similarity measures. To evaluate the properties of the proposed test, sensitivity and specificity were used where sensitivity is analogous to the power of the proposed test. While we report both sensitivity and specificity we are more concerned with sensitivity than specificity for the following reason. The measures of sensitivity and specificity are calculated by using the gold standard test as the "truth". Recall that the gold standard test is designed to reject the null hypothesis in favor of the alternative hypothesis of similarity. Failing to reject the null hypothesis that the two dose-response curves are different with respect to the specified similarity region is not the same as rejecting the null hypothesis and concluding that the two curves are different, as in the traditional hypothesis testing framework. This is to say that using the gold standard test as the truth for calculating specificity is not technically correct as we are only failing to conclude similarity and not concluding that the two curves are different. Given the

specified similarity region, failure to conclude similarity using the gold standard test could be due to sample size/power issues or study design. These types of issues are addressed in Chapter 5.

To gain perspective on the performance of the proposed similarity measure, UUD, when the reference and candidate mixtures have the same number of components, we compared its performance to the method described by Stork et al. (2008). The sensitivity for the unadjusted unweighted similarity measure is 0.70 whereas the sensitivity for the method provided by Stork et al. (2008) is 0.10. The proposed similarity measure out performs the method of Stork et al. (2008) with respect to sensitivity, illustrating an improvement in the proposed similarity measure.

Recall that the purpose of the simulation studies was to assess the performance of the proposed similarity measures for the different chemical/proportion situations. Given the original reference mixture of five pyrethroids (Table 4.1) and adding resmethrin to it and making the appropriate adjustments as described in Section 4.2 (creating four candidate mixtures; Table 4.11), the similarity measure was calculated for the different distance measures (Table 4.12). This attempts to summarize how well the measures perform as a whole and to demonstrate that even when a measure indicates sufficient similarity, it does not necessarily indicate the test performs well with respect to sensitivity.

Table 4.11 Reference mixture and the four candidate mixtures obtained from different regions.

Chemical	Reference Mixture(Red)	Candidate Mixture 1 (full)	Candidate Mixture 2 (full)	Candidate Mixture 3 (full)	Candidate Mixture 4 (full)
Permethrin	0.522	0.418	0.261	0.183	0.112
Cypermethrin	0.288	0.230	0.144	0.101	0.062
β-Cyfluthrin	0.129	0.103	0.065	0.045	0.028
Deltamethrin	0.034	0.027	0.017	0.012	0.007
Esfenvalerate	0.027	0.022	0.014	0.009	0.006
Generic Set	0.000	0.200	0.500	0.650	0.785

Table 4.12 Proportion of subthreshold chemicals added to the candidate mixture, distance measure and calculated similarity measure.

Proportion	Distance Measure	Similarity Measure	Sufficiently Similar
0.2	Unadjsuted Unweighted	1.23	YES
	Adjusted Unweighted	1.2	YES
0.5	Unadjsuted Unweighted	1.59	YES
	Adjusted Unweighted	1.5	YES
0.65	Unadjsuted Unweighted	1.76	NO
	Adjusted Unweighted	1.65	YES
	Adjusted Weighted	1.62	YES
0.785	Unadjsuted Unweighted	1.92	NO
	Adjusted Unweighted	1.79	NO
	Adjusted Weighted	1.62	YES

Note: The similarity bounds for h are (0.35, 1.65)

When the proportion of subthreshold chemicals added to the candidate mixture is 20 and 50 percent the unadjusted unweighted distance measure produces a similarity measure that is within the similarity bounds and thus it is concluded that the reference mixture and candidate

mixture are sufficiently similar in dose-response. As the proportion of subthreshold chemicals increases to 65 and 78.5 percent the adjusted weighted distance measure is needed to produce a similarity measure that is within the similarity bounds. We would like to know how often our test reveals sufficient similarity when in fact the reference and candidate dose-response curves are sufficiently similar. This was essentially evaluated through the sensitivity values calculated from the simulation studies. For candidate mixture 1 both distance measures (unadjusted and adjusted) perform well. The unadjusted unweighted distance measure yields a sensitivity of 0.68, while the adjusted unweighted distance measure shows a significant increase in sensitivity to 0.86. In candidate mixture 2, while both distance measures produce similarity measures within the bounds the adjusted unweighted distance measure has a drastically higher sensitivity (sensitivity=0.93) than the unadjusted unweighted distance measure (sensitivity=0.13). For candidate mixture 3 the adjusted weighted distance measure produced a similarity measure that was within the bounds and its associated sensitivity was 0.61. Although in candidate mixture 3 the adjusted unweighted distance measure produced a similarity measure within the bounds, utilizing the specified weighting scheme no measure of sensitivity was able to be obtained.

This example demonstrates that when the specified similarity measures indicate the reference and candidate mixtures are sufficiently similar in dose-response, the associated sensitivity is acceptable. Overall, the selected similarity measures perform well.

Chapter 5: Sufficient Similarity: Addressing the Technical Issues

Section 5.1 Introduction

In Chapters 3 and 4 we have presented a new method that consists of four similarity measures that are a function of Euclidean distance and multiple simulation studies that were conducted for the purpose of assessing how well the proposed similarity measures perform with respect to sensitivity and specificity. While these chapters contain detail on how to both utilize the similarity measure in practice (Chapter 3) and how to conduct simulation studies to assess performance (Chapter 4), many of the technical issues that one may encounter have not been addressed. In this chapter we will address:

- issues that may be present in the original data set, such as departure from additivity, and how to adjust for this problem with respect to simulating data
- problems that may be encountered when attempting to simulate candidate mixture mixing ratios and dose-response data, and
- technical issues that arise when fitting a non-linear mixed effects model will be addressed

For many of these problems there is most likely more than one plausible solution, however, in most instances we will provide suggestions that at the very least will provide a starting point given the problem/issue encountered.

Section 5.2 Technical Issues

Recall the original complex chemical mixture of five pyrethroids (permethrin, cypermethrin, β-cyfluthrin, deltamethrin, esfenvelerate) with complete dose-response data (Figure 3.4) and the following mixing ratios and total dose groups (Tables 3.2 and 3.3) as well as the simulated reference dose-response data set (generated under the assumption of additivity; Figure 3.6) that was used in both Chapters 3 and 4. For the purposes of demonstrating how to deal with certain technical issues, these dose-response data will be used and referred to as the 'reference mixture' or 'reference dose-response data' and 'simulated reference dose-response data'. Similar to Chapters 3 and 4 the similarity region was defined as 65% shifts in either direction for the both the ED(20) and ED(50).

Section 5.2.1 Dealing With the Assumption of Additivity With Respect to Simulations

In both Chapters 3 and 4 all the simulations are conducted under the assumption of additivity. To be able to generate dose-response data that is the result of utilizing individual mixing ratios from a (candidate) mixture, it is necessary to impose the assumption of additivity based on single chemical dose-response curves. Further recall that in Chapters 3 and 4 it is necessary to simulate dose-response data for the generated candidate mixtures that is subsequently compared to the reference dose-response data. Because we are comparing data that were generated under the assumption of additivity to the original reference dose-response data, for the purposes of making an accurate comparison, it is necessary to investigate whether there is departure from additivity in the reference dose-response data. If single chemical data are available on all of the chemicals of concern, the single chemical required (SCR) approach

described by Casey et al. (2004) can be used. In this case, the SCR approach was implemented (data not shown) and indicated that there was evidence of departure from additivity in the reference dose-response data set. If single chemical data are not available on all the chemicals in the reference mixture, but the respective slope estimates are available in the literature, etc., performing the following steps can act as a surrogate for assessing additivity.

1. Using the individual slope estimates for the chemicals and the common variance estimate of the reference dose-response data, simulate 1000 reference mixture dose-response data sets under the assumption of additivity using the original mixing ratios (Table 3.2).

2. Perform the "gold standard" test for sufficient similarity given the specified similarity bounds on the reference dose-response data and the simulated reference dose-response data sets

3. Calculate the percent of the time that the simulated reference mixture data were sufficiently similar to the reference dose-response data.

If additivity is a valid assumption then one would expect that the reference mixture data and simulated reference mixture data would be sufficiently similar a significant percentage of the time. Keep in mind that this is only a rule-of-thumb test in that there is no particular percent for which the researcher is looking. Intuition suggests that if the curves are not sufficiently similar at least 50% of the time (as this is representative of flipping a coin to determine additivity) then additivity is most likely not a suitable assumption. For the examples presented in Chapters 3 and 4, the simulated reference dose-response data were used and therefore any examples used to illustrate any points of interest will also use the simulated reference dose-response data set.

Section 5.2.2 Study Design

Another technical issue often encountered that was not addressed in the preceding chapters is the concept of study design for the reference mixture dose-response data of concern (recall that we are using the simulated reference dose-response data). Given the simulated reference mixture described in this example, it is of interest to determine how often this mixture is sufficiently similar to itself given the specified similarity bounds. If this mixture study were repeated 1000 times, it would be expected that the resulting dose-response curves should be sufficiently similar in dose-response at least 50% of the time if the study was designed well, however, as this is an indication of the power of the gold standard test it may be desired to conclude sufficient similarity at least 80% of the time. This is an integral step in testing for sufficient similarity in the data poor case because if it cannot be concluded that a mixture is sufficiently similar to itself, given the design of the study, then it does not make sense to think that it would be sufficiently similar to another comparable mixture with additional components. In essence the method evaluates the power of the "gold standard" test for sufficient similarity, as described in chapter 2. In order to assess the power of the equivalence test for sufficient similarity in the case when complete data are available, Monte Carlo simulation studies are conducted. In each simulation study, two data sets are created that contain dose-response data for two dose-response curves. The dose-response curves are generated using an appropriate nonlinear dose-response model for additivity

$$\mu = \begin{cases} \alpha + \gamma & t\sum_{i=1}^{c} a_i\beta_i \geq \delta \\ \alpha + \gamma \exp\left\{ t\sum_{i=1}^{c} a_i\beta_i - \delta \right\} & t\sum_{i=1}^{c} a_i\beta_i < \delta \end{cases}. \qquad (5.1)$$

Appropriate parameter and variance estimates are obtained from an existing data set or from a comparable study. For this example the slope and common threshold estimates were obtained from Wolansky et al (2005). For each i=1,…,N simulation studies two curves are generated under the assumed model. For purposes of this example, N=1000 simulation studies were conducted. A normal random deviate, e, where $e \sim N\left(0, \sigma^2\right)$ is added to the predicted values at each dose group, such that each curve has the same common variance, σ^2. The appropriate nonlinear mixed effects dose response model

$$\mu = \begin{cases} \alpha + \gamma & t \le \delta \\ \alpha + \gamma \exp\left(\beta h(t-\delta)\right) & t > \delta \end{cases} \tag{5.2}$$

is fit to the two curves in each of the N studies. The appropriate confidence ellipse is plotted and if the ellipse is contained in the similarity region, then it is concluded that the two curves are sufficiently similar in dose-response with respect to the designated shifts. The number of times that the ellipse generated for each data set is counted. Power is calculated as

$$Power = \frac{\#concluded\ \ sufficiently\ \ similar}{N}.$$

To assess the relationship between power and study design, simulations should be conducted at different combinations of sample size per dose group, number of dose groups, and size of similarity region. For each of the simulation studies conducted, a data set that produces an ellipse contained in the similarity region is given a 1 and a 0 otherwise. This is to say

$$x_i = \begin{cases} 1 & \text{if ellipse contained in similarity region} \\ 0 & \text{if ellipse contained in similarity region} \end{cases}.$$

Let $y = \sum_{i=1}^{N} x_i$,where:

$x_i \sim Bernoulli(1, p)$ with variance, $\sigma^2 = p(1-p)$

$y \sim Binomial(N, p)$ with variance, $\sigma^2 = Np(1-p)$.

Now, $Power = \frac{y}{N}$ and has asymptotic variance $\sigma^2 = \frac{p(1-p)}{N}$ which is obtained using the Delta Method (Casella and Berger, 2002; Agresti, 2002).

The results for the example with the simulated dose-response data are in Table 5.2. (The results for the original five pyrethroid mixture are in Appendix A.5.) The original study design is adequate as the power of the gold standard procedure is 90%. This is to say that 90% of the time the original dose-response study was sufficiently similar to itself. It is demonstrated (Table 5.2) that as sample size and similarity region decrease, power decreases.

Table 5.1. Number of dose groups and the specified dose levels for the proposed mixture studies in table 4.3.

	Dose Group							
	1	2	3	4	5	6	7	8
Mixture	**Dose Level**							
1	0.000	0.275	1.096	2.740	9.042	13.700	18.084	27.400
2	0.000	1.096	13.700	27.400				
3	0.000	9.042	18.084	27.400				

Table 5.2 Power values for the "new" reference mixture data set with 65% and 55% shifts in either direction for the ED(20) and ED(50), respectively.

Power With 65% Shifts in the ED(20) and ED(50)

Mixture	Dose Groups	n/Group	Power	Variance	Standard Error	# not converging/1000
1a	8	12	90.49	0.00009	0.0093	1
1b	8	6	47.50	0.00025	0.0158	0
2a	4	12	42.40	0.00024	0.0156	7
2b	4	6	10.62	0.00009	0.0097	21
3a	4	12	20.30	0.00016	0.0127	0
3b	4	6	9.80	0.00009	0.0094	0

Power With 55% Shifts in the ED(20) and ED(50)

Mixture	Dose Groups	n/Group	Power	Variance	Standard Error	# not converging/1000
1a	8	12	69.47	0.00021	0.0146	1
1b	8	6	22.60	0.00017	0.0132	0
2a	4	12	21.65	0.00017	0.0130	7
2b	4	6	4.29	0.00004	0.0064	21
3a	4	12	14.00	0.00012	0.0110	0
3b	4	6	1.10	0.00001	0.0033	0

Section 5.2.3 Generating Candidate Mixture Mixing Ratios

Perhaps one of the most challenging aspects in the process of evaluating sufficient similarity is coming up with methods to generate possible candidate mixtures (i.e. generating a_i's), which are needed to simulate candidate dose-response data for the purposes of risk assessment as in Chapter 3 or to evaluate the sensitivity of the proposed similarity measure as in Chapter 4. In the best case scenario, as is the case in the example presented in Chapter 3, there would be some form of exposure data available for the chemicals of concern. For the example in Chapter 4, it was assumed that the log transform of the available exposure data followed a multivariate normal distribution for ease of simulation. (See Appendix A.4 for the marginal distributions for the exposure data for the five chemicals in the original pyrethroid study.) Making the assumption of multivariate normality, was the naïve approach, however, in the absence of being able to identify the appropriate multivariate distribution this is a logical choice.

When exposure data are available, another possible option to generate candidate mixtures is to utilize bootstrap methodology. For the observed data presented in Chapter 4, 126 observations could be drawn with replacement and the average proportions could be calculated following the logic described in Section 4.2 of Chapter 4. This process could then be repeated 1000 times to create 1000 possible candidate mixtures. In the event that there are no raw exposure data available and no known distributional information available for the chemicals in the reference mixture, then one might use the Dirichlet distribution, denoted Dir(α) (Wikipedia, 2009). Named after Johann Peter Gustav Lejeune Dirichlet, this is a family of continuous multivariate probability distributions with parameter vector α where $\alpha \geq 0$. The Dirichlet distribution is the multivariate generalization of the beta distribution and is the conjugate prior to the multinomial

distribution (Wikipedia, 2009). There also exists a very close relationship between the Dirichlet distribution and the joint distribution of *K* gamma distributions. Utilizing this relationship, *K* proportions, a_i, can be generated with the constraint that $\sum_{i=1}^{K} a_i = 1$. In order to generate these proportions for the example presented here let *K*=5 (representing the five pyrethorids) and use the following algorithm

1. Draw five independent samples, y_i, from gamma distributions with densities

   $\Gamma(\alpha_i, \beta = 1) = \frac{y_i^{\alpha_i - 1} e^{-y_i}}{\Gamma(\alpha_i)}$ [$(\alpha_1, ..., \alpha_5)$ are from the 5-dimensional Dirichlet distribution].

2. Now set $a_i = \frac{y_i}{\sum_{j=1}^{5} y_i}$

Using this algorithm we can generate five proportions that sum to one and are centered around the proportions in table 5.1. One of the drawbacks to using the Dirichlet distribution or the bootstrap methodology is that both of these approaches, as presented, generate mixtures that are centered around the original mixture in Table 5.1. In centering around the proportions in the reference mixture, these two approaches tend to exclude extreme observations from the population and may thus yield results that are not representative of the entire population. In order to generate more extreme candidate mixtures, the scaling parameter, β, can be utilized in the Dirichlet distribution.

Section 5.2.4 Model Parameterization

Another important issue to consider when conducting the test for sufficient similarity, is the parameterization of the model as this parameterization is used in conducting the "gold standard" test when assessing the performance of the similarity measure. When conducting many simulations, problems that arise from parameterization can become tedious to deal with. Therefore, the parameterization of the model should be chosen with this in mind. In chapter 3 we chose to parameterize the model as in eq. (5.2)

$$\mu = \begin{cases} \alpha + \gamma & t \le \delta \\ \alpha + \gamma \exp\left(\beta h(t-\delta)\right) & t > \delta \end{cases}.$$

We could have just as easily parameterized the model in the following manner

$$\mu = \begin{cases} \alpha + \gamma & t \ge \delta \\ \alpha + \gamma \exp\left(\beta h t - \delta\right) & t < \delta \end{cases} \tag{5.4}$$

However, in our work this parameterization (eq. 5.4) had a tendency to produce variance-covariance matrices with unstable estimates. This parameterization (eq. 5.4) also had a tendency to produce problems with convergence and optimization when running the required simulations. This problem could potentially be the result of identifiability which stems from the structure of the model and the method of parameterization (Seber and Wild, 1989). This identifiability problem is signaled by the information matrix being singular or nearly singular (Seber and Wild, 1989). When deciding on the appropriate non-linear model it is of value to explore the different parameterizations to avoid these issues. Seber and Wild (1989) point out that many issues arising in non-linear models could be due to approximate nonidentifiability, correlated estimates,

and poor precision of estimation in certain directions (Seber and Wild, 1989). According to Seber and Wild (1989) these problems are not clearly distinguished and use the term "ill conditioning" to describe these problems on the grounds that a major symptom of the problem is an ill-conditioned information matrix.

Section 5.3 Summary

The purpose of this chapter is to identify some of the technical issues related to the process of evaluating sufficient similarity. It was demonstrated that assessing the assumption of additivity and the given study design are essential to performing a sound evaluation of sufficient similarity. In some instances information might exist a priori that suggests there is evidence of departure from additivity, however, as long as there are single chemical information available such as parameter estimates for single chemical data, additivity can be assessed through simulation. Being able to make the assumption of additivity is essential to simulating dose-response data based on the mixing ratios, a_i, of the candidate mixtures. If the assumption of additvity is violated or in question in the reference dose-response data set then it does not make sense to compare it to any data set simulated under the assumption of additivity, such as a candidate dose-response data set. To avoid this issue, when there is evidence of departure from additivity in the reference dose-response data set, the reference dose-response data should be simulated under the assumption of additivity. It is essential to mention that in order to simulate any dose-response data under the assumption of additivity, it is necessary to have single chemical slope estimates at the very least. In the absence of single chemical data/slope estimates it will be difficult to evaluate the process.

Being able to generate candidate mixing ratios (a_i's) is a key component to simulating dose-response data and to evaluating the performance of the similarity measure. However, it should be noted that no generated candidate mixtures are needed to evaluate sufficient similarity in dose-response. In the example presented in Chapter 4, raw exposure data were available which provided distributional information, although it was somewhat of a naïve assumption to assume multivariate normality of the log transform of the data. If the user is not comfortable with making this assumption when data are available, the bootstrap method is another option. In the case where there are no data, utilization of the relationship between the gamma and Dirichlet distribution can generate proportions.

