



VCU

Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations

Graduate School

2006

Statistical Methods and Experimental Design for Inference Regarding Dose and/or Interaction Thresholds Along a Fixed- Ratio Ray

Sharon Dziuba Yeatts
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/1296>

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.

© Sharon Dziuba Yeatts 2006

All Rights Reserved

STATISTICAL METHODS AND EXPERIMENTAL DESIGN FOR INFERENCE
REGARDING DOSE AND/OR INTERACTION THRESHOLDS ALONG A FIXED-
RATIO RAY

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

by

SHARON DZIUBA YEATTS
M.S., University of South Carolina, 2002
B.S., University of South Carolina, 2000

Director: CHRIS GENNINGS, PH.D.
PROFESSOR, DEPARTMENT OF BIostatISTICS

Virginia Commonwealth University
Richmond, Virginia
May, 2006

Acknowledgements

I would like to begin by thanking the committee members for their willingness to be involved with this project. As an advisor, Dr. Gennings is an endless resource of enthusiasm and knowledge. As a mentor, she is continually willing to provide encouragement and share her wisdom and experience. I am lucky to have had the opportunity to learn from her knowledge and example. Dr. Carter and Dr. Johnson have also been amazing teachers and friends. In the classroom, I have been caught up in their passion for learning. In the lunchroom, Dr. Johnson's jokes flow freely, helping the students release the stress of graduate school and research. Dr. Carter is always quick with a smile and words of encouragement, even during his recent retirement. They are both ready to help at a moment's notice, and I appreciate their time immensely.

I hope that I can find the right words to thank Brandon, my best friend of more than 10 years and my husband of four years, though I don't know that mere words would ever be sufficient. I could not have made it through the last four years without him. He graciously accepts my dedication to my education, which often involves working nights and weekends. This achievement belong as much to him as to me. I appreciate his never-ending love, support and sense of humor more than he could ever know. He can always make me laugh, even when it's the last thing that I want to do. I am a better person because of his love. I hope that I can return the favor someday.

I'd like to take this opportunity to thank my parents, whose unconditional love and support are always encouraging me. My father Stan is a very calming presence in my life. His sense of humor lightens even the most stressful of occasions. He is easy to talk to and quick to bring laughter to any situation. He has taught me not to take myself so seriously, which is a lesson I try to employ on a daily basis. His laidback manner makes him very approachable; he is always willing to talk, to laugh, and to help out in any way he can. I have always been, and will continue to be, Daddy's Little Girl.

My mother Eileen is a great source of strength for me. At any given moment, she knows exactly the right thing to say and how to help. Most recently, she has helped me through the job search process. I got married two weeks after completing my Master's program at the University of South Carolina; she helped me to plan a wedding from across the country. During my first semester of college, I blindsided her with a rambling letter about unfulfilled hopes and dreams. She has done all she can do to teach me that, as long as I have done my best, I will have succeeded. She is a truly amazing person and mother. I have often been told that I look like her; I hope that the similarity does not end with my appearance.

As my time at VCU ends, I am becoming more aware of what an amazing experience my education here has been. The students in the Department of Biostatistics are a close-knit group, and our daily lunch breaks provide much-needed stress relief. Who can forget Afternoons with Amy and H2-Whoa? I don't know who I'll eat lunch

with when I leave, but I have a suspicion that the group won't hold a candle to the Biostatistics students at VCU. Though not students, Yvonne and Denise should really be included in this group. They are always quick with a smile, always willing to listen, and always eager to help. The daily ins and outs of student life are much easier because of them.

Table of Contents

	Page
Acknowledgements.....	ii
List of Tables	x
List of Figures.....	xi
Chapter	
1 Introduction and Motivation	1
1.1 Background	1
1.2 The Ray Design	3
1.3 The Existence of a Threshold.....	5
1.4 The Definition of Additivity.....	6
1.5 Dose-Dependent Interactions	8
1.6 Statistical Design and Methodology.....	11
1.7 Prospectus.....	12
2 Detecting Departure from Additivity Along a Fixed-Ratio Mixture Ray with a Piecewise Model for Dose and Interaction Thresholds	15
2.1 Introduction and Motivation.....	15
2.2 Methodology	18
2.2.1 Data Description.....	18

