View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by VCU Scholars Compass

Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations

Graduate School

2010

# Bayesian and Frequentist Approaches for the Analysis of Multiple Endpoints Data Resulting from Exposure to Multiple Health Stressors.

Epiphanie Nyirabahizi
*Virginia Commonwealth University*

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Biostatistics Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/136

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

# Bayesian and Frequentist Approaches for the Analysis of Multiple Endpoints Data Resulting from Exposure to Multiple Health Stressors.

Epiphanie Nyirabahizi

A dissertation submitted to the Faculty of the

Virginia Commonwealth University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biostatistics

Chris Gennings, Co-Director

Edward L. Boone, Co-Director

Al M. Best

Paul J. Brooks

Nitai D. Mukhopadhyay

Roy Sabo

March , 2010

Richmond, Virginia

Keywords: Benchmark dose tolerable area, Multiple endpoints data, Benchmark dose tolerable region, Bayesian Hierarchical Models, MCMC, Bayesian Model averaging, Composite scores.

Copyright 2010, Epiphanie Nyirabahizi

# Bayesian and Frequentist approaches for the analysis of multiple endpoints data resulting from multiple stressors.

Epiphanie Nyirabahizi

(ABSTRACT)

In risk analysis, Benchmark dose ($BMD$) methodology is used to quantify the risk associated with exposure to stressors such as environmental chemicals. It consists of fitting a mathematical model to the exposure data and the $BMD$ is the dose expected to result in a pre-specified response or benchmark response ($BMR$). Most available exposure data are from single chemical exposure, but living objects are exposed to multiple sources of hazards. Furthermore, in some studies, researchers may observe multiple endpoints on one subject. Statistical approaches to address multiple endpoints problem can be partitioned into a dimension reduction group and a dimension preservative group. Composite scores using desirability function is used, as a dimension reduction method, to evaluate neurotoxicity effects of a mixture of five organophosphate pesticides (OP) at a fixed mixing ratio ray, and five endpoints were observed. Then, a Bayesian hierarchical model approach, as a single unifying dimension preservative method is introduced to evaluate the risk associated with the exposure to mixtures chemicals. At a pre-specified vector of $BMR$ of interest, the method estimates a tolerable area referred to as benchmark dose tolerable area ($BMDTA$) in multidimensional Euclidean plan. Endpoints defining the $BMDTA$ are determined and model uncertainty and model selection problems are addressed by using the Bayesian Model Averaging (BMA) method.

# Dedication

To my mother, my sons Arnel and Adellar.

To Al M. Best and his wife Barbara.

# Acknowledgments

I would like to take the opportunity to acknowledge my advisors, Dr. Chris Gennings and Dr. Edward L. Boone, for their contributions to the successful completion of my dissertation. I greatly appreciate the encouragement, consideration and belief in my ability to succeed they have shown me. I would not have made it without their guidance and support. I would like to thank Dr. Paul J. Brooks for his thoughtful suggestions on many important issues. I would like to thank Dr Al Best, Dr. Nitai Mukhopadhyay and Dr. Roy Sabo for their time and support.

Finally I would like to thank all my colleague students, all faculty members and staff in Biostatistics department.

*(This work was partially supported by NIEHS #T32ES007334.)*

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

In risk assessment of genotoxic carcinogens, it is considered that all levels of dose exposure can produce adverse effects. The assumption is based on the United States Environmental Protection Agency (USEPA, 1995) idea that a single molecule of genotoxicants may be sufficient to cause a DNA damage that eventually may result in the development of a tumor. Because of this assumption, cancer risk is treated as proportional to the dose. For agents that are not genotoxic, a threshold below which there is no significant biological effects is assumed. Classical risk analysis conducted in this scope consist of finding the $NOAEL$ (no observable adverses effect level) used to determine a point of departure(POD) for establishing tolerable exposures. The $NOAEL$ is usually derived from animal data, using an approach which is based on the statistical comparison of the mean response for each dose group against the control group. The $NOAEL$ is considered to be the highest dose for which the mean response does not differ significantly from the mean of the control. In general, an uncertainty factor is then applied to the $NOAEL$ for animal to man extrapolation, an uncertainty factor is also used to account for differences in sensitivity in the human population. These two factors are assigned a value of 10 (WHO, 1999). Thus, to obtain a guidance value, the $NOAEL$ is divided by a default value of $100 = 10\times10$. Using the $NOAEL$ to set standards for human exposure has been criticized in the literature (Kimmel and Gaylor, 1988; Leisenring and Ryan, 1992; USEPA, 1995) and the EPA adopted the benchmark dose ($BMD$) method introduced by Crump (1984). It consists of fitting a mathematical dose response to the data and the $BMD$ is the dose corresponding to a pre-specified response also called the benchmark dose response ($BMR$). The 95% lower bound on the $BMD$ estimate, or $BMDL$, has been accepted as a replacement of the $NOAEL$, (USEPA, 2000).

However, living objects are rarely exposed to one health hazard at a time, (Monosson, 2005). Exposure could be via food, water, or contaminated sites. Typical examples include mixtures of pesticides on/in the food that we consume, petroleum hydrocarbons, metals and metalloids in mining wastes materials. Specifically, health risk assessment of chemical mixtures can be complex and getting sufficient evidence-based data for proper evaluation can be expensive. In many cases, experimental data are available for single chemical exposure. But, the key features in risk assessment of chemical mixtures consist of understanding dose response relationship and interaction of agents in the mixture. The EPA recommends the use of a components approach, where the data for each individual chemical are combined in an additive(zero interaction) if there is no adequate data on a particular mixture (U.S. EPA 2000). In Chapter 2, we introduce a method that can be used to estimate the risk associated with exposure to a mixture of chemicals assuming additivity. In this Chapter, the question of whether single chemical exposure data have enough information for predicting the risk associated with the exposure to the mixtures, given chemicals in the mixtures do not interact is answered.

### Additivity

The term additivity (i.e., zero interaction) is based on the classical isobologram introduced by Fraser (1870-1871). Different authors including Loewe (1953), Berenbaum (1981, 1989), Wessinger (1976), later extended and reviewed the use of this method. Briefly, according to Berenbaum, if $E_i$ represents the concentration/dose of the $i^{th}$ component alone that yields a fixed response, and if $x_i$ represents the concentration/dose of the $i^{th}$ component in combination with the $c$ agents that yields the same response, if the chemicals combine in an additive fashion i.e., with zero interaction, then

$$\sum_{i=1}^{c} \frac{x_i}{E_i} = 1. \tag{1.1}$$

If the interaction index equals 1, then the chemicals in the mixture are said to interact additively or *additivity* is observed. When the left-hand side of (1.1) is less than 1, then the effect of the mixture is greater than additive and chemicals are said to interact synergistically, or *synergy* is observed; whereas when the left-hand side of (1.1) is greater than 1, the effect of the mixture is less than additive and the chemicals are said to interact antagonistically, or *antagonism* is observed.

In some toxicological studies, a single numeric value is often not sufficient to summarize the toxicity effects caused by the exposure. Several measurements may be required to attempt to describe a complex toxic effect like developmental, reproductive, neurobehavioral, clinical chemistry, organ weight, or pathology effects (e.g., Faes *et al.*, 2006; Moser, 2000; Crowell *et al.*, 2004; Reed *et al.*, 2004). These types of studies result in large amounts of data that poses a challenge in terms of statistical analysis and interpretation.

## 1.1 Analysis of Multiple endpoints studies in risk assessment

### 1.1.1 Dimension reduction method

For risk assessment purposes, in the presence of multiple endpoints data, some authors choose to use the analysis of the most sensitive endpoint as the base. In such a case, the researcher conducts separate and independent analysis on each endpoint. Guidance values can be derived following the analysis of the most sensitive endpoint, (Myers, 2001a; Reiss *et al.*, 2005; Moser *et al.*, 2005). Another alternative to performing statistical tests on several study endpoints is to combine multiple endpoints into one summary measure and perform one test on that summary statistic, (Moser *et al.*, 1995; Shih *et al.*, 2003; Coffey *et al.*, 2007). In Chapter 3, a composite scores method using desirability functions as described in Coffey *et al.*, 2007, is used to evaluate the toxicity effect of the exposure to a mixture of organophosphate pesticides (OP). The composite scores method is a dimension reduction tool flexible to allow combination of data of different types (e.g., continuous, categorical, binary).

### 1.1.2 Dimension preservative method

When dealing with multiple endpoints data, one can use multiple regression (Kleinbaum, Kupper, Muller, Nizam, 1998) or principal components analysis PCA (Johnson and Wichern, 2002). In risk assessment, the classical practice of using the analysis of the most sensitive endpoint as the base may fail to properly identify tolerable doses when multiple endpoints are measured from exposure to multiple sources of hazard. In fact, there may not be an endpoint which is the most sensible to all levels of hazards under study, or there may not be a single endpoint to consider as the most sensitive to all levels of hazard. The most sensitive endpoint practice may fall short to capture all dose levels that may cause harmful effects.

Let us assume five endpoints are measured as a result of simultaneous exposure to two chemicals or stressors (chem1 and chem2). Assume each endpoint is modeled separately and the $BMD$ is computed and plotted (Figure(1.1)). In Figure(1.1), let's designate five different areas by alphabetical letters A, B, C, D, and E. Let's apply the most sensitive endpoint practice to determine the area containing tolerable doses for both chemicals.

- At the $BMR_\alpha$ of interest, $BMD_\alpha$ of the first endpoint discards area D (Figure(1.1),1)
- $BMD_\alpha$ of the second endpoint discards area E (Figure(1.1),2)
- $BMD_\alpha$ of the third endpoint discards area C (Figure(1.1),3). By using the most

Figure 1.1: The tolerable region associated with the exposure to stressors: A demonstrating example

sensitive endpoints, area A and B are in the tolerable region. But plotting the $BMD_\alpha$ for the fourth endpoint shows that to use the analysis of the most sensitive endpoint method to determine the tolerable region may not be adequate in some circumstances like the one demonstrated here. In fact, (Figure(1.1),4) shows that, the entire tolerable region as determined by the most sensitive endpoint may not be tolerable considering the fourth endpoint.

- $BMD_\alpha$ of the fourth endpoint discards area B (Figure(1.1),4), but this endpoint is not the most sensitive neither for chem1 nor for chem2 as evidenced by the combination of three line segments that define the upper boundary for A.

The scenario demonstrated above shows that determining guidance values based on the most sensitive endpoint may fail in some circumstances like the one shown above. The scenario suggests that modeling endpoints simultaneously may borrow strength across endpoints and lead to a more accurate tolerable region. A method to simultaneously model multiple endpoints is introduced in Chapter 4 through the use of the Bayesian hierarchical model.

### Bayesian Framework

The difference in the philosophy behind the two statistical branches (Bayesian vs. frequentist) lays in model specification. The general problem to be considered is such that $\mathbf{y} = (y_1, ..., y_n)$ are observed data with probability distribution characterized by vector of unknown parameters $\boldsymbol{\theta}$. Interest is to draw inference on $\boldsymbol{\theta}$. For frequentists, unknown parameters, $\boldsymbol{\theta}$, are treated as fixed constants. Inference is based on repeatedly sampling from the distribution $f(\mathbf{y}|\boldsymbol{\theta})$ or equivalently the likelihood $L(\boldsymbol{\theta}|\mathbf{y})$, and to find the maximum likelihood estimate (MLE), the values $\hat{\boldsymbol{\theta}}^{\star}$ that maximize the likelihood function. Bayesian methods treat unknown parameters as random quantities that follow a certain distribution, a *prior* distribution, $\pi(\boldsymbol{\theta})$. The prior distribution summarizes available information about the unknown parameter before considering the data in hand. Inference about $\boldsymbol{\theta}$ is based on the *posterior* distribution which is the joint probability mass or density function of the prior distribution and the sampling distribution $p(\mathbf{y}|\boldsymbol{\theta})$. It is possible that a prior is *improper* (i.e., it does not integrate to 1 over its possible range). Such a prior may add to identifiability problems (Gelfand and Sahu, 1999) so many studies prefer to adopt minimally informative priors which are proper. The prior distribution may have unknown parameters called hyperparameters $\boldsymbol{\eta}$. If we assume that hyperparameters $\boldsymbol{\eta}$ are known, then the prior is written as $\pi(\boldsymbol{\theta}) \equiv \pi(\boldsymbol{\theta}|\boldsymbol{\eta})$. Inference about $\boldsymbol{\theta}$ is based on the posterior distribution:

$$\frac{p(\mathbf{y},\boldsymbol{\theta})}{p(\mathbf{y})} = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\theta}. \tag{1.2}$$

*Conjugate priors* lead to a posterior distribution belonging to the same family of distributions as the prior. In the case where the prior parameters $\boldsymbol{\eta}$ depend on unknown parameters, the Bayesian method quantifies the uncertainty in a next-stage prior distribution called the *hyperprior*. A third stage (and higher) prior distribution is possible if hyperprior parameters (or beyond) involve unknown quantities. The method of specifying a model over several levels is called *hierarchical modeling* (Carlin and Louis, 2008). The integral in (1.2) can be tedious due to high dimensionality. Markov chain Monte Carlo methods are popular methods to handle high-dimensional numerical integration.

## Sampling from the posterior distribution

The majority of Bayesian Markov Chain Monte Carlo (MCMC) computing is accomplished using one of two basic algorithms, the *Metropolis-Hastings(M-H)* algorithm (Metropolis *et al.*, 1953; Hastings, 1970) and the *Gibbs sampler* (Geman and Geman, 1984; Gelfand and Smith, 1990).

## Gibbs Sampling

Assume a model with $k$ parameters of interest, $\theta = (\theta_1, ..., \theta_k)$. Assume samples can be generated from each of the conditional distributions $p(\theta_i|\theta_{j\neq i}, \mathbf{y}), i = 1, ..., k$. Under mild conditions the collection of full conditional distributions uniquely determine the joint posterior distribution, $p(\theta|y)$, and hence all marginal posterior distributions $p(\theta_i|y), i = 1, ..., k$. Assume $\{\theta_2^{(0)}, ..., \theta_k^{(0)}\}$, are arbitrary starting values. The algorithm proceeds as follows:

For (t=1,..., T), repeat:

Step 1: Draw $\theta_1^{(t)}$ from $p(\theta_1|\theta_2^{(t-1)}, \theta_3^{(t-1)}, ..., \theta_k^{(t-1)}, \mathbf{y})$

Step 2: Draw $\theta_2^{(t)}$ from $p(\theta_2|\theta_1^{(t)}, \theta_3^{(t-1)}, ..., \theta_k^{(t-1)}, \mathbf{y})$

$\vdots$

Step k: Draw $\theta_k^{(t)}$ from $p(\theta_k|\theta_1^{(t)}, \theta_2^{(t)}, ..., \theta_{k-1}^{(t)}, \mathbf{y})$

It can be shown, under mild conditions, that the $k$-tuple obtained at the $t^{th}$ iteration, $(\theta_1^{(t)}, ..., \theta_k^{(t)})$ converges in distribution to the posterior distribution $p(\theta_1, ..., \theta_k|\mathbf{y})$.

### Metropolis Algorithm

When the prior $\pi(\boldsymbol{\theta})$ and the likelihood $f(\mathbf{y}|\boldsymbol{\theta})$ are not conjugates, some of the conditional distributions $p(\theta_i|\theta_{j\neq i}, y), i = 1, ..., k$ have no closed form. A rejection algorithm that requires a function proportional to the distribution to be sampled and requires a rejection step from a candidate density, namely, the Metropolis algorithm and Metropolis-Hastings extension are used. If the goal is to sample from the unnormalized posterior $p(\boldsymbol{\theta}|\mathbf{y}) \propto h(\boldsymbol{\theta}) \equiv f(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$. A proposal density $q(\boldsymbol{\theta}^\star|\boldsymbol{\theta}^{(t-1)})$ which satisfies $q(\boldsymbol{\theta}^\star|\boldsymbol{\theta}^{(t-1)}) = q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}^\star)$ is specified. Then for $t = (1, ..., T)$, repeat:

Step 1: Draw $\boldsymbol{\theta}^\star$ from $q(.|\boldsymbol{\theta}^{(t-1)})$

Step 2: Compute the ratio $r = h(\boldsymbol{\theta}^\star)/h(\boldsymbol{\theta}^{(t-1)}) = exp[log(h(\boldsymbol{\theta}^\star)) - log(h(\boldsymbol{\theta}^{(t-1)}))]$

Step 3: If $r \geq 1$, set $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^\star$;

if $r < 1$, set

$$\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^\star, \tag{1.3}$$

with probability $r$; and set

$$\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t-1)}, \tag{1.4}$$

with probability $1 - r$. Under mild conditions like those supporting the Gibbs sampler, a draw $\boldsymbol{\theta}^{(t)}$ converges in distribution to a draw from the true posterior density $p(\boldsymbol{\theta}|y)$

### Metropolis-Hastings

The M-H algorithm doesn't require the candidate density to be symmetric $q(\boldsymbol{\theta}^\star|\boldsymbol{\theta}^{(t-1)}) \neq q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}^\star)$. The acceptance ratio $r$ in Step 2 of the Metropolis algorithm is replace by

$r = \frac{h(\boldsymbol{\theta}^{\star})q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}^{\star})}{h(\boldsymbol{\theta}^{(t-1)})q(\boldsymbol{\theta}^{\star}|\boldsymbol{\theta}^{(t-1)})}$. Under mild conditions, a draw $\boldsymbol{\theta}^{(t)}$ converges in distribution to a draw from the true posterior density $p(\boldsymbol{\theta}|y)$ as $t \to \infty$.

## 1.2 Model uncertainty and Bayesian Model averaging

Scientists use multiple regression as a tool to determine the relationship between a response and a set of potential predictors. Often there are many candidate predictor variables, that may describe or predict the response of interest. In such a case, researchers need a proper way to search through all the possible models to determine an appropriate one to explain the relationship between the predictors and response. Common methods for performing model selection are Maximum adjusted-$R^2$, forward, backward and stepwise selection methods (Hocking 1976), Mallow's $C_p$, (Mallow 1973), Predicted REsidual Sum of Squares or PRESS, (Allen 1974), the Bayesian Information Criterion or BIC, (Schwarz 1978), Akaike's Information Criterion or AIC, (Akaike 1974) and many more. Some algorithms search for the simplest model or parsimonious model, others use criteria such as F-tests to allow covariates in a model. These methods select predictors to put in the model and not the appropriate model to use. The resulting model is considered as the true model. In such a case, the uncertainty associated with the model is ignored.

A method which has gained popularity in the literature to describe model uncertainty is Bayesian Model Averaging (BMA). It is based on using probabilistic arguments to determine the model and to average over the model space. This technique incorporates model uncertainty into the analysis using posterior model probabilities (see Kass and Raftery, 1995;

Madigan and York, 1995; Raftery, 1996; Raftery, Madigan and Hoeting, 1997; Hoeting, Madigan, Raftery and Volinsky, 1999).

### Bayesian Model Averaging

Many of the existing methods for model selection are designed to select a single model, which does not allow for model uncertainty. Hence, when the selection algorithm is completed, the model selected is treated as the true model. The researcher then treats the data as if it were generated by the selected model. This may lead to over-confident inferences and decisions. Bayesian Model Averaging (BMA), on the other hand, allows the researcher to account for model uncertainty and express this uncertainty in terms of probability (Madigan and Raftery, 1994; Kass and Raftery, 1995; Clyde, 1999; Hoeting *et al.*, 1999). Let $D$ be some data and let $\mathbf{M}$ be a model space. To determine the posterior probability of a model $P(M_i|D)$ we will use Bayes' Theorem. In order to use Bayes' Theorem we need to assign each model $M_i$ a prior probability $P(M_i)$

$$p(M_i|D) = \frac{p(D|M_i)p(M_i)}{\sum_{j=1}^{\mathbf{M}} p(D|M_j)p(M_j)}. \tag{1.5}$$

Hence, we need the quantity $p(D|Mi)$. This quantity can be obtained from:

$$p(D|M_i) = \int L(D|\theta_i, M_i)p(\theta_i|M_i)d\theta_i \tag{1.6}$$

where $L(D|\theta_i, M_i)$ is the likelihood of the data, $\theta_i$ is the vector of parameters of model $M_i$. Once $p(D|M_i)$ has been determined for all $Mi$, then we can average the models using the law of total probability. For any quantity of interest $\zeta$, the probability $p(\zeta|D)$ is given by:

$$p(\zeta|D) = \sum_{i \in \mathbf{M}} p(\zeta|D, M_i)p(M_i|D), \tag{1.7}$$

where

$$p(\zeta|D, M_i) = \int p(\zeta|D, \theta_i, M_i)p(\theta_i|D, M_i)d\theta_i \tag{1.8}$$

Notice that $p(\zeta|D)$ does not depend on a model. This is the *averaged* model.

The BMA algorithm can be tedious and time consuming to implement as the model space size can be very large since it grows as the number of covariates increases. Notice that $P(M_i|D)$ can be used to determine the model with the highest posterior probability if the goal is to select the best single model. However, the averaged model which incorporates the model uncertainty into the analysis can also be computed. It improves the validity of inferences and results based on the averaged model do not depend on the model selected but on the model space $\mathbf{M}$. Bayesian Model Averaging was used in Chapter 5 to account for model uncertainty in estimation of the risk associated with the exposure to mixtures of chemicals.

This dissertation includes five primary chapters that are generally written in manuscript format. Summaries of each chapter are provided in the following:

## 1.3 Prospectus

Chapter 2 introduces a method of estimating Benchmark doses for chemical mixtures and evaluates the method using a mixture of 18 PHAHs. Chapter 3 is concentrated on evaluating

Neurotoxicity of a Mixture of Five OP Pesticides Using a Composite Score and desirability functions. Chapter 4 focuses on using a Bayesian approach for estimation of a tolerable region in multiple endpoints and multiple hazards exposure. Chapter 5 addresses the problem of model uncertainty using Bayesian Model Averaging in estimating the tolerable area for multiple endpoints and multiple hazard exposure. Chapter 6 compares the analysis methods using organophosphate pesticides data as reported in Moser *et al.*, (2005), in Chapter 3 and in Chapter 4. Chapter 7 provides a conclusion and discussion of future work.

### 1.3.1 Benchmark doses for chemical mixtures: Evaluation of a mixture of 18 PHAHs.

Benchmark doses ($BMDs$), defined as doses of a substance that are expected to result in a pre-specified level of benchmark response ($BMR$), have been used for quantifying the risk associated with exposure to environmental hazards. The lower confidence limit of the $BMD$ is used as a basis for the point of departure (POD) in risk assessments, often with additional uncertainty factors included. This risk estimate, however, does not account for potential interaction of the substance with other chemicals included in human exposures. The present work developed and tested a methodology to estimate $BMDs$ for mixtures of chemicals at fixed mixing ratios. Comparisons were made to the $BMD$ under the assumption of additivity. Young female Long-Evans rats were dosed via gavage with 18 different polyhalogenated aromatic hydrocarbons [2 dioxins, 4 dibenzofurans, and 12 PCBs, including dioxin-like and non-dioxin-like PCBs], or a mixture with the ratio of the 18 chemicals based on environmental concentrations, for 4 consecutive days. Serum total T4 was measured via radioimmunoassay in samples collected 24 hours after the last dose. Analyses of these data (Crofton *et al.*, 2005) did not suggest departure from additivity in the low dose region of the mixture, implying that the $BMD$ for the mixture could be estimated via an additivity model based on single chemical data. Our objective was to evaluate this assumption. Three candidate nonlinear additivity models were considered and ranked with the AIC criteria. To calculate $BMD$, $BMR$ was selected to be a 5% or 10% shift in T4 (expressed relative to controls). The results of a Wald-type test revealed no statistically significant difference between the $BMD$ under additivity compared to that from the mixture data using a 5% $BMR$. We conclude that use of a $BMD$ for this mixture could be based on single chemical data in an additivity model.

### 1.3.2 Evaluating Neurotoxicity of a Mixture of Five OP Pesticides Using a Composite Score.

The evaluation of the cumulative effects of neurotoxic pesticides often involves the analysis of both neurochemical and behavioral endpoints. Multiple statistical tests on many endpoints can greatly inflate Type I error rates. Multiple comparison adjustments are often overly conservative leading to reduced power to detect effects of interest. Furthermore, identification of the most sensitive endpoint may be chemical dependent so that neurotoxicity may be most evident on a per animal basis by evaluating many endpoints. Use of a composite score focuses the inference and avoids inflated type I error rates. Coffey *et al.*, (2007) described the development of an overall score based on desirability functions for the many types of outcomes measured in neurobehavioral toxicology experiments. Our objective was to evaluate the neurotoxicity of a mixture of five pesticides (Moser *et al.*, 2005). In particular, the desirability functions for neurochemical (blood and brain cholinesterase activity) and behavioral (motor activity, gait score, tail-pinch response score) endpoints were calculated for single chemical (acephate, diazinon, dimethoate, chlorpyrifos, and malathion) and mixture dose response data, and a composite score of neurotoxicity was determined. Both an additivity model using single chemical data and a model for an environmentally-relevant fixed-ratio mixture were estimated. Focusing in the low dose region using 5%, 10% and 20% benchmark responses ($BMRs$), departure from additivity was found at the 20% $BMR$ with all five pesticides in the mixture; additivity was observed at the lower $BMRs$. This methodology is therefore useful in evaluating the overall neurotoxicity of pesticide mixtures. Using the same relevant mixing ratio but without the most common chemical, malathion, departure from additivity was not detected at all three $BMRs$. Finally, malathion significantly influenced the interaction of the remaining chemicals in the mixture. This methodology is therefore useful in evaluating the overall neurotoxicity of pesticide mixtures.

### 1.3.3 Bayesian approaches for estimation of tolerable region in multiple endpoints and multiple hazards exposure.

Determining benchmark dosages ($BMD$) is of interest to toxicologists and risk assessors. Often data come from experiments with multiple chemicals and multiple endpoints. Current methodology evaluates each chemical and endpoint separately resulting in multiple statistical tests consequently inflating Type I error rates. Methods to adjust for multiple comparisons such as Bonferroni correction are subject to reducing the power to detect effects of interest. We introduce a Bayesian approach for both multiple endpoints and multiple chemical data into a single unified model. This model is estimated using MCMC in WinBugs and R. Inferences on the endpoint model and chemical effects are done via Bayes' factors. Using this method, and the benchmark dose method, a 95% estimation of a Bayesian' tolerable dose region is computed.

### 1.3.4 Bayesian Model Averaging for estimation of tolerable area in multiple endpoints and multiple hazards exposure.

Risk assessors are frequently interested in estimating the dose associated with a pre-specified excess risk above the background. These estimates depend on the model used. Risk related endpoints are often estimated using animal toxicity studies where outcomes are modeled as a function of the dose considered and excess risk is determined from this dose-response model. In this setting, there often exist multiple dose response models that describe the data well and risk assessors may not have a prior reason, to prefer a given model over the other models considered. This problem becomes complex in the case where multiple endpoints are observed. In fact, there may not be a model that fits all the endpoints well. Model selection methods in regression do so by making decisions on which predictors should be in the model based on pre-existing criteria. Existing model selection methods are not adequate to select an appropriate model among multiple models. The method that has captured researcher's attention in the case of model uncertainty is the Bayesian model Averaging (BMA). Following the logic in Chapter 4 to determine the benchmark dose tolerable area or $BMDTA$, we propose the Bayesian Model averaging in estimating the tolerable area in the multiple endpoints and multiple hazards exposure setting. This model is estimated using MCMC in WinBugs, R, and with the help of Monte Carlo integration. A 95% estimation of a Bayesian Tolerable Area is computed.