Technical issues such as identifiability and the concept of reparameterizing the non-linear model were addressed as well. As Seber and Wild (1989) suggest, many of the problems that arise in non-linear models are not clearly distinguished and use the term "ill conditioning" to describe these problems on the grounds that a major symptom of the problem is an ill-conditioned information matrix.

Chapter 6 Summary and Extensions

Section 6.1 Summary

It is safe to say that the work of Stork et al. (2008) concerning the concept of sufficient similarity in dose-response was truly a significant advance with respect to introducing an empirical approach to the evaluation of sufficient similarity in chemical mixtures in that it utilized empirical dose-response data and not solely exposure data. While Feder et al. (2009) have introduced novel concepts for evaluating sufficient similarity employing accepted multivariate techniques, the proposed method does not make a connection between exposure and risk. The methodology developed by Stork et al. (2008) concerning sufficient similarity in dose-response motivated the methods and research contained in the chapters of this dissertation.

Stork et al. (2008) suggest using mixed model theory, equivalence testing logic, and the principle of confidence region inclusion to test for sufficient similarity in the data rich situation. The data rich situation exists when there are complete dose-response data on both the reference and candidate mixtures. In Chapter 2, an example is presented where there are complete dose-response data on both of the mixtures in the study; however, one mixture contains an additional chemical component, malathion, which is relatively inactive in this study design but constitutes a large portion (82.5%) of the mixture. Essentially at each dose group 82.5% of the mass is inactive. It does not make sense to test to see if these curves are sufficiently similar in dose-

response when the dose scales are so drastically different. This situation along with the work of Casey et al. (2004) motivated adding a dose adjustment factor. In the example presented in Chapter 2, utilizing the dose adjustment factor produced dose scales that were identical and more importantly had an effect on the conclusion of the test for sufficient similarity. When the dose adjustment factor was used it was concluded that the two curves were sufficiently similar in dose-response, and without using this factor, it could not be concluded that the two curves were sufficiently similar in dose-response. It was also determined, in a separate analysis, that there existed an interaction among the chemicals. Even in the presence of a statistically significant interaction, sufficient similarity can still be concluded. The important concepts to garnish from Chapter 2 are that dose scale matters and even in the presence of a statistically significant interaction sufficient similarity can be concluded.

The work of Stork et al. (2008) was the first empirical approach to evaluating sufficient similarity in dose-response. Chapter 3 addresses the major limitation in the method provided by Stork et al. In the method presented by Stork et al. (2008) if the reference and candidate mixture do not contain the same chemicals then the method cannot be applied. In Chapter 3 we suggest an extension to the method that allows for either the reference or candidate mixture to be a subset of the either. Four similarity measures based on Euclidean distance are presented. The associated similarity measures, h, are functions of the proposed distance measures.

The research presented in Chapter 3 proposed computing bounds on the similarity measure/random effect, h , in the data poor situation. If the computed similarity measure falls in the similarity bounds then the reference and candidate mixtures are considered to be sufficiently

similar in dose-response. An example was presented that demonstrates the applicability of the method in risk assessment.

In Chapter 4 simulation studies were conducted to evaluate the performance of the measure in different scenarios. Simulation studies revealed that when properly chosen, the methods perform very well with respect to sensitivity.

Chapter 5 addresses some of the technical issues that are encountered when implementing the methods proposed in Chapter 3. This chapter handles issues regarding departure from additivity, generating possible candidate mixtures, and evaluating study design.

Section 6.2 Extensions

Up to this point in the research, it has been assumed that the additional chemical(s) in the full mixture are present either in an inactive range but present in large amounts, or the additional chemicals are present in an active range but present in negligible amounts. The methods presented in Chapter 3 can handle these two types of situations. Assume that the additional chemicals in the full mixture are a combination of chemicals in active and inactive ranges. As of now, the capabilities of the proposed method to handle this situation have not been explored. Also, the issue of having a mixture that is comprised of different types of chemicals has not been addressed. The possibility of utilizing the weight matrix, W, to deal with these issues needs to be explored further as the full capabilities of the weight matrix are not utilized when its use is constrained to the diagonal elements.

The similarity regions defined in this work were defined as rectangles in two dimensions and hyper-rectangles in higher dimensions. However, the similarity region could be defined in

any geometric shape. The region could account for relationships among the model parameters. For example, the similarity region could require the lower bound of the ED(50) to be greater than the lower bound of the ED(20).

In the simulation studies presented in Chapter 4, regarding the issues of power and sample size, we the users are constrained to the initial study design. For example, if an expert toxicologist determined that 50% shifts in either direction of the ED(20) and ED(50) constituted a biologically significant region of similarity for the dose-response curve of the reference mixture but the associated 95% confidence region for the ED(20) and ED(50) is not contained within the similarity region, it is like saying the mixture is not similar to itself. It does not seem unlikely for a situation like this to arise in practice. Work needs to be done in the area of initial study design with respect to determined acceptable shifts. This is to say that an expert should be able to specify acceptable shifts in say, an ED(20) and ED(50) and the study is designed such that if the study were repeated numerous times, an acceptable number of these studies (e.g. 80% of the studies) would be considered sufficiently similar.

Lastly, use of the proposed similarity measures outside of environmental risk assessment needs to be explored, as this represents a tool for use in data reduction. Applications in the area of monitoring health and developing health indices is a possible interesting application.

List of References

List of References

Agresti, A. (2002). *Categorical Data Analysis*, New York: John Wiley & Sons, Inc.

Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley.

Berger, R. L., and Hsu, J. C. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statistical Sciences* **11**, 283-319.

Bull, R. J., Rice, G., Teuschler, L., Feder, P. (2009). Chemical Measures of Similarity Among Disinfection By-Product Mixtures. *Journal of Toxicology and Environmental Health, Part A* **72:7**, 482-493.

Carter, W. H., Jr., Chinchilli, V. M., Myers, R. H., and Campbell, E. D. (1986). Confidence intervals and an improved ridge analysis of response surfaces. *Technometrics* **28**, 339-346.

Casella, G., and Berger, R. L. (2002). *Statistical Inference*, California: Wadsworth Group.

Casey, M., Gennings, C., Carter, W. H., Jr., Moser, V. C., Simmons, J. E., (2004). Detecting Interaction(s) and assessing the impact of component subsets in a chemical mixture using fixed-ratio mixture ray designs. *Journal of Agricultural, Biological, and Environmental Statistics* **9**, 339-361.

EPA (1986). Guidelines for the health risk assessment of chemical mixtures. *Federal Register* **51**(185), 34014-34025, Washington, D.C.:U.S. Environmental Protection Agency.

EPA (1992). Guidelines for exposure assessment. *Federal Register* **57**(104),22888-22938, Washington, D.C.: Risk Assessment Forumk, U.S. Environmental Protection.

EPA (1996). Food quality protection act (FQPA). http://www.epa.gov/opp00001/regulating/laws/fqpa/.

EPA (1996). Safe drinking water act amendments of 1996 (SDWA): Strengthening protection for America's drinking water. http://www.epa.gov/ogwdw000/sdwa/theme.html.

EPA (2000). Supplementary guidance for conducting health risk assessment of chemical mixtures. Risk Assessment Forum, EPA/630/R-00/002, Washington, D.C.:U.S. Environmental Protection Agency.

EPA (2001). Peer Consultation "Workshop on approaches to polycyclic aromatic hydrocarbon (PAH) health assessment." EPA/635/R-02/005. Available at http://oaspub.epa.gov/eims/eimscomm.getfile?p_download_id=36313 [last accessed: July 16, 2006].

Feder, P. I., Ma, Z. J., Bull, R. J., Teuschler, L. K., Schenck, K. M., Simmons, J. E., Rice, G. (2009). Evaluating Sufficient Similarity for Disinfection By-Product (DBP) Mixtures: Multivariate Statistical Procedures. *Journal of Toxicology and Environmental Health, Part A*,**72:7**, 468 — 481.

Fritz, G. J., Herbarth, O. Asthmatic disease among urban preschoolers: an observational study. *Int. J. Hyg. Environ. Health* **2004**, *207* (1), 23-30.

Gennings, C., Carter, W.H., Jr., Casey, M., Moser, V., Carchman, R., Simmons, J.E. (2004). Analysis of functional effects of a mixture of five pesticides using a ray design, Environmental Toxicology and Pharmacology 18, 115-125.

Johnson, R. A., and Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis*, Prentice Hall.

Moser, V. C, Casey, M., Hamm, A., Carter, W. H., Jr., Simmons, J. E., Gennings, C. (2005). Neurotoxicological and statistical analyses of a mixture of five organophosphorus pesticides using a ray design. *Toxicological Sciences* **86**, 101-115.

Nafstad, P., Jaakkola, J. J., Skrondal, A., Magnus, P. Day care center characteristics and children's respiratory health. *Indoor Air* **2005**, *15* (2), 69-75.

Rao, C. R. (1973). *Linear Statistical Inference and Its Application*, New York: John Wile & Sons, Inc.

Rice, G. E., Teuschler, L. K., Bull, R. J., Simmons, J. E., Feder, P. L. Evaluating the Similarity of Complex Drinking-Water Disinfection By-Product Mixtures: Overview of the Issues. *Journal of Toxicology and Environmental Health, Part A* **72:7**,429 — 436.

Seber, G. A. F., and Wild, C. J. (1989). *Nonlinear Regression*, New York: John Wiley & Sons, Inc.

Seed, J., Brown, R. P., Olin, S. S., and Foran, J. A. (1995). Chemical mixtures: Current risk assessment methodologies and future directions. *Regulatory Toxicology and Pharmacology* **22**, 76-94.

Spjotvoll, E. (1972). Multiple comparison of regression function. *Annals of Mathematical Statistics* **43**, 1076-1078.

Stork, L.G., Gennings, C., Carter, Jr., W.H., Teuschler, L.K., Carney, E. (2008) Empirical evaluation of sufficient similarity in dose-response for environmental risk assessment of chemical mixtures. Journal of Agricultural, Biological, and Environmental Statistics, in press.

Tulve, S. N., Jones, P. A., Nishioka, M. G., Fortmann, R., C., Croghan, C. W, Zhou, J. Y., Fraser, F., Cave, C., Friedman, Warren. (2006). Pesticide Measurements from the First National Environmental Health Survey of Child Care Centers Using a Multi-Residue GC/MS Analysis Method. *Environmental Science and Technology* **40**, 6269-6274.

Vom Saal, Frederick S, Welshons, Wade V. Large effects from small exposures.II The importance of positive controls in low dose research on bisphenol A. *Environmental Research.* **2006;100**, 50-76.

Wikipedia (2009). http://en.wikipedia.org/wiki/Dirichlet_distribution.

Wilson, N. K., Chuang, J. C., Iachan, R.; Lyu, C., Gordon, S. M., Morgan,M.K., Ozkaynak, H., Sheldon, L. S. Design and sampling methodology for a large study of preschool children's aggregate exposures to persistent organic pollutants in their everyday environments. *J. Exposure Anal. Environ. Epidemiol.* **2004**, *14*(3), 260-74.

Wolansky, M. J., Gennings, C., and Crofton, K. M. (2005). Relative potencies for acute affects of pyrethroids on motor function in rats. *Toxicological Sciences*, 1-7.

Wolfinger, R. D., and Lin, X. (1997). Two Taylor-series approximation methods for nonlinear mixed models. *Computational Statistics & Data Analysis* **25**, 465-490.

Appendix A: Additional Chapter Information

Appendix A.2: Additional Tables and Figures for Chapter 2

Table A.2.1. Number with abnormal gait by dose group and Mixture.

Mixture	Total Dose (mg/kg)	Sample Size	Number with Abnormal Gait	Proportion with Abnormal Gait
1	0	14	0	0.00
1	10	12	3	0.25
1	55	12	3	0.25
1	100	12	8	0.67
1	200	12	10	0.83
1	300	12	11	0.92
1	450	12	12	1.00
2	0	8	0	0.00
2	1.75	12	0	0.00
2	9.6	12	2	0.17
2	17.5	12	2	0.17
2	35	12	6	0.50
2	52.5	12	12	1.00
2	78.8	12	11	0.92

Table A.2.2 Composition by Mixture.

Mixture Composition

Mixture	Acephate	Chlorpyrifos	Diazinon	Dimethoate	Malathion
1	0.04	0.031	0.002	0.102	0.825
2	0.229	0.177	0.011	0.583	0

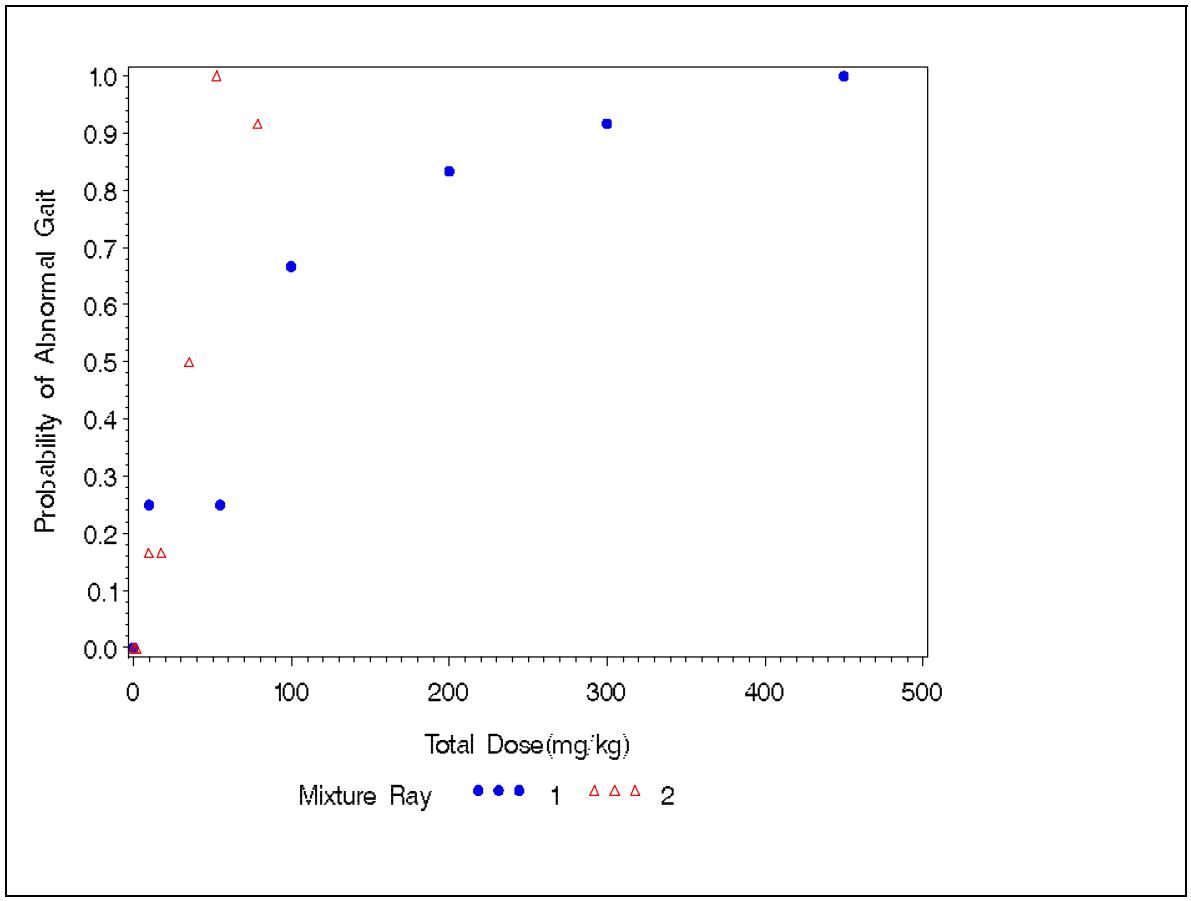


Figure A.2.1. Observed probability of abnormal gait for the associated dose groups in Mixtures 1 and 2.

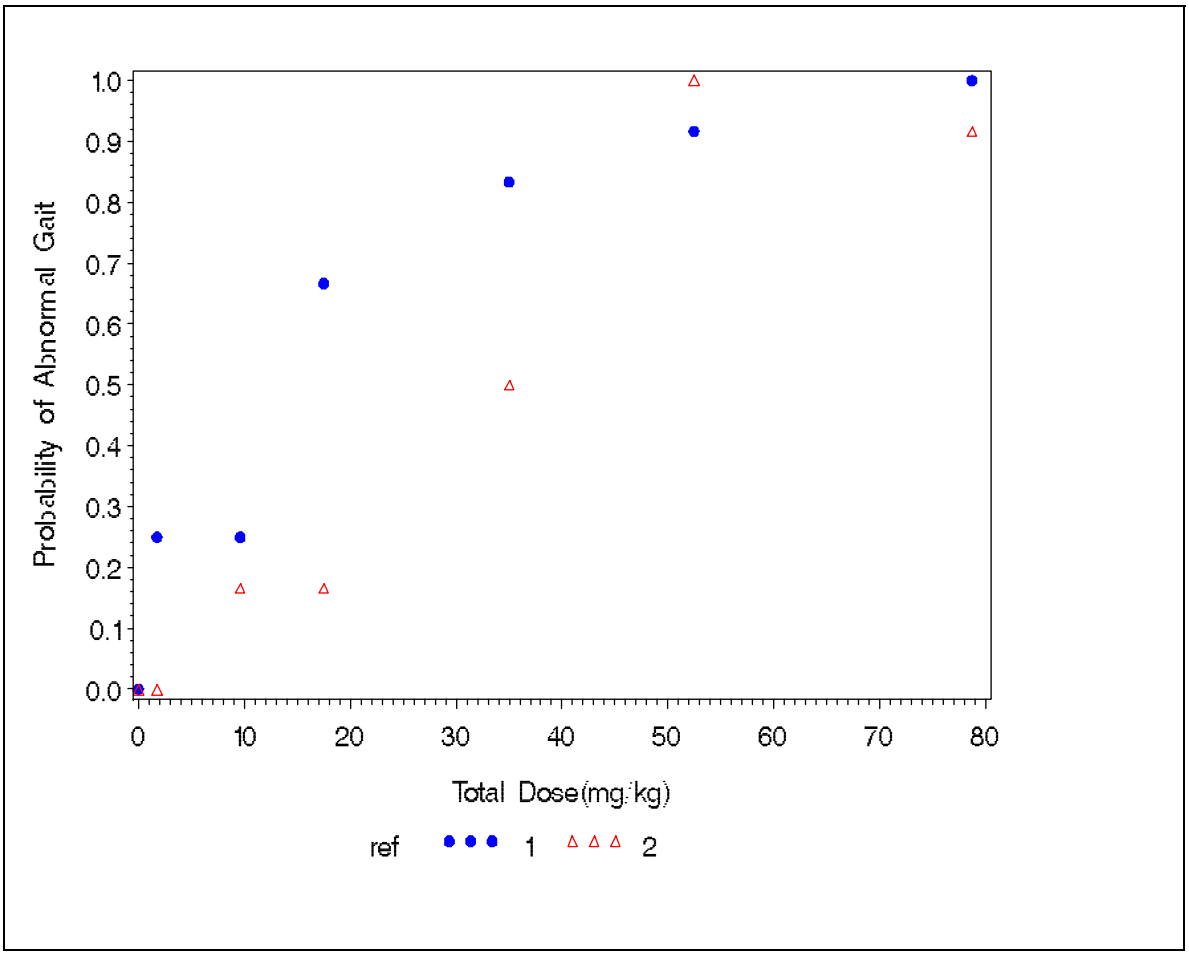


Figure A.2.2. Observed probability of abnormal gait for the associated adjusted (rescaled) dose groups in Mixtures 1 and 2.

Table A.2.3. Estimate of the (mean) random effect for the two mixtures on the original dose scales.

	Estimate: Random Effect(*h*)
Mixture 1	0.45
Mixture 2	1.52

Table A.2.4. Estimate of the (mean) random effect for the two mixtures on the adjusted (rescaled) dose scale.