2.2.2	The Dose and Interaction Thresholds Model	20
2.2.3	Quasi-likelihood Estimation.....	24
2.2.4	Inference	26
2.2.4.1	Test of Additivity	29
2.2.4.2	Test for a Region of Additivity	30
2.3	Chlorination Ray Example	30
2.3.1	Background	30
2.3.2	Model Selection.....	32
2.3.3	Results	36
2.3.4	Inference	38
2.3.5	Design Consideration	43
2.4	Optimal Experimental Designs	45
2.4.1	D- Optimality	51
2.4.2	Optimality with Regard to the Hypothesis Test of Additivity	54
2.4.3	Additivity Optimal Designs for the General Case.....	57
2.5	Conclusion.....	59
3	The Flexible Single Chemical Required (FSCR) Method for Detecting Departure from Additivity Along a Fixed-Ratio Mixture Ray	62
3.1	Introduction	62
3.2	Motivation	64
3.3	Methodology Review	64

3.3.1	Model Definitions.....	65
3.3.1.1	The Full Model.....	65
3.3.1.2	The Dose Threshold Additivity Model	66
3.3.1.3	The FSCR Additivity Model	67
3.3.1.4	The FSCR Interaction Threshold Model	69
3.3.2	Estimation.....	70
3.3.3	Inference	71
3.4	Example.....	74
3.4.1	Background	74
3.4.2	Model Selection.....	76
3.4.3	Inference	85
3.4.4	Design Consideration	86
4	Optimal Design for the Flexible Single Chemical Required (FSCR) Interaction Threshold Model.....	88
4.1	Introduction and Motivation.....	88
4.2	Motivating Example: The Estimation of an Interaction Threshold in a Mixture of Eighteen PHAHs	90
4.3	Construction of the Parameter Covariance Matrix for the Scaled FSCR Interaction Threshold Model.....	93
4.3.1	Differentiation of the Scaled FSCR Additivity Model for Two Agents.....	95
4.3.2	Differentiation of the Scaled FSCR Interaction Threshold Model with Combined Minimum Response Parameters	100

4.4	Optimal Experimental Design for Improving the Precision of the Interaction Threshold Estimate in the Mixture of 18 PHAHs.....	103
4.5	Penalized Optimal Design Methodology	114
4.5.1	Experimental Design Preferences for the Second Stage Fixed-Ratio Ray Mixture Experiment of 18 PHAHs.....	116
4.5.2	Penalized Optimal Design for Improving the Precision of the Interaction Threshold Estimate	124
4.6	Discussion	131
5	Determining Optimal Experimental Designs for Nonlinear Models Using Likelihood Ratio-Based Inference	133
5.1	Introduction and Motivation.....	133
5.2	Methods.....	136
5.3	Demonstration of the Quasi-Likelihood Ratio-Based Lower Confidence Bound as an Optimality Criterion	145
5.4	Comparison of the Second Stage Designs Associated with the D-, Ds-, and Quasi-Likelihood Ratio-Based Lower Confidence Bound Optimality Criteria	151
5.5	Comparison of the Second Stage Designs Associated with the D-, Ds-, and Quasi-Likelihood Ratio-Based Lower Bound Optimality Criteria under an Alternate Set of Hypothesized Parameter Values	156
5.6	Discussion	162
6	Summary and Extensions.....	166
6.1	Summary	166
6.2	Extensions	169

References.....173

Appendices.....178

 Appendix A.....179

 Appendix B.....198

 Appendix C.....219

 Appendix D.....239

Vita251

List of Tables

	Page
Table 2.1: Haloacetic Acid Byproducts of the Chlorination Process.	31
Table 2.2: Parameter Estimates Resulting from Fit of the Dose and Interaction Thresholds Model to the Single Chemical and Mixture Data Along the Ray.	41
Table 2.3: Threshold Parameter Estimates.	42
Table 2.4: D- Optimal Designs.	53
Table 2.5: Additivity Optimal Designs for the Hypothesized Dose and Interaction Thresholds Model.	56
Table 3.1: Parameter Estimates Resulting from the Fit of the Full FSCR Model.	79
Table 3.2: Parameter Estimates Resulting from the Fit of the FSCR Additivity Model. ..	81
Table 3.3: Parameter Estimates Resulting from the Fit of the FSCR Interaction Threshold Model.	83
Table 4.1: Ds- Optimal Second Stage Designs for the Estimation of the Interaction Threshold.	109
Table 4.2: Desirability Functions for the Penalized Optimal Second Stage Design.	118
Table 4.3: Optimal Second Stage Designs for the Estimation of the Interaction Threshold.	129
Table 5.1: Parameter Estimates Resulting from the Fit of the Model.	146
Table 5.2: Optimal Second Stage Designs for the Estimation of the Dose Threshold. ...	152
Table 5.3: Optimal Second Stage Designs for the Estimation of the Dose Threshold Under the Alternative Model.	159