### 1.3.5 Overall assessment of the analysis of organophosphate pesticides data using composite scores and Bayesian hierarchical modeling.

In toxicology, studies conducted to understand the health effect of stressors may require observations of many outcomes simultaneously. Furthermore, some studies involve assessing the simultaneous health effect of multiple stressors. Data resulting from these types of studies are multidimensional and pose a challenge in terms of methods of analysis. Some authors choose to analyze the most sensitive endpoint, others have considered models that combine multiple endpoints into a composite score and therefore avoid having to select the most sensitive endpoint. Statistical methods to analyze these types of data can be grouped into two main parts. The first part includes dimension reduction methods; those methods that collapse the data dimension (e.g., composite scores). The second part includes dimension preserving methods uses methods (e.g., multiple statistical testing, Bayesian hierarchical). In Chapter 6, we compare and contrast these methods from the two different groups as they were applied to OP data.

# Chapter 2

# Benchmark doses for chemical mixtures: Evaluation of a mixture of 18 PHAHs.

## 2.1 Introduction

The EPA is responsible for establishing guidance values (such as the acceptable daily intake (ADI)) governing exposure to hazardous environmental agents or chemical compounds. For noncarcinogenic agents, a toxicological point of departure (POD) is identified and forms the basis for derivation of the reference values for risk assessment. The POD may be defined as the highest dose at which no adverse effects are observed ($NOAEL$). The $NOAEL$ is usually derived from animal data, using an approach based on the statistical comparison of the mean response for each dose group against the control group. The $NOAEL$ is considered to be the highest dose for which the mean response does not differ significantly from the mean of the control. In general, uncertainty factors are then applied to the $NOAEL$ for animal to man extrapolation, and to account for differences in sensitivity in the human population. These two factors are assigned a value of 10 (WHO, 1999). Thus, to obtain a guidance value, the $NOAEL$ is divided by a default value of $100 = 10 \times 10$. Additional uncertainty factors may also be used.

Using the $NOAEL$ to set standards for human exposure has been criticized in the literature (Crump, 1984; Kimmel and Gaylor, 1988; Leisenring and Ryan, 1992; Allen *et al.*, 1994a; Barnes *et al.*, 1995; U.S.EPA, 1995). Some pitfalls associated with this approach for health risk assessment are summarized here:

1. The estimation of the $NOAEL$ is dependent on both the dose and the sample size. By definition, the $NOAEL$ has to be one of the experimental doses. Furthermore, since the $NOAEL$ is the highest dose for which there is no significant difference from background, the power of the experiment for demonstrating such differences is critical. For large sample sizes, it is possible to detect a significant difference that is not biologically meaningful; for small sample sizes, there may be insufficient power to conclude that a biologically meaningful difference is statistically significant.

2. The $NOAEL$ approach inefficiently uses available information since it ignores the dose response relationship.

3. In some studies, the $NOAEL$ cannot be found. If the toxicological data are insufficient,

particularly at low doses, the approach may fail to identify an $NOAEL$ even though one exists. If the chemical is considered to pose a risk at all levels of exposure to which the substance can realistically be controlled, the $NOAEL$ may be considered not to exist.

The above listed limitations identified with the $NOAEL$ approach have pushed researchers to consider the use of the benchmark dose ($BMD$) as an alternative approach in health risk assessment.

The $BMD$ method, introduced by Crump (1984), consists of fitting a mathematical model to dose response data. The $BMD$ is defined as the dose of a substance that is expected to result in a pre-specified level of benchmark response ($BMR$) and is estimated via a dose response model. The lower statistical confidence limit on the $BMD$ (the $BMDL$) has been proposed to replace the $NOAEL$ as a POD for determination of guidance values. This approach makes more efficient use of the data than the $NOAEL$ approach, because it incorporates information from the full dose-response curve in the estimation process and can be obtained even when the $NOAEL$ cannot be determined. In addition, the use of a lower confidence limit ($BMDL$) appropriately reflects the sample size of a study; larger studies tend to result in shorter confidence intervals and thus lower uncertainty. Although there is an increasing use of the benchmark dose approach in the health risk assessment of environmental contaminants in North America, European researchers have not yet embraced the method mainly because, in many cases, available data are from studies with a limited number of dose levels (Kuljus *et al.*, 2006).

### Chemicals and their mode of action.

Living objects are rarely exposed to one environmental contaminant at a time (Monosson, 2005). Exposure comes from many sources including food, water, and contaminated sites. Types of exposure include mixtures of pesticides on/in the food that we consume, disinfectant byproducts in drinking water, petroleum hydrocarbons, metals and metalloids in mining waste materials. But exposure to environmental chemicals is not necessarily associated with an increase in health risk. The key features in risk assessment of chemical mixtures consist of understanding the dose response relationship and joint action of agents in the mixture. In 1939, Bliss defined three main categories of joint chemical action that are still relevant today. Briefly, according to Bliss, chemicals can:

- Independently act and have different modes of action (i.e., different mechanisms). This mechanism is also termed response addition (U.S. EPA, 1986) or simple independent action (Finney, 1971). This form of joint action is noninteractive; chemicals in the mixture are not expected to affect the toxicity of one another and the combined toxicity can be predicted from knowledge of the independent chemicals.

- Have similar joint action, also termed dose addition or simple similar action (Finney, 1971). This form of similar joint action is noninteractive; the chemicals in the mixture do not affect the toxicity of one another and toxicity can be predicted with knowledge of the individual chemicals.

- Have synergistic or antagonistic action, where the toxicity of the mixture cannot be

> assessed from that of the individual chemicals but depends upon knowledge of their combined toxicity when used in different proportions.

Effects are antagonistic when the effect of one substance counteracts the adverse effect of another, such that exposure to the substances in combination has less effect than the effects of independent exposures. Effects are synergistic when the final effect observed is greater than the effects of separate exposure to each substance. For analysis purposes, the EPA recommends the use of a components approach, where the data for each individual chemical are combined in an additive manner if there are no adequate data on a particular mixture (U.S. EPA 2000).

### Additivity

The definition of additivity (i.e., zero interaction) used in the present analysis is given by Berenbaum (1981). It is based on the classical isobologram introduced by Fraser (1870-1871). Different authors including Loewe (1953), Berenbaum (1981, 1989), and Wessinger (1986), later extended and reviewed the use of this method. Briefly, according to Berenbaum (1981), let $E_i$ represent the concentration/dose of the $i^{th}$ component alone that yields a fixed response, and $x_i$ represent the concentration/dose of the $i^{th}$ component in combination with the $c$ agents that yields the same response, if the chemicals combine in an additive fashion i.e., with zero interaction, then

$$\sum_{i=1}^{c} \frac{x_i}{E_i} = 1. \tag{2.1}$$

When the left-hand side of (2.1) is less than 1, the effect of the mixture is greater than additive and chemicals are said to interact synergistically, or *synergy* is observed; when the left-hand side of (2.1) is greater than 1, the effect of the mixture is less than additive and the chemicals are said to interact antagonistically, or *antagonism* is observed.

### The ray design, departure from additivity

Various experimental designs have been used to support the estimation of the response surface of chemical mixtures. The classical approach is the factorial design where each level of every component is combined with all other components. However as the number of chemicals increase the experimental designs are infeasible. The ray design, described by Martin (1942), Mantel (1958), Finney (1964), and others, enables researchers to describe the relationship among multiple compounds using fixed mixing ratios of interest, is a method for assessing departure from additivity in a mixture of $c$ chemicals. Let the fixed-ratio ray for the mixture be defined by the mixing ratio, $\mathbf{a} = [a_1.a_2.a_3....a_c]$, where the $a_i$ represent the proportion of each chemical in the mixture with the constraint:

$$\sum_{i=1}^{c} a_i = 1. \tag{2.2}$$

Let $t$ represent the total dose of the mixture along the specified fixed-ratio ray. Hence, for any given total mixture dose along the mixing ray, $t$, the dose of the $i^{th}$ chemical in the mixture is $x_i = a_i t$, and the dose response of the mixture can be modeled as a function of total dose, allowing for the assessment of departure from additivity. In a study of a mixture of 18 chemicals, the mixing ratios are given in Table (2.1). We introduce new methodology for

Table 2.1: *Chemicals, dose range, and ratio of individual chemical in the mixture*

| *Chemical* | *Dose range*(mg/kg) | *Ratio*(*Total mass*) |
|---|---|---|
| $TCDD$ | $(0.0001, 10)$ | 0.000007 |
| $PCDD$ | $(0.003, 10)$ | 0.000007 |
| $TCDF$ | $(0.3, 100)$ | 0.000001 |
| $1 - PCDF$ | $(0.03, 100)$ | 0.000003 |
| $4 - PCDF$ | $(0.03, 90)$ | 0.000013 |
| $OCDF$ | $(0.1, 300)$ | 0.000032 |
| $PCB - 28$ | $(100, 90000)$ | 0.039237 |
| $PCB - 52$ | $(100, 90000)$ | 0.077523 |
| $PCB - 77$ | $(100, 30000)$ | 0.000988 |
| $PCB - 101$ | $(50, 30000)$ | 0.076814 |
| $PCB - 105$ | $(90, 90000)$ | 0.038282 |
| $PCB - 118$ | $(10, 10000)$ | 0.190302 |
| $PCB - 126$ | $(0.001, 100)$ | 0.000302 |
| $PCB - 138$ | $(100, 90000)$ | 0.190181 |
| $PCB - 153$ | $(100, 90000)$ | 0.190861 |
| $PCB - 156$ | $(10, 10000)$ | 0.006541 |
| $PCB - 169$ | $(1, 1000)$ | 0.000197 |
| $PCB - 180$ | $(100, 90000)$ | 0.188700 |

estimation of the $BMD$ for a mixture of $c$ chemicals. To demonstrate this new methodology, we reanalyze the mixture of 18 PHAHs described by Crofton *et al.*, (2005). We estimate the $BMD$ for the mixture as predicted under additivity; the $BMD$ is also estimated from the mixture data alone. The new methodology allows us to determine whether the additivity assumption is appropriate in the estimation of the $BMD$ for this particular mixture.

## 2.2 Methods

### Single chemicals and mixture data

As described in Crofton *et al.*, (2005), young female Long-Evans rats were dosed for 4 consecutive days via gavage with one or a mixture of 18 different polyhalogenated aromatic hydrocarbons: 2 dioxins, 4 dibenzofurans, and 12 PCBs, including dioxin-like and non-dioxin-like PCBs. Serum total T4 was measured via radioimmunoassay in samples collected 24 hours after the last dose. A mixture was custom synthesized with the ratio of chemicals based on concentrations found in breast milk, fish tissues and other sources of human exposure.

### Modeling the data and estimating $BMDs$

EPA suggests that the selection of the model to be used for $BMD$ calculation should be based on the ability of the model to describe the data (EPA 1995). We considered nonlinear sigmoid-shaped models of the form:

$$g(\mu;\omega) = \beta_0 + \sum_{i=1}^{c} \beta_i x_i, \tag{2.3}$$

where $g(\mu;\omega)$ is the link function that relates the doses of the chemicals to the mean response (McCullagh and Nelder, 1989), and $\omega$ is a vector of nonlinear parameters. Let $BMD_{add}$ represent the total dose of the mixture that is associated with the relevant $BMR$, calculated under the assumption of additivity of the chemicals in the mixture. Using the ray design, $BMD_{add}$ is a function of the mixing proportions and the single chemical $BMD_is$. In fact by setting $BMD_i = E_i$ and assuming additivity, equation (2.1) becomes:

$$\sum_{i=1}^{c} \frac{a_i t}{BMD_i} = 1; \tag{2.4}$$

where $t$ is the total mixture dose along the mixing ray, in this case, equivalent to the $BMD_{add}$, $BMD_i$ are the effective doses of each individual chemical that on their own produce the same effect as the mixture, and defines $a_i$ as the proportions of the corresponding individual effect doses present in the total mixture dose (see Table 1). Thus, the $BMD_{add}$, which is the mixture total dose under the additivity assumption, is given by:

$$BMD_{add} = \left[\sum_{i=1}^{c} \frac{a_i}{BMD_i}\right]^{-1}. \tag{2.5}$$

The $BMD_{add}$ is statistically compared to the $BMD_{mix}$, which is the total dose for the mixture at the given $BMR$, by testing the nonlinear hypothesis

$$H_0 : BMD_{add} = BMD_{mix}. \tag{2.6}$$

A Wald-type statistic can be used to test the nonlinear hypothesis, (Seber and Wild, 1989). The delta method can be used to estimate the variance of $BMD_{add}$ and $BMD_{mix}$ used in the estimation of the lower 95% confidence limits, $BMDLs$.

## 2.3 Application of the method to mixture of 18 PHAHs data

To illustrate our approach, three nonlinear sigmoid-shaped candidate models were considered namely, the Gompertz, logistic, and exponential. Model selection was based on goodness-of-fit as indicated by the Akaike's Information Criterion (AIC), (U.S. EPA 2000). The nonlinear function selected for the analysis was the logistic function. Using single chemical data and the mixture, we considered a nonlinear logistic function for decreasing dose-response curves of the form:

$$\mu_i = \alpha_i + \frac{\gamma_i}{1 + exp(-(\beta_0 + \sum_1^c \beta_i x_i))}, i = 1, ..., 19, \tag{2.7}$$

where $\alpha_i + \frac{\gamma_i}{1+exp(-\beta_0)}$ is the magnitude of the control response, $x_i$ is the dose for the $i^{th}$ chemical, $i = 1, ..., 18$ and $x_{19}$ is the total dose of the mixture, $\alpha_i$ is the maximum effect parameter for the $i^{th}$ chemical or mixture, $\gamma_i$ is the response range for the $i^{th}$ chemical or mixture, $i = 1, ..., 19$. $\beta_0$ is the unknown intercept parameter, and $\beta_i$ is the slope parameter for the $i^{th}$ chemical or mixture. For this model the link function is $g(\mu; \omega) = -log(\frac{\gamma+\alpha-\mu}{\mu-\alpha})$ where $\omega$ is given by $\omega = [\alpha, \gamma]$.

Model (2.7) was fit to single chemical and the mixture data simultaneously. For the sake of parsimony, a forward selection algorithm was used to find common estimates for the maximum effect parameter, $\alpha_i$, for groups of chemicals. Three groups were identified as (1) the mixture; (2) PCB105, PCB118, PCB138, PCB156; and (3) the remaining chemicals. Using the nonlinear logistic model in (2.7) where $\alpha_i + \frac{\gamma_i}{1+exp(-\beta_0)}$ was set to 100, and $\mu_0 =$

$BMR$

$$BMD_i = \frac{log\left[\frac{\mu_0-\alpha_i}{100-\mu_0}\right] - \beta_0}{\beta_i}, \tag{2.8}$$

and

$$BMD_{add} = \left[\sum_{i=1}^{18} \frac{a_i}{BMD_i}\right]^{-1}, \tag{2.9}$$

and

$$BMD_{mix} = \frac{log\left[\frac{\mu_0-\alpha_{mix}}{100-\mu_0}\right] - \beta_0}{\beta_{mix}} \tag{2.10}$$

A Wald statistic of the form:

$$W = \left(\frac{(B\hat{M}D_{add} - B\hat{M}D_{mix})^2}{Var(B\hat{M}D_{add}) + Var(B\hat{M}D_{mix}) - 2Cov(B\hat{M}D_{add}, B\hat{M}D_{mix})}\right). \tag{2.11}$$

was used to test (2.6), and the delta method was used to estimate the variance of the test statistic. The Wald statistic was compared to a F-distribution with $df = 1, 1341$. A quasi-Newton iterative algorithm (Proc NLMIXED in SAS; version 9.1) was used to estimate the dose-response curves.

## Results

The experimentally observed dose response data and additivity model predicted response for each single chemical or the mixture are shown on Figure (2.1). From these figures a decreasing relationship is evident between the mean response in T4 as single chemical doses increase. Note that the dose-response model for OCDF was reduced to background (100%) because the maximum effect parameter was not different from 100%. Figure (2.2) presents the dose-response curves for the mixture and the mixture as predicted under additivity. From Table

Figure 2.1: The dose-response curve for the each observed PHAH overlapped with predicted T4 under additivity

Figure 2.2: The dose-response curve for the each observed PHAH overlapped with predicted T4 under additivity

Figure 2.3: The dose-response curve for the each observed PHAH overlapped with predicted T4 under additivity

Figure 2.4: The dose-response curve (solid curve) for the mixture overlayed with the one of the mixture as predicted for T4 under additivity (dashed).

Table 2.2: *Test results for the hypothesis that the BMD for the mixture data or* $BMD_{mix}$ *values are equal to those predicted under the additivity model or* $BMD_{add}$ *using BMR of 5% and then 10%.*

| $BMR$ | $Parameter$ | $Estimate$ | $SE$ | $p-value$ | $BMDL$ |
|---|---|---|---|---|---|
| 5% | $BMD_{mix}$ | 165.82 | 32.17 | $< 0.001$ | 112.86 |
| | $BMD_{add}$ | 178.54 | 59.28 | $< 0.001$ | 80.96 |
| | $BMD_{add} - BMD_{mix}$ | 12.71 | 62.97 | 0.84 | $95\%CI : (-90.93, 116.37)$ |
| 10% | $BMD_{mix}$ | 275.94 | 52.96 | $< 0.001$ | 188.76 |
| | $BMD_{add}$ | 492.55 | 59.58 | $< 0.001$ | 394.48 |
| | $BMD_{add} - BMD_{mix}$ | 216.61 | 77.77 | 0.005 | $95\%CI : (88.59, 344.62)$ |

(2.2), the results of a Wald-type test revealed no statistically significant difference between the benchmark dose under the additivity assumption ($BMD_{add}$) compared to that from the mixture data ($BMD_{mix}$) in the low dose region using a $BMR$ of 5%, ($p$-value=0.84). However, at the $BMR$ of 10%, $BMD_{add}$ for the mixture as predicted by single chemicals was higher than that estimated from the mixture data (492.55 versus 275.94 mg/kg; $p$-value=0.0054). The general relationship is provided in Figure (2.2), where the dose-response curve for the mixture drops below that predicted under additivity (dotted line); but the curves stay similar in the low dose region. Predicted $BMD$ and $BMDL$ (95% lower one-sided confidence limit) from the single-chemical data in an additivity model or $BMD_{add}$, predicted $BMD$ and $BMDL$ from the PHAH mixture or $BMD_{mix}$, and their difference are shown in Table (2.2). At the $BMR = 10\%$, the $BMD_{mix}$ is significantly lower than the $BMD_{add}$ suggesting the $BMD_{add}$ underestimates the risk associated with the exposure to the mixture of 18 PHAHS.

## 2.4 Discussion

We introduce a method to estimate the $BMD$ for mixtures of chemicals under the assumption of additivity and have applied the method to the 18 PHAHs data described by Crofton *et al.*, (2005). Crofton demonstrated that the exposure to the 18 chemicals resulted in an additive effect in the low dose region. The ray design enables us to depict the relationship among the 18 PHAHs at a fixed mixing ratio, and to derive the $BMD_{add}$ as a function of the mixing proportions and single chemicals $BMD_i$, see equation (2.2). The $BMD_{add}$ was statistically compared to the benchmark dose from the mixture data or $BMD_{mix}$. Three candidate nonlinear models were considered and the Akaike's Information Criterion (AIC) was used to select the model that fits the best. The nonlinear logistic model was selected. The dose-response model for OCDF was reduced to the background (100%) and its maximum effect parameter was not different from 100%. The single-chemical and mixture data were modeled successfully using the nonlinear logistic model. A one sided lower confidence interval, $BMDL$, was obtained using the delta method to estimate the variance (see Table 2.2 for 95% CI). A Wald type test was used to compare the $BMD_{add}$ with the $BMD_{mix}$. The present work provides a methodology for researchers to estimate the risk associated with exposure to mixtures of chemicals. Single chemicals are used since all possible mixtures cannot be observed. For any specified mixing ratio, we can estimate $BMD_{add}$ which, in this case, is proven to be a good estimate of the observed $BMD_{mix}$ in the low dose region. Therefore the use of a $BMD$ for this mixture could be based on single chemical data in an additivity model.

Power and sample size have crucial roles in hypothesis testing and are determined at the study design level. Insufficient power may lead to accept the null hypothesis when it is false. In the case of testing departure from additivity at a given mixture point, additivity may be incorrectly claimed, and such a decision may lead to serious consequences. The problem of power and sample size calculation in toxicology studies has been discussed by Meadows-Shropshire *et al.*, (2005), who took the approach of computing sample size and power for testing departure from additivity at specific mixture points. Casey *et al.*, (2006) improved the method by introducing sample size and power calculation for detecting departure from additivity along multiple fixed-ratio rays simultaneously, which was proved to be appropriate for any hypothesis involving linear combinations of the model parameters. This work does not consider the power and sample size problem, however they can be considered in future work.

# Chapter 3

# Evaluating Neurotoxicity of a Mixture of Five OP Pesticides Using a Composite Score

## 3.1 Introduction

Toxicological studies generally gather observations from multiple endpoints on each subject. The large amount of data resulting from these types of studies constitute a challenge to statisticians in terms of methods of analysis and interpretation of the results. Researchers may perform multiple statistical tests, which on one hand inflate Type I error rates. On the other hand, adjustments for multiple testing such as the Bonferroni correction reduce the power of the test by making the test more conservative. Other methods (Holm's correction;

Hochberg's correction) have been shown to be less conservative but may still impact the power (Holm, 1979; Hocberg, 1988). Examples of studies with multiple endpoints include Moser *et al.*, (2005), who analyzed each endpoint independently, and Reiss *et al.*, (2005) who analyzed the most sensitive endpoint.

Alternatively, Coffey *et al.*, (2007) describe the development of an overall score based on desirability functions to use in the analysis of multiple endpoints data resulting from toxicology studies. Desirability functions were first introduced by Harrington (1965) in the engineering field to use in optimization of the quality of a manufactured product. Harrington's idea is based on evaluating the quality of a manufactured product by measuring multiple endpoints and using desirability functions to find the levels of the factors that optimize the overall quality as measured by the many endpoints (Derringer and Suich, 1980; Derringer, 1994).

The objective of this work is to apply the approach described by Coffey *et al.*, (2007), to data from five organophosphate pesticides (OP) described in Moser *et al.*, (2005), and to compare the results to an independent analysis of each endpoint as reported in Moser *et al.*, (2005). These authors evaluated the effect of exposure to an environmentally relevant fixed ratio of a mixture of five OPs, by performing separate analyses of each endpoint. The objective of each analysis was to test if the mixture effect can be estimated by single chemical data assuming the chemicals in the mixture combine in an additive fashion or if there is evidence of interaction (either synergy or antagonism).

Desirability scores as defined by Coffey *et al.*, (2007) (see Figure 3.1 ) for neurochemical (blood and brain cholinesterase ChE activity) and behavioral (motor activity, gait score, tail-pinch response score) endpoints were calculated for the single chemical and mixture data. The overall desirability function combines all five endpoints into a single toxicity index via the geometric mean. An additivity model from the single chemical data and a dose-response curve for the two fixed-ratio mixtures were simultaneously estimated using the toxicity index. The method of Casey *et al.*, (2005) could not be used to test additivity as the maximum effect for the mixtures were different from the maximum effect in the additivity model. The comparison of the $BMD$ under additivity to that estimated from the mixture data described in Chapter 2 is a test of additivity at the $BMR$ of interest. The benchmark doses ($BMD$) under additivity and for the observed mixtures data were calculated and compared using 5%, 10%, and 20% benchmark response from the mixture dose-response curve. The estimate for the $BMD$ of the mixture under additivity (with a one sided lower 95% confidence interval) provides exposure margins for the mixture at an environmentally relevant ratio using an index for overall neurotoxicity.

## 3.2 Methods

### 3.2.1 Experimental design

Data are described in Moser *et al.*, (2005). In short, behavioral measures (motor activity, gait score, and tail-pinch response) were evaluated in adult male Long-Evans rats at the

Figure 3.1: The form of desirability functions as predicted in Coffey *et al.*, (2007)

time of peak effect following an oral dose of a single chemical or mixture. Measurement for blood and brain cholinesterase were obtain from collected tissues. The mixing ratio for the full (with all five chemicals included) mixture was (0.040, 0.031, 0.002, 0.102, 0.825) for acephate (ACE), chlopyrifos (CPF), diazinon (DIA), dimethoate (DIM), and malathion (MAL), respectively. For the reduced ray, where MAL was removed from the mixture, the mixing ratio was (0.229, 0.011, 0.177, 0.583) for ACE, DIA, CPF, and DIM, respectively. Five concentrations and a vehicle control (0, 3, 10, 30, 60 and 120 mg/kg) of ACE were experimentally evaluated in 8 rats each (total 48 rats). Seven concentrations and a vehicle control (0, 5, 25, 50, 75, 125, 150 and 250 mg/kg) of DIA were experimentally evaluated in 16, 16, 16, 8, 16, 8, 8 and 8 rats, respectively (total 96 rats). Five concentrations and a vehicle control (0, 2, 10, 20, 30 and 50 mg/kg) of CPF were experimentally evaluated in 8 rats each (total 48 rats). Two concentrations and a vehicle control (0, 100 and 500 mg/kg) of MAL were experimentally evaluated in 7, 8, and 8 rats, respectively (total 23 rats). Five concentrations and a vehicle control (0, 5, 10, 25, 50 and 75 mg/kg) of DIM were experimentally evaluated in 8 rats each (total 48 rats). The fixed-ratio ray of the five pesticides in the mixture data reflected the relative dietary exposure estimates of the general population as projected by the US EPA Dietary Exposure Evaluation Model (DEEM).

### 3.2.2 Statistical methods

Desirability functions were used to transform observed measurements to unitless scores between 0 and 1. Scores depend on how the observed response is desirable. Different desirability functions were utilized depending on whether a given response was to be minimized, maximized, or given a target value (Coffey *et al.*, 2007). The individual scores were combined into a single composite score through the geometric mean. Standard statistical analysis are performed on the newly created composite scores variable.