	Estimate: Random Effect(*h*)
Mixture 1	1.14
Mixture 2	0.84

Transformation to Polar Coordinates

Following the logic of Carter (1986) to calculate the bounds of the confidence region, it is first necessary to identify the points on the boundary *C*. Anderson (1958) gives a transformation from rectangular coordinates to polar coordinates that permits identification of points on the boundary of *C* which expedites this search in *D* dimensions:

Let P be the $D \times D$ orthogonal matrix for which $P'\hat{\Omega}^{-1}P = \Lambda$
where Λ is the diagonal matrix of the eigenvalues of $\hat{\Omega}^{-1}$. Then,

$$\begin{aligned}
\chi^2_{1-\alpha,D} &= \left(\underline{\hat{\omega}} - \underline{\omega}\right)' \hat{\Omega}^{-1} \left(\underline{\hat{\omega}} - \underline{\omega}\right) \\
&= \left(\underline{\hat{\omega}} - \underline{\omega}\right)' PP'\hat{\Omega}^{-1}PP' \left(\underline{\hat{\omega}} - \underline{\omega}\right) \\
&= \left(\underline{\hat{\omega}} - \underline{\omega}\right)' PP'\Lambda PP' \left(\underline{\hat{\omega}} - \underline{\omega}\right) \\
&= \left(\underline{\hat{\omega}} - \underline{\omega}\right)' PP'\Lambda^{1/2}\Lambda^{1/2}PP' \left(\underline{\hat{\omega}} - \underline{\omega}\right) \\
&= zz'.
\end{aligned}$$

The confidence region about ω has been transformed to a *D*-dimensional hyper-sphere of radius $r = (\chi^2_{1-\alpha,D})^{1/2}$. The search for elements on the boundary of *C* can be restricted to this hyper-sphere. Anderson (1958) gives a transformation from rectangular coordinates that expedites this search in *D*-dimensions:

$$\begin{aligned}
z_1 &= r\sin(\psi_1) \\
z_2 &= r\cos(\psi_1)\sin(\psi_2) \\
&\vdots \\
z_{D-1} &= r\cos(\psi_1)\cos(\psi_2)\cdots\cos(\psi_{D-2})\sin(\psi_{D-1}) \\
z_D &= r\cos(\psi_1)\cos(\psi_2)\cdots\cos(\psi_{D-2})\cos(\psi_{D-1}).
\end{aligned}$$

By considering values of r in $0 < r \le (\chi^2_{1-\alpha,D})^{1/2}$ and ψ_i in $-\pi/2 \le \psi_i \le \pi/2$ $(i = 1, 2, \ldots, D-2)$, ψ_{D-1} in $-\pi \le \psi_{D-1} \le \pi$, any number of points of the boundary of C can be determined by $\underline{\omega} = \underline{\hat{\omega}} - P\Lambda^{-1/2} z$. Once the elements of C have been determined in this manner, the confidence region about $\underline{\omega}$ can be found in general by evaluating

$$C = \left\{ \underline{\omega} : \left(\underline{\hat{\omega}} - \underline{\omega} \right)' \hat{\Omega}^{-1} \left(\underline{\hat{\omega}} - \underline{\omega} \right) \le \chi^2_{1-\alpha,D} \right\}.$$

Appendix A.4: Additional Tables and Figures for Chapter 4

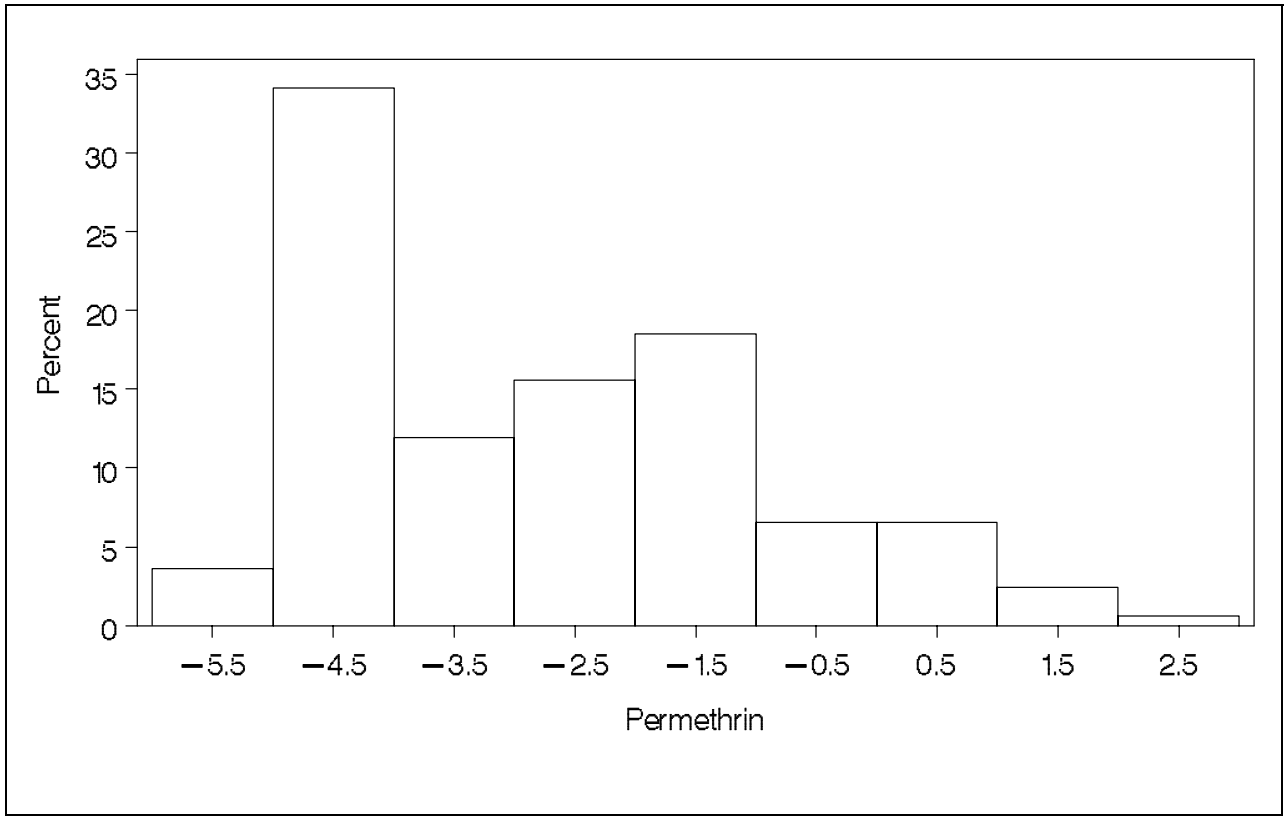


Figure A.4.1. Histogram of the log-transform of total loading for permethrin.

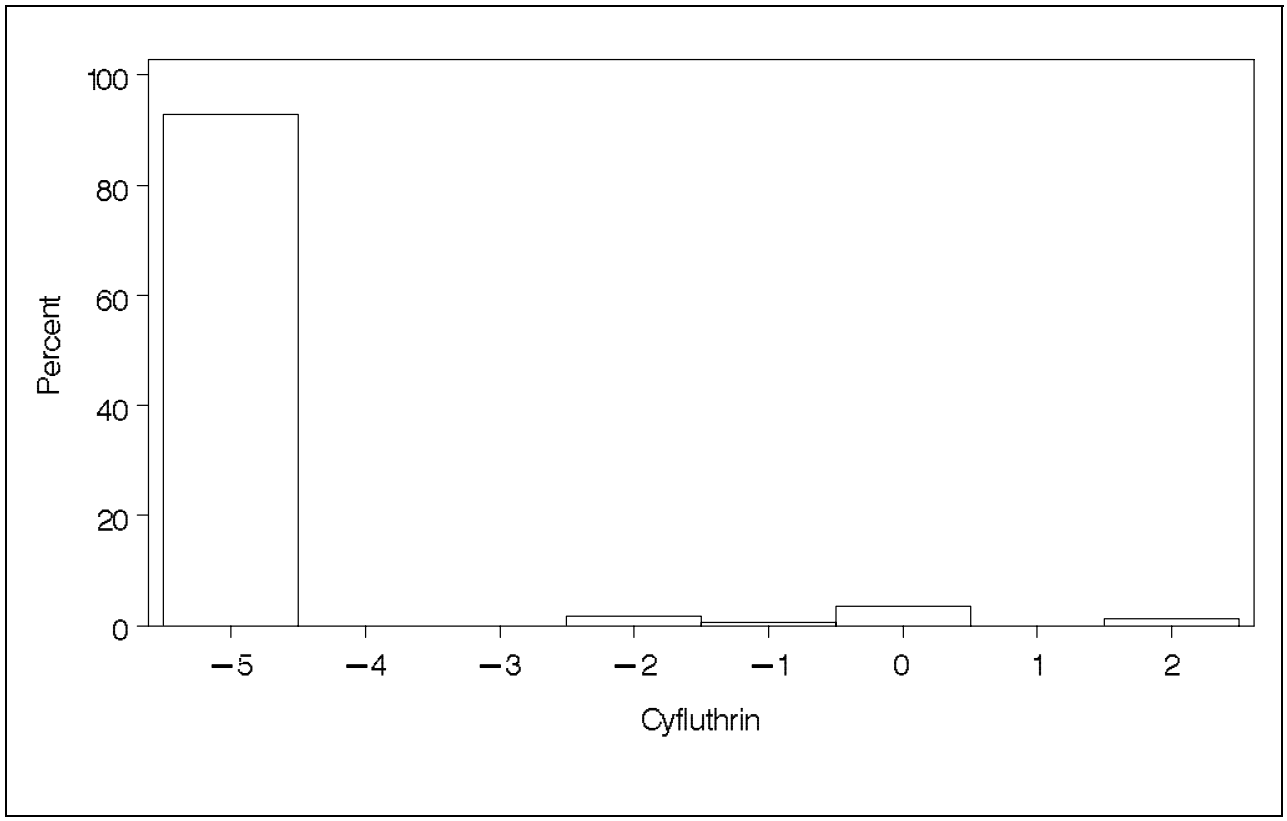


Figure A.4.2. Histogram of the log-transform of total loading for β-cyfluthrin.

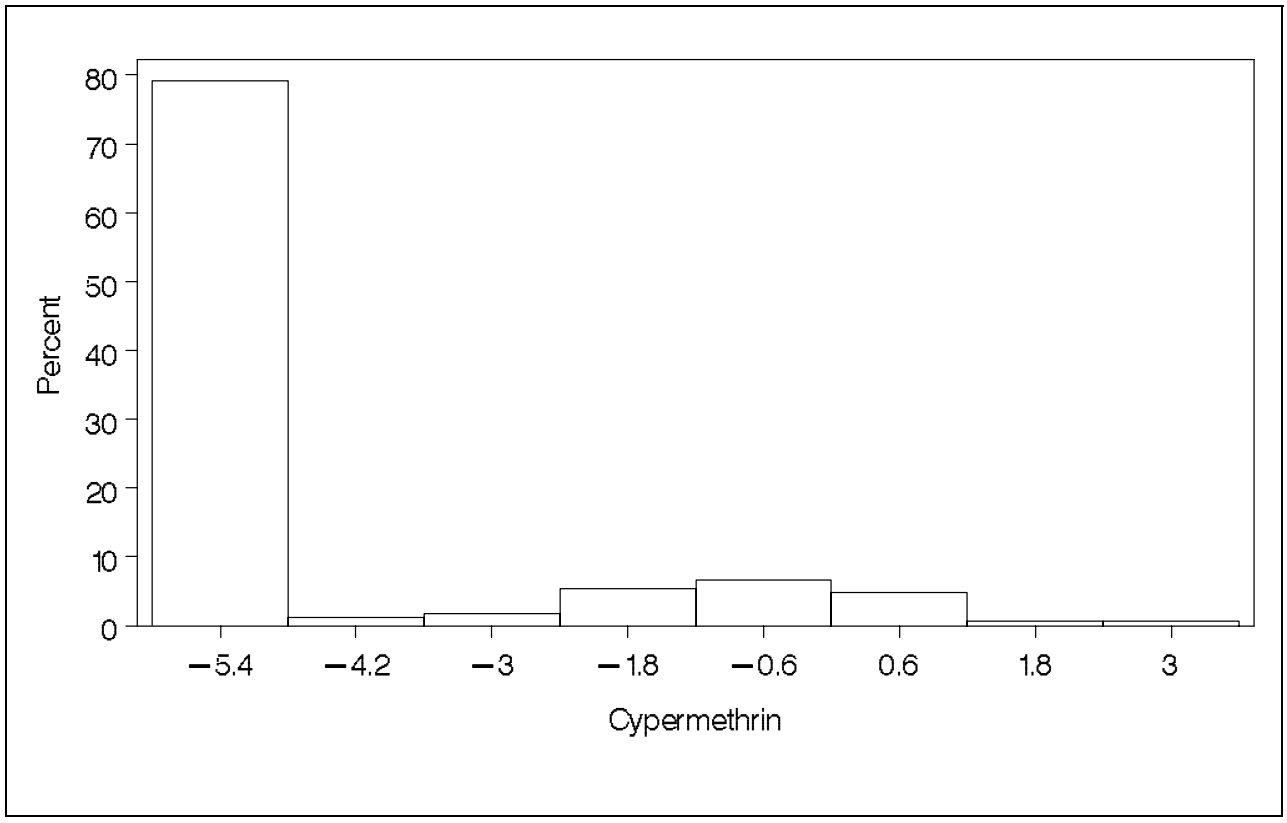


Figure A.4.3. Histogram of the log-transform of total loading for permethrin.

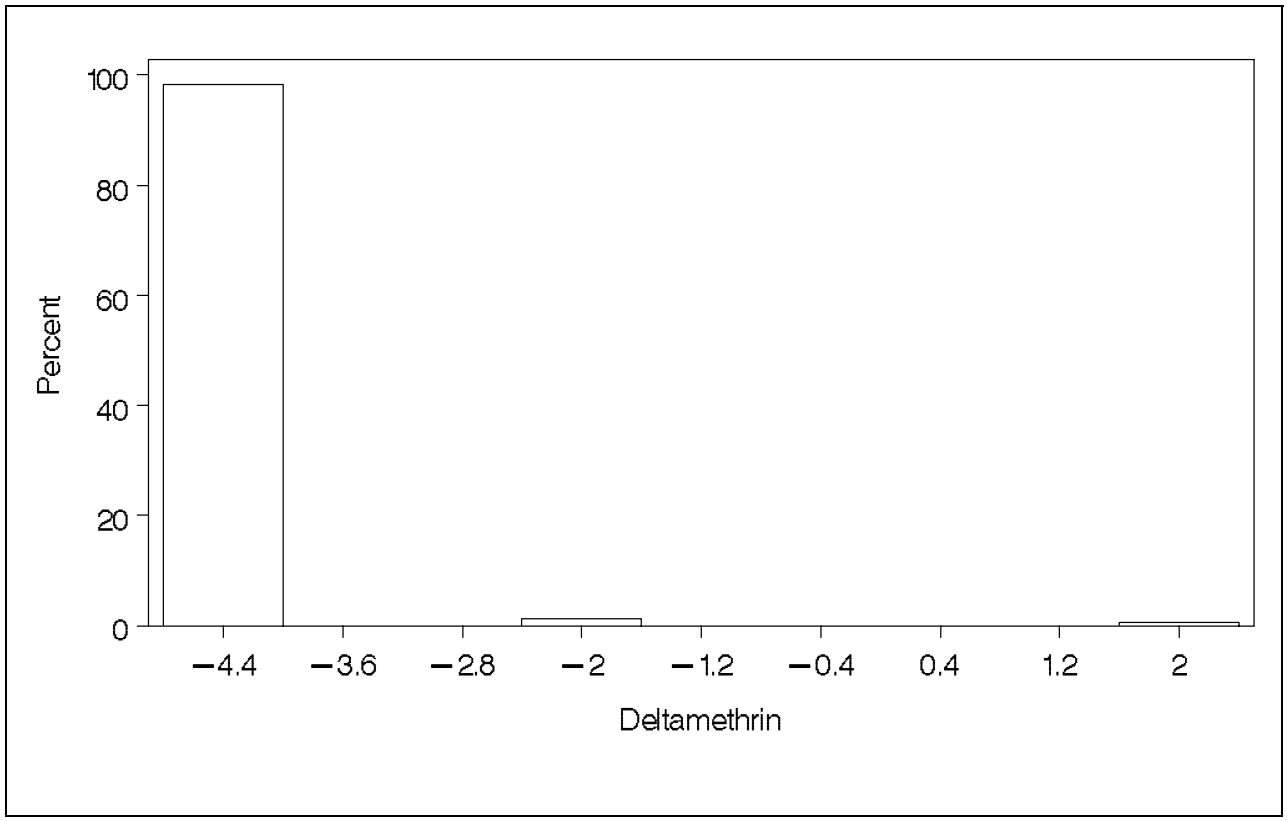


Figure A.4.4. Histogram of the log-transform of total loading for deltamethrin.

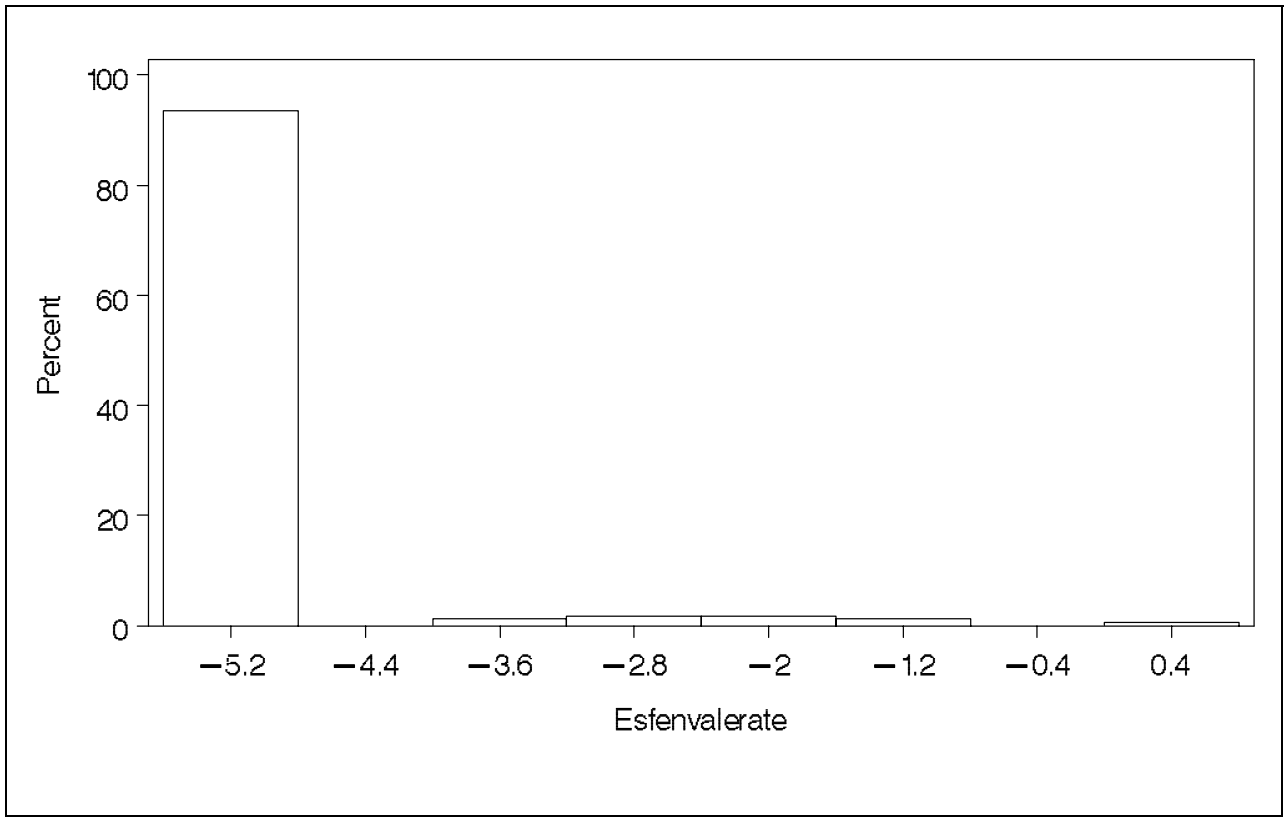


Figure A.4.5. Histogram of the log-transform of total loading for esfenvelerate.