List of Figures

	Page
Figure 2.1: Cell Viability vs. Total Concentration of the Mixture Along the Chlorination Ray	32
Figure 2.2: Predicted Cell Viability Under Additivity.....	34
Figure 2.3: Change in Variance as a Function of the Mean Response	36
Figure 2.4: Assessment of Model Fit.....	37
Figure 2.5: Hypothesized Relationship Between Cell Viability and Total Concentration.....	48
Figure 2.6: The D- Optimal Design for Supporting the Hypothesized Dose and Interaction Thresholds Model	54
Figure 2.7: The Additivity Optimal Design for the Hypothesized Dose and Interaction Thresholds Model	57
Figure 2.8: The Additivity Optimal Design in the General Antagonism Case.....	58
Figure 2.9: The Additivity Optimal Design in the General Synergism Case	59
Figure 3.1: T4 vs. Total Dose of the Mixture Along the Ray.....	75
Figure 3.2: Predicted Variance vs. Mean.....	76
Figure 3.3: Fit of FSCR Full Model (solid line) and FSCR Additivity Model (dashed line) to Mixture Data Along the Fixed-Ratio Mixture Ray.....	80
Figure 3.4: Fit of FSCR Full Model (dashed line) and FSCR Interaction Threshold Model (solid line) to Mixture Data Along the Fixed-Ratio Mixture Ray	84
Figure 3.5: Fit of FSCR Interaction Threshold Model (solid line) and FSCR Additivity Model (dashed line) to Mixture Data Along the Fixed-Ratio Mixture Ray ...	84
Figure 4.1: Predicted T4 vs. Total Dose of the Mixture	105

Figure 4.2: The Ds- Optimal Second Stage Design for the Estimation of the Interaction Threshold	113
Figure 4.3: Variance of the Hypothesized Interaction Threshold and Desirability vs. λ for the Eight Point Penalized Optimal Second Stage Design	126
Figure 4.4: Variance of the Hypothesized Interaction Threshold and Desirability vs. λ for the Six Point Penalized Optimal Second Stage Design.....	127
Figure 4.5: Variance of the Hypothesized Interaction Threshold and Desirability vs. λ for the Four Point Penalized Optimal Second Stage Design	128
Figure 4.6: The Penalized Optimal Second Stage Design for the Estimation of the Interaction Threshold.....	130
Figure 5.1: Motor Activity vs. Administered Dose of Deltamethrin.....	145
Figure 5.2: The Fit of the Exponential Threshold Model to the Deltamethrin Data	147
Figure 5.3: Quasi-likelihood Associated with a Fixed Value of the Threshold.....	148
Figure 5.4: Quasi-likelihood Associated with a Fixed Value of the Threshold for Fixed Minimum Response	150
Figure 5.5: The Second Stage Designs for the Estimation of the Dose Threshold.....	153
Figure 5.6: The Distribution of Lower Confidence Bounds on the Dose Threshold by Optimality Criterion.....	155
Figure 5.7: The Fit of the Alternative Model to the Deltamethrin Data	158
Figure 5.8: The Second Stage Designs for the Estimation of the Dose Threshold Under the Alternative Model	160
Figure 5.9: The Distribution of Lower Confidence Bounds on the Dose Threshold by Optimality Criterion.....	161

Abstract

Statistical Methods and Experimental Design for Inference Regarding Dose and/or Interaction Thresholds Along a Fixed-Ratio Ray

By Sharon Dziuba Yeatts

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

Virginia Commonwealth University, 2006

Major Director: Dr. Chris Gennings
Professor, Department of Biostatistics

An alternative to the full factorial design, the ray design is appropriate for investigating a mixture of c chemicals, which are present according to a fixed mixing ratio, called the mixture ray. Using single chemical and mixture ray data, we can investigate interaction among the chemicals in a particular mixture. Statistical models have been used to describe the dose-response relationship of the single agents and the mixture; additivity is tested through the significance of model parameters associated with the coincidence of the additivity and mixture models.