Coffey *et al.*, (2007), collected questionnaire data from independent subject matter experts neurotoxicologists that were used to estimate the shapes of the desirability curves in Figure (3.1). For continuous endpoints (e.g., motor activity, brain and blood ChE activity), the basic shape of the function was determined by whether one is trying to maximize or minimize the response. The mathematical form of a maximizing and a minimizing desirability function used in this analysis are described by Shih *et al.*, (2003). Using Figure (3.1: D, E), each category of a categorical endpoints (e.g., gait and tail-pinch responses) was assigned a value between 0 and 1. For example, in the analysis conducted here, a gait score of 1 was assigned a value of $d$=0.95; a gait score of 2 was assigned a value of $d$=0.86; a gait score of 3 was assigned a value of $d$=0.65; and a gait score of 4 was assigned a value of $d$=0.37. Individual desirability scores were then combined using the geometric mean to arrive at a composite measure of the overall desirability, $D$, such that:

$$D = (d_1 \times d_2 \times, ..., \times d_k)^{1/k}, \tag{3.1}$$

where $k$ denotes the number of the responses, here $k$=5.

Table 3.1: *Demonstration of the calculation of the toxicity index for three rats from the control group and two mixture dose groups. The desirability score can be read from Figure (3.1) for the observed response values. The observed responses are transformed to % control values for motor activity, brain ChE and blood ChE activity. The toxicity index is the geometric mean of the desirability scores from the five endpoints.*

| | Control Group | | Mixture:55 $mg/kg$ | | Mixture:300 $mg/kg$ | |
|---|---|---|---|---|---|---|
| | RatID = 2149 | | RatID = 2243 | | RatID =2251 | |
| Endpoint | Observed | Desirability | Observed | Desirability | Observed | Desirability |
| Motor Activity | 212 | 0.90 | 221 | 0.89 | 122 | 0.60 |
| ( $\%control$) | (106.3) | – | (110.8) | – | (61.2) | – |
| BrainChE | 5.02 | 0.89 | 3.54 | 0.57 | 1.22 | 0.07 |
| ( $\%control$) | (90.5) | – | (63.8) | – | (22.0) | – |
| BloodChE | 0.48 | 0.91 | 0.05 | 0.22 | 0.06 | 0.23 |
| ( $\%control$) | (98.1) | – | (9.6) | – | (11.5) | – |
| Gait | 1 | 0.95 | 1 | 0.95 | 2 | 0.86 |
| Tail pinch | 4 | 0.88 | 4 | 0.88 | 2 | 0.63 |
| Toxicity Index | – | 0.90 | – | 0.62 | – | 0.35 |

### Construction of an additivity model

Assuming a quasi-likelihood framework, the threshold additivity model relating the doses of the chemicals under study to the mean through a link function, $g(\mu;\alpha)$, can be expressed for decreasing dose-response relationships as:

$$g\left(\mu_{add}\right) = \alpha + \left(\mu_0 - \alpha\right) exp\left(\beta_i(x_i - \delta_i)(x_i > \delta_i)\right), i = 1, \cdots, c \tag{3.2}$$

where $g(\mu_{add})$ is the specified link function of the response of interest, $x_i$ is the dose of the $i^{th}$ chemical $i = 1, \cdots, c$, $\mu_0$ is the mean of the control group and $\alpha$ is the maximum effect parameter, $\beta_i$ is an unknown parameter associated with the slope of the $i^{th}$ pesticide, $\delta_i$ is an unknown threshold parameter associated with $i^{th}$ chemical. In the case where the dose threshold parameter is outside the experimental region, the threshold additivity model is replaced by the corresponding smooth model:

$$g(\mu_{add}) = \alpha + (\mu_0 - \alpha) exp\left(\beta_i x_i\right), i = 1, \cdots, c \tag{3.3}$$

Using the single chemical dose-response curves, the total dose, $t$, of a mixture with mixing proportions $(a_1 : a_2 : a_3 : a_4 : a_5)$ (where $\sum_{i=1}^{5} a_i = 1$) associated with response $\mu_{add}$ is given by

$$t_{add} = \left[ \sum \frac{a_i}{\frac{log\left(\frac{g(\mu_{add})-\alpha}{\mu_0-\alpha}\right)}{\beta_i}} \right]^{-1} \tag{3.4}$$

When the observed response along the fixed-ratio mixture ray is more extreme than that predicted under the additivity model, then *synergy* is claimed; if the response is less extreme than that predicted under additivity, then *antagonism* is claimed, if the curves coincide, then *additivity* is claimed (Gennings *et al.*, 2002).

The mixture data along the fixed-ratio ray were fit to a similarly parameterized model:

$$g(\mu_{mix}) = \alpha_{mix} + (\mu_0 - \alpha_{mix})exp\left(\theta_{mix}(t - \delta_{mix})(t > \delta_{mix})\right), \tag{3.5}$$

where $\theta_{mix}$ is the slope of the mixture and $t$ is the total mixture dose. When the threshold is estimated outside of the experimental region, the following model is used:

$$g(\mu_{mix}) = \alpha_{mix} + (\mu_0 - \alpha_{mix})exp(\theta_{mix}t) \tag{3.6}$$

Models in (3.2) and in (3.5) are fit simultaneously to accommodate a common estimate for $\alpha$. At $BMR$ of interest these models are used to estimate $BMD_{add}$ under additivity and $BMD_{mix}$ from the mixtures respectively. Applying the method developed in Chapter 2, the

$$BMD_{add} = \left[\sum_{i=1}^{5} \frac{a_i}{BMD_i}\right]^{-1}, \tag{3.7}$$

where

$$BMD_i = \frac{log\left[\frac{\mu_0 - \alpha}{g_{add} - \alpha}\right]}{\beta_i}, \tag{3.8}$$

and

$$BMD_{mix} = \frac{log\left[\frac{\mu_0 - \alpha_{mix}}{g_{mix} - \alpha_{mix}}\right]}{\beta_{mix} + \delta_{mix}}, \tag{3.9}$$

where $maxint$ is the maximum intercept.

The $BMD_{add}$ is statistically compared to the $BMD_{mix}$, the total dose for the mixture, by testing the following hypothesis:

$$H_0 : BMD_{add} = BMD_{mix}. \tag{3.10}$$

**Evaluate the effect of malathion on the remaining components of the mixture.**

Malathion held the largest proportion of the mixture (i.e., 82.5%). To evaluate the effect of malathion on the dose-response relationship of the other four pesticides, a reduced ray was evaluated experimentally in which the remaining four pesticides were fixed at the same relative ratios as given in the full ray, i.e.,

$$\frac{a_{i(full)}}{a_{j(full)}} = \frac{a_{i(reduced)}}{a_{j(reduced)}}, \tag{3.11}$$

where $a_{i(full)}$ , and $a_{j(reduced)}$ denotes values associated with the proportions of the five-chemical mixture and in the reduced mixture (where malathion was removed), respectively. The effect of malathion in the mixture was tested by comparing the dose-response curves for the two rays while noting that $t_{reduced} = t_{full}(1 - a_5)$ for $a_5 = 0.825$, the proportion of malathion in the full mixture. The hypothesis of no malathion effect on the mixture can be formulated as follows:

$$H_0 = \begin{cases} \frac{\theta_{full}}{\kappa} = \theta_{reduced} \\ \delta_{full} = \delta_{reduced} \end{cases} \tag{3.12}$$

where $\kappa$ represents the proportion of the mixing ratio for the full ray associated with chemicals that remain in the reduced ray study. The Wald type test was used to test the hypothesis in (3.10) and (3.12) and compared to a F-distribution. A quasi-Newton iterative algorithm (Proc NLMIXED in SAS; version 9.1) was used to estimate the dose-response curves. The delta method can be used to estimate the variance of $BMD_{add}$ and $BMD_{mix}$ in estimation of the lower 95% confidence limits, $BMDL$.

## 3.3 Application and Results

Using the desirability curves taken from the fit of the questionnaire data from subject-matter experts (see Coffey *et al.*, 2007) and reproduced in Figure (3.1), the observed data are transformed into desirability scores between 0 and 1. For example, in Table (3.1), a rat in the 300 mg/kg dose group of the mixture had an observed motor activity of 122 counts, which was transformed to 61.2 percent control units. From Figure (3.1), a motor activity of 61.2% of control is associated with a desirability score of 0.60. The calculations of the other four desirability scores were obtained in a similar way. The geometric mean of these five values resulted in a toxicity index of 0.35. Calculations are also demonstrated for a rat from the control group and for a rat from a moderate dose group of 55 mg/kg (see Table 3.1). In these three rats, the toxicity index decreased as the dose of the mixture increased, indicating the toxicity increased with dose.

Profile plots of the desirability scores for the full mixture study are provided in Figure (3.2), and (3.3). In general, the control group and lowest dose group (10 mg/kg) of the mixture did not exhibit toxicity in any of the endpoints. However, as the dose increased to 55 mg/kg, BrainChE and BloodChE activity are affected; motor activity is also affected at

100 mg/kg and higher; but tail-pinch response and gait score are not highly affected until the dose of about 200 mg/kg and higher.

The toxicity index was calculated for each rat in the single chemical and mixture by taking the geometric mean of the individual desirability scores. The nonlinear exponential model was fit to these data and $BMD_{05}$, $BMD_{10}$, and $BMD_{20}$ were computed. The common maximum-effect parameter for the single chemicals is 0.23, and that for both full and reduced mixtures is 0.31. There is a dose threshold parameter for DIM (estimated at 1.72 mg/kg) and for the full and reduced mixtures (estimated at 4.14 mg/kg and 8.89mg/kg, respectively). We compared the $BMDs$ for the mixtures under the assumption of additivity to that estimated from the mixture data. For the full mixture, there was not a significant difference in the $BMD_{05}s$ or $BMD_{10}s$ from the mixture compared to that predicted by the single chemicals under additivity ($p$-value=0.75 and $p$-value = 0.33, respectively). However, $BMD_{20}$ for the mixture as predicted by single chemicals overestimated the MOE compared to that estimated from the mixture data (36.02 versus 22.33 mg/kg; $p$-value = 0.001). The general relationship is provided in Figure (3.3 (Full)).

For the reduced mixing ratio ray, there was no departure from additivity in the low dose region, as measured by the $BMD_{05}$ and $BMD_{10}$; however, a departure from additivity was observed at the $BMD_{20}$. In fact, the $BMD_{20}$ for the mixture data underestimated the $BMD_{20}$ of the mixture as predicted by single chemical data (5.51mg/kg versus 8.06mg/kg). The general relationship is provided in Figure (3.3 (Reduced)). See Table (3.3), Table (3.4), and Table (3.5) for estimates of the $BMD_{05}$, the $BMD_{10}$, and the $BMD_{20}$ respectively; on each single chemical and on the full and reduced mixtures, and their corresponding $BMDL$. To the hypothesis of no malathion effect on the mixture, there is evidence that malathion has an effect on the mixture ($p$-value $\leq 0.001$)

## 3.4 Conclusion: Comparing the results to Moser *et al.*, (2005)

In Moser *et al.*, (2005) a threshold additivity model, as described in Gennings *et al.*, (1997), or a generalized linear model was used to fit each endpoint separately. Furthermore, the effect of malathion, the least potent yet most prevalent chemical among the five chemicals, was evaluated in the mixture. In particular, as described by Casey *et al.*, (2004), the effect of a subset of chemicals in a mixture can be evaluated by experimentally evaluating the fixed ratio mixture with and without the subset and testing the appropriate hypothesis of no effect of the subset. The method was applied to a full mixture of the five OPs compared to a similar mixture (reduced mixture) with malathion removed. Using the likelihood ratio, a test of additivity was rejected for both mixtures the full and the reduced when considering the brain cholinesterase (ChE) endpoint. Appropriate comparison of the full and the reduced ray, Casey *et al.*, (2006), revealed that the two mixtures were not significantly different, $p$-value $= 0.421$, which suggested that malathion did not interact with the remaining four chemicals. The same conclusions were drawn when considering the motor activity endpoint. For the blood ChE endpoint, the study did not find evidences of departure from additivity

Table 3.2: *Parameter estimates using the nonlinear threshold additivity model.*

| Chemical | Parameter | Estimate | SE | $P$-value |
|---|---|---|---|---|
| ALPHA | $\alpha_{1,2,3,4,5}$ | 0.23 | 0.021 | $\leq .001$ |
| ACE | $\beta_1$ | $-0.034$ | 0.004 | $\leq .001$ |
| CPF | $\beta_2$ | $-0.032$ | 0.004 | $\leq .001$ |
| DIA | $\beta_3$ | $-0.009$ | 0.001 | $\leq .001$ |
|  | $\delta_4$ | 3.06 | 1.5 | 0.465 |
| DIM | $\beta_4$ | $-0.032$ | 0.003 | $\leq .001$ |
| MAL | $\beta_5$ | $-0.0003$ | 0.0001 | 0.0005 |
|  | $\alpha_{mix}$ | 0.31 | 0.014 | $\leq .001$ |
| MIX | $\beta_{mix}$ | $-0.017$ | 0.002 | $\leq .001$ |
|  | $\delta_{mix}$ | 9.4 | 2.393 | .001 |
| MIXred | $\beta_{mixred}$ | $-0.04$ | 0.006 | $\leq .001$ |
|  | $\delta_{mixred}$ | 14.12 | 1.71 | $\leq .001$ |

Table 3.3: *$ED_5$ estimates, their corresponding Confidence Intervals, and the p-value for the additivity test. There is no departure from additivity at 5%$BMR$ for both full and reduced mixtures; p-value* $= \mathbf{0.75}$ *and p-value* $= \mathbf{0.74}$ *respectively.*

| Parameter | Estimate | $BMDL$ | $P$-value |
|---|---|---|---|
| $ED_{1(05)}$ | 1.29 | 1.01 | |
| $ED_{2(05)}$ | 1.36 | 1.18 | |
| $ED_{3(05)}$ | 4.87 | 4.19 | |
| $ED_{4(05)}$ | 3.05 | 0.37 | |
| $ED_{5(05)}$ | 126.44 | 48.46 | |
| $ED_{add(05)}$ | 10.78 | 6.87 | |
| $ED_{mix(05)}$ | 11.97 | 7.13 | |
| $ED_{add(05)}$ - $ED_{mix(05)}$ | $-1.18(SE = 3.76)$ | $[-7.42, 5.05]$ | $\mathbf{0.75}$ |
| $ED_{addred(05)}$ | 2.53 | 1.30 | |
| $ED_{mixred(05)}$ | 2.14 | .65 | |
| $ED_{addred(05)}$ - $ED_{mixred(05)}$ | $0.38(SE = 1.79)$ | $[-1.55, 2.32]$ | $\leq \mathbf{0.74}$ |

Table 3.4: *$ED_{10}$ estimates, their corresponding Confidence Intervals, and the p-value for the additivity test. There is no departure from additivity at 10%$BMR$ for both full and reduced mixtures; p-value* $=$ $\mathbf{0.33}$ *and p-value* $=$ $\mathbf{0.35}$ respectively.

| Parameter | Estimate | $BMDL$ | $P$-value |
|---|---|---|---|
| $ED_{1(10)}$ | 2.62 | 2.07 | |
| $ED_{2(10)}$ | 3.00 | 2.43 | |
| $ED_{3(10)}$ | 10.85 | 8.58 | |
| $ED_{4(10)}$ | 4.61 | 1.41 | |
| $ED_{5(10)}$ | 258.26 | 99.13 | |
| $ED_{add(10)}$ | 19.36 | 14.05 | |
| $ED_{mix(10)}$ | 15.21 | 10.59 | |
| $ED_{add(10)}$ - $ED_{mix(10)}$ | $3.37(SE = 4.32)$ | $[-3.74, 10.49]$ | $\mathbf{0.33}$ |
| $ED_{addred(10)}$ | 4.36 | 2.83 | |
| $ED_{mixred(10)}$ | 3.20 | 1.80 | |
| $ED_{addred(10)}$ - $ED_{mixred(10)}$ | $1.16(SE = 1.25)$ | $[-.921, 3.23]$ | $\leq \mathbf{0.35}$ |

Table 3.5: *$ED_{20}$ estimates, their corresponding Confidence Intervals, and the p-value for the additivity test.There is evidence of departure from additivity at 20%BMR for both full and reduced mixtures; p-value =* **0.001** *and p-value =* **0.003** *respectively.*

| Parameter | Estimate | $BMDL$ | $P$-value |
|---|---|---|---|
| $ED_{1(20)}$ | 6.01 | 4.35 | |
| $ED_{2(20)}$ | 6.34 | 5.10 | |
| $ED_{3(20)}$ | 22.86 | 18.04 | |
| $ED_{4(20)}$ | 7.99 | 5.06 | |
| $ED_{5(20)}$ | 543.15 | 207.98 | |
| $ED_{add(20)}$ | 36.02 | 29.72 | |
| $ED_{mix(20)}$ | 22.33 | 19.56 | |
| $ED_{add(20)}$ - $ED_{mix(20)}$ | $14.45(SE = 4.11)$ | $[7.18, 21.71]$ | **0.001** |
| $ED_{addred(20)}$ | 8.06 | 6.48 | |
| $ED_{mixred(20)}$ | 5.51 | 4.18 | |
| $ED_{addred(20)}$ - $ED_{mixred(20)}$ | $2.54(SE = 1.21)$ | $[0.54, 4.54]$ | $\leq$ **0.003** |

Figure 3.2: Profile plots of the desirability scores for the full mixture

Figure 3.3: Profile plots of the desirability scores for the full mixture

Figure 3.4: Observed (asterisk), and fitted (solid line) mean responses along the full five-pesticide and reduced (where MAL was removed from the mixture) fixed-ratio rays.

but this study found evidence that malathion did interact with the remaining four chemicals. For the tail pinch endpoint, because of the irregularity shown in the data the results were left inconclusive or not reliable. For the gait score endpoint, the comparison between the reduced and full ray revealed a difference between the two mixtures suggesting that malathion interacted with the remaining four chemicals. There was no evidence of departure from additivity, $p$-value is 0.314 and 0.053 for the full reduced and full rays respectively. The findings are summarized in Table (3.6). Moser *et al.*, (2005) focused on a more full dose-response range; the present analysis considers, only the low dose range.

In conclusion, for the full mixture ray, in the low dose region (defined by the 5% and 10% $BMR$) there is not a significant difference between the $BMD$s under additivity and that estimated from the mixture data. However, the $BMD$ associated with a $BMR$ of 20% as predicted under additivity from single chemical data overestimated the BMD from the mixture data (36.02 versus 22.33mg/kg). For the reduced mixture ray, the mixture is significantly greater than additive ($p$-value $\leq$ 0.003). There is not a significant difference between the additive estimate of the $BMD$s and that estimated from the reduced ray mixture data in the low dose region estimated by $BMD_{05}$ and $BMD_{10}$. Overall, in Moser *et al.*, (2005), there were deviations from additivity for some endpoints but not others for both mixtures, see Table (3.6). Using a composite scores approach, both the 5-OP mixture and the 4-OP mixture were significantly different from additivity. There is evidence that malathion influences the chemicals remaining in the reduced ray. Chapter 4 develops a new methodology based on Bayesian hierarchical model as a unique unifying model to simultaneously model

Table 3.6: *OP data, independent analysis for each endpoint: Overall results as they were published in* Moser *et al.*, (2005)

| Endpoint | Full Ray | Reduced Ray | Malathion effect |
|---|---|---|---|
| Motor Activity | Synergy | Synergy | No |
| BloodChE | Synergy | Synergy | Yes |
| BrainChE | Synergy | Synergy | No |
| Tail pinch | Synergy | Synergy | No |
| Gait Score | Additivity | Additivity | Yes |

multiple endpoints resulting from multiple sources of exposure.

# Chapter 4

# Bayesian Approach for Estimating the Tolerable Region

## 4.1 Introduction

Toxicologists are concerned about the exposure of living organisms to potentially toxic materials and how to reduce the risk resulting from the exposure. The USEPA (Environmental Protection Agency) is responsible to set up guidance values governing exposure to stressors by establishing the Point Of Departure (POD) such as $NOAELs$ (No Observable Adverse Effect Level) and $BMDs$ (Benchmark dose). The $NOAEL$ is the classical method, usually derived from animal data using an approach which is based on the statistical comparison of the mean response for each dose group against the control group. The $NOAEL$ is considered to be the highest dose for which the mean response does not differ significantly from the mean of the control. A number of pitfalls associated with the $NOAELs$ have been identified. Researchers have pointed out the problems with a single numerical value adequately reflecting study size and the shape of the underlying dose-response curves (Crump 1984; Slob 1999). $NOAELs$ are not fixed attributes of toxic substances; rather, they reflect features of experimental design. Larger experimental studies will detect effects at lower exposures and thus will yield lower $NOAELs$ (Crump 2002, Scholze and Kortenkamp 2007). Because of the abovementioned problems, the benchmark dose ($BMD$) has been introduced as a statistical tool used to determine acceptable exposures to a stressor (Crump 1984). The $BMD$ is a dose that causes a pre-specified effect above the background or $BMR$ and is estimated by fitting a mathematical dose-response model to experimental data. EPA has accepted the replacement of $NOAELs$ by $BMDs$ whenever appropriate quantitative data are available (USEPA 1994).

For the univariate case where measurements are taken on one endpoint, $y_i$, resulting from one source of exposure with different dose level, $x_i$, the $BMD$ is estimated by fitting a mathematical dose response model relating the mean response to the exposure, pre-specifying an admissible extra risk $BMR_\eta$ above the background, and inverse map the relationship to get the corresponding dose.

Example: Let

$$y_i = f(\boldsymbol{\beta} x_i),$$

where $f$ is a link function that relates the mean-observations to exposure and $\boldsymbol{\beta}$ is a vector of unknown parameters. Then, at the pre-specified response of interest $\eta$ or $BMR_\eta$ above the background, the $BMD$ is $x^*$ that satisfies the relationship

$$BMD_\eta = f^{-1}(\eta) = \boldsymbol{\beta}' x^*$$

where $f^{-1}$ is the inverse mapping of the function $f$ or the link function, $\boldsymbol{\beta}'$ is the unknown parameter estimates in $\boldsymbol{\beta}$ that satisfy the equality. A quasi-Newton iterative algorithm (Proc NLMIXED in SAS) can be used to estimate the dose-response curve and the delta method may be used to estimate the variance of the estimates as described in Chapter 2.

In some studies it is important to consider multiple endpoints when assessing the risk (Ryan, 1992). Data resulting from these types of studies are large and pose a challenge in terms of statistical analysis and interpretation. The use of multiple statistical test method results in Type I error rate inflation. Available methods to correct for Type I error when performing multiple statistical test, such as the Bonferroni correction, trade the power of the test to detect the effect of interest. Some authors proposed methods of combining multiple endpoints in the same model or in one single composite index, others choose the analysis of the most sensitive endpoint. Samuel *et al.*, (1997) used a latent-variable model for mixed discrete and continuous correlated outcomes. Coffey *et al.*, (2007) introduced the use of an overall score based on desirability function methods to construct a composite index for analysis of multiple outcomes in toxicology studies. Reiss *et al.*, (2005) chose the analysis of the most sensitive endpoint. Coffey *et al.*, (2007) used generalized estimating equations with nonlinear models to combine mixed data types. Other authors have used pseudolikelihood estimation when combining continuous and ordinal outcomes to simplify the numerical challenges of using a joint density (see for example, Faes *et al.*, 2004, Geys *et al.*, 1999). Regan and Catalano (1999a), who evaluated the developmental effects of Ethylene Glycol (EG) by focusing on fetus malformation and fetal weight, proposed the joint risk assessment method for two outcomes considering separate $BMR$ for each of the two responses. They used generalized estimating equation methodology to account for correlations. Yu and Catalano (2005) evaluated the parathion neurotoxic effects and proposed a likelihood based model that allows separate dose-response models for each outcome while accounting for the bivariate correlation and overall characterization of risk.

The Bayesian hierarchical method has been proposed as a method to analyze data with multiple endpoints. Choi *et al.*, (2004) used Bayesian Hierarchical methods for multiple endpoint data resulting from exposure to Perchlorate. The study shows that exposure to perchlorate inhibits the uptake of iodide in the thyroid thereby causing a reduction in the hormones thyroxine (T3) and thriiodothyronine (T4), and an increase in thyroid stimulating hormones (THS) in blood. Faes *et al.*, (2006) used Bayesian hierarchical method on multiple endpoints data resulting from exposure to EG on mice. The developmental toxicity effect studied include fetal, low birth weight, and malformation(external, visceral or skeletal). They proposed a two stage Bayesian hierarchical structure where the first stage models the probability that a fetus is non-viable and the second stage models the probability that a survivor fetus has malformation. Both Choi (2004) and Faes (2006) used Bayesian hierarchical model to model the toxicity from the exposure to one single chemical/stressor and each endpoint is modeled separately. But an assessment from joint risk may encompass greater overall sensitivity than evaluating the risk associated with each individual endpoint separately.

## 4.2 Motivating Example

The problem of statistical analysis and interpretation for multiple endpoints data become more complex for studies interested in multiple endpoints resulting from multiple sources of stressors. EPA researchers are interested in understanding the neurotoxicity effect of organophosphates pesticides (OP) commonly used in agriculture. The pesticides considered here are acephate (ACE), diazinon (DIA), malathion (MAL), chlorpyriphos (CPF), and dimethoate (DMI). Long-Evans rats were orally exposed to either a single chemical dose or a mixture dose. The mixing proportions for the full mixture was (0.040, 0.031, 0.002, 0.102, 0.825) for ACE, CPF, DIA, DIM, and MAL, respectively. For the reduced mixture (with no malathion), the mixing proportion was (0.229, 0.011, 0.177, 0.583) for ACE, DIA, CPF, and DIM respectively. Five outcomes were evaluated at the time of peak effect, motor activity, tail pinch, and gait score, and then tissues were collected for measurement of ChE activity (blood and brain ChE activity). The doses considered are as follow: Five concentrations and a vehicle control (0, 3, 10, 30, 60 and 120 mg/kg) of ACE were experimentally evaluated in 8 rats each (total 48 rats). Seven concentrations and a vehicle control (0, 5, 25, 50, 75, 125, 150 and 250 mg/kg) of DIA were experimentally evaluated in 16, 16, 16, 8, 16, 8, 8 and 8 rats, respectively (total 96 rats). Five concentrations and a vehicle control (0, 2, 10, 20, 30 and 50 mg/kg) of CPF were experimentally evaluated in 8 rats each (total 48 rats). Two concentrations and a vehicle control (0, 100 and 500 mg/kg) of MAL were experimentally evaluated in 7, 8, and 8 rats, respectively (total 23 rats). Five concentrations and a vehicle control (0, 5, 10, 25, 50 and 75 mg/kg) of DIM were experimentally evaluated in 8 rats each (total 48 rats).

In this report, we used data for two chemicals, ACE and DIA, to demonstrate our method. A two chemical study was chosen to accommodate graphical illustration of the methodology. The choice of which chemicals was arbitrary as all five chemicals have enough data in the region of interest (low dose region) except malathion which has only two concentrations and a control group. This work only considered four endpoints (BloodChE and BrainChE, motor activity and tail pinch). Figures (4.1) and (4.2) show each endpoint versus the doses of the two chemicals considered. We chose to model the probability of no tail pinch to achieve a mixed response behavior, i.e, as the doses increase some curves increase and others decrease. Later in Chapter 6, the method is applied to all five chemicals. That is, although the mixture data included 5 chemicals, the analysis considered the data to be a function of the 2 selected chemicals. All the data are described in Moser *et al.*, (2005).