.

Appendix A.5: Additional Tables and Figures for Chapter 5

Table A.5.1. Number of dose groups and the specified dose levels (mg/kg) for the proposed mixture studies in table 4.3.

	Dose Group							
	1	2	3	4	5	6	7	8
Mixture	**Dose Level (mg/kg)**							
1	0.000	0.275	1.096	2.740	9.042	13.700	18.084	27.400
2	0.000	1.096	13.700	24.000				
3	0.000	9.042	18.084	27.400				

Table A.5.2. Power values for the "new" reference mixture data set with 65% and 55% shifts in either direction for the ED(20) and ED(50).

Power With 50% Shifts in the ED(20) and ED(50)

Mixture	Dose Groups	n/Group	Power	Variance	Standard Error	# not converging/1000
1a	8	12	99.8	2E-06	0.001413	0
1b	8	6	91.39	7.87E-05	0.008871	1
2a	4	12	92.25	7.15E-05	0.008455	7
2b	4	6	70.67	0.000207	0.014397	25
3a	4	12	77.8	0.000173	0.013142	0
3b	4	6	43.4	0.000246	0.015673	0

Power With 30% Shifts in the ED(20) and ED(50)

Mixture	Dose Groups	n/Group	Power	Variance	Standard Error	# not converging/1000
1a	8	12	85.4	0.000125	0.011166	0
1b	8	6	51.45	0.00025	0.015805	1
2a	4	12	65.16	0.000227	0.015067	7
2b	4	6	28.21	0.000203	0.014231	25
3a	4	12	21.9	0.000171	0.013078	0
3b	4	6	8.6	7.86E-05	0.008866	0

Appendix B: SAS Code

Appendix B.2: SAS Code for Chapter 2

Appendix B.2.1: SAS Code Used to Rescale Total Dose and Perform Test of Sufficient Similarity

```
goptions colors=(black) htext=1.5 ftext=swiss;
/*This reads in the raw data and creates a
reference cell type coding for the two mixture rays*/
data mix;
set sasuser.response;
if mixture_ray=1 then do;ref=1;total_dose__mg_kg_ =total_dose__mg_kg_*.175;
end;
*if mixture_ray=2 then do;ref=2;*total_dose__mg_kg_
=total_dose__mg_kg_/.175;end;
run;

proc print data=mix;
run;
quit;

data mix1(keep=Sample_Size Number_with_Abnormal_Gait total_dose__mg_kg_
mixture_ray ref proportion_with_abnormal_gait);
set mix;
run;

/*produces the plot of the observed data*/
symbol1 v=dot i=none c=blue;
symbol2 v=triangle i=none c=red;
axis1 label=(a=90 "Probability of Abnormal Gait");
axis2 label=("Total Dose(mg/kg)");
proc gplot data=mix1;
plot proportion_with_abnormal_gait*total_dose__mg_kg_=ref/vaxis=axis1
haxis=axis2;
run;
quit;

data forplot1;
do total_dose__mg_kg_ =0 to 450 by 1;
ref=2;
output;
```

```
end;
run;

data forplot2;
do total_dose__mg_kg_ =0 to 450 by 1;
ref=1;
output;
end;
run;

data forplot;
set forplot1 forplot2;
run;

data mixanal;
set mix1 forplot;
if ref=2 then prop2=proportion_with_abnormal_gait;
if ref=1 then prop1=proportion_with_abnormal_gait;
run;

/*sorts the data to be prepared for analysis*/
proc sort data=mixanal;
by ref;
run;


/*Used to obtain estimates of initial starting values*/
/*
data start_value;
set mix;
if mixture_ray=1;
run;

proc logistic data=start_value;
model Number_with_Abnormal_Gait/Sample_Size=Total_Dose__mg_kg_/link=logit;
run;
quit;
*/

/*Fits the fixed effects model*/
proc nlmixed data=mixanal cov  hess;
parms b0=-1.9 b1=.03  s2u=0.49;
*s2u=su*su;
mu=1/(1+exp(-(b0+b1*(u+1)*Total_Dose__mg_kg_)));
estimate 'ED(20)' (log(.25)-b0)/b1;
estimate 'ED(50)' (log(1)-b0)/b1;
random u ~ normal(0,s2u) subject=ref out=randomeff;
model Number_with_Abnormal_Gait~binomial(sample_size,mu);
predict mu out=pred1;
ods output parameterestimates=vars1 covmatparmest=covs;
run;
quit;

proc sort data=pred1;
by dose ref;
```

```
run;

data pred1;
set pred1;
if ref=2 then do; pred2=pred;end;
if ref=1 then do; pred1=pred;end;
run;

/*produces the plots of observed with predicted overlayed*/

symbol1 v=dot  c=blue;
symbol2 v=triangle  c=red ;
symbol3 v=dot i=none  c=green;
symbol4 v=dot  i=none c=purple;
axis1 label=(a=90 "Probability of Abnormal Gait");
axis2 label=("Total Dose(mg/kg)");
proc gplot data=pred1;
plot (prop1 prop2 pred1 pred2)*Total_Dose__mg_kg_/ overlay legend vaxis=axis1
haxis=axis2;
label prop1='Prob. of Abnormal Gait: Mix. 1';
label prop2='Prob. of Abnormal Gait: Mix. 2';
label pred1='Pred. Prob. Ab. Gait: Mix. 1';
label pred2='Pred. Prob. Ab. Gait: Mix. 2';
run;
quit;


/*produces the confidence ellipse*/
proc iml;
use vars1;
read all var{estimate} where (parameter='b0') into b0;
use vars1;
read all var{estimate} where (parameter='b1') into b1;
g=b0//b1;
use covs;
read all var{b0} where (parameter='b0') into vb0;
use covs;
read all var{b1} where (parameter='b1') into vb1;
use covs;
read all var{b0} where (parameter='b1') into covb0b1;
var_asy=(vb0||covb0b1)//(covb0b1||vb1);
varinv=inv(var_asy);
s=nrow(g);
** transformations to polar coordinates;
call eigen(eval,p,varinv);
lambda=diag(eval);
lambdahalf=root(lambda);
lamhalfinv=inv(lambdahalf);
check=lambdahalf*lambdahalf;
test=p`*varinv*p;
test2=p*p`;
type1 = 0.05;
bign=14;
totp=2;
f=cinv(1-type1,s);*,bign-totp);
```

```
r = (f)**(1/2);
pi=constant('pi');
twopie=2*pi;
znew=j(2,1,0);
gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      znew=znew||z;
      gw=gw||(g-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
label={"gw1" "gw2"};
create gw1 from gw[colname=label];append from gw;
quit;

/*produces the plot of the confidence ellipse*/
symbol1 v=none i=join c=blue;
axis1 label=(a=90);

data plotter;
set gw1;
run;

data plot_adj;
set plotter;
if _n_=1;
run;

data plot_region;
set plotter plot_adj;
run;

proc gplot data=plot_region;
plot gw2*gw1/noframe vaxis=axis1;
label gw1='B0(beta not)';
label gw2='B1(beta one)';
run;
quit;

/*********Confidence Region in terms of ED's*******************/
proc iml;
use vars1;
read all var{estimate} where (parameter='b0') into b0;
use vars1;
read all var{estimate} where (parameter='b1') into b1;
use covs;
read all var{b0} where (parameter='b0') into vb0;
use covs;
read all var{b1} where (parameter='b1') into vb1;
use covs;
read all var{b0} where (parameter='b1') into covb0b1;
var_asy=(vb0||covb0b1)//(covb0b1||vb1);
```

```
omega=var_asy;
mu20=.20 ;
mu50=.50 ;
ed20=(log(mu20/(1-mu20)))/b1 - b0/b1;
ed50=(log(mu50/(1-mu50)))/b1 - b0/b1;
print ed20 ed50;
gomega=ed20//ed50;
big_g=(-1/b1||-log(mu20/(1-mu20))*b1**-2 + b0*b1**-2)//(-1/b1||-log(mu50/(1-
mu50))*b1**-2 + b0*b1**-2);
cov_gomega=big_g*omega*big_g`;
se_gomega = sqrt(vecdiag(cov_gomega));
var=cov_gomega;
s=nrow(gomega);
varinv=inv(var);
** transformations to polar coordinates;
call eigen(eval,p,varinv);
lambda=diag(eval);
lambdahalf=root(lambda);
lamhalfinv=inv(lambdahalf);
check=lambdahalf*lambdahalf;
test=p`*varinv*p;
test2=p*p`;
type1 = 0.05;
bign=14;
totp=2;
f=cinv(1-type1,s);*,bign-totp);
r = (f)**(1/2);
pi=constant('pi');
twopie=2*pi;
znew=j(2,1,0);
gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
label={"gw1" "gw2"};
create gw1_ed from gw[colname=label];append from gw;
quit;

/*produces the plot of the confidence ellipse*/
symbol1 v=none i=join c=red;
axis1 label=(a=90);

data box;
boxgw1=5.7; boxgw2=210.54; output;
boxgw1=108.3; output;
boxgw2=70.18; output;
boxgw1=5.7; output;
boxgw2=210.54; output;
```

```
proc gplot data=box;
plot boxgw2*boxgw1/noframe vaxis=axis1;
label boxgw1='ED20:Similarity Bounds';
label boxgw2='ED50:Similarity Bounds';
run;
quit;

data plotter_ed;
set gw1_ed box;
run;

data plot_adj_ed;
set plotter_ed;
if _n_=1;
run;

data plot_region_ed;
set plotter_ed plot_adj_ed;
run;

symbol1 v=none i=join c=blue;
proc gplot data=plot_region_ed;
plot gw2*gw1/overlay noframe vaxis=axis1;
label gw1='ED20';
label gw2='ED50';
run;
quit;

symbol1 v=none i=join c=blue;
symbol2 v=none i=join c=red;
proc gplot data=plot_region_ed;
plot gw2*gw1 boxgw2*boxgw1/overlay noframe vaxis=axis1;
label gw1='ED20';
label gw2='ED50';
run;
quit;
```

Appendix B.2.2: SAS Code Used Perform Test of Sufficient Similarity on Original Total Dose Scale

```
goptions colors=(black) htext=1.5 ftext=swiss;
/*This reads in the raw data and creates a
reference cell type coding for the two mixture rays*/
data mix;
set sasuser.response;
if mixture_ray=1 then do;ref=1;end;
if mixture_ray=2 then do;ref=2;end;
run;
```

```
proc print data=mix;
run;
quit;

data mix1(keep=Sample_Size Number_with_Abnormal_Gait total_dose__mg_kg_
mixture_ray ref proportion_with_abnormal_gait);
set mix;
run;

/*produces the plot of the observed data*/
symbol1 v=dot i=none c=blue;
symbol2 v=triangle i=none c=red;
axis1 label=(a=90 "Probability of Abnormal Gait");
axis2 label=("Total Dose(mg/kg)");

proc gplot data=mix1;
plot proportion_with_abnormal_gait*total_dose__mg_kg_=ref/legend vaxis=axis1
haxis=axis2;
label ref='Mixture Ray';

run;
quit;

data forplot1;
do total_dose__mg_kg_ =0 to 450 by 1;
ref=2;
output;
end;
run;

data forplot2;
do total_dose__mg_kg_ =0 to 450 by 1;
ref=1;
output;
end;
run;

data forplot;
set forplot1 forplot2;
run;

data mixanal;
set mix1 forplot;
if ref=2 then prop2=proportion_with_abnormal_gait;
if ref=1 then prop1=proportion_with_abnormal_gait;
run;

/*sorts the data to be prepared for analysis*/
proc sort data=mixanal;
by ref;
run;

/*Used to obtain estimates of initial starting values*/
/*
data start_value;
```

```
set mix;
if mixture_ray=1;
run;

proc logistic data=start_value;
model Number_with_Abnormal_Gait/Sample_Size=Total_Dose__mg_kg_/link=logit;
run;
quit;
*/

/*Fits the mixed-effects model*/
proc nlmixed data=mixanal cov  hess;
parms b0=-1.9 b1=.03  s2u=0.49;
*s2u=su*su;
mu=1/(1+exp(-(b0+b1*(u+1)*Total_Dose__mg_kg_)));
estimate 'ED(20)' (log(.25)-b0)/b1;
estimate 'ED(50)' (log(1)-b0)/b1;
random u ~ normal(0,s2u) subject=ref out=randomeff;
model Number_with_Abnormal_Gait~binomial(sample_size,mu);
predict mu out=pred1;
ods output parameterestimates=vars1 covmatparmest=covs;
run;
quit;

proc sort data=pred1;
by ref dose;
run;

data pred1;
set pred1;
if ref=2 then do; pred2=pred;end;
if ref=1 then do; pred1=pred;end;
run;

/*produces the plots of observed with predicted overlayed*/

symbol1 v=dot  c=blue;
symbol2 v=triangle  c=red ;
symbol3 v=dot i=none  c=green;
symbol4 v=dot  i=none c=purple;
axis1 label=(a=90 "Probability of Abnormal Gait");
axis2 label=("Total Dose(mg/kg)");
proc gplot data=pred1;
plot (prop1 prop2 pred1 pred2)*Total_Dose__mg_kg_/ overlay legend vaxis=axis1
haxis=axis2;
label prop1='Prob. of Abnormal Gait: Mix. 1';
label prop2='Prob. of Abnormal Gait: Mix. 2';
label pred1='Pred. Prob. Ab. Gait: Mix. 1';
label pred2='Pred. Prob. Ab. Gait: Mix. 2';
run;
quit;


/*produces the confidence ellipse in terms of the parameters*/
proc iml;
```

```
use vars1;
read all var{estimate} where (parameter='b0') into b0;
use vars1;
read all var{estimate} where (parameter='b1') into b1;
g=b0//b1;
use covs;
read all var{b0} where (parameter='b0') into vb0;
use covs;
read all var{b1} where (parameter='b1') into vb1;
use covs;
read all var{b0} where (parameter='b1') into covb0b1;
var_asy=(vb0||covb0b1)//(covb0b1||vb1);
varinv=inv(var_asy);
s=nrow(g);
** transformations to polar coordinates;
call eigen(eval,p,varinv);
lambda=diag(eval);
lambdahalf=root(lambda);
lamhalfinv=inv(lambdahalf);
check=lambdahalf*lambdahalf;
test=p`*varinv*p;
test2=p*p`;
type1 = 0.05;
bign=14;
totp=2;
f=cinv(1-type1,s);*,bign-totp);
r = (f)**(1/2);
pi=constant('pi');
twopie=2*pi;
znew=j(2,1,0);
gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      znew=znew||z;
      gw=gw||(g-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
label={"gw1" "gw2"};
create gw1 from gw[colname=label];append from gw;
quit;

/*produces the plot of the confidence ellipse*/
symbol1 v=none i=join c=blue;
symbol2 l=1;
symbol3 l=2;
axis1 label=(a=90);

data plotter;
set gw1;
run;

data plot_adj;
```

```
set plotter;
if _n_=1;
run;

data plot_region;
set plotter plot_adj;
run;

proc gplot data=plot_region;
plot gw2*gw1/noframe vaxis=axis1;
label gw1='B0(beta not)';
label gw2='B1(beta one)';
run;
quit;

/*********Confidence Region in terms of ED's*******************/
proc iml;
use vars1;
read all var{estimate} where (parameter='b0') into b0;
use vars1;
read all var{estimate} where (parameter='b1') into b1;
use covs;
read all var{b0} where (parameter='b0') into vb0;
use covs;
read all var{b1} where (parameter='b1') into vb1;
use covs;
read all var{b0} where (parameter='b1') into covb0b1;
var_asy=(vb0||covb0b1)//(covb0b1||vb1);
omega=var_asy;
mu20=.20 ;
mu50=.50 ;
ed20=(log(mu20/(1-mu20)))/b1 - b0/b1;
ed50=(log(mu50/(1-mu50)))/b1 - b0/b1;
print ed20 ed50;
gomega=ed20//ed50;
big_g=(-1/b1||-log(mu20/(1-mu20))*b1**-2 + b0*b1**-2)//(-1/b1||-log(mu50/(1-
mu50))*b1**-2 + b0*b1**-2);
cov_gomega=big_g*omega*big_g`;
se_gomega = sqrt(vecdiag(cov_gomega));
var=cov_gomega;
s=nrow(gomega);
varinv=inv(var);
** transformations to polar coordinates;
call eigen(eval,p,varinv);
lambda=diag(eval);
lambdahalf=root(lambda);
lamhalfinv=inv(lambdahalf);
check=lambdahalf*lambdahalf;
test=p`*varinv*p;
test2=p*p`;
type1 = 0.05;
bign=14;
totp=2;
f=cinv(1-type1,s);*,bign-totp);
r = (f)**(1/2);
```

```
pi=constant('pi');
twopie=2*pi;
znew=j(2,1,0);
gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
label={"gw1" "gw2"};
create gw1_ed from gw[colname=label];append from gw;
quit;

/*produces the plot of the confidence ellipse*/

data box;
boxgw1=2.00; boxgw2=73.97; output;
boxgw1=37.94; output;
boxgw2=24.66; output;
boxgw1=2.00; output;
boxgw2=73.97; output;

symbol1 v=none i=join c=red;
axis1 label=(a=90);
proc gplot data=box;
plot boxgw2*boxgw1/noframe vaxis=axis1;
label boxgw1='ED(20):Similarity Bounds';
label boxgw2='ED(50):Similarity Bounds';
run;
quit;

data plotter_ed;
set gw1_ed box;
run;

data plot_adj_ed;
set plotter_ed;
if _n_=1;
run;

data plot_region_ed;
set plotter_ed plot_adj_ed;
run;

symbol1 v=none i=join c=blue;
proc gplot data=plot_region_ed;
plot gw2*gw1/noframe vaxis=axis1;
label gw1='ED(20)';
label gw2='ED(50)';
run;
quit;
```

```
data box1;
boxgw1=2.00; boxgw2=73.97; output;
boxgw1=37.94; output;
boxgw2=24.66; output;
boxgw1=2.00; output;
boxgw2=73.97; output;

data plotter_ed1;
set gw1_ed box1;
run;

data plot_adj_ed1;
set plotter_ed1;
if _n_=1;
run;

data plot_region_ed1;
set plotter_ed1 plot_adj_ed1;
run;

symbol1 v=none i=join c=blue;
symbol2 v=none i=join c=red;
proc gplot data=plot_region_ed1;
plot gw2*gw1 boxgw2*boxgw1/overlay noframe vaxis=axis1;
label gw1='ED(20)';
label gw2='ED(50)';
run;
quit;

data box1;
boxgw1=-10; boxgw2=110; output;
boxgw1=50; output;
boxgw2=-10; output;
boxgw1=-10; output;
boxgw2=110; output;

axis1 label=(a=90);
proc gplot data=box1;
plot boxgw2*boxgw1/noframe vaxis=axis1;
label boxgw1='ED(20):Similarity Bounds';
label boxgw2='ED(50):Similarity Bounds';
run;
quit;

data plotter_ed_sim;
set gw1_ed box1;
run;

data plot_adj_ed_sim;
set plotter_ed_sim;
if _n_=1;
run;

data plot_region_ed_sim;
```

```
set plotter_ed_sim plot_adj_ed_sim;
run;

symbol1 v=none i=join c=blue;
symbol2 v=none i=join c=red;
axis1 label=(a=90);
proc gplot data=plot_region_ed_sim;
plot gw2*gw1 boxgw2*boxgw1/overlay  vaxis=axis1 noframe;
label gw1='ED(20)';
label gw2='ED(50)';
run;
quit;
```