It is often assumed that a chemical or mixture must be administered above an unknown dose threshold in order to produce an effect different from background. Risk

assessors often assume that interactions are a high-dose phenomenon, indicating that doses below the unknown interaction threshold are associated with additivity. We developed methodology that allows the user to simultaneously estimate the dose threshold and the interaction threshold. This methodology allows us to test for interaction and, secondarily, to test for a region of additivity. The methodology and optimal design characteristics were illustrated using a mixture of nine haloacetic acids.

The application of statistical optimality criteria to the development of experimental designs is vital to the successful study of complex mixtures. Since the optimal design depends on the model of interest and the planned method of analysis, developments in statistical methodology should necessarily correspond to consideration of the experimental design characteristics necessary to implement them. The Flexible Single Chemical Required methodology is based on an implicit statement of additivity. We developed a method for constructing the parameter covariance matrix, which forms the basis of many alphabetic optimality criteria, for the implicit FSCR models. The method was demonstrated for a fixed-ratio mixture of 18 chemicals; the original mixture experiment comprises the first stage data, and the optimal second stage design was presented. Wald-type procedures for hypothesis testing in nonlinear models are based on a linear approximation. As a result, likelihood ratio-based procedures may be preferred over Wald-type procedures. We developed a procedure for using the likelihood ratio-based lower confidence bound as an optimality criterion, which can be used to find the optimal second stage design for improving the inference on a particular model parameter. The method was

demonstrated for a single agent, as a means of improving the inference on the dose threshold.

Chapter 1

Introduction and Motivation

1.1 Background

Humans are exposed to mixtures of chemicals on a daily basis, and the number of chemicals involved in the mixture may be quite large (Monosson, 2005; Teuschler et al., 2002). Simmons (1995), however, observed that, “in contrast to the environmental reality that human exposure is to mixtures, the vast majority of toxicology studies examine the cancer and noncancer health effects of single chemicals.” This is likely because the study of a mixture is often difficult when the number of chemicals involved is moderate to large. In addition, the number of potential mixtures for study is seemingly endless. As a result, dose-response and toxicity information often exist for the single chemicals in the mixture, but not for the mixture itself (Teuschler and Hertzberg, 1995).

Full factorial experiments are traditionally used to study chemical mixtures, because they allow the experimenter to investigate the relationship among the chemicals throughout the combination space. However, the factorial experiment that is generally used to investigate interaction may not be experimentally feasible when the number of chemicals in the mixture is large. As the number of chemicals involved in the mixture increases, the experiment quickly becomes impractical due to the amount of time and

resources required. Therefore, Teuschler et al. (2002) comment on the importance of considering “efficient experimental designs ... [that] can be employed to provide information on toxicity and interaction effects without implementing a full factorial design.” In this regard, the ray design is a useful alternative to the factorial design and an important development in the design and analysis of chemical mixtures.

As defined by the United States Environmental Protection Agency, a complex mixture is a mixture containing “so many components that any estimation of its toxicity based on its components’ toxicities contains too much uncertainty and error to be useful” (U.S. EPA, 2000). Often, such a mixture is the result of a process; the potential to improve the process may come from knowledge regarding the behavior of the individual components and the mixture. For example, the disinfection of drinking water results in the addition of disinfection byproducts (DBPs) to the treated water. The result is a complex mixture for which the relative proportions of the constituents are only partially known. As will be discussed in Chapter 2, the EPA has selected a class of DBPs to study in a component-based mixture, where the proportion of the components in the mixture is specified by a particular disinfection process.

According to Teuschler et al. (2002), methods for the risk assessment of chemical mixtures fall into one of two categories. Whole mixtures approaches are those in which the mixture is evaluated as if it were a single entity rather than a mixture of agents. Component-based approaches are those in which the mixture is evaluated as a combination of the components, often under the default assumption of additivity. The focus of this dissertation is on the development of methods that use information from

both the individual components and an experimentally constructed mixture to test for interaction with statistical rigor. For mixtures with a large number of components, we are particularly interested in the use of the fixed-ratio ray design, since these rays may be determined such that the mixing ratio is environmentally relevant in terms of human exposure.