Using the most sensitive endpoint as a method to determine the point of departure, each endpoint is separately modeled by a simple exponential model of form

$$\mu_i = exp(\beta_0 + \beta_1 x_i) \tag{4.1}$$

and the $BMDL_{50}$ computed. The delta method is used to estimate the variance of the estimate. A quasi-Newton iterative algorithm (Proc NLMIXED in SAS; version 9.1) was used for estimation. In Table (4.1), BloodChE is the most sensitive endpoint because of its

Figure 4.1: Endpoints versus single chemicals: BrainChE and Tail pinch.

Figure 4.2: Endpoints versus single chemicals: BloodChE and Motor Activity.

Table 4.1: $BMDL_{50}$ *resulting from an independent analysis of each endpoint resulting from exposure to DIA or ACE*

| Endpoint | DIA | ACE |
|---|---|---|
| $BrainChE$ | 111.51 | 21.73 |
| $BloodChE$ | 7.92 | 23.79 |
| $MotorActivity$ | 131.23 | 24.33 |
| $Tailpinch$ | 160.44 | 56.44 |
| $min(\text{Endpoint}_i)$ | 7.92 | 21.73 |

lower $BMDL_{50}$ values considering exposure to DIA. Whereas, BrainChE is the most sensitive considering exposure to ACE. In both cases, tail pinch is the least sensitive endpoint with higher values for $BMDL_{50}$. To determine guidance values, such as point of departure (POD), the most sensitive endpoint method can be used. In this case, the POD for DIA is 7.92 mg/kg and 21.73 mg/kg for ACE and the shaded area in Figure (4.3) shows where all dose combinations considered to be tolerable for exposure are located considering independent analysis and the most sensitive endpoint. In figure (4.4) tolerable doses are shown considering combination effect, i.e, assuming additivity.

Determining POD or tolerable doses in risk assessment as explained above may results in unacceptable levels of toxicity, especially if the correlation that may exist among multiple measurements taken on the same subject is ignored, and simultaneous exposures

Figure 4.3: Tolerable region as defined by the $BMDL_{50}$ obtained using the analysis of the most sensitive endpoint method and assuming independent endpoints.

Figure 4.4: Tolerable doses considering combination effect: From inside out, BloodChE, BrainChE, Motor Activity, and Tail pinch.

are considered to be independent of exposures.

This work proposes a single unifying Bayesian hierarchical model to evaluate dose-response relationships among multiple health endpoints that may be correlated resulting from multiple exposures. Our method allows for determining a Benchmark Dose Tolerable Area ($BMDTA$) which is the area covering all possible dose combinations considered to be tolerable for exposure at a given $BMR$ of interest. The method is based on Bayesian hierarchical models to simultaneously model all four endpoints. Simultaneous modeling borrows strength across endpoints by allowing correlation between endpoints and may lead to a more conservative tolerable region. Furthermore we can determine which endpoints are sensitive,in terms of which endpoints bound this area.

## 4.3 Method

Let $y_{ij}$ be the response of the $i^{th}$ subject on the $j^{th}$ endpoint. For $J$ endpoints this can be formed into the vector $Y_i = (y_{i1}, ..., y_{iJ})$ for all $n$, the total number of subjects. Let $K$ be the number of stressors considered, and let $x_{ik}$ be the $k^{th}$ stressor on the $i^{th}$ subject, where $k = 1, 2, .., K$. This can be formed into a vector $X_i = (x_{i1}, ..., x_{iK})$. The relationship between the $x_{ik}$, $y_{ij}$, and a function $f = (f_1, ..., f_J)$ can be formed where each $f_j$ is an invertible function. Let $\boldsymbol{\beta}_j$ denote a vector of parameters corresponding to the $j^{th}$ endpoint. Hence the model can be formed as:

$$Y_i = f(X_i\boldsymbol{\beta}) = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iJ} \end{pmatrix} = \begin{pmatrix} f_1(X_i\boldsymbol{\beta}_1) \\ f_2(X_i\boldsymbol{\beta}_2) \\ \vdots \\ f_J(X_i\boldsymbol{\beta}_J) \end{pmatrix}. \tag{4.2}$$

This formulation allows for each endpoint to have its own model with parameters $\boldsymbol{\beta}_j$. For consistency, in this formulation we have all stressors $X_i$ in each of the endpoint models. For a given $BMR\eta = (\eta_1, ..., \eta_J)$ the inveribility of the $f_j$ allow for simultaneous determination of the $BMD$ or tolerable region, $T$, using the following inequality:

$$f^{-1}(\eta) = \begin{pmatrix} f_1^{-1}(\eta_1) \\ f_2^{-1}(\eta_2) \\ \vdots \\ f_J^{-1}(\eta_J) \end{pmatrix} \geq \begin{pmatrix} \beta_1' \\ \beta_2' \\ \vdots \\ \beta_J' \end{pmatrix} \begin{pmatrix} \mathbf{X}^* \end{pmatrix}, \mathbf{X}^* \geq 0, \tag{4.3}$$

where $\beta_j$ are the estimates of parameters $\beta$s in the model (4.2). Any stressors combination $X^*$ that satisfies the above inequality is considered tolerable with respect to the $BMR\eta$.

### 4.3.1 Bayesian structure

#### Likelihood

The likelihood depends on the type of data corresponding to each endpoint. Since the data may be of differing mixture types (discrete and continuous) the likelihood should be appropriate to the endpoint. Let $g_j(y_{ij}|f_j(X_i\beta_j), \gamma_j)$ be the probability distribution for each endpoint $j = 1, ..., J$) where $\gamma_j$ are additional parameters that may be necessary to define the distribution. Given $\beta_j$, the endpoints are conditionally independent and hence the likelihood can be constructed as:

$$L(Y|\beta X) = \prod_{i=1}^{n}\prod_{j=1}^{J} g_j(y_{ij}|f_j(X_i\beta_j), \gamma_j). \tag{4.4}$$

While the likelihood does not explicitly incorporate a correlation structure, the hierarchical nature of the model captures the possible correlations among observations through the $\boldsymbol{\beta}$'s that are allowed to be correlated.

### Prior specification

Proper prior distributions are employed to ensure the resulting posterior distribution is proper (i.e. $\int p(\theta|D)d\theta = 1$ where $\theta$ represents all the parameters in the model). The following prior distributions are used:

$$\begin{aligned}\boldsymbol{\beta} &\sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})\\ \boldsymbol{\Omega} &\sim Wishart(R, \rho)\\ \gamma_j &\sim p(\gamma_j)\\ \boldsymbol{\mu} &\sim N(a, A)\end{aligned}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ are the mean vector and precision matrix, respectively, for $\boldsymbol{\beta}$. Here R is specified as the $J(K+1) \times J(K+1)$ identity matrix. The prior distribution for $\gamma_j$, $p(\gamma_j)$ is an appropriate distribution and $a$ and $A$ are the hyperparmaters governing $\boldsymbol{\mu}$. To fit the model Markov Chain Monte Carlo techniques can be used via WinBUGS or OpenBUGS.

The samples generated can be analyzed in R to ensure convergence of the chains and quality of the samples. For more on MCMC methods and diagnostics see Gelman *et al.*, (2005).

### 4.3.2 Evaluation of the Benchmark dose Tolerable Area

At $BMR_\eta$ of interest, the inequality in (4.3) defines $T_J$ tolerable regions where $J$ corresponds to the $j^{th}$ endpoints. To define the tolerable region for all $J$ endpoints, is to find the intersection of all $T_J$ tolerable regions; $T = T_1 \cap T_2 \cap \cdots \cap T_J$. The $T$ defines the benchmark tolerable region $BMDT$, the analogue to the tradition $BMD$. The $T$ is used to evaluate the benchmark dose tolerable area analogue to the traditional $BMDL$. Using MCMC samples, for each MCMC sample there are $T^{(m)}$ tolerable regions where $m$ corresponds to the $m^{th}$ MCMC sample. $A_{T^{(m)}}$, the tolerable area for each MCMC sample is evaluated as:

$$A_{T^{(m)}} = \int_{T^{(m)}} d\mathbf{X}$$

the tolerable area can be ranked and a lower bound on the tolerable area is defined by the $\beta$ MCMC sample that correspond to the $q^{th}$ quantile of the $A_{T^{(m)}}$. This gives a $100 \times q\%$ credible region for the lower bound for the $BMDTA$. This is analogous to the traditional $BMDL$.

### 4.3.3 Hyper-sensitive, co-sensitive, hypo-sensitive endpoints, and feasible probabilities

A given endpoint $i$ is said to be *Hyper-sensitive* if the following condition is met:

$$\cap_{j=1}^{J} T_j = T_i. \tag{4.5}$$

Likewise, we define *Hypo-sensitive* endpoints, as those that satisfy the following condition:

$$\cap_{j \neq i}^{J} T_j = \cap_{j=1}^{J} T_j. \tag{4.6}$$

For $Z$, a set of endpoints such that

$$Z = \{j | j \in J\} \tag{4.7}$$

endpoints in $Z^*$, a subset of $Z$ are *co-sensitive* if

$$\cap_{j \in Z^*} T_j = \cap_{j=1}^{J} T_j, \tag{4.8}$$

Briefly, a *Hyper-sensitive* endpoint is when one endpoint defines the tolerable region, *Hypo-sensitive* endpoint is when an endpoint has no role in defining the tolerable region, and when more than one endpoints define the tolerable region, those endpoints defining the tolerable region are said to be *co-sensitive*.

In addition to determining a credible region, the MCMC samples can be used to determine which endpoints define the tolerable area and their corresponding probabilities.

This can be achieved using the GNU Linear Programming Kit (GLPK) available as a contributed package in R. The GNU Linear Programming Kit (GLPK) is a library of routines that use simplex algorithms, among others, to optimize (maximize or minimize) linear programming problems of the form:

$$z = C^T X \tag{4.9}$$

subject to:

$$\mathbf{A}x \leq b, LB \leq x \leq UB, \tag{4.10}$$

where $z$ is the objective function, $C$ is a vector of objective function coefficients, b is a vector of each constraint in the constraint matrix, UB and LB are upper and lower bounds on variables $x$. In geometric context, each inequality in (4.10) specifies a half-space in multidimensional Euclidean space, and their intersection region which is a polytope is the set of all feasible values the variables can take. The resulting region is a polytope and the goal is to determine which endpoints bound the polytope. To do so, let's construct a linear program (LP) as follows:

- First set $C = 0$
- Second, each inequality in 4.10 corresponds to

$$(\boldsymbol{\beta}'_j)^T X \leq f_j^{-1}(\eta_j), j = 1, ..., J \tag{4.11}$$

- Third, for each endpoint in turn, define linear programming $LP^j$ by setting the $j^{th}$

inequality to equality

$$(\boldsymbol{\beta}'_j)^T X = f_j^{-1}(\eta_j) \tag{4.12}$$

for j= 1,...,J and check for feasibility.

If $Lp^j$ is feasible, then the endpoint defines or bounds the polytope. Apply the routine to MCMC samples. Let $n_s$ be the number of MCMC samples, and let $Z$ be the set of endpoints that bounds the polytope. The feasible probability of each endpoint can be estimated by $\hat{P}_j$:

$$\hat{P}_j = \frac{\sum_{\iota=1}^{n_s} I_{E_{j\iota} \in Z}(E_{j\iota})}{n_s}, \tag{4.13}$$

where $E_{j\iota}$ designates the $j^{th}$ endpoint for the $\iota^{th}$ sample, $I$ is an indicator function that takes values 1 if the endpoint $j$ bounds the polytope and 0 otherwise. The feasible probability for an endpoint, estimated by $\hat{P}_j$, is the importance of that endpoint to define the polytope. If the feasible probability for a given endpoint is estimated to be 1, then corresponding endpoint is inevitable in defining the polytope. If on the other hand the feasibility probability for a given endpoint is estimated to be 0, then corresponding endpoint does not have any role in defining the polytope hence, the endpoint does not need further consideration.

## 4.4 Simulated data and results

To illustrate our method we simulated data with five endpoints resulting from two stressors. Our goal is to find a tolerable area corresponding to a pre-specified $BMR$. Furthermore we would like to find the feasible probability of each endpoint. A dataset of $n = 100$ observations is simulated.

Let, $x_{i1} \sim Unif(0,1)$, $x_{i2} \sim Unif(0,1)$ and let:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \end{pmatrix} = \begin{pmatrix} \exp(-6x_{i1} - 1x_{i2}) \\ \exp(-8x_{i1} + 0x_{i2}) \\ \exp(0x_{i1} - 8x_{i2}) \\ \exp(-7x_{i1} - 7x_{i2}) \\ \exp(-1x_{i1} - 6x_{i2}) \end{pmatrix} + \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \epsilon_{i3} \\ \epsilon_{i4} \\ \epsilon_{i5} \end{pmatrix} \tag{4.14}$$

where $\epsilon_{ij} \sim N_{(0,0.05)}$.

Notice that among the five endpoints, $Y_1$, $Y_4$, and $Y_5$ are affected by both stressors, and exhibit a decreasing relationship with respect to both stressors. Endpoint $Y_2$ is only affected by stressor 2 and $Y_3$ is only affected by stressor 1.

The hierarchical structure is specified as follow:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \end{pmatrix} \sim \begin{pmatrix} N\left(\exp(\beta_{11}x_{i1} + \beta_{12}x_{i2}), \tau_1\right) \\ N\left(\exp(\beta_{21}x_{i1} + \beta_{22}x_{i2}), \tau_2\right) \\ N\left(\exp(\beta_{31}x_{i1} + \beta_{32}x_{i2}), \tau_3\right) \\ N\left(\exp(\beta_{41}x_{i1} + \beta_{42}x_{i2}), \tau_4\right) \\ N\left(\exp(\beta_{51}x_{i1} + \beta_{52}x_{i2}), \tau_5\right) \end{pmatrix} \tag{4.15}$$

To fully specify the model the following prior distributions are used:

$$\begin{aligned} \boldsymbol{\beta} &\sim N(\boldsymbol{\mu}, \boldsymbol{\Omega}) \\ \boldsymbol{\Omega} &\sim Wishart(I_{10}, 10) \\ \boldsymbol{\mu} &\sim N(0, 100) \end{aligned}$$

$$\tau_j \sim Gamma(1,1)$$

where $I_{10}$ is the $10 \times 10$ identity matrix. While the prior distributions are proper they should be relatively vague. This formulation leads to the following posterior distribution:

$$\begin{aligned} P(\boldsymbol{\beta}, \boldsymbol{\Omega}, \boldsymbol{\mu}, \boldsymbol{\tau}|Y) &\propto e^{-\sum_{j=1}^{5}\tau_j^2} \times e^{-\frac{1}{200}\sum_{j=1}^{5}\mu_j^2} \\ &\times \frac{|\boldsymbol{\Omega}|^{(-5)} e^{-trace(I\boldsymbol{\Omega}^{-1})/2}}{2^{50}\Gamma_{10}(5)} \\ &\times |\boldsymbol{\Omega}|^{-1/2} e^{-\frac{1}{2}(\beta-\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\beta-\boldsymbol{\mu})} \\ &\times \left(\prod_{j=1}^{5}\tau_j^2\right)^{-n/2} e^{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{5}\tau_j^{-2}(Y_{ij}-e^{x_i\beta_j})^2}. \end{aligned}$$

**Results**

WinBUGS generated 4 chains of 1,100,000 MCMC samples from the posterior distribution. Convergence of the chains is investigated. The first 1,000 samples from each chain were discarded as burn-in samples. The remaining 1,099,000 samples were thinned by 10 to minimize autocorrelation in the samples, which resulted in 109,900 samples from each chain. The 109,900 samples from the first chain were used to draw inferences. Computation took 7 hours on The 109,900 samples from the first chain were used to draw inferences.

Table 4.2 shows the true value for $\beta_{ik}$ and the 2.5%, 50% and 97.5% quantiles as well as the $\hat{R}$ value for each parameter. All of the posterior credible intervals captured the true value that was used to generate the data. Hence the use of WinBUGS is reasonable for this problem. Furthermore, the estimation of this model is feasible.

Table 4.2: The quantiles of the posterior samples of parameters for simulated dataset. Based on 109,900 posterior samples.

| Endpoint | Parameter | True | 2.5% | 50% | 97.5% | $\hat{R}$ |
|---|---|---|---|---|---|---|
| $y_1$ | $\beta_{11}$ | -6 | -7.768 | -5.547 | -4.185 | 1.000 |
|  | $\beta_{12}$ | -1 | -1.502 | -1.036 | -0.623 | 1.000 |
| $y_2$ | $\beta_{21}$ | -8 | -10.612 | -8.089 | -6.314 | 1.000 |
|  | $\beta_{22}$ | 0 | -0.237 | 0.033 | 0.281 | 1.000 |
| $y_3$ | $\beta_{31}$ | 0 | -0.376 | 0.078 | 0.527 | 1.000 |
|  | $\beta_{32}$ | -8 | -11.028 | -8.601 | -6.701 | 1.000 |
| $y_4$ | $\beta_{41}$ | -7 | -14.530 | -8.951 | -5.929 | 1.100 |
|  | $\beta_{42}$ | -7 | -16.737 | -7.660 | -3.543 | 1.100 |
| $y_5$ | $\beta_{51}$ | -1 | -1.505 | -1.039 | -0.629 | 1.000 |
|  | $\beta_{52}$ | -6 | -8.091 | -5.835 | -4.275 | 1.000 |

### Estimating the $BMDTA$

The goal is to determine the $BMDTA$ that corresponds to a $BMR_{50}$ on each endpoint. Hence $\eta = (1/2, 1/2, 1/2, 1/2, 1/2)^T$ and by our model definition $f_j^{-1} = ln(\eta_j)$. For each of the MCMC samples $A_{T(m)}$ was determined and the $50^{th}$ quantile was found and the corresponding $\boldsymbol{\beta}$ was determined. This gives the $BMDTA_{50}$ as any dose combination point $\mathbf{x}^*$ that satisfies the following system of inequalities:

$$\begin{pmatrix} ln(1/2) \\ ln(1/2) \\ ln(1/2) \\ ln(1/2) \\ ln(1/2) \end{pmatrix} \geq \begin{pmatrix} -5.826 & -1.239 \\ -8.487 & -0.313 \\ -0.972 & -8.507 \\ -8.146 & -7.907 \\ -0.918 & -6.526 \end{pmatrix} \begin{pmatrix} \mathbf{x_1^*} \\ \mathbf{x_2^*} \end{pmatrix}, \mathbf{x_1^*} \geq 0, \mathbf{x_2^*} \geq 0. \tag{4.16}$$

To determine the level of sensitivity of each endpoint and feasible probability, GLPK routine is applied to MCMC samples and feasible probabilities are recorded. The results shows that, $Y_5$ has the lowest feasible probability 0.006 followed by $Y_1$ with 0.001 feasible probability. Endpoint $Y_2$ and $Y_3$ have 0.670 and 0.613 feasible probabilities respectively. $Y_4$ is essential to determine the tolerable area with feasible probability of 1. Endpoints $Y_2$, $Y_3$, and $Y_4$ are co-sensitive endpoints.

### Sensitivity Analysis

A sensitivity analysis is conducted to investigate the robustness of our $BMDTA$ estimate. This is done by changing the precision ($\frac{1}{\sigma^2}$, where $\sigma^2$ is the variance of $\mu$s) from 100 to 50

Figure 4.5: The $BMDTA_{50}$ for simulated data

then, to 10. The results shows that changing the precision from 100 to 50 does not affect the $BMDTA$ estimate. However, changing the precision from 100 to 10 results in area that agrees with the previous one by 97%. These results indicate that the $BMDTA$ estimates is not affected by the assumption made on the prior distribution.

## 4.5 Application on OP pesticides data and results

This method of determining a tolerable region when there are multiple sources of exposure and multiple measurements are taken on each subject is applied to the OP pesticides data. Each endpoint measurement was transformed to percent control by dividing each measurement by the mean of the control group. We used a linearizable nonlinear exponential model to fit the data simultaneously. The following model was used to model the data:

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \\ Y_{4i} \end{pmatrix} \sim \begin{pmatrix} N\left(\exp(\beta_{10} + \beta_{11}x_{1i} + \beta_{12}x_{2i}), \sigma_1\right) \\ N\left(\exp(\beta_{20} + \beta_{21}x_{1i} + \beta_{22}x_{2i}), \sigma_2\right) \\ N\left(\exp(\beta_{30} + \beta_{31}x_{1i} + \beta_{32}x_{2i}), \sigma_3\right) \\ Binom\left(n_i, \frac{1}{1+\exp(\beta_{40}+\beta_{41}x_{1i}+\beta_{42}x_{2i})}\right) \end{pmatrix} \tag{4.17}$$

The hierarchical aspect of the model was specified as follow:

$$\mathbf{Y}_i = \begin{pmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \\ y_{4i} \end{pmatrix} \sim \begin{pmatrix} f_1(x_i)|\boldsymbol{\beta} \\ f_2(x_i)|\boldsymbol{\beta} \\ f_3(x_i)|\boldsymbol{\beta} \\ f_4(x_i)|\boldsymbol{\beta} \end{pmatrix}$$

To fully specify the model the following prior distributions are used:

$$\begin{aligned} \boldsymbol{\beta} &\sim N(\boldsymbol{\mu}, \boldsymbol{\Omega}) \\ \boldsymbol{\Omega} &\sim Wishart(I_{12}, 12) \\ \boldsymbol{\mu} &\sim N(0, 100) \\ \sigma_j &\sim Gamma(1, 1) \end{aligned}$$

where $I_{12}$ is the $12 \times 12$ identity matrix.

The resulting posterior distribution is of the form:

$$\begin{aligned}
P(\boldsymbol{\beta}, \boldsymbol{\Omega}, \boldsymbol{\mu}, \boldsymbol{\tau}|D) \quad \propto \quad & e^{-\sum_{j=1}^{3}\sigma_j^2} \times e^{-\frac{1}{200}\sum_{j=1}^{4}\mu_j^2} \\
\times \quad & \frac{|\boldsymbol{\Omega}|^{(-6)} e^{-trace(I\boldsymbol{\Omega}^{-1})/2}}{2^{72}\Gamma_{12}(6)} \\
\times \quad & |\Omega|^{-1/2} e^{-\frac{1}{2}(\beta-\boldsymbol{\mu})'\Omega^{-1}(\beta-\boldsymbol{\mu})} \\
\times \quad & \prod_{i=1}^{n} \begin{pmatrix} n_i \\ Y_{4i} \end{pmatrix} \left(\frac{1}{1+e^{x_i\beta_4}}\right)^{Y_{4i}} \left(1-\frac{1}{1+e^{x_i\beta_4}}\right)^{n_i-Y_{4i}} \\
\times \quad & \left(\prod_{j=1}^{3}\sigma_j^2\right)^{-n/2} \quad e^{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{3}\sigma_j^{-2}(Y_{ij}-e^{x_i\beta_j})^2}
\end{aligned}$$

WinBUGS was used to generate 4 chains of 101,000 MCMC samples from the posterior distribution. To verify convergence of the chains traceplots were examined for good mixing and $\hat{R}$ was verified to be less than 1.005 for each chain. The first 1,000 samples from each chain were discarded as burn-in samples. The remaining 100,000 samples were thinned by 10 to minimize autocorrelation in the samples, which resulted in 10,000 samples from each chain. The 10,000 samples from the first chain were used to draw inferences. This sampling took 3 hours on an intel centrino Duo Pentium D620 Computer with 2GHz processor.

### Results

Table 4.3 show the 2.5%, 50% and 97.5% quantiles as well as the $\hat{R}$ value for each regression parameter. Notice the coefficients indicate a decreasing relationship with both ACE and DIA for all endpoints.

Table 4.3: *The quantiles of the posterior samples for parameters.*

| *Endpoint* | *Parameter* | 2.5% | 50% | 97.5% | $\hat{R}$ |
|---|---|---|---|---|---|
| Brain ChE | $\beta_{10}$ | $-0.252$ | 0.203 | $-0.161$ | 1.000 |
| | $\beta_{21}$ | $-0.005$ | $-0.003$ | $-0.002$ | 1.000 |
| | $\beta_{32}$ | $-0.126$ | $-0.094$ | $-0.069$ | 1.000 |
| Blood ChE | $\beta_{40}$ | $-0.585$ | $-0.508$ | $-0.435$ | 1.000 |
| | $\beta_{51}$ | $-0.061$ | $-0.032$ | $-0.018$ | 1.000 |
| | $\beta_{62}$ | $-0.131$ | $-0.080$ | $-0.044$ | 1.000 |
| Motor Activity | $\beta_{70}$ | $-0.193$ | $-0.148$ | $-0.107$ | 1.000 |
| | $\beta_{81}$ | $-0.003$ | $-0.002$ | $-0.001$ | 1.000 |
| | $\beta_{92}$ | $-0.054$ | $-0.039$ | $-0.027$ | 1.000 |
| Tail pinch | $\beta_{100}$ | 0.787 | 0.869 | 0.955 | 1.000 |
| | $\beta_{111}$ | $-0.0009$ | $-0.001$ | $-0.003$ | 1.000 |
| | $\beta_{122}$ | $-0.002$ | $-0.006$ | $-0.010$ | 1.000 |

**Evaluating** $BMDTA$

To determine the $BMDTA_\eta$, for each of the MCMC samples $A_{T(m)}$ was determined and the $5^{th}$ quantile of $A_{T(m)}$ and the corresponding $\boldsymbol{\beta}$ were found. This results in a $BMDTA_\eta$ that is any point $(x_1^*, x_2^*)$ that satisfies the corresponding inequalities. Equations 4.18, 4.19, and 4.20 give the system of inequalities for $BMDTA_{10}$, $BMDTA_{25}$ and $BMDTA_{50}$, respectively.

The $BMDTA_{10}$ is any point $(x_1^*,\, x_2^*)$ that satisfies the following:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{10}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{10}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{10}\right) \\ ln\left(\frac{p_4(\mathbf{0})/10}{1-p_4(\mathbf{0})/10}\right) \end{pmatrix} \geq \begin{pmatrix} -0.193 & -0.004 & -0.086 \\ -0.547 & -0.023 & -0.073 \\ -0.158 & -0.002 & -0.054 \\ -0.867 & -0.003 & -0.006 \end{pmatrix} \begin{pmatrix} 1 \\ x_1^* \\ x_2^* \end{pmatrix}, \tag{4.18}$$

$$x_1^* \geq 0, x_2^* \geq 0.$$

The $BMDTA_{25}$ is any point $(x_1^*,\, x_2^*)$ that satisfies the following:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{4}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{4}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{4}\right) \\ ln\left(\frac{p_4(\mathbf{0})/4}{1-p_4(\mathbf{0})/4}\right) \end{pmatrix} \geq \begin{pmatrix} -0.216 & -0.004 & -0.081 \\ -0.581 & -0.030 & -0.082 \\ -0.115 & -0.002 & -0.047 \\ -0.860 & -0.001 & -0.008 \end{pmatrix} \begin{pmatrix} 1 \\ x_1^* \\ x_2^* \end{pmatrix}, \tag{4.19}$$

$$x_1^* \geq 0, x_2^* \geq 0.$$

The $BMDTA_{50}$ is any point ($x_1^*$, $x_2^*$) that satisfies the following:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{2}\right) \\ ln\left(\frac{p_4(\mathbf{0})/2}{1-p_4(\mathbf{0})/2}\right) \end{pmatrix} \geq \begin{pmatrix} -0.181 & -0.004 & -0.069 \\ -0.550 & -0.039 & -0.090 \\ -0.135 & -0.003 & -0.042 \\ -0.861 & -0.001 & -0.003 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_1^*} \\ \mathbf{x_2^*} \end{pmatrix}, \mathbf{x_1^*} \geq 0, \mathbf{x_2^*} \geq 0. \quad (4.20)$$

The 50% $BMDTA$ associated with exposure to ACE and DIA is a polytope shown on Figure (4.6). It corresponds to all possible dose combinations from ACE and DIA that are considered to be tolerable for simultaneous exposure.