Appendix B.3: SAS Code for Chapter 3

Appendix B.3.1: SAS Code for Example Part I

```
libname ver 'C:\Sufficient Similarity Research\verification_code';
libname pyr 'C:\Sufficient Similarity Research\verification';
libname perm 'C:\Sufficient Similarity
Research\verification_code\sens_spec_data';
goptions colors=(black) htext=1.8 ftext=swiss;

data fivechem;
set ver.newmix;
ptotal1=ptotal/100;
constant=1;
run;

proc sort data=fivechem;
by dose;
run;

proc print data=fivechem;
run;

proc means data=fivechem;
var ptotal1;
by dose;
run;

/*calculates the benchmark response of interest*/
proc iml;
control_mean=1.00;
constrol_std_dev=0.199;
bench_response=1.00-2*0.199;
print bench_response;
quit;

proc gplot data=fivechem;
plot ptotal*dose;
run;
quit;

proc print data=fivechem;
run;

/*calculates the adjusted total dose*/
proc iml;
    rel_pot={.059 .235 1.136 1 2.092};
   * rel_pot={.059 .235 1.136 1 2.092 0.009}; /*use when doing the weighted
analysis*/
    numchem=ncol(rel_pot);
```

```
*** 3 cases of weights;

*************** relative potencies define weight;
  *  w=numchem#rel_pot`/sum(rel_pot);

*************** unweighted analysis ;
    w= j(numchem,1,1);
*************** adjusted total dose;
  *  w= j(numchem,1,1);

***************  downweighted resmethrin;
  *  w= j(numchem,1,1);
  *  w[6]=.009;
  *   wi=(numchem-w[6])/(numchem-1);
  *   w=wi//wi//wi//wi//wi//w[6];
******;

      sum_w=sum(w);
      print w sum_w;
     a={0.522 0.288 0.129 0.034 0.027}; ** relative proportions in reference
mixture;
       *a={0.28188      0.15552     0.06966     0.01836     0.01458 0}; **
relative proportion in candidate mixture;

     aw=a*w || 1;
    label={"aw" "constant"};
    print aw;
    create aw from aw[colname=label]; append from aw;
quit;

data fivechem;
   merge aw fivechem ; by constant;
   drop constant;

data forplot;
set aw; drop constant;
do dose=0 to 30 by 0.5;output;end;
run;

data fivechem_anal;
set fivechem forplot;
adose=aw*dose;
run;

/*this is the fixed effects model using adjusted total dose*/
data initparms;
   set aw;
   keep b del s2;
   b=-0.07/aw;
   del=10*aw;
   *s2=650;
   s2=.0650;

/*
proc print data=fivechem;
```

```
run;
*/


/*
data fivechem_start;
set fivechem;
new_ptotal=log((ptotal-20)/(100-20));
run;

proc reg data=fivechem_start;
model new_ptotal=dose;
run;
*/


proc nlin data=fivechem_anal ;
parms b=-0.07 del=10;
a=20;
g=100-a;
mu = a+g*exp(b*(adose-del)*(adose>del));
y=ptotal;
model y=mu;
output out=pred p=pred;
run;
quit;

/*fits the fixed effects non-linear exponential threshold model and
calculates
BMD's and other estimates of interest*/
proc nlmixed data=fivechem_anal cov;
    parms /data=initparms;/* b=-.07 del=10 s2=650;*/
    a=.25;
    g=1-a;
    mu = a+g*exp(b*(adose-del)*(adose>del));
    model ptotal1 ~ normal(mu,s2);
    mu20= a + g*0.8; *85.3;
    mu50= a + g*0.5; *63.3;
      mu_bench= a + g*0.60;
    ed20=((log((mu20-a)/g))/b)+del;
    ed50=((log((mu50-a)/g))/b)+del;
      ed_bench=((log((mu_bench-a)/g))/b)+del;

    estimate "ED20" ((log((mu20-a)/g))/b)+del   alpha=0.001;
    estimate "ED50" ((log((mu50-a)/g))/b)+del   alpha=0.001;
      estimate "ED_Bench" ((log((mu_bench-a)/g))/b)+del   alpha=0.10;



    ED20low=0.5*ed20;  **********sets bounds based on percentages
of ED20 amd ED50 estimates;
    ED20high=1.5*ed20;
    ED50low=.5*ed50;
    ED50high=1.5*ed50;
```

```
    /*
    ED20low=0.70*ed20;  ********************sets bounds based on percentages
of ED20 amd ED50 estimates;
    ED20high=1.30*ed20;
    ED50low=.70*ed50;
    ED50high=1.30*ed50;
    */

    logmu20=log((mu20-a)/g);
    logmu50=log((mu50-a)/g);
    quot=logmu50/logmu20;
    denom_l=ed20low-(ed50low-quot*ed20low)/(1-quot);
    blow=logmu20/denom_l;
    dellow=(ed50low-quot*ed20low)/(1-quot);
    denom_h=ed20high-(ed50high-quot*ed20high)/(1-quot);
    bhigh=logmu20/denom_h;
    delhigh=(ed50high-quot*ed20high)/(1-quot);
    mulow=a+g*exp(blow*(dose-dellow)*(dose>dellow));
    muhigh=a+g*exp(bhigh*(dose-delhigh)*(dose>delhigh));
    id quot mulow muhigh ed20low ed20high ed50low ed50high;

    predict  mu out=pred;
    ods output parameterestimates=vars covmatparmest=covs;
*proc print data=pred;
run;
quit;

/*calculates the individual benchmark doses*/
proc iml;
mix_ratios={0.522 0.288 0.129 0.034 0.027};
bench_dose=14.35;
ind_bench_dose=bench_dose*mix_ratios;
print ind_bench_dose;
quit;


*******************************************
      * Compute minimum and maximum mixing ratios, *
      * ai_l (5.9.2) and ai_u (5.9.3)              *
      *********************************************;
proc iml;
      a={0.522, 0.288, 0.129, 0.034, 0.027};
      print a;
      c=nrow(a);  *c=number of chemicals*;
      print c;
      ai_l=j(c,1,0);
      ai_u=j(c,1,0);
      i_c=j(c,1,0);
      do i=1 to c;
          ai=a[i,1];
            i_c[i,1]=i;
            *ai_l[i,1]=(ai*hlf)/(ai*hlf+(1-ai)*huf);
            *ai_u[i,1]=(ai*huf)/(ai*huf+(1-ai)*hlf);
            ai_l[i,1]=(ai*0.35)/(ai*0.35+(1-ai)*1.65);
            ai_u[i,1]=(ai*1.65)/(ai*1.65+(1-ai)*0.35);
```

```
      end;
      ** Candidate mixing ratios for Cl mixture **;
      *cand_a={0.03, 0.30, 0.26, 0.06, 0.03, 0.02, 0.12, 0.04, 0.14};
      *ais = i_c ||a||ai_l||ai_u||cand_a;
      ais = i_c ||a||ai_l||ai_u;
      *aislab = {"Chemical_i" "a" "a_low" "a_up" "cand_a"};
      aislab = {"Chemical_i" "a" "a_low" "a_up"};
      print ais[colname=aislab];
      create ais from ais[colname=aislab];
      append from ais;
quit;

proc print data=ais;
run;
```

Appendix B.3.2: SAS Code for Example Part II

```
libname ver 'C:\Sufficient Similarity Research\verification_code';
libname pyr 'C:\Sufficient Similarity Research\verification';
libname perm 'C:\Sufficient Similarity
Research\verification_code\sens_spec_data';
goptions colors=(black) htext=1.8 ftext=swiss;

/*creates the dose groups*/
data dose;
  input dose;
  constant=0;
  cards;
0
0.275
1.096
2.740
9.042
13.70
18.084
27.400
;
run;
quit;

/*creates the reference data set under the assumption of additivity*/
%macro refdata;
      %do _i_=1 %to 1;

data ref_ais;
      constant=0;
      a1=0.522;
      a2=0.288;
      a3=0.129;
      a4=0.034;
      a5=0.027;
```

```
run;

data ref_data;
 merge ref_ais dose;
 by constant;
    seed=102679;
      aa=0.2521;
      b1= -0.0139;
      b2= -0.0554;
      b3= -0.2686;
      b4= -0.2364;
      b5= -0.4959;
      del= -0.2359;
      drop b1 b2 b3 b4 b5 constant;
      ba=b1*a1+b2*a2+b3*a3+b4*a4+b5*a5;
      put ba;
      curve=1;
            do k=1 to 12;
           term=ba*dose;
               mu = aa+(1-aa)*exp((term-del)*(term<del));
             pact=mu+sqrt(0.0648)*rannor(seed);
               if pact<0 then do;pact=mu;end;
               output;
            end;
run;

%end;
%mend refdata;

%refdata;

/*calculates the benchmark response*/
proc means data=ref_data;
by dose;
var pact;
run;

proc iml;
control_mean=1;
bench_mark_response=1-2*0.22;
print bench_mark_response;
quit;

/* this is fitting the non-linear exponential threshold model to the
generated
   reference data set and plotting the ellipse with random effect variance of
0
   and with random effect variance increasing, as well as calculating the hl
and hu*/
ods trace off;
proc nlmixed data=ref_data cov ecov;
    parms b=-.02 to -.10 by -.02 delta=1 to 5 by 1 s2=0.0648;
a= 0.25;
g=1-a;
    y = a+g*exp(b*(dose-delta)*(dose>delta));
```

```
    model pact ~ normal(y,s2);
    mu20= a + g*0.8; *85.3;
    mu50= a + g*0.5; *63.3;
      mu_bench= a + g*0.56;
    ed20=((log((mu20-a)/g))/b)+delta;
    ed50=((log((mu50-a)/g))/b)+delta;
      ed_bench=((log((mu_bench-a)/g))/b)+delta;

    estimate "ED20" ((log((mu20-a)/g))/b)+delta   alpha=0.001;
    estimate "ED50" ((log((mu50-a)/g))/b)+delta   alpha=0.001;
      estimate "ED_Bench" ((log((mu_bench-a)/g))/b)+delta   alpha=0.10;


     ED20low=0.35*ed20;  **********sets bounds based on
percentages of ED20 amd ED50 estimates;
    ED20high=1.65*ed20;
    ED50low=.35*ed50;
    ED50high=1.65*ed50;

    /*
    ED20low=0.70*ed20;  **********sets bounds based on percentages
of ED20 amd ED50 estimates;
    ED20high=1.30*ed20;
    ED50low=.70*ed50;
    ED50high=1.30*ed50;
    */

    logmu20=log((mu20-a)/g);
    logmu50=log((mu50-a)/g);
    quot=logmu50/logmu20;
    denom_l=ed20low-(ed50low-quot*ed20low)/(1-quot);
    blow=logmu20/denom_l;
    dellow=(ed50low-quot*ed20low)/(1-quot);
    denom_h=ed20high-(ed50high-quot*ed20high)/(1-quot);
    bhigh=logmu20/denom_h;
    delhigh=(ed50high-quot*ed20high)/(1-quot);
    mulow=a+g*exp(blow*(dose-dellow)*(dose>dellow));
    muhigh=a+g*exp(bhigh*(dose-delhigh)*(dose>delhigh));
    id quot mulow muhigh ed20low ed20high ed50low ed50high;

    predict  y out=pred;
    ods output parameterestimates=vars covmatparmest=covs
CovMatAddEst=add_cov;

run;
quit;

data pyr;
set pyr.pyr_data;
run;

proc sort data=pyr;
by descending total_loading ;
run;
quit;
```

```
/*calculates the mixing ratios*/
data ratios;
set pyr;
a1=cis_trans_p/total_loading;a2=cyperme/total_loading;a3=cyfluth/total_loadin
g;a4=delta_t/total_loading;a5=esfenva/total_loading;
check=a1+a2+a3+a4+a5;
constant=0;
run;

/*selects the top 20 candidate mixtures based on total loading*/
data top_ratios;
set ratios;
if _n_<=20;
run;

data id;
constant=0;
do id=1 to 20;
output;end;
run;

data ais_final;
merge top_ratios id;
by constant;
run;

proc print data=ais_final;
run;

/*creates 20 data sets each having a unique candidate mixture*/
%macro ais;
%do _i_=1 %to 20;

data ais_&_i_;
set ais_final;
where id=&_i_;
constant=0;
run;

%end;

%mend ais;

%ais;

data dist;
run;

/*calculates the distance between the candidate mixtures and the simulated
reference mixture*/
%macro distance;
%do _i_=1 %to 20;

proc iml;
```

```
use ais_&_i_;
read all var {a1} into a1;
read all var {a2} into a2;
read all var {a3} into a3;
read all var {a4} into a4;
read all var {a5} into a5;
read all var {total_loading} into total_loading;

a_ref=0.522//0.288//0.129//0.034//0.027;
a_cand=a1//a2//a3//a4//a5;

diffsq=(a_ref-a_cand)`*(a_ref-a_cand);
dist=1+sqrt(diffsq);

in=(dist>=0.35)*(dist<=1.65);

if in=1 then do;pow=1;end;
if in=0 then do;pow=0;end;

rank=&_i_;

distances=rank||a1||a2||a3||a4||a5||pow||dist||total_loading;

label={"id" "a1" "a2" "a3" "a4" "a5" "pow" "distance" "total loading"};
create distance from distances[colname=label];append from distances;

data dist;
set dist distance;
run;

%end;

%mend distance;

%distance;

data dist;
set dist;
if id=. then delete;
run;

proc sort data=dist;
by distance;
run;

proc print data=dist;
run;

symbol1 v=dot i=none;
axis1 label=(a=90 "Total Loading(ng/cm^2)");
axis2 label=("Similarity Measure");
proc gplot data =dist;
plot total_loading*distance/vaxis=axis1 haxis=axis2;
run;
quit;
```

```
/*simulates 1000 data sets for the selected candidate mixture and creates
1000
studies*/
%macro cand_data;
      %do _i_=1 %to 1000;

data cand_ais;
      constant=0;
      a1=0.995;
      a2=0.0009;
      a3=0.0009;
      a4=0.002;
      a5=0.0012;
run;

data cand_generate&_i_;
 merge cand_ais dose;
 by constant;
    seed=102679+&_i_;
      aa=0.2521;
      b1= -0.0139;
      b2= -0.0554;
      b3= -0.2686;
      b4= -0.2364;
      b5= -0.4959;
      del= -0.2359;
      drop b1 b2 b3 b4 b5 constant;
      ba=b1*a1+b2*a2+b3*a3+b4*a4+b5*a5;
      put ba;
      curve=2;
            do k=1 to 12;
           term=ba*dose;
               mu = aa+(1-aa)*exp((term-del)*(term<del));
             pact=mu+sqrt(0.0648)*rannor(seed);
               if pact<0 then do;pact=mu;end;
               output;
            end;
run;

data anal&_i_;
set ref_data cand_generate&_i_ ;
run;
quit;

data anal&_i_;
set anal&_i_;
if curve=2 then do; pact2=pact;end;
if curve=1 then do; pact1=pact;end;
run;

%end;
%mend cand_data;
```

```
%cand_data;

data yesno;
run;

data yesno1;
run;

data all;
run;

data all1;
run;

data powers;
run;

data outers;
run;

data benchmark_doses;
run;

data benchmark_doses_lower;
run;

data ind_benchmark_doses;
run;

data ind_benchmark_doses_lower;
run;

/*performs the gold standard test for sufficient similarity for the 1000
simulated
candidate mixtures and the simulated reference mixture; calculates the
benchmark dose as well as the lower 95% CI for the benchmark dose*/
%macro analysis;
%do _i_=1 %to 1000;

*footnote 'sample='&_i_;

proc sort data=anal&_i_;
by curve dose;
run;
quit;

proc nlmixed data=anal&_i_ cov hess tech=trureg method=firo;
parms b=0 to -.16 by -.02 delta=1 to 2 by .05 s2=0.0648 su=0.0001 to .1 by
.05;
a= 0.25;
g=1-a;
s2u=su*su;
    *y = a_term+g*exp(b*(u+1)*(dose-delta)*(dose>delta));
      y = a+g*exp(b*(1+u)*(dose-delta)*(dose>delta));
      estimate 'ed20' (log(.8)+b*delta)/(b);
```

```
    estimate 'ed50' (log(.5)+b*delta)/(b);
    random u ~ normal(0,s2u) subject=curve out=randomest&_i_;
    model pact ~ normal(y,s2);
    predict y out=pred&_i_;
    ods output parameterestimates=vars covmatparmest=covs
CovMatAddEst=add_cov;
run;
quit;

data covs;
set covs;
where parameter='b' or parameter='delta';
run;

data add_cov;
set add_cov;
where Label='ED_Bench';
run;

data noest;
set vars;
if df=. then do;
flag=1;end;
if df=1 then do;
flag=0; end;
run;

/*produces box and confidence region*/
proc iml;
  use vars;
      read all var{estimate} where (parameter='b') into beta;
      read all var{estimate} where (parameter='delta') into delta;
      read all var{estimate} where (parameter='s2') into mse;
      read all var{estimate} where (parameter='s2u') into sigmah;

use covs;
      read all var{b delta} into covs;

      aa=.25; * 25.21;
      g=1-aa;
      mu20=0.8#g+aa;
      mu50=0.5#g+aa;
      ED20=(log((mu20-aa)/g)/beta)+delta;
      ED50=(log((mu50-aa)/g)/beta)+delta;
      *print ed20 ed50;
      gomega=ed20//ed50;
      big_g=(-(beta**-2)*log((mu20-aa)/g)||1)//(-(beta**-2)*log((mu50-
aa)/g)||1);
      cov_gomega=big_g*covs*big_g`;
*     print covs cov_gomega;
      varinv=inv(cov_gomega);
      se_gomega = sqrt(vecdiag(cov_gomega));
      *print  cov_gomega varinv se_gomega;

** transformations to polar coordinates;
```

```
      call eigen(eval,p,varinv);
      lambda=diag(eval);
      lambdahalf=root(lambda);
      lamhalfinv=inv(lambdahalf);
      check=lambdahalf*lambdahalf;
      *print lambda p;
*     print lambda check;  ** should equal lambda;
      test=p`*varinv*p;  * test should equal lambda;
*     print lambda test;
      test2=p*p`;
*     print test2;  *test2 shoould equal the identity matrix;
      type1 = 0.05;
      bign=192;
      totp=2;
      s=nrow(gomega);
      f=finv(1-type1,s,bign-totp);
      r = (s*f)**(1/2);
      pi=constant('pi');
      twopie=2*pi;

      znew=j(2,1,0);
      gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      *print z;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
      label={"gw1" "gw2"};
      create gw1 from gw[colname=label];append from gw;
*     labelz={"z1" "z2"};
*     znew=znew`;
*     create znew from znew[colname=labelz];
*    append from znew;
quit;

data box;
      boxgw1=1.35; boxgw2=2.82; output;
      boxgw1=6.35;output;
      boxgw2=13.32;output;
      boxgw1=1.35;output;
      boxgw2=2.82;output;

data plotrefa;
      set gw1 box;

data plotref1a;
      set plotrefa;
      if _n_=1;
      run;
```

```
data plotref2a;
      set plotrefa plotref1a;
run;

proc means data=plotref2a max min noprint;
      var gw1 gw2;
      output out=maxmin max= max1 max2 min= min1 min2;
run;

data yesno;
      set maxmin;
      id=&_i_;
      yesno=0;
      if min1>1.35 and max1<6.35 and min2>2.82 and max2<13.32 then yesno=1;
      if min1=0 and max1=0 and min2=0 and max2=0 then noest=1;
run;

data all;
set all yesno;
drop _type_ _freq_;
if yesno=. then delete;
run;

data bmd_std;
set cand_generate&_i_;
if dose=0;
run;

proc means data=bmd_std;
var pact;
output out=stand;
run;

proc iml;
use cand_ais;
read all var {a1} into a1_c;
read all var {a2} into a2_c;
read all var {a3} into a3_c;
read all var {a4} into a4_c;
read all var {a5} into a5_c;

use noest;
read all var {flag} into flagger;

use stand;
read all var{pact}where (_STAT_='STD')into std;

use vars;
read all var{estimate} where (parameter='b') into beta;

use vars;
read all var{estimate} where (parameter='delta') into delta;

use add_cov;
read all var{Cov3} into bench_var;
```

```
if flagger=1 then do; flag=1;end;
if flagger=0 then do; flag=0;end;

bench_std=sqrt(bench_var);


a_ref=0.522//0.288//0.129//0.034//0.027;
a_cand=a1_c//a2_c//a3_c//a4_c//a5_c;

diffsq=(a_ref-a_cand)`*(a_ref-a_cand);
dist=1+sqrt(diffsq);

id=&_i_;

if flag=1 then do;out=id||flag;end;
if flag=0 then do;out=id||flag;end;

in=(dist>0.35)*(dist<1.65)*(flag=0);

if in=1 then do;pow=1;end;
if in=0 then do;pow=0;end;

power=id||pow||dist||flag;

label={"id" "yesno" "distance" "nopt"};
create pows from power[colname=label];append from power;

create outs from out;append from out;

a=0.25;
g=1-a;
control_mean=1.00;
mu_bench_response=1.00-2*std;
mu_bench=a+g*mu_bench_response;
bmd=(log((mu_bench-a)/g)/beta)+delta;
bench_dose=id||bmd;
*print bench_dose;

z=1.645;
lower_bmd=bmd-z*bench_std;
*print lower_bmd;
bench_dose_lower=id||lower_bmd;
*print bench_dose_lower;

ind_bmd=bmd*a_cand`;
ind_bench_dose=id||ind_bmd;
ind_bmd_lower=lower_bmd*a_cand`;
ind_bench_dose_lower=id||ind_bmd_lower;

label={"id" "benchmark_dose"};
create bench_doses from bench_dose[colname=label];append from bench_dose;

label={"id" "lower_benchmark_dose"};
```

```
create bench_doses_lower from bench_dose_lower[colname=label];append from
bench_dose_lower;


label={"id" "permethrin" "cypermethrin" "betacyfluthrin" "deltamethrin"
"esfenvalerate"};
create ind_bench_doses from ind_bench_dose[colname=label];append from
ind_bench_dose;

label={"id" "permethrin" "cypermethrin" "betacyfluthrin" "deltamethrin"
"esfenvalerate"};
create ind_bench_doses_lower from ind_bench_dose_lower[colname=label];append
from ind_bench_dose_lower;

quit;

data benchmark_doses;
set benchmark_doses bench_doses;
run;

data benchmark_doses_lower;
set benchmark_doses_lower bench_doses_lower;
run;

data ind_benchmark_doses;
set ind_benchmark_doses ind_bench_doses;
run;

data ind_benchmark_doses_lower;
set ind_benchmark_doses_lower ind_bench_doses_lower;
run;

data powers;
set powers pows;
run;

data outers;
set outers outs;
run;

%end;
%mend analysis;

%analysis;

/*the next data steps clean the data and calculate the power (how often the
gold standard
test and the proposed similarity measure agreed*/
data benchmark_doses_anal3;
set benchmark_doses;
if id=. then delete;
run;

data benchmark_doses_lower_anal3;
set benchmark_doses_lower;
```

```
if id=. then delete;
run;

data ind_benchmark_doses_anal3;
set ind_benchmark_doses;
if id=. then delete;
run;

data ind_benchmark_doses_lower_anal3;
set ind_benchmark_doses_lower;
if id=. then delete;
run;

data all_anal3;
set all;
*if noest=1 then delete;
run;

data power_anal3(drop=yesno);
set powers;
if id=. then delete;
in=yesno;
run;

data comp_power3;
merge power_anal3 all_anal3;
by id;
run;

data bmd_anal3;
merge comp_power3 benchmark_doses_anal3;
by id;
if nopt=1 then delete;
if noest=1 then delete;
run;

data bmd_lower_anal3;
merge comp_power3 benchmark_doses_lower_anal3;
by id;
if nopt=1 then delete;
if noest=1 then delete;
run;

data ind_bmd_anal3;
merge comp_power3 ind_benchmark_doses_anal3;
by id;
if nopt=1 then delete;
if noest=1 then delete;
run;

data ind_bmd_lower_anal3;
merge comp_power3 ind_benchmark_doses_lower_anal3;
by id;
if nopt=1 then delete;
if noest=1 then delete;
```

```
run;

proc copy in=work out=perm;
select  benchmark_doses_anal3 benchmark_doses_lower_anal3
ind_benchmark_doses_anal3 ind_benchmark_doses_lower_anal3
all_anal3 power_anal3 comp_power3 bmd_anal3 bmd_lower_anal3 ind_bmd_anal3
ind_bmd_lower_anal3;
run;

proc print data=ind_bmd_lower_anal3;
run;

proc freq data=ind_bmd_lower_anal3;
tables yesno;
run;

/*creates the histograms for benchmark dose for the individual chemicals*/
proc univariate data=ind_bmd_lower_anal3;
histogram;
var permethrin cypermethrin betacyfluthrin deltamethrin esfenvalerate;
run;

data tag;
constant=0;
do tag=1 to 963;
output;end;
run;

data final_ind_bmd_lower_anal;
merge tag ind_bmd_lower_anal3;
run;

/*creates data sets containing all the calculated information for the
candidate mixtures
that converged*/
%macro bench_dose;
%do _i_=1 %to 963;

data ind_bmd_lower_&_i_;
set final_ind_bmd_lower_anal;
where tag=&_i_;
constant=0;
run;

%end;

%mend bench_dose_sim;

%bench_dose;

data bmd_low;
run;

/*calculated how often the individual chemicals BMD's were within the
proposed
```

```
similarity bounds for each mixture*/
%macro bmd_sim;
%do _i_=1 %to 963;

proc iml;
use ind_bmd_lower_&_i_;
read all var{permethrin} into bmd1_perm;
read all var{cypermethrin} into bmd2_cyperm;
read all var{betacyfluthrin} into bmd3_betacyf;
read all var{deltamethrin} into bmd4_delta;
read all var{esfenvalerate} into bmd5_esfen;

bm_doses=bmd1_perm||bmd2_cyperm||bmd3_betacyf||bmd4_delta||bmd5_esfen;

bmd1_sim=(bmd1_perm>0.999)*(bmd1_perm<4.710);
bmd2_sim=(bmd2_cyperm>0.551)*(bmd2_cyperm<2.599);
bmd3_sim=(bmd3_betacyf>0.247)*(bmd3_betacyf<1.164);
bmd4_sim=(bmd4_delta>0.065)*(bmd4_delta<0.307);
bmd5_sim=(bmd5_esfen>0.052)*(bmd5_esfen<0.244);
bench_sim=bmd1_sim*bmd2_sim*bmd3_sim*bmd4_sim*bmd5_sim;
if bench_sim=1 then do;bmd_sim=1;end;
if bench_sim=0 then do;bmd_sim=0;end;

if bmd1_sim=1 then do; bmd_sim1=1;end;
if bmd1_sim=0 then do; bmd_sim1=0;end;
if bmd2_sim=1 then do; bmd_sim2=1;end;
if bmd2_sim=0 then do; bmd_sim2=0;end;
if bmd3_sim=1 then do; bmd_sim3=1;end;
if bmd3_sim=0 then do; bmd_sim3=0;end;
if bmd4_sim=1 then do; bmd_sim4=1;end;
if bmd4_sim=0 then do; bmd_sim4=0;end;
if bmd5_sim=1 then do; bmd_sim5=1;end;
if bmd5_sim=0 then do; bmd_sim5=0;end;
ind_bench_dose=bmd1_perm||bmd2_cyperm||bmd3_betacyf||bmd4_delta||bmd5_esfen||
bmd_sim||bmd_sim1||bmd_sim2||bmd_sim3||bmd_sim4||bmd_sim5;

label={"permethrin" "cypermethrin" "betacyfluthrin" "deltamethrin"
"esfenvalerate" "bmd_sim" "bmd_sim1" "bmd_sim2" "bmd_sim3" "bmd_sim4"
"bmd_sim5"};
create bench_doses from ind_bench_dose[colname=label];append from
ind_bench_dose;

quit;

data bmd_low;
set bmd_low bench_doses;
run;

%end;

%mend bmd_sim;

%bmd_sim;

/*cleans the data and creates histogram for selected candidate mixture*/
```

```
data bmd_low_anal;
set bmd_low;
if permethrin=. then delete;
run;

proc print data=bmd_low_anal;
run;

proc freq data=bmd_low_anal;
tables bmd_sim bmd_sim1 bmd_sim2 bmd_sim3 bmd_sim4 bmd_sim5;
run;

data bmdl_dist1;
set perm.bmd_lower_anal3;
run;

proc print data=bmdl_dist1;
run;

proc univariate data=bmdl_dist1;
histogram;
var lower_benchmark_dose;
run;
```