To begin this dissertation, we will briefly review some concepts that are important to the development of statistical methodology for investigating the toxicology of complex chemical mixtures. In Section 1.2, we describe the ray design and its implications for the analysis of interactions among the chemicals in a complex mixture. The concept of a dose threshold is introduced in Section 1.3. In Section 1.4, we define additivity and interaction; support for the concept of a dose-dependent interaction is presented in Section 1.5. In Section 1.6, we briefly discuss current statistical methodologies for detecting interaction among the chemicals in a complex mixture. Finally, in Section 1.7, we provide an overview of the relationship between these topics and the prospectus for the dissertation.

1.2 The Ray Design

As described by Mantel (1958), Brunden and Vidmar (1989) and others, ray designs can be used to explore a response surface. Following the work of Gennings et al. (2002) and Casey et al. (2004), however, we can also use the ray design to focus interest on a particular ratio of chemicals. The ray design is appropriate for investigating the relationship among c chemicals when the chemicals are present in the mixture according

to a fixed mixing ratio, called the fixed-ratio mixture ray. Using single chemical data and data from the fixed-ratio mixture ray, we can investigate interaction among the chemicals in a particular mixture, without requiring that we consider the entire combination space. According to Teuschler et al. (2002), because the study of all possible mixture combinations is not possible, “the mixture components that are tested and their relative proportions in the mixture should reflect those seen in environmental samples.” As a result, the use of the ray design is an important advancement in the evaluation of mixture toxicology.

To set notation, let the first c rays represent the single chemical data, and let the last ray represent the mixture data along the fixed-ratio ray, for a total of $c + 1$ rays. The fixed-ratio ray for the mixture is defined by the mixing ratio $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_c]$, where $0 \leq a_i \leq 1$ for all $i = 1, \dots, c$ and $\sum_{i=1}^c a_i = 1$. The i^{th} single chemical is represented by a fixed-ratio ray where the mixing ratio is defined such that the i^{th} element of \mathbf{a} is 1, and the remaining elements are zero. In either case, a_i is the proportion of the i^{th} chemical in the mixture, and $x_i = a_i t$ is the dose of the i^{th} chemical associated with total dose t .

Because the amount of the individual chemicals in the mixture can be determined, the use of the ray design allows the investigator to look at questions regarding both the interaction among the chemicals and the main effects associated with the chemicals. Others have shown that, when single chemical data are available for all of the chemicals involved in the mixture, the ray design is a useful tool for efficiently investigating the

interaction among the chemicals along a given ray (Gennings et al., 2002; Casey et al., 2004; Hamm, 2004).

1.3 The Existence of a Threshold

As a result of complicated biotransformation, elimination and/or repair processes, many biological systems exhibit some level of tolerance to a toxic insult (Cox, 1987). In such instances, in order for the chemical to cause an observable change in response, the body must be exposed to quantities sufficient to overwhelm these processes. It is often reasonable, therefore, to assume that a chemical must be administered above an unknown dose threshold if it is to produce an effect that is different from background. To account for this, Cox (1987) and Ulm (1991) described single chemical dose response relationships using models in which the dose threshold is a parameter to be estimated.

Schwartz, Gennings and Chinchilli (1995) extended the concept and methodology of the dose threshold to the estimation of the response surface associated with a mixture of chemicals. Gennings et al. (1997) applied the dose threshold to the estimation of an additivity surface using single chemical data only. More recently, Casey et al. (2004) developed methodology for including the dose threshold in the analysis of a chemical mixture along a fixed-ratio ray. As these recent advances indicate, the continued development of statistical methods and experimental design for analyzing the dose-response relationship of a mixture in the low-dose region is of particular importance to risk assessors. In particular, Teuschler et al. (2002) encourage the development of

methods for detecting thresholds, since the dose-response relationship evident at high doses may not be applicable in the low-dose region.

1.4 The Definition of Additivity

Throughout this dissertation, the definition of additivity used is that given by the interaction index (Berenbaum, 1985). If the relationship among the c chemicals in a mixture is additive, then

$$\sum_{i=1}^c \frac{x_i}{E_i} = 1, \quad (1.1)$$

where x_i represents the concentration/dose of the i^{th} chemical in the mixture combination that yields response y , and E_i represents the concentration/dose of the i^{th} chemical that individually yields the same response. The left-hand side of equation (1.1) is referred to as the interaction index.

The basic concept of an interaction is simple. An interaction implies that the presence of one agent alters the slope of the dose-response relationship associated with another agent (Gennings et al., 2005; Teuschler et al, 2002). If the slope of the dose-response relationship is not affected by the presence of another agent, then the relationship is said to be additive. For a given mixture, if the interaction index is equal to one, then the relationship among the chemicals is said to be additive. If the interaction index is different from one, then an interaction exists among the chemicals in the mixture.