### Endpoint Probability

To determine the level of sensitivity of each endpoint and feasible probability, GLPK routine is applied to MCMC samples and feasible probabilities are recorded for $\eta = 0.1$, $\eta = 0.25$ and $\eta = 0.5$ and are given in Table 4.4. Notice that the probabilities of BrainChE and BloodChE are quite high indicating that each endpoint is important and hence would be deemed co-sensitive. Furthermore, the probability for Tail pinch is always 0, indicating it is hypo-sensitive.

### Sensitivity Analysis

A sensitivity analysis is conducted to investigate the robustness of our $BMDTA$ estimate. This is done by changing the precision from 100 to 50 then, to 10. The results shows that changing the precision from 100 to 50 does not affect the $BMDTA$ estimate. However, changing the precision from 100 to 10 results in area that agrees with the previous one by

Figure 4.6: $BMDTA$ associated with exposure to ACE and DIA

Table 4.4: Posterior endpoint sensitivity probabilities, for the OP data with exposure to ACE and DIA. Based on 10,000 posterior samples.

| Endpoint | $BMR_{10}$ | $BMR_{25}$ | $BMR_{50}$ |
|---|---|---|---|
| BrainChE | 0.674 | 0.705 | 0.775 |
| BloodChE | 0.998 | 1 | 1 |
| Motor Activity | 0.001 | 0 | 0 |
| Tail pinch | 0 | 0 | 0 |

86%. These results indicate that the $BMDTA$ estimates is not affected by the assumption made on the prior distribution.

## 4.6 Discussion

Multiple endpoints data pose a challenge to toxicologists. Adequate statistical methods to analyze these types of data are still in need. This work proposes a Bayesian hierarchical model as a single unifying method to analyze data with multiple endpoints resulting from multiple source of exposure. In terms of risk analysis, this method can determine the minimum tolerable area associated with pre-specified $BMR$. The tolerable area is seen as the area corresponding to all possible dose combination from stressors under consideration that one can be exposed to before bad things happen. The area depends on sensible endpoints and take into account correlation among endpoints. The $BMDTA$ method leads to more concise results because the method conserves the dimension of the data, but its flexibility to adopt different model types such as threshold models, interacting stressors, needs a closer investigation. Determining endpoints that define the $BMDTA$ is crucial, it may lead to cost and time reduction in future research. Vice versa, it may be important to determine which stressor is/are more dangerous (in terms of producing greatest adverse health effects) than the others, but this topic is beyond the scope of the current study. Among advantages of the method developed here, is to take into account the correlation between observations through the specification of a single unifying Bayesian hierarchical structure of the model. With this method, we can ensure that a good part of the correlation in the data is being accounted for, but we can not guarantee that all the correlation is accounted for and further research is needed to capture the entire correlation. It is important to notice that the method developed here leads to a partial and an overall interpretation of the results. It provides information on how endpoints are sensitive to the exposure partially and in overall. In the simulated data study, we used 50% $BMR$, the EPA recommend up to 10% $BMR$. The higher $BMR$ is used for proof of concept purposes, lower $BMR$ values were considered for the real data. Sensitivity analysis is conducted by considering different values of the precision because we believe precision is a major governor of data spread and variability.

The method developed here concluded that BloodChE is the most sensitive endpoint with 1 feasible probability, followed by BrainChE with 0.006 feasible probability. These findings confirm the results from the independent analysis by Moser *et al.*, (2005) as reproduced in Table (4.4). In this table, BloodChE endpoint has the lowest estimated values for $ED_{20}$ and $ED_{50}$ followed by BrainChE.

Table 4.5: *$ED_{20}$ and $ED_{50}$ Values for the predicted (under additivity) and Experimental Mixture data, for full and reduced rays*: Moser *et al.*, (2005)

| | | Full ray | | Reduced Ray | |
|---|---|---|---|---|---|
| Endpoint | | Additivity | Mixture | Additivity | Mixture |
| Brain ChE | $ED_{20}$ | 17.2 | 8.4 | 3.4 | 2.2 |
| | $ED_{50}$ | 47.6 | 26.6 | 10.3 | 7.8 |
| Blood ChE | $ED_{20}$ | 11.5 | 6.6 | 4.5 | 2.0 |
| | $ED_{50}$ | 32.2 | 17.8 | 11.8 | 5.4 |
| Motor Activity | $ED_{20}$ | 68 | 52.5 | – | – |
| | $ED_{50}$ | 210 | 81 | – | – |
| Gait Score | $ED_{20}$ | 65.2 | 21.7 | 14.8 | 6.7 |
| | $ED_{50}$ | 94.4 | 31.4 | 21.5 | 9.7 |
| Tail pinch | $ED_{20}$ | 224.4 | 63.3 | – | – |
| | $ED_{50}$ | 295 | 83.2 | – | – |

In this work, we only considered linearizable functions to model each endpoint to prove our method. We begin with a combination of simple and linearizable nonlinear models to model the four endpoint simultaneously. It may be interesting to know if the sample model chosen is the best model and how good or bad discarded models would have performed, which leads to the next Chapter that covers model uncertainty and Bayesian model averaging.

# Chapter 5

# Bayesian Model Averaging for estimation of tolerable area in multiple endpoints and multiple hazards exposure

## 5.1 Introduction

Multiple regression is a tool used by scientists to determine relationship between a response and its predictors. Often there are many candidates predictors variables with which they wish to describe or predict the response of interest. In such a case, researchers need a proper way to search through all the possible models to determine an appropriate one to explain the relationship between the predictors and the responses. Common methods for performing model selection are Maximum adjusted-$R^2$, forward, backward and stepwise selection methods (Hocking 1976), Mallow's $C_p$ (Mallow 1973), Predicted REsidual Sum of Squares (PRESS, Allen 1974), the Bayesian Information Criterion BIC, (Schwarz 1978); Akaike's Information Criterion AIC, (Akaike 1974) and many more. The algorithms cited above use different criteria to determine the best model. Some search for a parsimonious model, others use criteria such as F-tests to allow covariates to enter in the model. These algorithms select predictors to put in the model relieving the uncertainty about the relevant predictors to be included in the mode but leave intact the uncertainty regarding the choice of the appropriate model.

Toxicologist and risk assessors are concerned about how much environmental health hazards (such as chemicals) living objects are exposed to; from food, drinking water, household daily use products. The exposure to these stressors causes multiple health defects as most of them have the capability to disrupt the normal functionality of organs, such as thyroid, in living objects body, Crofton *et al.*, (2005); Desaulniers *et al.*, (2003). Risk assessors researchers in toxicology studies evaluate the risk associated with the exposure to environmental health hazards by estimating the benchmark dose ($BMD$) associated with a pre-specified benchmark dose response ($BMR$). Risk related endpoints are estimated using studies where outcomes are modeled as a function of the dose considered and $BMD$ is determined using this dose-response model. In this setting, there often exist multiple dose response models that describe the data well and risk assessors may not have a prior reason, to prefer a given model over the other models considered, on the other hand, in studies where multiple outcomes are observed, there may not be a single model that fits all endpoints best. All model selection methods aforementioned are impaired as they are based on selecting predictors to put in the model among multiple predictors; they cannot be used to select appropriate models from multiple models. Bayesian Model Averaging (BMA) is one method that has gained popularity in the literature to describe model uncertainty. It is based on using probabilistic arguments to determine the model and to average over all possible models. It was introduced in statistics in the mid-1990s and it incorporates model uncertainty into the analysis using posterior model probabilities (see Kass and Raftery (1995), Madigan and York (1995), Raftery (1996), Raftery, Madigan and Hoeting (1997), Hoeting, Madigan, Raftery and Volinsky (1999)). Since its introduction, the use of BMA in application has been expanding into different fields: Clyde (2000), Lamon and Clyde (2000); Viallefont, Raftery and Richardson (2001) employ these methods on case-control studies; Murphy and Wang (2001) use BMA in infant survival studies; Fernandez, Ley, and Steel (2001) applied the BMA in economics; Yeung, Bumgarner, and Raftery (2005) used BMA in biology; Morales *et al.*, (2006) used BMA in public health; Koop and Tole (2004) used BMA in toxicology.

The objective of this report is to address the problem of model uncertainty in analysis of multiple endpoints in toxicological data by applying BMA.

## 5.2 The data

Following the logic developed in Chapter 4 about determination of the benchmark dose tolerable area ($BMDTA$) as an estimate of the risk associated with stressors exposure, the goal here is to use BMA to account for model uncertainty in estimation of the neurotoxicity effect from simultaneous exposure to organophosphate pesticides. The study covered five pesticides commonly used in agriculture: Acephate (ACE), diazinon (DIA), malathion (MAL), chlorpyriphos (CPF), dimethoate (DMI). Long-Evans rats were orally exposed to single chemical dose or mixtures. The mixing proportions for the full mixture was (0.040, 0.031, 0.002, 0.102, 0.825) for ACE, CPF, DIA, DIM, and MAL, respectively. For the reduced mixture (with no malathion), the mixing proportion was (0.229, 0.011, 0.177, 0.583) for ACE, DIA, CPF, and DIM respectively. Five outcomes were evaluated at the time of peak effect, motor activity, tail pinch, and gait score, and then tissues were collected for measurement of ChE activity (blood and brain ChE activity). The doses considered are as follow: Five concentrations and a vehicle control (0, 3, 10, 30, 60 and 120 mg/kg) of ACE were experimentally evaluated in 8 rats each (total 48 rats). Seven concentrations and a vehicle control (0, 5, 25, 50, 75, 125, 150 and 250 mg/kg) of DIA were experimentally evaluated in 16, 16, 16, 8, 16, 8, 8 and 8 rats, respectively (total 96 rats). Five concentrations and a vehicle control (0, 2, 10, 20, 30 and 50 mg/kg) of CPF were experimentally evaluated in 8 rats each (total 48 rats). Two concentrations and a vehicle control (0, 100 and 500 mg/kg) of MAL were experimentally evaluated in 7, 8, and 8 rats, respectively (total 23 rats). Five concentrations and a vehicle control (0, 5, 10, 25, 50 and 75 mg/kg) of DIM were experimentally evaluated in 8 rats each (total 48 rats).

Although the mixture data included 5 chemicals, the analysis considered the data to be functions of two selected chemicals and the responses were averaged over the remaining chemicals in the mixture data. A two chemical study was chosen to accommodate graphical illustration. The choice of which chemicals was arbitrary as all five chemicals have enough data in the region of interest (low dose region) except malathion which has only two concentrations and a control group. This work only considered four endpoints (blood and brain ChE, motor activity and tail pinch). To achieve a mixed response behavior, i.e, as the doses increase some curves increase and others decrease we modeled the probability of no tail pinch. In Chapter 6, the method is applied to all five chemicals. All the data are described in Moser *et al.*, (2005).

## 5.3 Bayesian Model Averaging

Existing model selection methods select covariates to put in the model. At the end of the selection algorithm, the model selected is treated as the best fit model among a given set of predictors. This process does not account for model uncertainty. Bayesian Model Averaging (BMA), on the other hand, allows the researcher to account for model uncertainty and express this uncertainty in terms of probability (Madigan and Raftery 1994, Kass and Raftery 1995, Clyde 1999, Hoeting *et al.*, 1999). Let $\mathbf{M} = (M_1, M_2, M_3 \cdots M_U)$ be the set of models under consideration. Let $M_u$ denote the $u^{th}$ model in the set $\mathbf{M}$. The cardinality or size of $M_u$s is denoted by $|\mathbf{M}|$. Following Bayes Rule, the posterior probability for model $M_u$ given some data $D$ is given by:

$$p(M_u|D) = \frac{p(D|M_u)p(M_u)}{\sum_{v=1}^{\mathbf{M}} p(D|M_v)p(M_v)}. \tag{5.1}$$

To compute $P(M_u|D)$ we need to compute the probability of the data given model $M_u$, $P(D|M_u)$ for all $M_u \in \mathbf{M}$. We also need to specify the prior probability that the model $M_u$ is the correct model $P(M_u)$ for each model $M_u \in \mathbf{M}$. The marginal likelihood of the data $D$ given the model

$$p(D|M_u) = \int p(D|\theta_u, M_u)p(\theta_u|M_u)d\theta_u \tag{5.2}$$

is the integrated likelihood of model $M_u$, $\theta_u$ is the vector of parameters of model $M_u$, $p(\theta_u|M_u)$ is the prior density of the parameters under model $M_u$, $p(D|\theta_u, M_u)$ is the likelihood, and $p(M_u)$ is the prior probability that $M_u$ is the true model. After determining $p(D|M_u)$ for all model, the *law of total probability* is used to *average* the models. The higher dimensionality aspect of equation (5.2) makes its evaluation a very difficulty task. Monte Carlo Integration method is one approach among others to numerically evaluate equation (5.2) (e.g., George and McCulloch, 1997, and Raftery, Madigan, and Hoeting, 1997). A simple Monte Carlo estimate of an integral is

$$\int p(D|\theta_u, M_u)p(\theta_u|M_u)d\theta_u \approx \frac{1}{t}\sum_{u=1}^{t} p(D|\theta_u^{(j)}, M_u) \tag{5.3}$$

where $\theta_u^{(j)}, u = 1, \cdots, t$ are samples from the posterior distribution.

Let $\Delta$ be the quantity of interest, such as an effect size, a future observable, or a model parameter, then its posterior distribution given data $D$ is:

$$p(\Delta|D) = \sum_{u=1}^{U} p(\Delta|D, M_u)p(M_u|D), \tag{5.4}$$

which is an average of the posterior predictive distribution for $\Delta$ under each of the models considered, weighted by the corresponding posterior model probability.

To evaluate the posterior model probabilities using equation (5.1) we need to specify the prior probability of model $M_u$ being the true model. In the absence of any information we can set $P(M_u) = \frac{1}{|\mathbf{M}|}$. For more information on selecting $P(M_u)$ see Clyde (1999), Madigan, Gavrin, and Raftery (1995), George (1999). If prior information about the model space is available, we should incorporate this into our analysis. Once we have obtained $P(M_u|D)$ for all $M_u \in \mathbf{M}$ we can estimate the posterior distribution any quantity $\Delta$ given the data D by (5.4). Researchers have shown that the averaged has better predictive ability than any single model (Hoeting, Madigan, Raftery and Volinsky 1999). The number of models to be averaged can be high, which may result in computationally issues. Madigan and Raftery (1994) introduced the idea of using *Occam's Window* to shrink the model space to $m*$. Occam's window is:

$$m* = [M_u : \frac{max_{M_j \in \mathbf{M}} p(M_j|D)}{p(M_u|D)} \leq C], \tag{5.5}$$

where $C$ is set to some appropriate value.

### Model specification

Let $n$ denote the total number of subjects in the study. Let assume, on each subject, the number of measurements taken is $J$. If $y_{ij}$ is the response of the $i^{th}$ subject on the $j^{th}$ endpoint, then $Y_i = (y_{i1}, ..., y_{iJ})$ vector of all measurements taken on $i^{th}$ subject can be formed for all $n$ subjects. Let $K$ be the number of stressors considered, and let $x_{ik}$ represents the $k^{th}$ stressor on the $i^{th}$ subject. The relationship between the $x_{ik}$ and $y_{ij}$, a function $f = (f_1, ..., f_J)$ can be formed where each $f_j$ is invertible function. Let $\boldsymbol{\theta}_j$ denotes a vector parameters corresponding to the $j^{th}$ endpoint. Hence the model can be formed as:

$$Y_i = f(X_i\boldsymbol{\theta}) = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iJ} \end{pmatrix} = \begin{pmatrix} f_1(X_i\boldsymbol{\theta}_1) \\ f_2(X_i\boldsymbol{\theta}_2) \\ \vdots \\ f_J(X_i\boldsymbol{\theta}_J) \end{pmatrix}. \tag{5.6}$$

Where $X_i = (x_{i1}, ..., x_{iK})$ is a vector of all chemical dose applied on $i^{th}$ subject. In this setting each endpoint has its own model$f_j(X_i\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ are parameters of interest and $X_i$ is a vector of chemical doses.

### Likelihood

Let $g_j(y_{ij}|f_j(X_i\theta_j), \gamma_j)$ be the probability distribution for each endpoint $(j = 1, ..., J)$ where $\gamma_j$ are additional parameters that may be necessary to define the distribution. The likelihood depends on the type of data corresponding to each endpoint. Since the data may have of mixed types (continuous, discrete, categorical,...), the likelihood should be appropriate to the endpoint data-type. Given $\boldsymbol{\theta}_J$, the endpoints are conditionally independent and hence the likelihood can be constructed as:

$$L(Y|\theta X) = \prod_{i=1}^{n} \prod_{j=1}^{J} g_j(y_{ij}|f_j(X_i\theta_j), \gamma_j). \tag{5.7}$$

**Prior distributions**

Proper prior distributions are employed to ensure the resulting posterior distribution is proper (i.e. $\int p(\theta|D)d\theta = 1$ where $\boldsymbol{\theta}$ represents all the parameters in the model). To specify prior distributions on the hierarchical structure of the model we assume $\boldsymbol{\theta} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ and $\boldsymbol{\Omega} \sim Wishart(R, \rho)$ in return. Furthermore, we assume $\gamma_j \sim p(\gamma_j)$ and $\boldsymbol{\mu} \sim N(a, A)$ where $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ are the mean vector and precision matrix, respectively, for $\boldsymbol{\theta}$. Here R is specified as the $J(K+1) \times J(K+1)$ identity matrix. The prior distribution for $\gamma_j$, $p(\gamma_j)$ is an appropriate distribution and $a$ and $A$ are the hyperparmaters governing $\boldsymbol{\mu}$.

### 5.3.1 From $BMDL$ to $BMDTA$

For this work, we only considered linearizable nonlinear models. Each sample model is made of 4 combinations of endpoints models.

Example of sample model:

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \\ Y_{4i} \end{pmatrix} \sim \begin{pmatrix} N\left(\exp(\theta_{10} + \theta_{11}x_{1i} + \theta_{12}x_{2i}), \sigma_1\right) \\ N\left(\exp(-\exp(-(\theta_{20} + \theta_{21}x_{1i} + \theta_{22}x_{2i}))), \sigma_2\right) \\ N\left(\frac{1}{1+\exp(-(\theta_{30}+\theta_{31}x_{1i}+\theta_{32}x_{2i}))}, \sigma_3\right) \\ Binom\left(n_i, \frac{1}{1+\exp(\theta_{40}+\theta_{41}x_{1i}+\theta_{42}x_{2i})}\right) \end{pmatrix} \tag{5.8}$$

Table 5.1: *Linearizable nonlinear endpoint models considered for continuous endpoints (brain ChE, blood ChE and motor activity), and for binary endpoint (tail pinch).*

| *Endpoint Models* | *Linearizable nonlinear models* |
|---|---|
| Continuous | $exp(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i})$ |
| | $(1 - \exp(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i}))$ |
| | $exp(-\exp(-(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i})))$ |
| | $1 - \exp(-(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i}))$ |
| | $\frac{1}{1+\exp(-(\theta_0+\theta_1 x_{1i}+\theta_2 x_{2i}))}$ |
| Binary | $1 - \exp(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i})$ |
| | $\frac{1}{1+\exp(\theta_0+\theta_1 x_{1i}+\theta_2 x_{2i})}$ |

In total, there are 250 candidate models, $|\mathbf{M}| = 250$. To specify prior distributions on the hierarchical structure of the model we assume $\boldsymbol{\theta} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ and $\boldsymbol{\Omega} \sim Wishart(I_{12}, 12)$, where $I$ is the identity matrix. Furthermore, we assume $\sigma_j \sim Gamma(1,1)$ and $\boldsymbol{\mu} \sim N(0, 100)$. For each of the 250 models, WinBUGS generated 4 chains of 11,000 MCMC samples from the corresponding posterior distribution. The chains convergence was diagnosed. The first 1,000 samples from each chain were discarded as burn-in samples. The 10,000 samples from the first chain were used to draw inferences. For more on MCMC methods and diagnostics seen Gelman *et al.*, (2005).

For a given $BMR\eta = (\eta_1, ..., \eta_J)$, the tolerable region is obtained by solving the following inequality:

$$f^{-1}(\eta) = \begin{pmatrix} f_1^{-1}(\eta_1) \\ f_2^{-1}(\eta_2) \\ \vdots \\ f_J^{-1}(\eta_J) \end{pmatrix} \geq \begin{pmatrix} \theta_1' \\ \theta_2' \\ \vdots \\ \theta_J' \end{pmatrix} \begin{pmatrix} \mathbf{X}^* \end{pmatrix}, \mathbf{X}^* \geq 0, \tag{5.9}$$

where $\theta_j'$ are the estimates of parameters $\theta$s in the model (5.8). We assume $f$ is an invertible function. Any chemical dose combination $X^*$ that satisfies the above inequality is considered tolerable with respect to the $BMR\eta$. At $BMR_\eta$ of interest, the inequality in (5.9) defines $T_J$ tolerable regions where $J$ corresponds to the $j^{th}$ endpoints. To define the tolerable region for all $J$ endpoints, is to find the intersection of all $T_J$ defined by inequality in (5.9); $T = T_1 \cap T_2 \cap \cdots \cap T_J$. The $T$ defines the benchmark tolerable region $BMDT$, the analogue to the tradition $BMD$. The region $T$ is used to evaluate the benchmark dose tolerable area analogue to the traditional $BMDL$.Using MCMC samples, for each MCMC sample there are $T^{(m)}$ tolerable regions where $m$ corresponds to the $m^{th}$ MCMC sample. $A_{T^{(m)}}$, the tolerable area for each MCMC sample is evaluated as:

$$A_{T^{(m)}} = \int_{T^{(m)}} d\mathbf{X}$$

the tolerable area can be ranked and a lower bound on the tolerable area is defined by the $\beta$ MCMC sample that correspond to the $q^{th}$ quantile of the $A_{T^{(m)}}$. This gives a $100 \times q\%$ credible region for the lower bound for the $BMDTA$.

### 5.3.2 Hyper-sensitive, co-sensitive, hypo-sensitive endpoints, and feasible probability

An endpoint $i$ is said to be *Hyper-sensitive* if the following condition is met:

$$\cap_{j=1}^{J} T_j = T_i, \tag{5.10}$$

Likewise, we define *Hypo-sensitive* endpoints, all endpoint that satisfy the following condition:

$$\cap_{j\neq i}^{J} T_j = \cap_{j=1}^{J} T_j, \tag{5.11}$$

For any $Z$, a set of endpoints such that:

$$Z = \{j | j \in J\}, \tag{5.12}$$

endpoints in $Z^*$, a subset of $Z$ are *co-sensitive* if

$$\cap_{j\in Z^*} T_j = \cap_{j=1}^{J} T_j, \tag{5.13}$$

The method of determining important endpoints for the $BMDTA$ and feasible probability has been introduced in Chapter 4. It consists of using a simplex algorithm, a part of GLPK (GNU Linear Programming Kit) routine in R, applied to MCMC to determine which endpoints define the tolerable area and their corresponding feasible probabilities. In this process, a linear program problem subject to linear constraints is solved. The linear constraint is constructed from the estimates of the parameter defining the $BMDTA$ and the inequality in (5.9). The general linear program problem is constructed as follow:

$$z = C^T X \tag{5.14}$$

subject to:

$$\mathbf{A}x \geq b, LB \leq x \leq UB. \tag{5.15}$$

In this settings, UB and LB are upper and lower bounds on variables $x$, $b$ is a vector of each constraint in the constraint matrix $A$, $z$ is the objective function, $C$ is a vector of objective function coefficients. In geometric context, each inequality in (5.15) specifies a half-space in multidimensional Euclidean space, and their intersection region which is a polytope is the set of all feasible values the variables $x$ can take. The goal is to determine which endpoints bound the polytope. A linear program (LP) is constructed such that $C = 0$. Then, each inequality in (5.15) corresponds to:

$$(\boldsymbol{\theta}'_j)^T X \geq f_j^{-1}(\eta_j), j = 1, ..., J \tag{5.16}$$

For each endpoint in return, define linear programming $LP^j$ by setting the $j^{th}$ inequality to equality

$$(\boldsymbol{\theta}'_j)^T X = f_j^{-1}(\eta_j) \tag{5.17}$$

for j= 1,...,J and check for feasibility. If $Lp^j$ is feasible, then its corresponding endpoint bounds the polytope. Applied this method to all MCMC samples, then the proportion of time an endpoint is reported as feasible defines the estimate of its feasible probability. That proportion can be estimated by: $\hat{P}_j$

$$\hat{P}_j = \frac{\sum_{\iota=1}^{n_s} I_{E_{j\iota} \in Z^*}(E_{j\iota})}{n_s}, \tag{5.18}$$

where $E_{j\iota}$ designates the $j^{th}$ endpoint for the $\iota^{th}$ sample, $I$ is an indicator function that takes values 1 if the endpoint $j$ bounds the polytope and 0 otherwise, $n_s$ denotes the number of MCMC samples, and $Z^*$ denotes the set of endpoints that bound the polytope. The feasible probability for an endpoint is the importance of the endpoint to define the polytope. If the feasible probability for a given endpoint is 1 then corresponding endpoint is inevitable to define the polytope. If on the other hand the feasibility probability for a given endpoint is 0, then corresponding endpoint does not have any role in defining the polytope, therefore no further consideration.

## 5.4 Application to OP data

Since we have no information *a priori* about the models, we chose $P(M_u) = 1/250$. To implement BMA, we computed $P(M_u|D)$ for each model $M_u \in \mathbf{M}$ using Monte Carlo integration. To reduce the number of models in which to average, we used an Occam's window approach, we only considered models $M_u$ such that $P(M_u|D) > 0.00005$. This reduced our averaged model space to 4 models and their corresponding posterior probabilities are 0.990, 0.005, 0.0004, 0.0003. These 4 sample models account for almost 100% of the total posterior probability and are given in Table (5.2) with their respective posterior probabilities.