Appendix B.4: SAS Code for Chapter 4

Appendix B.4.1: SAS Code to Evaluate Sensitivity and Specificity for the Unadjusted Unweighted Similarity Measure and to Compare to Stork et al. (2008) Method

```
libname pyr 'C:\Sufficient Similarity Research\verification';
libname ver 'C:\Sufficient Similarity Research\verification_code';
libname perm 'C:\Sufficient Similarity
Research\verification_code\sens_spec_data';
goptions colors=(black) htext=1.8 ftext=swiss;

/*reads in the original observed exposure data and takes the log
transformation*/
data pyr;
set pyr.pyr_data;
logcyfluth=log(cyfluth);
logcyperme=log(cyperme);
logdelta_t=log(delta_t);
logesfenva=log(esfenva);
logcis_trans_p=log(cis_trans_p);
run;

proc univariate data=pyr;
var logdelta_t logesfenva;
histogram;
run;

proc sort data=pyr;
by descending total_loading ;
run;
quit;

/*selects the top 20 observations/child care centers by total loading*/
data pyr;
set pyr;
if _n_<=20;
run;

proc iml;
use pyr;

read all var{logcis_trans_p} into cistp;
read all var{logcyperme} into cype;
read all var{logcyfluth} into cyfl;
read all var{logdelta_t} into deltat;
read all var{logesfenva} into esfen;
```

```
x=cistp||cype||cyfl||deltat||esfen;
Ident=I(20);
ones=j(20,1,1);
samp_cov=(1/(20-1))*x`*(ident-1/20*ones*ones`)*x;
n=nrow(cistp);
aones=j(1,n,1);
avg_cistp=aones*cistp/n;
avg_cype=aones*cype/n;
avg_cyfl=aones*cyfl/n;
avg_deltat=aones*deltat/n;
avg_asfen=aones*esfen/n;
print avg_cistp
avg_cype
avg_cyfl
avg_deltat
avg_asfen;
print samp_cov;
quit;


/* Sample from Multivariate Normal distribution with (Mean, Cov).
N           is the number of desired observations sampled from the
            multivariate normal distribution.
Mean  is a 1xp vector of means.
Cov         is a pxp symmetric positive definite variance-covariance matrix.

Each row of the returned matrix is a row vector
sampled from the multivariate normal distribution.
*/

proc iml;
start RANDNORMAL( N, Mean , Cov );
      /* Algorithm:
      1. By a Cholesky factorization, get T such that T`T = variance.
      2. Generate a p-vector of iid N(0,1) random variables z = (z_1, ...
,z_p).
      3. x = T`z + mean.
      4. Then x follows the multivariate normal distribution with Mean, Cov.

      Reference:
      Gentle, J.E. (2003), Random Number Generation and Monte Carlo Methods,
            New York: Springer-Verlag, Inc., 197-198.
*/
      /* check parameters */
      if N<1 then do;
      print "The requested number of observations should be at least 1:" N;
stop;
      end;

      mMean = rowvec(Mean);
      p = ncol(mMean);
      /* Upper triangular matrix T from the Cholesky decomposition:
            variance = symmetric positive-definite matrix.
         This call will fail if the matrix is not sym. pos. def. */
```

```
      T = root( Cov );
      Z=j(N,p,.);
      call randgen(Z,'NORMAL');
      outX = Z*T + repeat(mMean,N,1);
      return(outX);
finish;
store module=RANDNORMAL;
quit;

/*utilizes the module randnormal to simulate observed exposure data for the
purposes of generating possible candidate mixtures to assess performance of
the similarity measures*/
proc iml;
load module=RANDNORMAL;
/* doc example 1: how to use a correlation matrx and vector of variances
to create the required covariance matrix in RANDNORMAL function*/
call randseed(1);
N=1000;
Mean = {-0.03 -1.30 -3.61 -4.02 -4.27}; /*from log-normal transformation*/

Cov = {  3.6990065 -0.251427 -3.251512 -1.831635 -0.050669,
        -0.251427 8.3298461 0.8294449 0.8971857 1.6237268,
        -3.251512 0.8294449 8.0506249 2.3486282  -0.30063,
        -1.831635 0.8971857 2.3486282 2.3926647 -0.081364,
        -0.050669 1.6237268  -0.30063 -0.081364 3.2804322};/*from log-normal
transformation*/

x = RANDNORMAL( N, Mean, Cov );
print x;

label={"cis_trans_p" "cyperme" "cyfluth" "delta_t" "esfenva"};
create expose from x[colname=label]; append from x;
quit;


/*uses the log-normal transformation and converts back by exponentiating and
calculates 1000 possible candidate mixtures*/
data expose;
set expose;
cis_trans_p=exp(cis_trans_p);
cyperme=exp(cyperme);
cyfluth=exp(cyfluth);
delta_t=exp(delta_t);
esfenva=exp(esfenva);
constant=0;
total_loading=cyfluth+cyperme+delta_t+esfenva+cis_trans_p;
a1=cis_trans_p/total_loading;a2=cyperme/total_loading;a3=cyfluth/total_loadin
g;a4=delta_t/total_loading;a5=esfenva/total_loading;
run;

data id;
constant=0;
do id=1 to 1000;
output;end;
run;
```

```
data ais_final;
merge expose id;
by constant;
run;

data dose;
  input dose;
  constant=0;
  cards;
0
0.275
1.096
2.740
9.042
13.70
18.084
27.400
;
run;
quit;

data sample;
run;

/*creates 1000 candidate data sets using the generated candidate mixtures
under the assumption of additivity*/
%macro ais;
%do _i_=1 %to 1000;

data ais_&_i_;
set ais_final;
where id=&_i_;
constant=0;
run;

data new;
merge ais_&_i_ dose;
by constant;
run;

data generate&_i_;
set new;
    seed=100597+&_i_;
    *aa=0.2521;
      aa=0.2521;
      b1= -0.0139;
      b2= -0.0554;
      b3= -0.2686;
      b4= -0.2364;
      b5= -0.4959;
      del= -0.2359;
      drop b1 b2 b3 b4 b5 constant check id;
      ba=b1*a1+b2*a2+b3*a3+b4*a4+b5*a5;
      put ba;
```

```
      curve=2;
      group=&_i_;
            do k=1 to 12;
           term=ba*dose;
               *pact = aa+(1-aa)*exp((term-del)*(term<del));
               mu = aa+(1-aa)*exp((term-del)*(term<del));
           pact=mu+sqrt(0.0648)*rannor(seed);
               *pact=mu+sqrt(0.0348)*rannor(seed);
               *pact=mu+sqrt(0.18)*rannor(seed);
             if pact<0 then do;pact=mu;end;
               output;
            end;
run;

data sample;
set sample generate&_i_;
run;
quit;
%end;
%mend ais;

%ais;

symbol1 i=none v=dot;
proc gplot data =generate1;
plot pact*dose;
run;
quit;

data sasuser.sample;
set sample;
if dose=. then delete;
run;
quit;

data sample;
set sasuser.sample;
run;

/*creates the reference data set under the assumption of additivity using
the mixing proportions from the Crofton et al. study*/
%macro refdata;
      %do _i_=1 %to 1;

data ref_ais;
      constant=0;
      a1=0.522;
      a2=0.288;
      a3=0.129;
      a4=0.034;
      a5=0.027;
run;

data ref_data;
 merge ref_ais dose;
```

```
 by constant;
    seed=102679;
      aa=0.2521;
      b1= -0.0139;
      b2= -0.0554;
      b3= -0.2686;
      b4= -0.2364;
      b5= -0.4959;
      del= -0.2359;
      drop b1 b2 b3 b4 b5 constant;
      ba=b1*a1+b2*a2+b3*a3+b4*a4+b5*a5;
      put ba;
      curve=1;
            do k=1 to 12;
           term=ba*dose;
               mu = aa+(1-aa)*exp((term-del)*(term<del));
             pact=mu+sqrt(0.0648)*rannor(seed);
               if pact<0 then do;pact=mu;end;
               output;
            end;
run;

%end;
%mend refdata;

%refdata;

data forplot;
do dose=0 to 30 by 0.5;output;end;
run;

data ref_data_plot;
set ref_data forplot;
run;


/*this is fitting the non-linear exponential threshold model to the generated
   reference data set and plotting the ellipse with random effect variance of
0
   and with random effect variance increasing, as well as calculating the hl
and hu*/

proc nlmixed data=ref_data_plot cov;
    parms b=-.02 to -.10 by -.02 delta=1 to 5 by 1 s2=0.0648;
a= 0.25;
g=1-a;
    y = a+g*exp(b*(dose-delta)*(dose>delta));
    model pact ~ normal(y,s2);
    mu20= a + g*0.8; *85.3;
    mu50= a + g*0.5; *63.3;
    ed20=((log((mu20-a)/g))/b)+delta;
    ed50=((log((mu50-a)/g))/b)+delta;

    estimate "ED20" ((log((mu20-a)/g))/b)+delta   alpha=0.001;
    estimate "ED50" ((log((mu50-a)/g))/b)+delta   alpha=0.001;
```

```
     ED20low=0.35*ed20;  ********************sets bounds based on
percentages of ED20 amd ED50 estimates;
    ED20high=1.65*ed20;
    ED50low=.35*ed50;
    ED50high=1.65*ed50;

    logmu20=log((mu20-a)/g);
    logmu50=log((mu50-a)/g);
    quot=logmu50/logmu20;
    denom_l=ed20low-(ed50low-quot*ed20low)/(1-quot);
    blow=logmu20/denom_l;
    dellow=(ed50low-quot*ed20low)/(1-quot);
    denom_h=ed20high-(ed50high-quot*ed20high)/(1-quot);
    bhigh=logmu20/denom_h;
    delhigh=(ed50high-quot*ed20high)/(1-quot);
    mulow=a+g*exp(blow*(dose-dellow)*(dose>dellow));
    muhigh=a+g*exp(bhigh*(dose-delhigh)*(dose>delhigh));
    id quot mulow muhigh ed20low ed20high ed50low ed50high;

    predict  y out=pred;
    ods output parameterestimates=vars covmatparmest=covs;
*proc print data=pred;
run;
quit;

proc sort data=pred;
by dose;
run;

symbol1 v=dot i=none;
symbol2 v=none i=join l=1;
symbol3 v=none i=join l=2 c=red;
symbol4 v=none i=join l=15 c=red;
axis1 label=(a=90 "Motor Activity(% control)");
axis2 label=("Total Dose(mg/kg)");
proc gplot data=pred;
plot (pact pred mulow muhigh)*dose/overlay  vaxis=axis1 haxis=axis2;
run;
quit;

proc gplot data=pred;
plot (pact pred)*dose/overlay vaxis=axis1 haxis=axis2;
run;
quit;

data ellipse;
set ref_data_plot;
if pact=. then delete;
run;

proc print data=ellipse;
run;
```

```
proc print data=covs;
run;

/*calculates the 95% confidence ellipse for the simulated reference data
set*/
proc iml;
use vars;
read all var{estimate} where (parameter='b') into beta;
use vars;
read all var{estimate} where (parameter='delta') into delta;
use vars;
read all var{estimate} where (parameter='s2') into mse;
use ellipse;
read all var{dose} into dose;
a=.20;
g=1-a;
*print dose;
dfdbeta=g#(dose[,1]-delta)#exp(beta#(dose[,1]-delta))#(dose[,1]>delta);
dfdel=-g#beta#exp((beta#(dose[,1]-delta)))#(dose[,1]>delta);
dfdb=g#(beta#dose[,1]-beta#delta)#exp(beta#(dose[,1]-
delta))#(dose[,1]>delta);
x0=dfdbeta||dfdel;
z0=dfdb;
d=0;
n=nrow(z0);
*print n;
i=I(96);
R= (mse)*i;
Rinv=inv(R);
Rinv_z0_D=rinv*z0*D;
z0t_Rinv_z0_D=z0`*Rinv_z0_D;
z0t_Rinv=z0`*Rinv;
v0inv=rinv - Rinv_z0_D*inv(1+z0t_Rinv_z0_D)*z0t_Rinv;
omega=inv(x0`*v0inv*x0);
print omega;
mu20= a + g*0.8; *85.3;
mu50= a + g*0.5; *63.3;
ED20=(log((mu20-a)/g)/beta)+delta;
ED50=(log((mu50-a)/g)/beta)+delta;
*print ed20 ed50;
gomega=ed20//ed50;
big_g=(-(beta**-2)*log((mu20-a)/g)||1)//(-(beta**-2)*log((mu50-a)/g)||1);
cov_gomega=big_g*omega*big_g`;
se_gomega = sqrt(vecdiag(cov_gomega));
var=cov_gomega;
s=nrow(gomega);
varinv=inv(var);
*print var d varinv;
** transformations to polar coordinates;
call eigen(eval,p,varinv);
lambda=diag(eval);
lambdahalf=root(lambda);
lamhalfinv=inv(lambdahalf);
check=lambdahalf*lambdahalf;
*print lambda check;  ** should equal lambda;
```

```
test=p`*varinv*p;  * test should equal lambda;
*print lambda test;
test2=p*p`;
*print test2;  *test2 shoould equal the identity matrix;
type1 = 0.05;
bign=96;
totp=2;
f=finv(1-type1,s,bign-totp);
r = (s*f)**(1/2);
pi=constant('pi');
twopie=2*pi;

znew=j(2,1,0);
gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      *print z;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
      *print gw;
label={"gw1" "gw2"};
create gw1 from gw[colname=label];append from gw;
quit;

/*this creates the plot of the confidence ellipse with the
appropriate similarity bounds*/
symbol1 v=none i=join;
 symbol2 l=1;
 symbol3 l=2;
 axis1 label=(a=90);* order=(0 to 160 by 1);
 *axis2 order=(0 to 160 by 1);
data box;
  set pred;
  if _n_=1;
  keep boxgw1 boxgw2 ed20low ed20high ed50low ed50high;
boxgw1=ed20low; *0.001; boxgw2=ed50low; *0.31; output;  **two low values;
boxgw1=ed20high; *0.53; output;  ***ed20hi;
boxgw2=ed50high; *.62; output;   **ed50hi;
boxgw1=ed20low; *.001; output;  ***ed20low;
boxgw2=ed50low; *.31; output;   **ed50low;
*proc print data=box; run;
data plotrefa;
set gw1 box;

data plotref1a;
set plotrefa;
if _n_=1;
run;

data plotref2a;
```

```
set plotrefa plotref1a;
run;

symbol1 v=none i=join c=blue;
symbol2 v=none i=join c=red;
proc gplot data=plotref2a;
plot  gw2*gw1 boxgw2*boxgw1/overlay vaxis=axis1 noframe;
label gw2='ED50 (Total Dose (mg/kg))';
label gw1='ED20 (Total Dose (mg/kg))';
run; quit;

/*this allows sigma-squared-h to increase until it touches the edge of the
similarity bounds*/
proc iml;
use vars;
  read all var{estimate} where (parameter='b') into beta;
  read all var{estimate} where (parameter='delta') into delta;
  read all var{estimate} where (parameter='s2') into mse;
use pred;
   read first var{ed20low} into ed20low;
   read first var{ed20high} into ed20high;
   read first var{ed50low} into ed50low;
   read first var{ed50high} into ed50high;

 a=.20;
 g=1-a;
 b=0;
 use ellipse;
   read all var{dose} into dose;
D=0;
sigh=0;
flag=0;
test=0;
do d=0 to 1 by .0001 while(flag=0);
dfdbeta=g#(dose[,1]-delta)#exp(beta#(dose[,1]-delta))#(dose[,1]>delta);
dfdel=-g#beta#exp((beta#(dose[,1]-delta)))#(dose[,1]>delta);
dfdb=g#(beta#dose[,1]-beta#delta)#exp(beta#(dose[,1]-
delta))#(dose[,1]>delta);
x0=dfdbeta||dfdel;
z0=dfdb;
n=nrow(z0);
i=I(96);
R= (mse)*i;
Rinv=inv(R);
Rinv_z0_D=rinv*z0*D;
z0t_Rinv_z0_D=z0`*Rinv_z0_D;
z0t_Rinv=z0`*Rinv;
v0inv=rinv - Rinv_z0_D*inv(1+z0t_Rinv_z0_D)*z0t_Rinv;
omega=inv(x0`*v0inv*x0);
mu20= a + g*0.8;
mu50= a + g*0.5;

ED20=(log((mu20-a)/g)/beta)+delta;
ED50=(log((mu50-a)/g)/beta)+delta;
gomega=ed20//ed50;
```

```
big_g=(-(beta**-2)*log((mu20-a)/g)||1)//(-(beta**-2)*log((mu50-a)/g)||1);
cov_gomega=big_g*omega*big_g`;

se_gomega = sqrt(vecdiag(cov_gomega));

var=cov_gomega;
s=nrow(gomega);
varinv=inv(var);
** transformations to polar coordinates;
call eigen(eval,p,varinv);
lambda=diag(eval);
lambdahalf=root(lambda);
lamhalfinv=inv(lambdahalf);
check=lambdahalf*lambdahalf;
*print lambda check;  ** should equal lambda;
test=p`*varinv*p;  * test should equal lambda;
*print lambda test;
test2=p*p`;
*print test2;  *test2 shoould equal the identity matrix;
type1 = 0.05;
bign=96;
totp=2;
f=finv(1-type1,s,bign-totp);
r = (s*f)**(1/2);
pi=constant('pi');
twopie=2*pi;

znew=j(2,1,0);
gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      *print z;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
      d1=(gw[,1]-ed20low)>.0001;
      d2=(ed20high-gw[,1])>.0001;
      d3=(gw[,2]-ed50low)>.0001;
      d4=(ed50high-gw[,2])>.0001;
      test=d1#d2#d3#d4;
      if test=1 then do;flag=0;end;
      else do; flag=1;sigh=sigh//d;end;
end;
print  sigh;
label={"gw1" "gw2"};
create gw from gw[colname=label];append from gw;
quit;

*this plots the expanded confidence ellipse;
 symbol1 v=none i=join;
 symbol2 l=1;
```

```
 symbol3 l=2;
 axis1 label=(a=90);* order=(0 to 160 by 1);
 *axis2 order=(0 to 160 by 1);

data plotref;
set gw box;

data plotref1;
set plotref;
if _n_=1;
run;

data plotref2;
set plotref plotref1;
run;

proc gplot data=plotref2;
plot  gw2*gw1 boxgw2*boxgw1/overlay vaxis=axis1 noframe;
label gw2='ED50 (Total Dose (mg/kg))';
label gw1='ED20 (Total Dose (mg/kg))';
run; quit;

/*this calculates zmax and hl and hu*/
proc iml;
use vars;
   read all var{estimate} where (parameter='b') into beta;
   read all var{estimate} where (parameter='delta') into delta;
use pred;
   read first var{ed20low} into ed20low;
   read first var{ed20high} into ed20high;
   read first var{ed50low} into ed50low;
   read first var{ed50high} into ed50high;
a=.20;
g=1-a;
b=0;
use ref_data;
read all var{dose} into dose;
mu20= a + g*0.8;
mu50= a + g*0.5;
ED20=(log((mu20-a)/g)/beta)+delta;
ED50=(log((mu50-a)/g)/beta)+delta;
zmax=0;
sigsqh=.0141;
sigh=sqrt(sigsqh);
flag=0;
hlf=0;
huf=0;
zm=0;
delta1= ed20low; *.001;
delta2=ed20high; *.53;
delta3=ed50low; *.31;
delta4=ed50high; *.62;

do zmax=0 to 10 by .001 while(flag=0); ****maybe not true;
```

```
hl=1-zmax*sigh;
hu=1+zmax*sigh;
lowdel1=ed20*hl;
updel1=ed20*hu;
lowdel2=ed50*hl;
updel2=ed50*hu;
t1=-(delta1-lowdel1);
t2=-(updel1-delta2);
t3=-(delta3-lowdel2);
t4=-(updel2-delta4);

eps=0.01;
if ((t1<eps)*(t1 >0))=1 then do; max1=1;end;
else do; max1=0;end;

if ((t2<eps)*(t2 >0))=1 then do; max2=1;end;
else do; max2=0;end;

if ((t3<eps)*(t3 >0))=1 then do; max3=1;end;
else do; max3=0;end;

if ((t4<eps)*(t4 >0))=1 then do; max4=1;end;
else do; max4=0;end;

test=max1+max2+max3+max4;

if test=0 then do;flag=0;end;
if test>=1 then do;flag=1;zm=zm//zmax;hlf=hlf//hl;huf=huf//hu;
 print zmax;
 print hlf huf;
 print t1 t2 t3 t4;
 print max1 max2 max3 max4;
end; end;
quit;

proc sort data=ref_data;
by dose;
run;

proc gplot data=ref_data;
plot pact*dose;
run;
quit;

data yesno;
run;

data yesno1;
run;

data all;
run;

data all1;
run;
```

```
data powers;
run;

data outers;
run;

/*this macro creates 1000 studies and conducts the gold standard test for
sufficient
similarity and calculates sensitivity for the unadjusted unweighted
similarity
measure and for the method proposed by Strok et al. (2008)*/
%macro subsets;
%do _i_=1 %to 1000;

footnote 'sample='&_i_;

data subset&_i_;
set sample;
where group=&_i_;
run;

data evaluate&_i_;
set ref_data subset&_i_;
run;

data evaluate&_i_;
set evaluate&_i_;
if curve=2 then do; pact2=pact;end;
if curve=1 then do; pact1=pact;end;
run;

proc sort data=evaluate&_i_;
by curve dose;
run;
quit;

proc nlmixed data=evaluate&_i_ cov hess method=firo;
parms b=0 to -.16 by -.02 delta=1 to 2 by .1 s2=0.0648 su=0.0001 to .1 by
.05;
a= 0.2521;
g=1-a;
s2u=su*su;
    *y = a_term+g*exp(b*(u+1)*(dose-delta)*(dose>delta));
      y = a+g*exp(b*(1+u)*(dose-delta)*(dose>delta));
      estimate 'ed20' (log(.8)+b*delta)/(b);
    estimate 'ed50' (log(.5)+b*delta)/(b);
    random u ~ normal(0,s2u) subject=curve out=randomest&_i_;
    model pact ~ normal(y,s2);
    predict y out=pred&_i_;
    ods output parameterestimates=vars covmatparmest=covs;
run;
quit;
```

```
data covs;
set covs;
where parameter='b' or parameter='delta';
run;

data noest;
set vars;
if df=. then do;
flag=1;end;
if df=1 then do;
flag=0; end;
run;

/*produces box and confidence region*/
proc iml;
  use vars;
      read all var{estimate} where (parameter='b') into beta;
      read all var{estimate} where (parameter='delta') into delta;
      read all var{estimate} where (parameter='s2') into mse;
      read all var{estimate} where (parameter='s2u') into sigmah;

use covs;
      read all var{b delta} into covs;

      aa=.2521; * 25.21;
      g=1-aa;
      mu20=0.8#g+aa;
      mu50=0.5#g+aa;
      ED20=(log((mu20-aa)/g)/beta)+delta;
      ED50=(log((mu50-aa)/g)/beta)+delta;
      *print ed20 ed50;
      gomega=ed20//ed50;
      big_g=(-(beta**-2)*log((mu20-aa)/g)||1)//(-(beta**-2)*log((mu50-
aa)/g)||1);
      cov_gomega=big_g*covs*big_g`;
*     print covs cov_gomega;
      varinv=inv(cov_gomega);
      se_gomega = sqrt(vecdiag(cov_gomega));
      *print  cov_gomega varinv se_gomega;

** transformations to polar coordinates;
      call eigen(eval,p,varinv);
      lambda=diag(eval);
      lambdahalf=root(lambda);
      lamhalfinv=inv(lambdahalf);
      check=lambdahalf*lambdahalf;
      *print lambda p;
*     print lambda check;  ** should equal lambda;
      test=p`*varinv*p;  * test should equal lambda;
*     print lambda test;
      test2=p*p`;
*     print test2;  *test2 shoould equal the identity matrix;
      type1 = 0.05;
      bign=192;
      totp=2;
```

```
      s=nrow(gomega);
      f=finv(1-type1,s,bign-totp);
      r = (s*f)**(1/2);
      pi=constant('pi');
      twopie=2*pi;

      znew=j(2,1,0);
      gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      *print z;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
      label={"gw1" "gw2"};
      create gw1 from gw[colname=label];append from gw;
*     labelz={"z1" "z2"};
*     znew=znew`;
*     create znew from znew[colname=labelz];
*    append from znew;
quit;
*proc gplot data=znew;
*   plot z2*z1;
*   run;
*   quit;


symbol1 v=none i=join;
symbol2 l=1;
symbol3 l=2;
axis4 label=(a=90 'ED50 (mg/kg)');* order=(0 to 100 by 20);
axis5 label=('ED20 (mg/kg)') ;*order=(0 to 100 by 20);
data box;
      boxgw1=1.35; boxgw2=2.79; output;
      boxgw1=6.37;output;
      boxgw2=13.15;output;
      boxgw1=1.35;output;
      boxgw2=2.79;output;

data plotrefa;
      set gw1 box;

data plotref1a;
      set plotrefa;
      if _n_=1;
      run;

data plotref2a;
      set plotrefa plotref1a;
run;
```

```
/*
symbol2 i=join v=none;
proc gplot data=plotref2a;
      plot  gw2*gw1 boxgw2*boxgw1/ overlay vaxis=axis4 haxis=axis5 noframe;
*     plot  gw2*gw1 / overlay vaxis=axis4 haxis=axis5 noframe;
*     plot  boxgw2*boxgw1/  vaxis=axis4 haxis=axis5 noframe;
      label gw2='ED50 (mg/kg)';
      label gw1='ED20 (mg/kg)';
      run; quit;
*/

proc means data=plotref2a max min noprint;
      var gw1 gw2;
      output out=maxmin max= max1 max2 min= min1 min2;
run;

data yesno;
      set maxmin;
      id=&_i_;
      yesno=0;
      if min1>1.35 and max1<6.37 and min2>2.79 and max2<13.15 then yesno=1;
      if min1=0 and max1=0 and min2=0 and max2=0 then noest=1;
run;


/*
data yesno1;
      set maxmin;
      id=&_i_;
      yesno=0;
      if min1>9.32 and max1<17.30 and min2>13.16 and max2<24.45 then yesno=1;
      if min1=0 and max1=0 and min2=0 and max2=0 then noest=1;
run;
*/

data all;
set all yesno;
drop _type_ _freq_;
if yesno=. then delete;
run;

/*
data all1;
set all1 yesno1;
drop _type_ _freq_;
if yesno=. then delete;
run;
*/

proc iml;
use ais_&_i_;
read all var {a1} into a1;
read all var {a2} into a2;
read all var {a3} into a3;
read all var {a4} into a4;
```

```
read all var {a5} into a5;


use noest;
read all var {flag} into flagger;
if flagger=1 then do; flag=1;end;
if flagger=0 then do; flag=0;end;

a_ref=0.522//0.288//0.129//0.034//0.027;
a_cand=a1//a2//a3//a4//a5;

diffsq=(a_ref-a_cand)`*(a_ref-a_cand);
dist=1+sqrt(diffsq);

id=&_i_;

if flag=1 then do;out=id||flag;end;
if flag=0 then do;out=id||flag;end;

in=(dist>0.35)*(dist<1.65)*(flag=0);

if in=1 then do;pow=1;end;
if in=0 then do;pow=0;end;

t1=(a1>0.19)*(a1<0.84);
t2=(a2>0.08)*(a2<0.66);
t3=(a3>0.03)*(a3<0.41);
t4=(a4>0.007)*(a4<0.14);
t5=(a5>0.006)*(a5<0.12);
la=t1*t2*t3*t4*t5;
if la=1 then do;leanna=1;end;
if la=0 then do;leanna=0;end;

power=id||pow||dist||flag||leanna;

label={"id" "yesno" "distance" "nopt" "leanna"};
create pows from power[colname=label];append from power;

create outs from out;append from out;
quit;

data powers;
set powers pows;
run;

data outers;
set outers outs;
run;

%end;
%mend;

%subsets;

/*cleans data and calculates sensitivity and specificity*/
```

```
proc copy in=work out=perm;
select  sample components comp_power ref_data;
run;

data all_anal;
set all;
*if noest=1 then delete;
run;

data power_anal(drop=yesno);
set powers;
if id=. then delete;
in=yesno;
run;

data comp_power;
merge power_anal all_anal;
by id;
run;

data comp_power;
set comp_power;
if nopt=1 then delete;
if noest=1 then delete;
if leanna=0 then do; leanna_out=1;end;
if leanna=1 then do; leanna_out=0;end;
run;

proc print data=comp_power;
run;

data sensitivity;
set comp_power;
if yesno=1;
run;

data specificity;
set comp_power;
if yesno=0;
run;


data specificity_no;
set comp_power;
if yesno=0 then do;no=1;end;
if yesno=1 then do;no=0;end;
if in=0 then do; out=1;end;
if in=1 then do; out=0;end;
run;


proc iml;
use sensitivity;
read all var{yesno} into yesno;
read all var{in} into in;
```

```
n=nrow(yesno);
ones=j(1,n,1);
sumyesno=ones*yesno;
sumin=ones*in;
sensitivity=sumin/sumyesno;
print sumyesno sumin sensitivity;
quit;

proc iml;
use specificity_no;
read all var {no} into no;
read all var {out} into notin;
n_out=nrow(no);
ones=j(1,n_out,1);
sumno=ones*no;
sumout=ones*notin;
specificity=sumout/sumno;
print sumno sumout specificity;
quit;

proc iml;
use sensitivity;
read all var{yesno} into yesno;
read all var{leanna} into in;
n=nrow(yesno);
ones=j(1,n,1);
sumyesno=ones*yesno;
sumin=ones*in;
sensitivity=sumin/sumyesno;
print sumyesno sumin sensitivity;
quit;

proc iml;
use specificity_no;
read all var {no} into no;
read all var {leanna_out} into notin;
n_out=nrow(no);
ones=j(1,n_out,1);
sumno=ones*no;
sumout=ones*notin;
specificity=sumout/sumno;
print sumno sumout specificity;
quit;

proc copy in=work out=perm;
select  expose_log;
run;
```