Averaged model coefficients were obtained by systematically sampling from the coef-

Table 5.2: *Sample models and their respective posterior probabilities considering endpoints were modeled in the following respective order: BrainChE, BloodChE, Motor Activity, Tail pinch*

| *Sample models* | *Posterior probability* |
|---|---|
| $N\left(\exp(\theta_{10}+\theta_{11}x_{1i}+\theta_{12}x_{2i}),\sigma_1\right)$<br>$N\left(\exp(\theta_{20}+\theta_{21}x_{1i}+\theta_{22}x_{2i}),\sigma_2\right)$<br>$N\left(\exp(\theta_{30}+\theta_{31}x_{1i}+\theta_{32}x_{2i}),\sigma_3\right)$<br>$Binom\left(n_i,\frac{1}{1+\exp(\theta_{40}+\theta_{41}x_{1i}+\theta_{42}x_{2i})}\right)$ | 0.994 |
| $N\left(\exp(\theta_{10}+\theta_{11}x_{1i}+\theta_{12}x_{2i}),\sigma_1\right)$<br>$N\left(\exp(\theta_{20}+\theta_{21}x_{1i}+\theta_{22}x_{2i}),\sigma_2\right)$<br>$N\left(\exp(\theta_{30}+\theta_{31}x_{1i}+\theta_{32}x_{2i}),\sigma_3\right)$<br>$Binom\left(1-\exp(\theta_{40}+\theta_{41}x_{1i}+\theta_{42}x_{2i})\right)$ | 0.0004 |
| $N\left(\exp(\theta_{10}+\theta_{11}x_{1i}+\theta_{12}x_{2i}),\sigma_1\right)$<br>$N\left(\exp(-\exp(-(\theta_{20}+\theta_{21}x_{1i}+\theta_{22}x_{2i}))),\sigma_2\right)$<br>$N\left(\frac{1}{1+\exp(-(\theta_{30}+\theta_{31}x_{1i}+\theta_{32}x_{2i}))},\sigma_3\right)$<br>$Binom\left(n_i,\frac{1}{1+\exp(\theta_{40}+\theta_{41}x_{1i}+\theta_{42}x_{2i})}\right)$ | 0.0003 |
| $N\left(\exp(\theta_{10}+\theta_{11}x_{1i}+\theta_{12}x_{2i}),\sigma_1\right)$<br>$N\left(\exp(-\exp(-(\theta_{20}+\theta_{21}x_{1i}+\theta_{22}x_{2i}))),\sigma_2\right)$<br>$N\left(\frac{1}{1+\exp(-(\theta_{30}+\theta_{31}x_{1i}+\theta_{32}x_{2i}))},\sigma_3\right)$<br>$Binom\left(1-\exp(\theta_{40}+\theta_{41}x_{1i}+\theta_{42}x_{2i})\right)$ | 0.005 |

Table 5.3: *Feasible probability for each endpoint using the averaged model*

| *Endpoint* | *Feasible Probability* |
|---|---|
| *MotorActivity* | 0 |
| *BloodChE* | 1 |
| *BrainChE* | 0.004 |
| *Tailpinch* | 0 |

ficients of the four candidate sample models with respect to their corresponding posterior probabilities. These coefficients do not depend on a model. The averaged model is consistent with biological expectations. The averaged model parameters are used to define $BMDTA$ by:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{2}\right) \\ ln\left(\frac{p_4(\mathbf{0})/2}{1-p_4(\mathbf{0})/2}\right) \end{pmatrix} \geq \begin{pmatrix} -0.184 & -0.004 & -0.092 \\ -0.605 & -0.032 & -0.046 \\ -0.118 & -0.002 & -0.043 \\ -7.224 & -0.038 & -3.942 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_1^*} \\ \mathbf{x_2^*} \end{pmatrix}, \mathbf{x_1^*} \geq 0, \mathbf{x_2^*} \geq 0. \quad (5.19)$$

Figure(5.1) shows the tolerable area considering simultaneous exposure to ACE and DIA.

Table (5.3) shows BloodChE is a hyper-sensitive endpoint defining the $BMDTA$ with 1 feasible probability. Motor Activity, Tail pinch and BrainChE endpoints are hypo-sensitive and have 0, 0, 0.004 feasible probability respectively therefore they are not important outcomes to define the $BMDTA$. These results reflect the importance of BloodChE endpoint in studies aimed to assess the risk associated with the exposure to five OPs.

Figure 5.1: $BMDTA$, under BMA, associated with exposure to DIA and ACE

## 5.5 Discussion and Conclusion

The $BMDTA$ as introduced in Chapter 4 is an estimate of the risk associated with multiple endpoints and multiple exposure. In classical model selection procedures, the model that best fits the data is selected. But when multiple endpoints are measured, there may not be a single model that best fits all the endpoints. BMA is proven to be handy in such a situation. In the current study, linearizable nonlinear models are considered to fit 4 endpoints simultaneously equivalently 250 sample models. With one sample model that has almost 100% posterior probability, we cannot conclude that the BMA was beneficial in this specific case, but it is also imperative to notice that some use BMA to get the best model and do not have to *average*. The goal of BMA is to account for the model uncertainty in the analysis. But it is still difficult to take into account every possible uncertainty in the model especially when some type of approximation has to be used through out the process. Approximations always have some uncertainty associated with them. Due to this approximation uncertainty we have not removed model uncertainty entirely. Instead, we have diminished its impact. Another critical issue with the BMA approach is how to select the prior probability of the model, the number of models to be included in the predictive distribution and the sampling distribution. The Occam's window approach proposed by Madigan and Raftery (1994) averages over a set of good models, the selection of size of Occam's razor is still unclear.

# Chapter 6

# Composite Scores, Bayesian Hierarchical Models, and Independent analysis of endpoint: Application to Five Organophosphate Pesticides

## 6.1 Introduction

The analysis of multiple endpoints data has received attention in the literature with a variety of approaches. Available statistical methods of analysis can be placed into two groups. The first group consists of methods that reduce the dimensions of the data before the analysis, we label it as *dimension reduction methods.* This group contains methods such as the composite score method, the analysis of the most sensitive endpoint, principle component analysis, and any other simple or complex method based on reducing the dimensionality aspect of the data before the analysis is conducted. The second group comprises those methods that maintain the multi-dimensionality aspect of the data and may/ may not adjust the results accordingly. We refer to the methods in this group as *dimensions preservative* methods. This group contains methods such as multiple statistical tests, a unifying Bayesian hierarchical method (as it is introduced in Chapter 4), or any other method that conserves the multidimensionality aspect of the data. In the present work, two methods, one from each group have been applied. In fact, in Chapter 3, a dimension reduction method, using a composite score of multiple endpoint is used to analyze organophosphate pesticides data. In Chapter 4, a single unifying Bayesian based dimension preservative method is introduced. In Chapter 4, the method of estimating the risk associated with the exposure to a mixture of five organophosphate pesticides considered only two chemicals chosen arbitrarily. In the present Chapter, we consider all five chemicals as they were considered in Moser *et al.*, (2005) study. Furthermore we compare and contrast the results from these two different analysis methods and the analysis conducted by Moser *et al.*, (2005) modeling each endpoint separately.

To understand the neurotoxicity effect resulting from the exposure to organophosphate pesticides, Moser *et al.*, (2005) published the results of a study in which multiple endpoints were observed from exposure to a mixture of five organophosphate pesticides. Long-Evans rats were exposed to single oral dose of a single chemical or mixtures of five OP pesticides: acephate(ACE), chlorpyrifos (CPF), diazinon (DIA), diamethoate (DIM), and malathion

(MAL). The proportions of each chemical in the full mixture are as follows: (0.040, 0.031, 0.002, 0.102, 0.825) for ACE, CPF, DIA, DIM, and MAL, respectively, i.e., the same relative proportions as in the full mixture. For the reduced mixture, where malathion was removed from the mixture, the proportion of each chemical in the reduced mixture was (0.229, 0.011, 0.177, 0.583) for ACE, DIA, CPF, and DIM respectively. These proportions reflect the relative dietary exposure estimates of the general population as projected by the US EPA Dietary Exposure Evaluation Model (DEEM). Single chemical dose-response studies were initially conducted. Following single oral dose, measurements were taken on five endpoints at the time of peak effect: Behavioral measures (motor activity, gait score, and tail-pinch). Tissues were collected for measurements of blood and brain cholinesterase (ChE) activity. In the present work, only four endpoints were considered: BloodChE and BrainChE, motor activity and tail pinch.

### 6.1.1 An overview of the results from an independent analysis on OP data by Moser *et al.*, (2005)

The results published by Moser *et al.*, (2005) are based on separate and independent analysis of each of the five endpoints. The objective of each analysis was to test if the mixture effect can be estimated by single chemical data assuming the chemicals in the mixture interact in an additive fashion or if there is evidence of interaction (either synergy or antagonism). Furthermore, the effect of malathion on the chemicals remaining in the reduced mixture was assessed. Motor Activity and BrainChE were modeled by a threshold additivity model as described in Gennings *et al.*, (1997). For brain ChE, the likelihood ratio test of additivity was rejected for both mixtures. The experimental mixture model showed greater than additivity effect in the low dose range. The comparison between the full and the reduced ray as described in Casey *et al.*, (2006) revealed that the two mixtures were not significantly different ($p = 0.421$). This shows that there is no evidence that malathion had an effect on the four remaining chemicals. For the motor activity endpoint, in the low dose range, there was evidence to reject the test of additivity ($p = 0.001$) and the comparison between the full and reduced rays showed no difference ($p = 0.378$). This shows that malathion did not interact with the four chemicals. In Moser *et al.*, (2005) analysis, blood ChE was fit by a generalized linear model because the data fit a threshold outside the experimental range. The likelihood ratio test of additivity was rejected ($p = 0.001$) suggesting evidence of synergy in the low dose range for both full and reduced mixtures. There were indications that the two mixtures ray differ ($p = 0.001$) which suggested that malathion interacted with the remaining chemicals. These results are summarized in Table (6.1). Chapter 3 covers the use of composite scores method using desirability function to analyze OP data.

### 6.1.2 A brief summary of the results from the composite scores method applied to OP data

In this analysis a desirability function is used to transform the observed response to a unitless score based on the appropriateness of the response. Then, the geometric mean is used to combine the individual scores into an overall composite score. Statistical analysis is based

Table 6.1: *OP data, independent analysis for each endpoint: Overall results as they were published in* Moser *et al.*, (2005)

| Endpoint | Full Ray | Reduced Ray | Malathion effect |
|---|---|---|---|
| Motor Activity | Synergy | Synergy | No |
| BloodChE | Synergy | Synergy | Yes |
| BrainChE | Synergy | Synergy | No |
| Tail pinch | Synergy | Synergy | No |
| Gait Score | Additivity | Additivity | Yes |

on the overall composite score. We used the threshold additivity model to fit the overall composite score. For the full mixture data the analysis concluded that there is evidence of departure from additivity. In fact, the mixture is significantly greater than additive ($p = 0.001$) suggesting synergy. In the low dose region (defined by the 5% and 10% $BMR$) there is not a significant difference between the $BMD$s under additivity and that estimated from the mixture data. However, the $BMD$ associated with a $BMR$ of 20% as predicted under additivity from single chemical data overestimated the BMD from the mixture data. For the reduced mixture data we observed that the mixture is significantly greater than additive ($p \leq 0.001$). There is not a significant difference between the additive estimate of the $BMD$s and that estimated from the reduced ray mixture data at $BMR_{20}$. Overall, in Moser *et al.*, (2005), deviations from additivity were reported for some endpoints but not others for both mixtures (see Table 6.1). Using this method, both the 5-OP mixture and the

4-OP mixture were significantly different from additivity. There was evidence that malathion has an influence on the four chemicals in the reduced ray. Chapter 4 introduces a unique unifying Bayesian hierarchical model to simultaneously estimate the risk associated with the exposure to OP pesticides. In the next section, the dimension preservative methodology developed in Chapter 4 is applied to the mixture of all five pesticides. Issues raised by model uncertainty are resolved using the Bayesian model averaging method, as described in the next section.

## 6.2 Dimension preservative method as applied to five OP data

In the context of dimension preservative methods, in Chapter 4 we introduce a unique unifying Bayesian hierarchical model to analyze the OP data (Moser *et al.*, 2005). The main objective in this study was to estimate the benchmark dose tolerable area, $BMDTA$, as an estimate of the risk associated with the exposure to a mixture of OPs. To demonstrate our method, in Chapter 4, we only used data from two chemicals selected arbitrarily; now we apply the method to data from all five chemicals, i.e., dose response data from each single chemical and two fixed ratio mixtures.

### Bayesian structure

For each observation, four endpoints were measured and considered here as a result of exposure to five chemicals. Let $y_{ij}$ be the response of the $i^{th}$ subject on the $j^{th}$ endpoint. For each subject, the response is a vector $Y_i = (y_{i1}, ..., y_{i4})$ for all $n$ subjects. And let $x_{ik}$ be the $k^{th}$ chemical on the $i^{th}$ subject, where $k = 1, 2, .., 5$. This can be formed into a vector $X_i = (x_{i1}, ..., x_{i5})$. The relationship between the $x_{ik}$ and $y_{ij}$, a function $f = (f_1, ..., f_4)$ can be formed where each $f_j$ is linearizable and invertible function. Let $\boldsymbol{\psi}_j$ denote a vector parameters corresponding to the $j^{th}$ endpoint. Hence the model can be formed as:

$$Y_i = f(X_i\boldsymbol{\psi}) = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{i4} \end{pmatrix} = \begin{pmatrix} f_1(X_i\boldsymbol{\psi}_1) \\ f_2(X_i\boldsymbol{\psi}_2) \\ \vdots \\ f_4(X_i\boldsymbol{\psi}_4) \end{pmatrix}. \tag{6.1}$$

In this setting, each endpoint has its own model$f_j(X_i\boldsymbol{\psi}_j)$ where $\boldsymbol{\psi}_j$ are parameters of interest and $X_i$ is a (matrix) vector of stressors. Proper likelihood taking into account the mixed types aspect of the data is considered.

### Likelihood

The likelihood can be constructed as:

$$L(Y|\psi X) = \prod_{i=1}^{n}\prod_{j=1}^{J} g_j(y_{ij}|f_j(X_i\psi_j), \gamma_j). \tag{6.2}$$

$j = 1, ..., 4$ assuming that, given $\boldsymbol{\psi}_j$, the endpoints are conditionally independent, where

$g_j(y_{ij}|f_j(X_i\psi_j),\gamma_j)$ is the probability distribution for each endpoint and $\gamma_j$ are additional parameters that may be necessary to define the distribution.

**Prior distribution specification**

The prior on the hierarchical structure of the model is specified as follow: we assume $\boldsymbol{\psi} \sim N(\boldsymbol{\mu},\boldsymbol{\Omega})$ and $\boldsymbol{\Omega} \sim Wishart(I_{24},24)$. Furthermore, we assume $\sigma_j \sim Gamma(1,1)$ and $\boldsymbol{\mu} \sim N(0,100)$.

$$\begin{aligned}
\boldsymbol{\psi} &\sim N(\boldsymbol{\mu},\boldsymbol{\Omega}) \\
\boldsymbol{\Omega} &\sim Wishart(I_{24},24) \\
\boldsymbol{\mu} &\sim N(0,100) \\
\sigma_j &\sim Gamma(1,1)
\end{aligned}$$

where $I_{24}$ is the $24 \times 24$ identity matrix.

### 6.2.1 Application and Results

We fit the data with a linearizable nonlinear model:

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \\ Y_{4i} \end{pmatrix} \sim \begin{pmatrix} N\left(\exp(\psi_{10} + \sum_{k=1}^{5} \psi_{jk} x_{ik}), \sigma_1\right) \\ N\left(\exp(\psi_{20} + \sum_{k=1}^{5} \psi_{jk} x_{ik}), \sigma_2\right) \\ N\left(\exp(\psi_{30} + \sum_{k=1}^{5} \psi_{jk} x_{ik}), \sigma_3\right) \\ Binom\left(n_i, \frac{1}{1+\exp(\psi_{40}+\sum_{k=1}^{5} \psi_{jk} x_{ik})}\right) \end{pmatrix} \tag{6.3}$$

WinBUGS generated 4 chains of 11,000 MCMC samples from the corresponding posterior distribution. The first 1,000 samples from each chain were discarded as burn-in samples. The 10,000 samples from the first chain were used to draw inferences.

### Results

At any $BMR_\eta$ of interest, following the logic of determining $BMDTA$ developed in Chapter 4, the benchmark dose tolerable area is obtained by solving the following inequality assuming $f_j$ is invertible function:

$$f^{-1}(\eta) = \begin{pmatrix} f_1^{-1}(\eta_1) \\ f_2^{-1}(\eta_2) \\ \vdots \\ f_J^{-1}(\eta_J) \end{pmatrix} \geq \begin{pmatrix} \psi'_1 \\ \psi'_2 \\ \vdots \\ \psi'_J \end{pmatrix} \begin{pmatrix} \mathbf{X}^* \end{pmatrix}, \mathbf{X}^* \geq 0. \tag{6.4}$$

where $\psi'_j$ are the estimates of the parameters $\psi$ in the model (6.3).

Table 6.2: *The quantiles of the posterior samples for parameters.*

| *Endpoint* | *Parameter* | 2.5% | 50% | 97.5% | $\hat{R}$ |
|---|---|---|---|---|---|
| Brain ChE | $\beta_{10}$ | $-0.039$ | $-0.004$ | $-0.030$ | 1.000 |
| | $\beta_{21}$ | $-0.007$ | $-0.006$ | $-0.005$ | 1.000 |
| | $\beta_{32}$ | $-0.115$ | $-0.089$ | $-0.067$ | 1.000 |
| | $\beta_{43}$ | $-0.049$ | $-0.037$ | $-0.028$ | 1.000 |
| | $\beta_{54}$ | 0.000 | 0.000 | 0.000 | 1.000 |
| | $\beta_{65}$ | $-0.022$ | $-0.018$ | $-0.014$ | 1.000 |
| Blood ChE | $\beta_{70}$ | $-0.072$ | $-0.032$ | $-0.007$ | 1.000 |
| | $\beta_{81}$ | $-0.111$ | $-0.076$ | $-0.055$ | 1.000 |
| | $\beta_{92}$ | $-0.045$ | $-0.033$ | $-0.025$ | 1.000 |
| | $\beta_{103}$ | $-0.050$ | $-0.038$ | $-0.030$ | 1.000 |
| | $\beta_{114}$ | $-0.002$ | $-0.002$ | $-0.001$ | 1.000 |
| | $\beta_{125}$ | $-0.045$ | $-0.614$ | $-0.479$ | 1.000 |
| Motor Activity | $\beta_{130}$ | $-0.045$ | $-0.009$ | $-0.025$ | 1.000 |
| | $\beta_{141}$ | $-0.004$ | $-0.003$ | $-0.002$ | 1.000 |
| | $\beta_{152}$ | $-0.035$ | $-0.026$ | $-0.019$ | 1.000 |
| | $\beta_{163}$ | $-0.017$ | $-0.014$ | $-0.010$ | 1.000 |
| | $\beta_{174}$ | $-0.003$ | $-0.000$ | $-0.000$ | 1.000 |
| | $\beta_{185}$ | $-0.030$ | $-0.023$ | $-0.018$ | 1.000 |
| Tail pinch | $\beta_{190}$ | $-0.877$ | $-0.781$ | $-0.687$ | 1.000 |
| | $\beta_{201}$ | $-0.004$ | $-0.002$ | $-0.000$ | 1.000 |
| | $\beta_{212}$ | $-0.011$ | $-0.006$ | $-0.001$ | 1.000 |
| | $\beta_{223}$ | $-0.009$ | $-0.003$ | $-0.001$ | 1.000 |

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{2}\right) \\ ln\left(\frac{p_4(\mathbf{0})/2}{1-p_4(\mathbf{0})/2}\right) \end{pmatrix} \geq \begin{pmatrix} -0.009 & -0.005 & -0.095 & -0.040 & -0.001 & -0.021 \\ -0.057 & -0.092 & -0.027 & -0.049 & -0.002 & -0.591 \\ -0.018 & -0.004 & -0.033 & -0.021 & -0.001 & -0.026 \\ -0.294 & -0.001 & -0.002 & -0.003 & -0.001 & -0.000 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_k^*} \end{pmatrix} \tag{6.5}$$

where $\mathbf{x_k^*} \geq 0, k = 1, ..., 5$.

The $BMDTA$ is defined by the inequality in (6.5).

Geometrically, each of the above inequality results in a half Euclidean plan. The intersection of all half Euclidean plan form a polytope that we call the benchmark tolerable region. Using the quantile method as described in Chapter 4, we define the benchmark tolerable area ($BMDTA$) for $\eta = 10$, $\eta = 25$, and $\eta = 50$. Due to the multidimensionality aspect of the data, we are unable to plot $BMDTA$.

The $BMDTA_{10}$ is any point $\mathbf{x_k^*} \geq 0, k = 1, ..., 5$. that satisfies the following:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{10}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{10}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{10}\right) \\ ln\left(\frac{1/10 p_4(\mathbf{0})}{1-1/10 p_4(\mathbf{0})}\right) \end{pmatrix} \geq \begin{pmatrix} -0.022 & -0.005 & -0.081 & -0.037 & -0.0002 & -0.016 \\ -0.049 & -0.095 & -0.034 & -0.037 & -0.0021 & -0.652 \\ -0.003 & -0.004 & -0.024 & -0.014 & -0.00001 & -0.024 \\ -0.761 & -0.002 & -0.004 & -0.004 & -0.00006 & -0.017 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_k^*} \end{pmatrix} \tag{6.6}$$

where $\mathbf{x_k^*} \geq 0, k = 1, ..., 5$.

The $BMDTA_{25}$ is any point $\mathbf{x_k^*} \geq 0, k = 1, ..., 5.$ that satisfies the following:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{4}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{4}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{4}\right) \\ ln\left(\frac{1/4p_4(\mathbf{0})}{1-1/4p_4(\mathbf{0})}\right) \end{pmatrix} \geq \begin{pmatrix} -0.016 & -0.005 & -0.098 & -0.035 & -0.0002 & -0.016 \\ -0.053 & -0.067 & -0.029 & -0.041 & -0.0017 & -0.568 \\ -0.008 & -0.003 & -0.023 & -0.013 & 0.00002 & -0.024 \\ -0.754 & -0.003 & -0.005 & -0.002 & -0.0010 & -0.006 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_k^*} \end{pmatrix} \tag{6.7}$$

where $\mathbf{x_k^*} \geq 0, k = 1, ..., 5.$

The $BMDTA_{50}$ is any point $\mathbf{x_k^*} \geq 0, k = 1, ..., 5.$ that satisfies the following:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{2}\right) \\ ln\left(\frac{p_4(\mathbf{0})/2}{1-p_4(\mathbf{0})/2}\right) \end{pmatrix} \geq \begin{pmatrix} -0.012 & -0.005 & -0.094 & -0.023 & -0.0002 & -0.018 \\ -0.053 & -0.104 & -0.032 & -0.038 & -0.0016 & -0.590 \\ -0.014 & -0.003 & -0.026 & -0.012 & 0.00003 & -0.027 \\ -0.749 & -0.001 & -0.007 & -0.002 & -0.0007 & -0.012 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_k^*} \end{pmatrix} \tag{6.8}$$

where $\mathbf{x_k^*} \geq 0, k = 1, ..., 5.$

### 6.2.2 Endpoint Probability

To determine the level of sensitivity of each endpoint and feasible probability, GLPK routine is applied to MCMC samples and feasible probabilities are recorded for $\eta = 10$, $\eta = 25$ and $\eta = 50$ and are given in Table 6.3. Notice that the probabilities of BrainChE and BloodChE are quite high indicating that each endpoint is important and therefore would be deemed co-sensitive. Furthermore, the probability for Tail pinch and Motor Activity is always 0, indicating it is hypo-sensitive.

Table 6.3: Posterior endpoint sensitivity probabilities, for the OP data with exposure to five OP. Based on 10,000 posterior samples.

| Endpoint ($y_j$) | $ED_{10}$ | $ED_{25}$ | $ED_{50}$ |
|---|---|---|---|
| BrainChE | 0.997 | 1 | 1 |
| BloodChE | 1 | 1 | 1 |
| Motor Activity | 0 | 0 | 0 |
| Tail pinch | 0 | 0 | 0 |

Given the criticism surrounding model uncertainty when modeling a given dataset, and the statistical improvement brought in this field by Bayesian model averaging, we used the BMA method to estimate $BMDTA$ associated with the exposure to five OP pesticides.

### 6.2.3 Bayesian Model Averaging as applied to data resulting from the exposure to five OPs

We used the combination of the following linearizable nonlinear models to create different sample models.

The sample models are all possible combination of the above linearizable nonlinear mod-

Table 6.4: *Linearizable nonlinear endpoint models considered for continuous endpoints (BrainChE, BloodChE and Motor Activity), and for binary endpoint (Tail pinch) as result of five chemicals exposure.*

| *Endpoint Models* | *Linearizable nonlinear models* |
|---|---|
| Continuous | $1-\exp(-(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik}))$<br>$\exp(-\exp(-(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik})))$<br>$1-\exp(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik})$<br>$\frac{1}{1+\exp(-(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik}))}$<br>$\exp(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik})$ |
| Binomial | $(1-\exp(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik}))$<br>$\frac{1}{1+\exp(\psi_0+\sum_{k=1}^{5}\psi_{jk}x_{ik})}$ |

els. In total there are 250 sample models, $|\mathbf{M}| = 250$. Example of sample model:

$$\begin{pmatrix} Y_{1i} \\ \\ Y_{2i} \\ \\ Y_{3i} \\ \\ Y_{4i} \end{pmatrix} \sim \begin{pmatrix} N\left(\exp(-\exp(-(\psi_0 + \sum_{k=1}^{5} \psi_{jk} x_{ik}))), \sigma_1\right) \\ \\ N\left(1 - \exp(\psi_0 + \sum_{k=1}^{5} \psi_{jk} x_{ik}), \sigma_2\right) \\ \\ N\left(\frac{1}{1+\exp(-(\psi_0 + \sum_{k=1}^{5} \psi_{jk} x_{ik}))}, \sigma_3\right) \\ \\ Binom\left(n_i, (1 - \exp(\psi_0 + \sum_{k=1}^{5} \psi_{jk} x_{ik}))\right) \end{pmatrix} \tag{6.9}$$

To specify prior distributions on the hierarchical structure of the model we assume $\boldsymbol{\psi} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ and $\boldsymbol{\Omega} \sim Wishart(I_{24}, 24)$, where $I$ is the identity matrix. Furthermore, we assume $\sigma_j \sim Gamma(1, 1)$ and $\boldsymbol{\mu} \sim N(0, 100)$. For each of the 250 models, WinBUGS generated 4 chains of 11,000 MCMC samples from the corresponding posterior distribution. The chains convergence was diagnosed. The first 1,000 samples from each chain were discarded as burn-in samples. The 10,000 samples from the first chain were used to draw inferences. Since we have no information a priori about the models, we chose $P(M_u) = 1/250$. To implement BMA, we computed $P(M_u|D)$ for each model $M_u \in \mathbf{M}$ using Monte Carlo integration. In this process, the sample model considered in equation 6.3 has almost 100% posterior probability and other models have extremely very low posterior probabilities. We did not have to *average*. At 50%$BMR$, the $BMDTA$ is given by the following inequality:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{2}\right) \\ ln\left(\frac{p_4(\mathbf{0})/2}{1-p_4(\mathbf{0})/2}\right) \end{pmatrix} \geq \begin{pmatrix} -0.009 & -0.005 & -0.092 & -0.038 & -0.001 & -0.023 \\ -0.061 & -0.089 & -0.030 & -0.050 & -0.002 & -0.571 \\ -0.019 & -0.004 & -0.041 & -0.019 & -0.001 & -0.028 \\ -0.309 & -0.001 & -0.002 & -0.003 & -0.001 & 0.000 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{x_k^*} \end{pmatrix} \tag{6.10}$$

where $\mathbf{x_k^*} \geq 0, k = 1, ..., 5.$

Geometrically, each of the above inequality results in a half Euclidean plan. The intersection of all half Euclidean plan form a polytope that we call the benchmark tolerable region. Due to the multidimensionality aspect of the data, we are unable to plot $BMDTA$. However, it is important to know which endpoints define the polytope. Following the logic of determining the endpoints probability i.e., which endpoints define the $BMDTA$, as described in Chapter 4, we determined that BloodChE and BrainChE have each, 1 feasible probability and are both co-sensitive endpoints, whereas Tail pinch and Motor Activity have each 0 feasible probability and are both hypo-sensitive endpoints.