Appendix B.4.2: SAS Code to Evaluate Sensitivity and Specificity for the Remaining Similarity Measures when Resmethrin or λ-cyhalothrin is Added

```
libname pyr 'C:\Sufficient Similarity Research\verification';
libname ver 'C:\Sufficient Similarity Research\verification_code';
```

```
libname perm 'C:\Sufficient Similarity
Research\verification_code\sens_spec_data';

/*Pulls in the log-transformed raw exposure data for the selected
pyrethroids*/
data expose_log;
set perm.expose_log;
run;

/*creates the original dose groups*/
data dose;
  input dose;
  constant=0;
  cards;
0
0.275
1.096
2.740
9.042
13.70
18.084
27.400
;
run;
quit;

/*creates the new dose groups resulting from adding resmethrin/l-cyhalothrin
at different
specified proportions; creates 1000 candidate mixture rays*/
%macro subsetprops(comp);

data expose_log;*(drop=a1 a2 a3 a4 a5 check);
set expose_log;

a1_c=a1*(1-&comp);
a2_c=a2*(1-&comp);
a3_c=a3*(1-&comp);
a4_c=a4*(1-&comp);
a5_c=a5*(1-&comp);
a6_c=&comp;
total=a1_c+a2_c+a3_c+a4_c+a5_c+a6_c;
constant=0;
run;

data new_dose;
set dose;

/*
dose=dose+(dose*&comp)/(1-&comp);
*/


adose=dose+(dose*&comp)/(1-&comp);
dose=(1-&comp)*adose;
```

```
run;

%mend subsetprops;

%subsetprops (0.785);

proc print data=new_dose;
run;

data id;
constant=0;
do id=1 to 1000;
output;end;
run;

data ais_final;
merge expose_log id;
by constant;
run;

/*simulates the reference data under the assumption of additivity*/
%macro refdata;
      %do _i_=1 %to 1;

data ref_ais;
      constant=0;
      a1=0.522;
      a2=0.288;
      a3=0.129;
      a4=0.034;
      a5=0.027;
run;

data ref_data;
 merge ref_ais dose;
 by constant;
    seed=102679;
      aa=0.2521;
      b1= -0.0139;
      b2= -0.0554;
      b3= -0.2686;
      b4= -0.2364;
      b5= -0.4959;
      del= -0.2359;
      drop b1 b2 b3 b4 b5 constant;
      ba=b1*a1+b2*a2+b3*a3+b4*a4+b5*a5;
      put ba;
      curve=1;
            do k=1 to 12;
           term=ba*dose;
               mu = aa+(1-aa)*exp((term-del)*(term<del));
             pact=mu+sqrt(0.0648)*rannor(seed);
               if pact<0 then do;pact=mu;end;
               output;
```

```
            end;
run;

%end;
%mend refdata;

%refdata;

/*Resmethrin slope=-0.0020*/

/*simulates 1000 data sets using the 1000 generated candidate mixtures under
the
assumption of additivity*/
%macro curves(slope);
%do _i_=1 %to 1000;

data ais_&_i_;
set ais_final;
where id=&_i_;
constant=0;
run;

data cand_new;
merge ais_&_i_ new_dose;
by constant;
run;

data cand_generate&_i_;
set cand_new;
    seed=100598+&_i_;
    *aa=0.2521;
      aa=0.2521;
      b1= -0.0139;
      b2= -0.0554;
      b3= -0.2686;
      b4= -0.2364;
      b5= -0.4959;
      b6=&slope;
      del= -0.2359;
      drop b1 b2 b3 b4 b5 b6 constant check id;
      ba=b1*a1_c+b2*a2_c+b3*a3_c+b4*a4_c+b5*a5_c+b6*a6_c;
      put ba;
      curve=2;
      group=&_i_;
            do k=1 to 12;
           term=ba*adose;
               *pact = aa+(1-aa)*exp((term-del)*(term<del));
               mu = aa+(1-aa)*exp((term-del)*(term<del));
           pact=mu+sqrt(0.0648)*rannor(seed);
               *pact=mu+sqrt(0.0348)*rannor(seed);
               *pact=mu+sqrt(0.18)*rannor(seed);
             if pact<0 then do;pact=mu;end;
               output;
            end;
run;
```

```
data anal&_i_;
set ref_data cand_generate&_i_ ;
run;
quit;

data anal&_i_;
set anal&_i_;
if curve=2 then do; pact2=pact;end;
if curve=1 then do; pact1=pact;end;
run;

%end;
%mend curves;

%curves (-0.0020);

data yesno;
run;

data yesno1;
run;

data all;
run;

data all1;
run;

data powers;
run;

data outers;
run;

/*creates 1000 studies; performs the gold standard test for sufficient
similarity;
evaluates sufficient similarity using the proposed methods*/
%macro analysis;
%do _i_=1 %to 1000;

footnote 'sample='&_i_;

proc sort data=anal&_i_;
by curve dose;
run;
quit;

proc nlmixed data=anal&_i_ cov hess tech=trureg method=firo;
parms b=0 to -.16 by -.02 delta=1 to 2 by .05 s2=0.0648 su=0.0001 to .1 by
.05;
a= 0.25;
g=1-a;
s2u=su*su;
    *y = a_term+g*exp(b*(u+1)*(dose-delta)*(dose>delta));
```

```
      y = a+g*exp(b*(1+u)*(dose-delta)*(dose>delta));
      estimate 'ed20' (log(.8)+b*delta)/(b);
    estimate 'ed50' (log(.5)+b*delta)/(b);
    random u ~ normal(0,s2u) subject=curve out=randomest&_i_;
    model pact ~ normal(y,s2);
    predict y out=pred&_i_;
    ods output parameterestimates=vars covmatparmest=covs;
run;
quit;

data covs;
set covs;
where parameter='b' or parameter='delta';
run;

data noest;
set vars;
if df=. then do;
flag=1;end;
if df=1 then do;
flag=0; end;
run;

/*produces box and confidence region*/
proc iml;
  use vars;
      read all var{estimate} where (parameter='b') into beta;
      read all var{estimate} where (parameter='delta') into delta;
      read all var{estimate} where (parameter='s2') into mse;
      read all var{estimate} where (parameter='s2u') into sigmah;

use covs;
      read all var{b delta} into covs;

      aa=.25; * 25.21;
      g=1-aa;
      mu20=0.8#g+aa;
      mu50=0.5#g+aa;
      ED20=(log((mu20-aa)/g)/beta)+delta;
      ED50=(log((mu50-aa)/g)/beta)+delta;
      *print ed20 ed50;
      gomega=ed20//ed50;
      big_g=(-(beta**-2)*log((mu20-aa)/g)||1)//(-(beta**-2)*log((mu50-
aa)/g)||1);
      cov_gomega=big_g*covs*big_g`;
*     print covs cov_gomega;
      varinv=inv(cov_gomega);
      se_gomega = sqrt(vecdiag(cov_gomega));
      *print  cov_gomega varinv se_gomega;

** transformations to polar coordinates;
      call eigen(eval,p,varinv);
      lambda=diag(eval);
      lambdahalf=root(lambda);
      lamhalfinv=inv(lambdahalf);
```

```
      check=lambdahalf*lambdahalf;
      *print lambda p;
*     print lambda check;  ** should equal lambda;
      test=p`*varinv*p;  * test should equal lambda;
*     print lambda test;
      test2=p*p`;
*     print test2;  *test2 shoould equal the identity matrix;
      type1 = 0.05;
      bign=192;
      totp=2;
      s=nrow(gomega);
      f=finv(1-type1,s,bign-totp);
      r = (s*f)**(1/2);
      pi=constant('pi');
      twopie=2*pi;

      znew=j(2,1,0);
      gw=j(2,1,0);
  do theta=0 to twopie by .1;
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      *print z;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
      label={"gw1" "gw2"};
      create gw1 from gw[colname=label];append from gw;
*     labelz={"z1" "z2"};
*     znew=znew`;
*     create znew from znew[colname=labelz];
*    append from znew;
quit;
*proc gplot data=znew;
*   plot z2*z1;
*   run;
*   quit;


data box;
      boxgw1=1.35; boxgw2=2.82; output;
      boxgw1=6.35;output;
      boxgw2=13.32;output;
      boxgw1=1.35;output;
      boxgw2=2.82;output;

data plotrefa;
      set gw1 box;

data plotref1a;
      set plotrefa;
      if _n_=1;
      run;
```

```
data plotref2a;
      set plotrefa plotref1a;
run;

proc means data=plotref2a max min noprint;
      var gw1 gw2;
      output out=maxmin max= max1 max2 min= min1 min2;
run;

data yesno;
      set maxmin;
      id=&_i_;
      yesno=0;
      if min1>1.35 and max1<6.35 and min2>2.82 and max2<13.32 then yesno=1;
      if min1=0 and max1=0 and min2=0 and max2=0 then noest=1;
run;

data all;
set all yesno;
drop _type_ _freq_;
if yesno=. then delete;
run;

proc iml;
use ais_&_i_;
read all var {a1_c} into a1_c;
read all var {a2_c} into a2_c;
read all var {a3_c} into a3_c;
read all var {a4_c} into a4_c;
read all var {a5_c} into a5_c;
read all var {a6_c} into a6_c;

use noest;
read all var {flag} into flagger;
if flagger=1 then do; flag=1;end;
if flagger=0 then do; flag=0;end;

a_ref=0.522//0.288//0.129//0.034//0.027//0;
a_cand=a1_c//a2_c//a3_c//a4_c//a5_c//a6_c;

w={1.01 0 0 0 0 0,
   0 1.01 0 0 0 0,
   0 0 1.01 0 0 0,
   0 0 0 1.01 0 0,
   0 0 0 0 1.01 0,
   0 0 0 0 0 0.95};

*diffsq=(a_ref-a_cand)`*(a_ref-a_cand);
*adj_unw_diffsq=(a_ref*.215-a_cand)`*(a_ref*.215-a_cand);
 adj_w_diffsq=(a_ref*.215-a_cand)`*w*w*(a_ref*.215-a_cand);
dist=1+sqrt(adj_w_diffsq);

id=&_i_;
```

```
if flag=1 then do;out=id||flag;end;
if flag=0 then do;out=id||flag;end;

in=(dist>0.35)*(dist<1.65)*(flag=0);

if in=1 then do;pow=1;end;
if in=0 then do;pow=0;end;


power=id||pow||dist||flag;

label={"id" "yesno" "distance" "nopt"};
create pows from power[colname=label];append from power;

create outs from out;append from out;
quit;

data powers;
set powers pows;
run;

data outers;
set outers outs;
run;

%end;
%mend analysis;

%analysis;

/*the next steps clean the data and calculate sensitivity and specificity*/
data all_anal;
set all;
*if noest=1 then delete;
run;

data power_anal(drop=yesno);
set powers;
if id=. then delete;
in=yesno;
run;

data comp_power;
merge power_anal all_anal;
by id;
run;

data comp_power;
set comp_power;
if nopt=1 then delete;
if noest=1 then delete;
run;

data sensitivity;
set comp_power;
```

```
if yesno=1;
run;

data specificity;
set comp_power;
if yesno=0;
run;

proc print data=sensitivity;
run;

data specificity_no;
set comp_power;
if yesno=0 then do;no=1;end;
if yesno=1 then do;no=0;end;
if in=0 then do; out=1;end;
if in=1 then do; out=0;end;
run;

proc print data=specificity_no;
run;

proc print data=specificity_no;
run;

proc iml;
use sensitivity;
read all var{yesno} into yesno;
read all var{in} into in;
n=nrow(yesno);
ones=j(1,n,1);
sumyesno=ones*yesno;
sumin=ones*in;
sensitivity=sumin/sumyesno;
print sumyesno sumin sensitivity;
quit;

proc iml;
use specificity_no;
read all var {no} into no;
read all var {out} into notin;
n_out=nrow(no);
ones=j(1,n_out,1);
sumno=ones*no;
sumout=ones*notin;
specificity=sumout/sumno;
print sumno sumout specificity;
quit;
```

Appendix B.5: SAS Code for Chapter 5

Appendix B.5.1: SAS Code to Evaluate Study Design and Power of the Gold Standard

```
goptions colors=(black) htext=1.8 ftext=swiss;
libname ver 'C:\Sufficient Similarity Research\verification_code';

/*this allows for the input/construction of different dose groups which in
return
creates different study designs*/
/*
data dose;
  input dose;
  constant=0;
  cards;
0
0.275
1.096
2.740
9.042
13.70
18.084
27.400
;
run;
quit;
*/

/*
data dose;
  input dose;
  constant=0;
  cards;
0
9.042
18.084
27.400
;
run;
quit;
*/

/*
data dose;
  input dose;
  constant=0;
  cards;
0
1.096
```

```
13.70
27.40
;
run;
quit;
*/

/*
data dose;
  input dose;
  constant=0;
  cards;
0
0.275
9.042
18.084
;
run;
quit;
*/

data all;
run;

data all1;
run;

data anal;
run;

/*this macro simulates the dose-response study under different constrainst on
the number of dose groups and the number allocated to each dose group and
performs the
gold standard test*/

%macro simulate(size,dosegroups);
      %do _i_=1 %to 1000;

      footnote 'sample='&_i_;

data parms;
constant=0;
a= 0.25;
beta= -0.1113 ;
del= 1.8450;
run;

data new_ref;
merge parms dose;
by constant;
run;

data ref_generate&_i_;
set new_ref;
    seed=100597+&_i_;
```

```
      constant=0;
      curve=1;
                  do k=1 to &size;
              mu = a+(1-a)*exp(beta*(dose-del)*(dose>del));
             pact=mu+sqrt(0.0648)*rannor(seed);
              if pact<0 then do;pact=0;end;
              output;
            end;
run;

data new_cand;
merge parms dose;
by constant;
run;

data cand_generate&_i_;
set new_cand;
    seed=100598+&_i_;
      curve=2;
                  do k=1 to &size;
              mu = a+(1-a)*exp(beta*(dose-del)*(dose>del));
             pact=mu+sqrt(0.0648)*rannor(seed);
              if pact<0 then do;pact=0;end;
              output;
            end;
run;

data anal&_i_;
set ref_generate&_i_ cand_generate&_i_ ;
run;
quit;

data anal&_i_;
set anal&_i_;
if curve=2 then do; pact2=pact;end;
if curve=1 then do; pact1=pact;end;
run;

proc nlmixed data=anal&_i_ cov hess method=firo tech=truereg itdetails;
parms b=-.1  delta=1 su=0 to 0.1 by 0.005 s2=0 to .07 by .01;
a_term= 0.2;
g=1-a_term;
s2u=su*su;
    mu = a_term+g*exp(b*(1+u)*(dose-delta)*(dose>delta));
    estimate 'ed20' (log(.8)+b*delta)/(b);
    estimate 'ed50' (log(.5)+b*delta)/(b);
    random u ~ normal(0,s2u) subject=curve out=randomest&_i_;
    model pact ~ normal(mu,s2);
    predict mu out=pred&_i_;
    ods output parameterestimates=vars covmatparmest=covs;
run;
quit;
```

```
data covs;
set covs;
where parameter='b' or parameter='delta';
run;

/*produces box and confidence region*/
proc iml;
  use vars;
      read all var{estimate} where (parameter='b') into beta;
      read all var{estimate} where (parameter='delta') into delta;
      read all var{estimate} where (parameter='s2') into mse;
      read all var{estimate} where (parameter='s2u') into sigmah;
  use covs;
      read all var{b delta} into covs;

      aa=.25; * 25.21;
      g=1-aa;
      mu20=0.8#g+aa;
      mu50=0.5#g+aa;
      ED20=(log((mu20-aa)/g)/beta)+delta;
      ED50=(log((mu50-aa)/g)/beta)+delta;
      *print ed20 ed50;
      gomega=ed20//ed50;
      big_g=(-(beta**-2)*log((mu20-aa)/g)||1)//(-(beta**-2)*log((mu50-
aa)/g)||1);
      cov_gomega=big_g*covs*big_g`;
*     print covs cov_gomega;
      varinv=inv(cov_gomega);
      se_gomega = sqrt(vecdiag(cov_gomega));
      *print  cov_gomega varinv se_gomega;

** transformations to polar coordinates;
      call eigen(eval,p,varinv);
      lambda=diag(eval);
      lambdahalf=root(lambda);
      lamhalfinv=inv(lambdahalf);
      check=lambdahalf*lambdahalf;
      *print lambda p;
*     print lambda check;  ** should equal lambda;
      test=p`*varinv*p;  * test should equal lambda;
*     print lambda test;
      test2=p*p`;
*     print test2;  *test2 shoould equal the identity matrix;
      type1 = 0.05;
      bign=&size*&dosegroups*2;
      totp=2;
      s=nrow(gomega);
      f=finv(1-type1,s,bign-totp);
      r = (s*f)**(1/2);
      pi=constant('pi');
      twopie=2*pi;

      znew=j(2,1,0);
      gw=j(2,1,0);
  do theta=0 to twopie by .1;
```

```
    z1=r*cos(theta);
      z2=r*sin(theta);
      z=z1//z2;
      *print z;
      znew=znew||z;
      gw=gw||(gomega-p*lamhalfinv*z);
  end;
      gw=gw`;
    gw=gw[2:64,];
      label={"gw1" "gw2"};
      create gw1 from gw[colname=label];append from gw;
*     labelz={"z1" "z2"};
*     znew=znew`;
*     create znew from znew[colname=labelz];
*    append from znew;
quit;
*proc gplot data=znew;
*   plot z2*z1;
*   run;
*   quit;


symbol1 v=none i=join;
symbol2 l=1;
symbol3 l=2;
axis4 label=(a=90 'ED50 (mg/kg)');* order=(0 to 100 by 20);
axis5 label=('ED20 (mg/kg)') ;*order=(0 to 100 by 20);
data box;
      boxgw1=1.35; boxgw2=2.79; output;
      boxgw1=6.37;output;
      boxgw2=13.15;output;
      boxgw1=1.35;output;
      boxgw2=2.79;output;

data plotrefa;
      set gw1 box;

data plotref1a;
      set plotrefa;
      if _n_=1;
      run;

data plotref2a;
      set plotrefa plotref1a;
run;

proc means data=plotref2a max min noprint;
      var gw1 gw2;
      output out=maxmin max= max1 max2 min= min1 min2;
run;

data yesno;
      set maxmin;
      sample=&_i_;
      yesno=0;
```

```
      if min1>1.35 and max1<6.37 and min2>2.79 and max2<13.15 then yesno=1;
      if min1=0 and max1=0 and min2=0 and max2=0 then noest=1;
run;


data yesno1;
      set maxmin;
      sample=&_i_;
      yesno=0;
      if min1>1.73 and max1<5.97 and min2>3.63 and max2<12.51 then yesno=1;
      if min1=0 and max1=0 and min2=0 and max2=0 then noest=1;
run;


data all;
set all yesno;
drop _type_ _freq_;
if yesno=. then delete;
run;

data all1;
set all1 yesno1;
drop _type_ _freq_;
if yesno=. then delete;
run;

%end;
%mend;

%simulate(6,4);

/*cleans the data and calculates the power for the simulated given study
design*/
data all_anal;
set all;
if noest=1 then delete;
run;

proc means data=all_anal sum;
var yesno;
run;


proc freq data=all_anal;
tables yesno;
run;

proc means data=all sum;
var noest;
run;


data all_anal1;
set all1;
if noest=1 then delete;
```

```
run;

proc means data=all_anal1 sum;
var yesno;
run;


proc freq data=all_anal1;
tables yesno;
run;

proc means data=all1 sum;
var noest;
run;
```

Vita

Scott Marshall was born on October 26, 1979 in Long Beach, CA. He attended Farragut High School in Knoxville, TN and graduated in 1998. Scott Marshall then moved on to the University of Tennessee at Chattanooga where he graduated with a B.S. in Applied Mathematics in 2003. Scott continued his studies at the University of Alabama at Birmingham where he earned an M.S. in Biostatistics in 2006. He continued his studies in Biostatistics at Virginia Commonwealth University in Richmond, VA where he pursued his Ph.D. under the direction of Professor Chris Gennings, Ph.D. Scott served as a Trainee on a T32 Taining Grant sponsored by the National Institute of Environmental Health Sciences for three years. Upon successfully comleting his Ph.D. Scott will pursue a career in industry at BioStat Solutions in Mt. Airy, MD.