## 6.3 Sensitivity Analysis

In terms of total mixture dose $t$, given the proportion of each chemical in the full mixture (0.040, 0.031, 0.002, 0.102, 0.825) for ACE, CPF, DIA, DIM, and MAL, respectively, we determined that the tolerable region intersects the full mixing ratio ray line at 131.73mg/kg,

Table 6.5: *Feasible probability for each endpoint using the averaged model.*

| *Endpoint* | *Probability* |
|---|---|
| $MotorActivity$ | 0 |
| $BloodChE$ | 1 |
| $BrainChE$ | 0.9 |
| $Tailpinch$ | 0 |

26.62mg/kg, 72.7mg/kg, and 312.56mg/kg for activity, BloodChE, BrainChE, and tail respectively (see Table 6.5). These values were obtained by independently solving for $t$ in each equality in the following system of equations:

$$\begin{pmatrix} ln\left(\frac{f_1(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_2(\mathbf{0})}{2}\right) \\ ln\left(\frac{f_3(\mathbf{0})}{2}\right) \\ ln\left(\frac{p_4(\mathbf{0})/2}{1-p_4(\mathbf{0})/2}\right) \end{pmatrix} = \begin{pmatrix} -0.009 & -0.005 & -0.095 & -0.040 & -0.001 & -0.021 \\ -0.057 & -0.092 & -0.027 & -0.049 & -0.002 & -0.591 \\ -0.018 & -0.004 & -0.033 & -0.021 & -0.001 & -0.026 \\ -0.294 & -0.001 & -0.002 & -0.0036 & -0.001 & -0.0001 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ \mathbf{a_i t} \end{pmatrix} \tag{6.11}$$

where $a_i, i = 1, 2, 3, 4, 5$, are the proportions of chemicals in the mixture.

The minimum, 1.21mg/kg, 2.09mg/kg, 6.47mg/kg, 9.18mg/kg, 26.6mg/kg, is considered as the tolerable dose for the mixture at 5%$BMR$, 10%$BMR$, 20%$BMR$, 25%$BMR$, 50%$BMR$ respectively. They all correspond to BloodChE endpoint, the most sensitive endpoint. The maximum, 60.21mg/kg, 61.24mg/kg, 113.21mg/kg, 171.16mg/kg, 312.5mg/kg, is considered as the tolerable dose for the mixture at 5%$BMR$, 10%$BMR$, 20%$BMR$,

Table 6.6: *The total mixture dose where the tolerable regions intersect the full mixing ratio ray.*

| *Endpoint* | $ED_{05}$ | $ED_{10}$ | $ED_{20}$ | $ED_{25}$ | $ED_{50}$ |
|---|---|---|---|---|---|
| *BrainChE* | 6.6 | 17.26 | 40.67 | 53.30 | 72.7 |
| *BloodChE* | 1.21 | 2.09 | 6.47 | 9.10 | 26.60 |
| *MotorActivity* | 4.56 | 10.35 | 23.06 | 30.10 | 131.7 |
| *Tailpinch* | 60.20 | 61.20 | 113.2 | 171.1 | 312.50 |

$25\%BMR$, $50\%BMR$ respectively. They all correspond to Tail pinch endpoint, the least sensitive endpoint.

## 6.4 Independent analysis, Composite scores, and Bayesian hierarchical model, side by side

Table (6.6) summarizes some characteristics associated with each method. The independent method consists of separate analysis of each endpoint is flexible and easy to implement. Complicated models such as threshold models can be implemented using this method. The method conserves the multidimensionality aspect of the data and does not require data transformation methods. Furthermore, the method allows researchers to determine which endpoint is most sensitive to the exposure. However, this method does not account for combined effects of the chemicals (NRC 2008) which may lead to underestimat-

Table 6.7: *Independent analysis on each endpoint, Composite Scores, and unifying Bayesian approach , side by side*

| Criteria | Independent | Composite Scores | Bayesian Hierarchical |
|---|---|---|---|
| Easy to implement | $X$ | $X$ | |
| Flexible | $X$ | $X$ | |
| Countable number of dimensions | $X$ | $X$ | $X$ |
| Easy to comprehend | $X$ | $X$ | |
| Recommended for mixture ray data | $X$ | $X$ | $X$ |
| Recommended for no ray mixture | | | $X$ |
| Partial or overall conclusion | $X$ | $X$ | $X$ |

Table 6.8: *$ED_{20}$ and $ED_{50}$ Values for the predicted (under additivity) and Experimental Mixture data, for full and reduced rays*: Moser *et al.*, (2005)

| | | Full ray | | Reduced Ray | |
|---|---|---|---|---|---|
| Endpoint | | Additivity | Mixture | Additivity | Mixture |
| BrainChE | $ED_{20}$ | 17.2 | 8.4 | 3.4 | 2.2 |
| | $ED_{50}$ | 47.6 | 26.6 | 10.3 | 7.8 |
| BloodChE | $ED_{20}$ | 11.5 | 6.6 | 4.5 | 2.0 |
| | $ED_{50}$ | 32.2 | 17.8 | 11.8 | 5.4 |
| Motor Activity | $ED_{20}$ | 68 | 52.5 | – | – |
| | $ED_{50}$ | 210 | 81 | – | – |
| Gait Score | $ED_{20}$ | 65.2 | 21.7 | 14.8 | 6.7 |
| | $ED_{50}$ | 94.4 | 31.4 | 21.5 | 9.7 |
| Tail pinch | $ED_{20}$ | 224.4 | 63.3 | – | – |
| | $ED_{50}$ | 295 | 83.2 | – | – |

ing the risk. Independent analysis method lacks the *whole animal evaluation* because the results are specific to one endpoint at a time. Omitting the dependence among observations may also lead to inadequate estimate of the risk associated with the exposure. For assessing the mixture effect, independent analysis method is implemented on mixtures with one fixed mixing ratio ray data.

The side by side comparison of the independent analysis and the composite score methods shows that both methods evaluate one or more fixed environmentally relevant ratio of the mixture. Independent action is a dimension preservative method, composite score is dimension reduction method. Although no specific correlation structure is used the composite score can intuitively evaluate correlated observations; independent action method ignores the possible correlation between observations. The composite score method can evaluate combined effects of the mixture, independent action method ignore the combined effect. Contrary to the independent analysis and composite score methods the $BMDTA$ can evaluate the risk at many environmentally relevant ratios. Furthermore, the $BMDTA$ accounts for the correlation among observations.

# Chapter 7

# Conclusions and Future work

We proposed the Bayesian hierarchical model as a dimension preservative method to analyze multidimensional data. Specifically, the method was used to compute the $BMDTA$ associated with exposure to multiple stressors and multiple observations were taken. It allows the parameters in the model to be correlated in order to capture the possible correlation among observation taken on one subject. Its main advantage is the reliability of the results. Because the method allow the use of the raw data without any transformation, all important information are conveyed into the analysis. It can be applied to any number of dimensions but due to computation requirement, this method is very time consuming and expensive. It is possible that stressors interact between them, in the Bayesian hierarchical model used here to compute the $BMDTA$, we did not account for possible interaction between stressors, this method can be extended to consider interaction models. For proof of concept, we considered simple models to start up with, but not all the endpoints are well fit with the simple model considered and specific individual models that are more complex such as threshold models can be considered. The Bayesian hierarchical model we proposed is capable of capturing the correlation among measurements through the hierarchical model specification which allow parameters to be correlated. We cannot guarantee that, in this way, all the correlation is being accounted for. Specifying correlation structure can be complex because observations may be of different types (binary, categorical, continuous). A method that captures 'all' the correlation is needed. It is assumed that all stressors can cause negative effects on each endpoint, but the amount of negative effects differs depending on endpoint or stressor being considered. To determine the $BMDTA$, it may be important to know which stressor is more dangerous than the other in the sense of causing negative effects. The method presented here can be extended to include methods that can allow stressor selection based on the significance of their biological effects. In all experimental studies the design problem is very crucial, in this case of determining the $BMDTA$ it is important to have a proper design in the low dose region where more emphasis is payed to.

# Bibliography

[1] Akaike, Hirotugu (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6), 716-723.

[2] Allen, B. C., Kavlock, R. J., Kimmel, C. A., Faustman, E. M., (1994a). Dose- Response assessment for developmental toxicity II. Comparison of generic benchmark dose estimate with no observed adverse effect levels. *Fundamental Applied Toxicology*, **23**, 487-495.

[3] Allen, D. M., (1974). The relationship between variable selection and prediction. *Technometrics*, **16**, 125-127.

[4] Barnes, D. G., Daston, G. P., Evans, T. S., Jarabek, A. M., Kavlock, R. J., Kimmel. C. A., Park, C., Spitzer, H. L., (1995). Benchmark dose workshop: criteria for use of a benchmark dose to estimate a reference dose. *Regulatory Toxicology and Pharmacology*, **21**, 296-306.

[5] Berenbaum, M. C., (1981). Criteria for analyzing interactions between biologically active agents. *Journal of Theoretical Biology*, **114**, 413-431.

[6] Berenbaum, M.C., (1989). What is Synergy? *Pharmacology review*, **41**, 93-141.

[7] Bliss, C.I., (1939). The toxicity of poisons applied jointly. *Annals of Applied Biology*, **26**, 585-615.

[8] Carlin, B. P., and Louis, T. A., (1996). *Bayes and Emperical Bayes Methods for data analysis* (First Edition).

[9] Carlin, B. P., and Louis, T. A., (2008). *Bayesian methods for data analysis* (Third Edition)

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

[11] Casey, M., Gennings, C., Carter Jr., H. W., Moser, V. C., Simmons, E. J., (2006). Power and sample size calculations for linear hypotheses associated with mixtures of many components using fixed-ratio ray designs. *Environmental Ecological Statistics*, **13**, 11-23.

[12] Choi, T., Schervish, M. J., Schmitt, K. and Small, M. J., (2009). Bayesian Hierarchical Analysis for Multiple Endpoints in a Toxicity Study. *Journal of Agricultural, Biological, and Environmental Statistics*

[13] Clyde, M.A., (1999). Bayesian Model Averaging and Model Search Strategies Bayesian. *Statistics 6, J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (Eds.),*Oxford University Press, pp. 157-185.

[14] Clyde, M., (2000). Model Uncertainty and Health effect Studies for Particulate Matter. *Environmetrics*, **6**, 745-763.

[15] Coffey, J. T., Gennings, C., Moser, V. C., (2007b). The simultaneous analysis of discrete and continuous outcomes in a dose-response study: Using desirability functions. *Regulatory Toxicology and Pharmacology*, **48**, 51-58.

[16] Coffey, J. T., Gennings, C., (2007a) The Simultaneous Analysis of Mixed Discrete and Continuous Outcomes Using Nonlinear Threshold Models *Journal of Agricultural, Biological, and Environmental Statistics*, **45**, 51-58.

[17] Congdon, P., (2006). Bayesian model choice based on Monte Carlo estimates of posterior model probabilities. *Computational Statistics & Data Analysis*, **50**, 346-357.

[18] Crofton, M. K., Craft, S. E., Hedge, M. J., Gennings, C., Simmons, E. J., Carchman, A. R., Carter Jr., H. W., DeVito, J. M., (2005). Thyroid Hormone-Disrupting Chemicals: Evidence for Dose-Dependent Additivity or Synergism. *Environmental Health Perspectives*, **113**, 1549-1554.

[19] Crowell, J. A., Korytko, P. J., Morrissey, R. L., Booth, T. D., Levine, B. S., (2004). Resveratrol-associated renal toxicity. *Toxicological Sciences*, **82**, 614-619.

[20] Crump, K. S., (1984). A new method for determining allowable daily intakes. *Fundamental Applied Toxicology*, **4**, 854-871.

[21] Crump, K. S., (1995). Calculation of benchmark doses for continuous data. *Risk Analysis*, **15**, 79-89.

[22] Crump, K.S., (2002). Critical issues in benchmark calculations from continuous data. *Critical Reviews Toxicology*, **32**, 133-153.

[23] Derringer, G., (1994). A balancing act: optimizing a product's properties. *Quality Progress*, **27**, 51-58.

[24] Derringer, G. and Suich, R., (1980). Simultaneous Optimization of Several Response Variables. *Journal of Quality Technology*, **12**, 214-219.

[25] Desaulniers, D., Leingartner, K., Musicki, B., Yagminas, A., Xiao, G. H., Cole J. *et al.*, (2003). Effects of postnatal exposure to mixtures of non-ortho-PCBs, PCDDs, and PCDFs in prepubertal female rats. *Toxicological Sciences*, **75(2)**, 468-480.

[26] Faes, C., Geys, H., Aerts, M., Molenberghs, G., Catalano, J. P., (2004). Modeling combined continuous and ordinal outcomes in a clustered setting. *Journal of Agricultural, Biological, Environmental Statistics*, **9(4)**, 515-530.

[27] Faes, C., Geys, H., Aerts, M., Molenberghs, G., (2006). A hierarchical modeling approach for risk assessment in developmental toxicity studies. *Computational Statistics & Data Analysis*, **51**, 1848-1861.

[28] Finney, D. J., (1964). *Statistical Methods in Biological assay (2nd ed.)*, London: Charles Griffin Co.

[29] Finney, D. J., (1964, 1971). *Probit Analysis*, London: Cambridge University Press.

[30] Foster, P. M., and Auton, T. R., (1995). Application of benchmark dose risk assessment methodology to developmental toxicity: an industrial view. *Journal Toxicology Letters*, **82-83**, 555-559.

[31] Fraser, T. R., (1870-1871). An experimental research on the antagonism between the action of thysostigma and atropia. *Proc. R. Soc. Edinburgh*, **7**, 506-511.

[32] Gaylor, D., Ryan, L., Krewski, D., Zhus, Y., (1998). Procedures for calculating benchmark Doses for health risk assessment. *Regulatory Toxicology and Pharmacology*, **28**, 150-164.

[33] Geman, S., and Geman, D., (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **6**, 721-741.

[34] Gelfand, A. E., and Smith. A. F. M., (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398-409.

[35] Gelfand, A. E., and Sahu, S. K., (1999). Identifiability, improper priors and Gibbs sampling for generalized linear models. *Journal of the American Statistical Association*, **94**, 247-253.

[36] Gelman A., Carlin J.B., Stern H.S. and Rubin D.B., (1995). *Bayesian Data Analysis.* Chapman Hall and CRC Press, Boca Raton, FL.

[37] Gennings, C., Schwartz, P., Carter Jr., W.H., Simmons, J.E., (1997). Detection of departures from additivity in mixtures of many chemicals with a threshold model. *Journal of Agricultural, Biological, Environmental Statistics*, **2**, 198-211.

[38] Gennings, C., Carter Jr., H. W., Campain, J. A., Bae, D., Yang, R. S. H., (2002). Statitical analysis of interactive cytotoxicity in human epidermal keratinocytes following exposure to a mixture of four metals. *Journal of Agricultural, Biological, and Environmental Statistics*, **7(1)**, 58-73.

[39] George, E. I., (1993). Variable Selection Via Gibbs Sampling. *Journal of the American Statistical Association*, **88**, 881-889.

[40] Geys, H., Molenberghs, G., and Ryan, M. L., (1999). pseudolikelihood modeling of multivariate outcomes in developmental toxicology *Journal of the American Statistical Association*, **94**, 734-745.

[41] Hasting, W. H., (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrica*, **57**, 97-109.

[42] Harrington, E. C. Jr., (1965). The Desirability Function. *Industrial Quality Control*, **21**, 494-498.

[43] Hocking, R. R., (1976). The Analysis and Selection of Variables in Linear Regression. *Biometrics*, **32**,150.

[44] Hochberg, Y., (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, **75**, 800-802.

[45] Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T., (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, **14**, 382-417.

[46] Holm, S. (1979). A Simple Sequentially Rejective Bonferroni Test Procedure. *Scandinavian Journal of Statistics*, **6**, 65 -70.

[47] Johnson,A. Richard, and Wichern, W. Dean, (2002). *Applied Multivariate Statistical Analysis* (Fifth Edition).

[48] Kass, R., and Raftery, A. E, (1995). Bayes Factors. *Journal of the American Statistical Association*, **90**, 773-795.

[49] Kleinbaum,G. David, Kupper, L. Lawrence, Muller, E. Keith, Nizam, Azhar (1998). *Applied Regression Analysis and Other Multivariate Methods* (Third Edition).

[50] Kimmel, C. A., and Gaylor,D., (1988). Issues in qualitative and quantitative risk analysis for developmental toxicology. *Risk Analysis*, **8**, 15-21.

[51] Krewski, D., (1990). Measuring carcinogenic potency. *Risk Analysis* , **10(4)**, 615-617.

[52] Kuljus, K., Rosen, D., Sand, S., Victorin, K., (2006). Comparing Experimental Designs for Benchmark Dose Calculations for Continuous Endpoints. *Risk Analysis*, **26(4)**, 1031-1043

[53] Lamon, E. C. and Clyde, M. A., (2000). Accounting for Model Uncertainty in Prediction of Chlorophyll A in Lake Okeechobee. *Journal of Agricultural Biological and Environmental Statistics*, **5**, 297-322.

[54] Leamer, E. E., (1978). *Specification Searches: Ad Hoc Inference with Non-Experimental Data.* John Wiley & Sons, New York.

[55] Leisenring, W., and Ryan, L. ,(1992). Statistical properties of the NOAEL. *Regulatory Toxicology and Pharmacology*, **15**, 161-171.

[56] Loewe, S., (1953). The problem of synergism and antagonism of combined drugs. *Arzeimittle forshung*, **3**, 285-290.

[57] Madigan, D. M., Gavrin, J. and Raftery, A. E., (1995). Eliciting Prior Information to Enhance the Predictive Performance of Bayesian Graphical Models. *Communications in Statistics - Theory and Methods*, **24**, 2271-2292.

[58] Madigan, D. M. and Raftery, A. E., (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, **89**, 1535-1546.

[59] Madigan, D. M. and York, J., (1995). Bayesian Graphical Models for Discrete Data. *International Statistical Review*, **63**, 215-232.

[60] Mallows, C. L., (1973). Some comments on Cp. *Technometrics*, **15(4)**, 661-675.

[61] Mantel, N., (1958). An experimental design in combination chemotherapy. *Annals of the New York academy of sciences*, **76**, 909.

[62] Martin, J. T., (1942). The problem of the evaluation of rotenone-containing plants. VI. The toxicity of 1-elliptone and of Poisons applied jointly, with further observations on the rotenone equivalent method of assessing the toxicity of Derris root. *Annals of applied Biology*, 29, 69-81.

[63] McCullagh, P. and Nelder, J. A., (1989). *Generalized Linear Models*, second ed. Chapman and Hall, New York.

[64] Meadows-Shropshire, S. L., Gennings, C., Carter, W. H. Jr, (2005). Sample size and power determination for detecting interactions in mixtures of chemicals. *Journal of the American Statistical Association*, **10**, 104-117.

[65] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E., (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**, 1087-1091.

[66] Monosson, E., (2005). Chemical Mixtures: Considering the Evolution of Toxicology and Chemical Assessment. *Environmental Health Perspectives*, **113**, 383-389. *Environmental Health Perspectives*, **113**, 383-389.

[67] Moser, V. C., (2000). Observational batteries in neurotoxicity testing. *International Journal of Toxicology*, **19**, 407-411.

[68] Moser V. C, Cheek, B. M., and MacPhail, R. C., (1995). A multidisciplinary approach to toxicological screening. III. Neurobehavioral toxicity. *Journal of Toxicology and Environmental Health*, **45**, 173-210.

[69] Moser, V. C., Casey, M., Hamm, A., Carter, Jr. W. H., 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.

[70] Murphy, M. and Wang, D., (2001). Do Previous Birth Interval and Maternal Education Influence Infant Survival? A Bayesian Model Averaging Analysis of Chinese Data. *Population Studies*, **55**, 37-48.

[71] Murrell, J. A., Portier, C. J., & Morris, R. W., (1998). Characterizing dose-response I: Critical assessment of the benchmark dose concept. *Risk Analysis*, **18**, 1326.

[72] Myers, D. P., (2001a). Dimethoate developmental neurotoxicity study in the CD rat by oral gavage administration. *Study conducted by Huntingdon Life Sciences, Ltd under international GLP.*

[73] National Research Council. Phthalates and Cumulative Risk Assessment: The Task Ahead. *Available: http://www.nap.edu/catalog/12528.html*

[74] Raftery, A. E, (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, **25**, 111-163.

[75] Raftery, A. E., (1996). Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika*, **83**, 251-266.

[76] Raftery, A. E., Madigan, D. and Hoeting, J. A., (1997). Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, **92**, 179-191.

[77] Regan, M. M., & Catalano, P. J., (1999a). Bivariate dose-response modeling and risk estimation in developmental toxicology. *Journal of Biological, Agricultural, and Environmental Statistics*, **4**, 217-237.

[78] Reed, M. D., Gigliotti, A. P., McDonald, J. D., Seagrave, J. C., Seilkop, S. K., Mauderly, J. L., (2004). Health effect of subchronic exposure to environmental levels of diesel exhaust. *Inhalation Toxicology*, **16**, 177-193.

[79] Reiss, R., Gaylor, D., (2005). Use of benchmark dose and meta-analysis to determine the most sensitive endpoint for risk assessment for dimethoate. *Regulatory Toxicology and Pharmacology*, **43**, 55-65.

[80] Ryan, L. M., (1992). Quantitative risk assessment for developmental toxicity. *Biometrics*, **48**, 163-174.

[81] Sammel, M. D., Ryan, M. L., and Legler, J. M., (1997). Latent variable models for mixed discrete and continuous outcomes. *Journal of Royal Statistics Society B*, **59**, 667-678.

[82] Scholze, M., and Kortenkamp, A., (2007). Statistical power considerations show the endocrine disrupter low dose issue in a new light. *Environmental Health Perspectives*, **115**, 84-90.

[83] Schwarz, G., (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.

[84] Shih, M., Gennings, C., Chinchilli, V. M., and Carter, W. H., Jr., (2003). Titrating and evaluating multi-drug regimens within subjects. *Statistics in Medicine*, **22**, 2257-2279.

[85] Slob, W., (1999). Thresholds in toxicology and risk assessment. *International Journal of Toxicology*, **18**, 259-268.

[86] Slob, W., (2002). Dose-Response Modeling of continuous Endpoints. *Toxicology Sciences*, **66**, 298-312.

[87] Slob, W., & Pieters, M. N., (1998). A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. *Risk Analysis*, **18**, 787-798.

[88] Teuschler, L., Klaunig, J., Carney, E., Chambers, J., Conolly, R., Gennings, C., Giesy, J., Hertzberg, R., Klaassen, C., Kodell, R., Paustenbach, D., and Yang, R., (2002). Support of Science-Based Decisions Concerning the Evaluation of the Toxicology of Mixtures: A New Beginning. *Regulatory Toxicology and Pharmacology* , **36**, 34-39.

[89] U.S. Environmental Protection Agency (1986). *Guidelines for the Health Risk Assessment of Chemical Mixtures*

[90] U.S. Environmental Protection Agency (1995). *The Use of the Benchmark Dose Method in Health Risk Assessment*, Report EPA/630/R-94/007. Washington DC: U.S. Environmental Protection Agency.

[91] U.S. Environmental Protection Agency (1996). *Summary of State Biological Assessment Programs for Streams and Rivers*, Office of Planning, Policy and Evaluation, EPA-230-R-96-007, Washington D.C.

[92] U.S. Environmental Protection Agency (2000). *Benchmark Dose Technical Guidance Document EPA/630/R-00/001. October 2000*, External Review Draft. Risk Assessment Forum, Washington DC.

[93] Viallefont, V., Raftery, A. E., and Richardson, S., (2001). Variable Selection and Bayesian Model Averaging in Case Control Studies. *Statistics in Medicine*, **20**, 3215-3230.

[94] Wessinger, W. D., (1976). Approaches to the study of Drug Interaction in Behavioral Pharmacology. *Neuroscience behavioral review*, **10**, 103-113.

[95] World Health Organization (WHO), (1999). *Principles for the assessment of risks to human health from exposure to chemicals*, Environmental Health Criteria 210, World health Organization, Geneva.

[96] Wu, F. C, (2005). Optimization of correlated multiple Quality Characteristics using Desirability Function. *Quality Engineering*, **17**, 119-126.

[97] Yu, Z. F., & Catalano, P. J., (2005). Quantitative risk assessment for multivariate continuous outcomes with application to neurotoxicology: The bivariate case. *Biometrics*, **61**, 757-766.

[98] Yu, Z. F., and Catalano, P. J., (2008). A Simulation Study of Quantitative Risk Assessment for Bivariate Continuous Outcomes. *Risk Analysis*, **28**, 1415-1430

# Appendix: SAS code

```
{SAS code for chapter 2}

goptions device=cgmof97L gsfname=pic gsfmode=replace colors=(black) ftext=swiss htext=1.

title ’Analysis of Dioxin Dose-Response Data’;

data paper;

%infile ’E:\paper.txt’;

input obs block $ chemical $ dose perT4 chem;

run;

axis1 label=(a=90 ’T4 (%Control)’) order=(0 to 180 by 20);

axis2 label=(’Dose (ug/kg)’);

symbol1 v=star i=none;

symbol2 v=none i=join;

symbol3 v=none i=join l=2;

%filename BMD ’C:\Documents and Settings\nyirabahizie\Desktop\BMD’;

goptions colors=(black)device=cgmof97L gsfname=BMD gsfmode=replace colors=(black) ftext=
proc gplot data=paper;

title ’Dioxin Dose-Response Data’;

plot perT4*dose/vaxis=axis1 haxis=axis2;

run;

quit;

data mixture;

set paper;

where chem=19;

run;

proc print data=mixture;

run;

proc means data=mixture mean n std noprint;

   by chemical dose;

   var pert4;

   output out=meansmix mean=pt4mean std=pt4std var=pt4var n=pt4n;

data forplot;

   chemical=’Mixture’;

   do dose=0 to 2020 by 20;

      output;

  end;

   do dose=5 to 70 ;
```

```
        output;

    end;

    data mix;

     set mixture forplot;

     total=dose;

     maxdose=2100;

data alldat;

     set paper;

     where chem ne 19;

     run;

proc sort data=alldat; by chemical dose;


data t4;

     set alldat; by chemical dose;

     retain chem;

    % if _n_=1 then chem=0;

     if first.chemical then chem=chem+1;

    % log_pert4=log(pert4);

     keep block chemical chem pert4 dose $log_pert4$;


proc format;
```

```
   value chem 1=’1-PCDF’ 2=’4-PCDF’ 3=’PCB101’ 4=’PCB105’ 5=’PCB118’ 6=’PCB126’
              7=’PCB138’ 8=’PCB153’ 9=’PCB156’ 10=’PCB169’ 11=’PCB180’ 12=’PCB28’
              13=’PCB52’ 14=’PCB77’ 15=’PCDD’ 16=’TCDD’ 17=’TCDF’ 18=’OCDF’ 19=’Mixture’
   value $ ais a1=’1-PCDF’ a2=’4-PCDF’ a3=’PCB101’ a4=’PCB105’ a5=’PCB118’ a6=’PCB126’
              a7=’PCB138’ a8=’PCB153’ a9=’PCB156’ a10=’PCB169’ a11=’PCB180’ a12=’PCB28’
              a13=’PCB52’ a14=’PCB77’ a15=’PCDD’ a16=’TCDD’ a17=’TCDF’ a19=’Mixture’;
run;

data forplot;
   do chem=1 to 14, 17;
     do dose=0 to 99 by .25;
         output;
       end;
     do dose=100 to 300 by 10, 300 to 20000 by 100;
    output;
end;
end;
    do chem=15, 16;
    do dose= 0 to 1 by .05, 1 to 3 by .1;
   output;
   end;
```

```
         end;
data temp;
   set mix;
   keep chemical dose total maxdose mix add chem pert4  ;
   total=dose; maxdose=2100; mix=1; add=0; chem=19;
data ais;
   a1= .000003;
   a2= .000013;
   a3= .076814;
   a4= .038282;
   a5= .190302;
   a6= .000302;
   a7= .190181;
   a8= .190861;
   a9= .006541;
   a10=.000197;
   a11=.188700;
   a12=.039237;
   a13=.077523;
   a14=.000988;
   a15=.000007;
   a16=.000007;
   a17=.000010;
proc transpose data=ais out=aisfile;
   $format _name_ ais.$;
run;
data two;
   set t4 forplot temp;
   x01=0; x02=0; x03=0; x04=0; x05=0; x06=0; x07=0; x08=0; x09=0;
   x010=0; x011=0; x012=0; x013=0; x014=0; x015=0; x016=0; x017=0; x018=0; x019=0;
   x1=0; x2=0; x3=0; x4=0; x5=0; x6=0; x7=0; x8=0; x9=0;
   x10=0; x11=0; x12=0; x13=0; x14=0; x15=0; x16=0; x17=0; x18=0; x19=0;
   a1=.000003;  a2=.000013;   a3=.076814;   a4=.038282;   a5=.190302;   a6=.000302;   a7
   a8=.190861;   a9=.006541;   a10=.000197;   a11=.188700;   a12=.039237;   a13=.077523;
   a16=.000007;   a17=.000010; a18=.000032;
   maxmix=2000;
   mixdose1=a1*maxmix;   mixdose2=a2*maxmix;   mixdose3=a3*maxmix;   mixdose4=a4*maxmix;
   mixdose6=a6*maxmix;   mixdose7=a7*maxmix;   mixdose8=a8*maxmix;   mixdose9=a9*maxmix;
   mixdose11=a11*maxmix;   mixdose12=a12*maxmix;   mixdose13=a13*maxmix;   mixdose14=a14
   mixdose16=a16*maxmix;   mixdose17=a17*maxmix;   mixdose18=a18*maxmix;
   if chem<19 then total=0;
   if chem=1 then do; x01=1; x1=dose; maxdose=100;      mixdose=mixdose1; end;
if chem=2 then do; x02=1; x2=dose; maxdose=90;*20;      mixdose=mixdose2; end;
if chem=3 then do; x03=1; x3=dose; maxdose=30000;   mixdose=mixdose3;  end;
if chem=4 then do; x04=1; x4=dose; maxdose=90000; *1200;    mixdose=mixdose4; end;
if chem=5 then do; x05=1; x5=dose; maxdose=10000; *1200;    mixdose=mixdose5; end;
if chem=6 then do; x06=1; x6=dose; maxdose=100; *4;       mixdose=mixdose6; end;
if chem=7 then do; x07=1; x7=dose; maxdose=90000;   mixdose=mixdose7;  end;
if chem=8 then do; x08=1; x8=dose; maxdose=90000;   mixdose=mixdose8;  end;
if chem=9 then do; x09=1; x9=dose; maxdose=10000; *1000;    mixdose=mixdose9; end;
if chem=10 then do; x010=1; x10=dose; maxdose=1000; *120;  mixdose=mixdose10;  end;
if chem=11 then do; x011=1; x11=dose; maxdose=90000; mixdose=mixdose11;  end;
if chem=12 then do; x012=1; x12=dose; maxdose=90000; *20000;mixdose=mixdose12;  end;
if chem=13 then do; x013=1; x13=dose; maxdose=90000; *20000;mixdose=mixdose13;  end;
if chem=14 then do; x014=1; x14=dose; maxdose=30000; mixdose=mixdose14; end;
if chem=15 then do; x015=1; x15=dose; maxdose=10; *1.2;  mixdose=mixdose15; end;
if chem=16 then do; x016=1; x16=dose; maxdose=10; *.4;   mixdose=mixdose16; end;
if chem=17 then do; x017=1; x17=dose; maxdose=100; *5;    mixdose=mixdose17; end;
if chem=18 then do; x018=1; x18=dose; maxdose=300;  mixdose=mixdose18;  end;
if mix=1 then do; chem=19; x019=1; x19=dose; maxdose=2100; end;
if chem<=18 then do; add=1; mix=0;end;
if dose<= max(maxdose,mixdose);
mixdoseresponse=0;
```

```
  yp = (pert4-40)/60;

  transy = log( yp/(1-yp) );

 format chem chem.;

 run;

 proc sort data=two; by chem dose;

 proc reg data=two;

  model transy = x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19;

  run;

proc nlmixed data=two  ALPHA=.1;

   where pert4 ne .;

   parms alpha=0  alpha4579=0   alphamix=50

        b0=1.5 b1=-.08 b2=-.06 b3=-.0003 b4=-.0008 b5=-.001 b6=-.3 b7=-.0003 b8=-.0001

        b10=-.007 b11=-.00005 b12=-.00003 b13=-.00003 b14=-.001 b15=-.8 b16=-2 b17=-.2

        thetamix=-0.002 sig=310;

bigalpha = alpha*(chem=1) + alpha*(chem=2) + alpha*(chem=3) + alpha4579*(chem=4) + alph

      +  alpha*(chem=6) + alpha4579*(chem=7)

      +  alpha*(chem=8) + alpha4579*(chem=9) + alpha*(chem=10) + alpha*(chem=11) + alph

      +  alpha*(chem=13) + alpha*(chem=14) + alpha*(chem=15) + alpha*(chem=16)

      +  alpha*(chem=17) + alpha*(chem=18)

  +  alphamix*(chem=19);

  g=100-bigalpha;
```

```
term= b1*x1 + b2*x2 + b3*x3 + b4*X4 + b5*x5 +b6*x6 + b7*x7 + b8*x8 +b9*x9 + b10*x10 +
      + b13*x13 + b14*x14 + b15*x15 + b16*x16 + b17*x17 ;
 mu    = bigalpha+  g/(1+exp(-(b0+ term*(chem<19) + thetamix*dose*(chem=19) )));
 rangemix = (100-alphamix)/(1+exp(-b0));
 muforED = alphamix+0.9*rangemix;
 estimate ’interceptmix’ alphamix+rangemix;
 range = (100-alpha)/(1+exp(-b0));
 estimate ’intercept’ alpha+range;
 range4579 = (100-alpha4579)/(1+exp(-b0));
 estimate ’intercept4579’ alpha4579+range4579;
 estimate ’mufored’ mufored;
 edmu1 = (log( (mufored-alpha)/(100-alpha - mufored+alpha)) -b0)/b1;
 edmu2 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0)/b2;
 edmu3 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b3;
 edmu4 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b4;
 edmu5 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b5;
 edmu6 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b6;
 edmu7 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b7;
 edmu8 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b8;
 edmu9 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b9;
 edmu10 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b10;
```

```
   edmu11 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b11;
   edmu12 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b12;
   edmu13 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b13;
   edmu14 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b14;
   edmu15 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b15;
   edmu16 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b16;
   edmu17 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b17;
   edmix = (log( (muforED-alphamix)/(100-alphamix- mufored+alphamix)) -b0 )/thetamix;
   estimate ’EDmix’ edmix;
   edadd = 1/( a1/edmu1 + a2/edmu2 + a3/edmu3 + a4/edmu4 + a5/edmu5 + a6/edmu6 + a7/edmu
             + a8/edmu8 + a9/edmu9 + a10/edmu10 + a11/edmu11 + a12/edmu12 + a13/edmu13
             + a14/edmu14 + a15/edmu15 + a16/edmu16 + a17/edmu17 );
  estimate ’EDadd’ edadd;
  estimate ’EDadd-EDmix’ edadd-edmix;
   model pert4 ~ normal(mu, sig);
   predict mu out=pred;
   ods output parameterestimates=vars;
run;
*proc sort data=pred;
proc sort data=prednlin out=pred; by chem dose;
axis1 label=(a=90 ’T4 (%Control)’) order=(0 to 180 by 20);
```

```
axis2 label=('Dose (ug/kg)');

symbol1 v=star i=none;

   edmu5 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b5;

   edmu6 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b6;

   edmu7 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b7;

   edmu8 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b8;

   edmu9 = (log( (mufored-alpha4579)/(100-alpha4579- mufored+alpha4579)) -b0 )/b9;

   edmu10 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b10;

   edmu11 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b11;

   edmu12 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b12;

   edmu13 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b13;

   edmu14 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b14;

   edmu15 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b15;

   edmu16 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b16;

   edmu17 = (log( (mufored-alpha)/(100-alpha- mufored+alpha)) -b0 )/b17;

   edmix = (log( (muforED-alphamix)/(100-alphamix- mufored+alphamix)) -b0 )/thetamix;

   estimate 'EDmix' edmix;

   edadd = 1/( a1/edmu1 + a2/edmu2 + a3/edmu3 + a4/edmu4 + a5/edmu5 + a6/edmu6 + a7/edmu

             + a8/edmu8 + a9/edmu9 + a10/edmu10 + a11/edmu11 + a12/edmu12 + a13/edmu13

             + a14/edmu14 + a15/edmu15 + a16/edmu16 + a17/edmu17 );

  estimate 'EDadd' edadd;
  estimate 'EDadd-EDmix' edadd-edmix;
   model pert4 ~ normal(mu, sig);
   predict mu out=pred;
   ods output parameterestimates=vars;
run;
*proc sort data=pred;
proc sort data=prednlin out=pred; by chem dose;
axis1 label=(a=90 'T4 (%Control)') order=(0 to 180 by 20);
axis2 label=('Dose (ug/kg)');
symbol1 v=star i=none;
symbol2 v=none i=join;
symbol3 v=none i=join l=2;
proc gplot data=pred;
   by chem;
   plot (pert4 pred)*dose mixdoseresponse*mixdose/overlay vaxis=axis1 haxis=axis2;
   title ' ';
   run; quit;
  proc transpose data=vars out=betas;
   var estimate;
   id parameter;
   data muadd;
   merge betas ais;
   do muadd = 53 to 98;
   edmu1 = (log( (muadd-alpha)/(100- muadd)) -b0)/b1;
   edmu2 = (log( (muadd-alpha)/(100- muadd)) -b0)/b2;
   edmu3 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b3;
   edmu4 = (log( (muadd-alpha4579)/(100- muadd)) -b0 )/b4;
   edmu5 = (log( (muadd-alpha4579)/(100- muadd)) -b0 )/b5;
   edmu6 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b6;
   edmu7 = (log( (muadd-alpha4579)/(100- muadd)) -b0 )/b7;
   edmu8 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b8;
   edmu9 = (log( (muadd-alpha4579)/(100- muadd)) -b0 )/b9;
   edmu10 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b10;
   edmu11 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b11;
   edmu12 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b12;
   edmu13 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b13;
   edmu14 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b14;
   edmu15 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b15;
   edmu16 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b16;
   edmu17 = (log( (muadd-alpha)/(100- muadd)) -b0 )/b17;
   edadd = 1/( a1/edmu1 + a2/edmu2 + a3/edmu3 + a4/edmu4 + a5/edmu5 + a6/edmu6 + a7/edmu
end;
```

```
data muadd;
   set muadd;
   keep edadd muadd;
   array edmu edmu1-edmu17;
   do over edmu;
    if edmu<0  then delete;
   end;
run;
data atzero;
  set betas;
  muadd= alphamix+ (100-alphamix)/(1+exp(-b0));  edadd=0;
  keep muadd edadd;
data muaddplus;
  set muadd atzero;
  chem=19;
data pred2;
  set pred muaddplus;
*proc print data=pred2;
*   where chem=19;
*   var chem pert4 pred dose muadd edadd;
axis3 label=('Total Dose of Mixture (ug/kg)') order=(0 to 2100 by 300) ;
```

```
proc gplot data=pred2;

   by chem;

   where chem=19;

   plot (pert4 pred )*dose muadd*edadd /overlay vaxis=axis1 haxis=axis3;

   title ' ';

   run; quit;


{SAS code for chapter 3}

proc format;

   value chem 1='acephate' 2='diazinon' 3='CPF' 4='malathion' 5='dimethoate';

proc contents data=allpest.sas7bdat;

/*proc print data=library.allpest;;

run;*/

*title 'OP pesticides and toxicity score';

data analysis;

 set library.allpest;

if pc=1 then delete;

 if mixray=2 and dose=100 then delete;

 /*proc print data=analysis;

 where mixray=2;

run;*/
```

```
 proc sort data=analysis;by group ;run;

data analysis;

 set analysis;

act=activity;

 activity=100*(activity/199.4285714);

brnche=100*(brainche/5.5410714);

bldche=100*(bloodche/0.4882856);

gait=gaitscr;

tail=tailpinch;

actmin=1/(1+exp((activity-143.0)/12.2703));

actmax=1/(1+exp(-(activity-53.2243)/18.7851));

actdes=actmin*actmax;

brnmin=1/(1+exp((brnche-154.4)/10.9430));

brnmax=1/(1+exp(-(brnche-59.3959)/14.8480));

brndes=brnmin*brnmax;

bldmin=1/(1+exp((bldche-155.3)/7.2806));

bldmax=1/(1+exp(-(bldche-41.5173)/24.9715));

blddes=bldmin*bldmax;

tailmin=1/(1+exp((tail-5.0897)/0.4846));

tailmax=1/(1+exp(-(tail-1.6547)/0.6368));

taildes=tailmin*tailmax;
```

```
gaitdes=1/(1+exp((gait-3.5469)/0.8582));
k=5-nmiss(actdes, brndes, blddes, gaitdes, taildes);
logactdes = log(actdes);
logbrndes = log(brndes);
logblddes = log(blddes);
loggaitdes = log(gaitdes);
logtaildes = log(taildes);
logd = sum(logactdes, logbrndes, logblddes, loggaitdes, logtaildes)/k;
D = exp(logd);
if pc=0;
 proc sort data=analysis; by mixray group pc dose;
 proc print;
 run;
 data forplot;
  group='MX1';
  do dose=0 to 10 by .2, 10 to 100 by 1, 100 to 450 by 10;
     mixray=1;
      output;
  end;
   group='MX2';
  do dose=0 to 10 by .2, 10 to 78.8 by .5;
```

```
    mixray=2;
      output;
  end;
 mixray=0;
  group='ACE';
  do dose=0 to 10 by .2, 10 to 120 by 1;
      output;
  end;
 group='CPF';
 do dose=0 to 10 by .2, 10 to 50 by 1;
      output;
  end;
  group='DIA';
  do dose=0 to 10 by .2, 10 to 100 by 1, 100 to 250 by 10;
      output;
  end;
  group='DIM';
 do dose=0 to 10 by .2, 10 to 75 by 1;
      output;
  end;
data analysis2;
```

```
   set analysis forplot;
 x1=0; x2=0; x3=0; x4=0; x5=0; t1=0; t2=0;
 x01=0; x02=0; x03=0; x04=0; x05=0; t01=0; t02=0;
 if group='ACE' then do; x01=1; x1=dose; end;
 if group='CPF' then do; x02=1; x2=dose;end;
 if group='DIA' then do; x03=1; x3=dose;end;
 if group='DIM' then do; x04=1; x4=dose;end;
 if group='MAL' then do; x05=1; x5=dose;end;
 if group='CON' then group='MIX';
 if mixray=1 then group='MX1';
 if mixray=2 then group='MX2';
 if group='MX1' then do; t01=1; t1=dose;end;
 if group='MX2' then do; t02=1; t2=dose;end;
 run;
proc sort data=analysis; by group dose;
proc gplot data=analysis;
 by dose;
 plot (actdes brndes blddes gaitdes taildes D)*dose;
run; quit;
 proc sort data=analysis2; by group dose;
proc means data=analysis2;
```

```
   where dose=0;
   var d;
   run;
   data analysis3;
   set analysis2;
   if d ne .;
ods output ParameterEstimates=betas;
   proc nlmixed data=analysis2 technique=quanew alpha=.1;
 * where k=5;
   parms  alpha=.23  b1=-.034 b2=-.032 b3=-.009 b4=-.13 b5=-.0001 theta1=-.018 amix=.34
          del4=4.7   sig2=0.001 delmix2=14 theta2=-.0417 delmix=10;
   int=0.88;
   mu1 = alpha + (int-alpha)*exp(b1*dose);
   mu2 = alpha + (int-alpha)*exp(b2*dose);
   mu3 = alpha + (int-alpha)*exp(b3*dose);
  mu4 = alpha + (int-alpha)*exp(b4*(dose-del4)*(dose>del4));
   mu5 = alpha + (int-alpha)*exp(b5*dose);
   *mumix = amix + (int-amix)*exp(theta1*dose);
   mumix = amix + (int-amix)*exp(theta1*(dose-delmix)*(dose>delmix));
   mumixred= amix + (int-amix)*exp(theta2*(dose-delmix2)*(dose>delmix2));
  * mumixred= amix + (int-amix)*exp((theta1/.175)*(dose-delmix2)*(dose>delmix2));
```

```
   mu = mu1*x01 + mu2*x02 + mu3*x03 + mu4*x04 + mu5*x05 + mumix*t01 + mumixred*t02;
   model D ~ normal(mu,sig2);
*run;
   mu0=.95*(int-amix)+amix;
   estimate 'mu0' mu0;
   ED1 = log( (mu0-alpha)/(int-alpha))/b1;
  ED2 = log( (mu0-alpha)/(int-alpha))/b2;
   ED3 = log( (mu0-alpha)/(int-alpha))/b3;
   ED4 = log( (mu0-alpha)/(int-alpha))/b4 + del4;
   ED5 = log( (mu0-alpha)/(int-alpha))/b5;
   edmix = log( (mu0-amix)/(int-amix))/theta1+delmix;
   *edmix = log( (mu0-amix)/(int-amix))/theta1;
   edmixred = log( (mu0-amix)/(int-amix))/(theta2) + delmix2;
  estimate 'ed1' ed1;
  estimate 'ed2' ed2;
  estimate 'ed3' ed3;
  estimate 'ed4' ed4;
  estimate 'ed5' ed5;
  estimate 'edmix' edmix;
  *estimate 'edmixadj' edmixadj;
  a1=0.040; a2=0.031; a3=0.002; a4=0.102; a5=0.825;
```

```
   t0add = 1/(a1/ED1 + a2/ED2 + a3/ED3 + a4/ED4 + a5/ED5);

   estimate 'edadd' t0add;

   estimate 'edadd-edmix' t0add-edmix;

   a1red=0.229; a2red=0.011; a3red=0.117; a4red=0.583; *a5=0;

   t0addred = 1/(a1red/ED1 + a2red/ED2 + a3red/ED3 + a4red/ED4 );

   estimate 'edmixred' edmixred;

   estimate 'edaddred' t0addred;

   estimate 'edaddred-edmixred' t0addred-edmixred;

   thetafulladj=theta1*(1/.175);

   estimate 'thetafulladj' thetafulladj;

   *edmixadj= log(mu0-amix)/(int-amix)/thetafulladj;

   *estimate 'edmixadj' edmixadj;

   *contrast 'fulladja' theta2-theta1*.825,delmix2-delmix/.825;

  contrast 'full' theta2-theta1*(1/.175);

  contrast 'addel' delmix2-delmix;

   predict mu out=prednl;

run;

proc transpose data=betas out=betast;

   var estimate;

   id parameter;

proc print data=betast; run;
```

```
data add;
 set betast;
  mixray=0;
  int=0.88;
  group='ACE';
  do dose=0 to 10 by .2, 10 to 120 by 1;
     mu = alpha + (int-alpha)*exp(b1*dose);
      output;
  end;
  group='CPF';
  do dose=0 to 10 by .2, 10 to 50 by 1;
      mu = alpha + (int-alpha)*exp(b2*dose);
      output;
  end;
  group='DIA';
  do dose=0 to 10 by .2, 10 to 100 by 1, 100 to 250 by 10;
      mu = alpha + (int-alpha)*exp(b3*dose);
      output;
 end;
  group='DIM';
  do dose=0 to 10 by .2, 10 to 75 by 1;
```

```
   mu= alpha + (int-alpha)*exp(b4*(dose-del4)*(dose>del4));
      *mu = alpha + (int-alpha)*exp(b4*dose);
     output;
  end;
   group='MAL';
   do dose = 0 to 500 by 100;
     mu = alpha + (int-alpha)*exp(b5*dose);
  output;
  end;
data add;
   set analysis add;
    set prednl add;
   if group='CON' then group='MIX';
  if mixray=0;
symbol1 c=black i=none v=star;
symbol2 c=black i=join v=none l=1;
proc sort; by group dose;
proc gplot data=add;
   by group;
 *where group='MIX1';
  plot (d mu)*dose /haxis=axis1 vaxis=axis2 overlay;
```

```
   title ' ';
   run; quit;
data muadd;
   set betast;
   group='MX1';
   a1=0.040; a2=0.031; a3=0.002; a4=0.102;  a5=0.825;
   mu0=0.88;
    do muadd=0.35 to 0.88 by .01;
      ED1 = log( (muadd-alpha)/(mu0-alpha) )/b1;
      ED2 = log( (muadd-alpha)/(mu0-alpha) )/b2;
      ED3 = log( (muadd-alpha)/(mu0-alpha) )/b3;
      ED4 = log( (muadd-alpha)/(mu0-alpha) )/b4;
      ED5 = log( (muadd-alpha)/(mu0-alpha) )/b5;
     tadd = 1/(a1/ED1 + a2/ED2 + a3/ED3 + a4/ED4 + a5/ED5);
     output;
 end;
   muadd=.; tadd=.; ed1=.; ed2=.; ed3=.; ed4=.; ed5=.;
   do dose=0 to 300 by .5;
   mufull = amix+(mu0-amix)*exp(theta1*dose);
   *mufull = amix+(mu0-amix)*exp(theta1*(dose-delmix)*(dose>delmix));
   *muredadjfull =  mu0*exp((theta2*.175)*(dose-(delmix2*.175)))*(dose>(delmix*.175));
```

```
    muredadjfull =amix+ (mu0-amix)*exp((theta2*(1/.175))*(dose-(delmix1))*(dose>(delmix1

      output;

 end;

*proc print;

*var muadd tadd mu dose;

*run;

data pred;

 set prednl muadd;

  where group='MX1';


proc sort data=pred; by group dose;

*proc print data=pred;

 *var group d pred dose mu muadd  ed1 ed2 ed3 ed4 tadd;

*run;

symbol1 c=black i=none v=star;

symbol2 c=black i=join v=none l=1;

symbol3 c=black i=join v=none l=2;

symbol4 c=black i=join v=none l=4;

legend1 value=('observed' 'predicted' 'additive')

        shape=symbol(4,2)

       position=(bottom left inside)
```

```
        mode=share ;
data muaddred;
   set betast;
   group='MX2';
   a1red=0.229; a2red=0.011; a3red=0.117; a4red=0.583; *a5=0;
  mu0=0.88;
  do muaddred=0.35 to 0.88 by .01;
      ED1red = log( (muaddred-alpha)/(mu0-alpha) )/b1;
      ED2red = log( (muaddred-alpha)/(mu0-alpha) )/b2;
      ED3red = log( (muaddred-alpha)/(mu0-alpha) )/b3;
      ED4red = log( (muaddred-alpha)/(mu0-alpha) )/b4;
      taddred = 1/(a1red/ED1red + a2red/ED2red + a3red/ED3red + a4red/ED4red);
     output;
 end;
   muaddred=.; taddred=.; ed1red=.; ed2red=.; ed3red=.; ed4red=.;
   do dose=0 to 10 by .2, 10 to 78.8 by .5;
    mured = amix+ (mu0-amix)*exp(theta2*(dose-delmix2)*(dose>delmix2));
   *mufulladjred =amix+ (mu0-amix)*exp((theta1*(1/.825))*(dose-delmix/(1/.825))*(dose>de
   *mufulladjred =amix+ (mu0-amix)*exp((theta1*(1/.825))*(dose-(delmix*.825))*(dose>(del
    mufulladjred =amix+ (mu0-amix)*exp((theta1*(1/.175))*(dose-(delmix2))*(dose>(delmix2
   *mufulladjred =amix+ (mu0-amix)*exp((theta1*(1/.825)*dose));
   *mured = amix+ (mu0-amix)*exp(theta2*(dose-delmix2)*(dose>delmix2));

       output;

 end;

*proc print;

*var mured dose muaddred taddred  mufulladjred;

*run;

 data predred;

 set prednl muaddred;

 where group='MX2';

 *if dose>80 then delete;

 *if tadd>450 then delete;

   proc sort data=predred; by group dose;

*proc print data=pred;

*var group d pred dose mu muadd  ed1 ed2 ed3 ed4 tadd;

*run;
```

# Vitae

Epiphanie Nyirabahizi was born in Rwanda. She was awarded a Pacem in Terris Scholarship in 1999. She attended La Roche College in 2000-2003 and graduated with a Bachelor of Science Degree. She then started her graduate studies in 2004 at the University of Pittsburgh before she was awarded a NIEHS training at Virginia Commonwealth University, Biostatistics Department from 2006 to 2010.