When the chemicals in a mixture interact, it is of interest to characterize the interaction as synergistic or antagonistic. Antagonism indicates that the effect of the

agents in combination is less than what is predicted under additivity; that is, the effect of a given agent is diminished by the presence of the other agents in the mixture. Conversely, synergism indicates that the effect of the agents in combination is greater than what is predicted under additivity; that is, the effect of a given agent is enhanced by the presence of the other agents in the mixture.

The magnitude of the interaction index, relative to the additivity value of 1, indicates whether the interaction is synergistic or antagonistic. If the interaction index is larger than 1 for a given mixture, then the interaction among the chemicals in the mixture is described as an antagonism. Conversely, if the interaction index is smaller than 1 for a given mixture, the interaction among the chemicals in the mixture is described as a synergism. While the interaction index has been used to determine the presence of an interaction in the toxicological literature, the statistical significance of the interaction index is rarely addressed.

Statistical models for assessing interaction have an advantage over Berenbaum's interaction index. It can be shown that Berenbaum's interaction index is algebraically equivalent to the statistical additivity model, which involves only linear terms. As a result, the statistical significance of cross-product terms is directly related to the presence of an interaction. In statistical practice, the null hypothesis of additivity is tested through the significance of these terms, rather than through computation of the interaction index.

In the case of the ray design, we can characterize an interaction as an antagonism or a synergism graphically. For decreasing dose-response curves, if the mixture model falls above the additivity curve, then the relationship among the chemicals is antagonistic. If

the mixture model falls below the additivity curve, then the relationship among the chemicals is synergistic. For increasing dose-response curves, if the mixture model falls below the additivity curve, then the relationship among the chemicals is antagonistic. If the mixture model falls above the additivity curve, the relationship among the chemicals is synergistic. The significance of the interaction corresponds to the significance of the model parameter(s) associated with the interaction.

1.5 Dose-Dependent Interactions

The interaction among the chemicals in a mixture may be dose-dependent (Könemann and Pieters, 1996). In fact, according to Gennings et al. (2004), “mechanisms of ... chemical interaction are rarely uniform across the entire dose-response range and instead tend to be highly dose-dependent.” For example, Gennings et al. (2002) found that interaction in the mixture of arsenic, cadmium, chromium and lead was synergistic for some dose combinations and antagonistic for others, while no interaction was detected for still other dose combinations. Gessner and Cabana (1970) studied a mixture of ethanol and chloral hydrate; their analysis suggests that a synergism exists “when low to moderate levels of ethanol are combined with moderate to high levels of chloral hydrate ... [whereas] for all other combinations, there does not appear to be a departure from additivity” (Carter, 1995). There are a number of other examples in recent peer-reviewed literature.

To account for the dose-dependent nature of interactions, El-Masri, Tessari and Yang (1996) developed the concept of an interaction threshold using Physiologically Based

Pharmacokinetic (PBPK) models. Hamm, Carter and Gennings (2005) incorporated the concept of an interaction threshold in the modeling of a dose response surface using an interaction boundary. Hamm (2004) also developed a statistical model for describing the dose response relationship of a chemical mixture along a fixed-ratio ray, in which the interaction threshold is a parameter to be estimated. The interaction threshold model allows the interaction among the chemicals to vary, depending on the location of the total dose of the mixture with respect to the unknown interaction threshold. In particular, Hamm's interaction threshold model describes additivity for total doses smaller than the interaction threshold, while the model allows the mixture to exhibit departure from additivity for larger total doses.

This assumption of low-dose additivity is acknowledged by the U.S. EPA (2000) and supported by Carpy, Kobel and Doe (2000). A literature review conducted by Carpy, Koble and Doe (2000) demonstrated that interactions are rare at low exposure levels, though the reason for this is unclear. It may be that additivity truly prevails in the low dose region. On the other hand, it may be that the study design was insufficient to detect an interaction, or that the interaction effect is of such small magnitude that it is difficult to detect or of no practical significance. However, Könemann and Pieters (1996) support the assumption of low dose additivity, stating that interactions "may not be very relevant in the low-dose region." When a mixture is under investigation for regulatory purposes, the analysis is often based on the assumption of additivity among the chemicals in the mixture. However, Teuschler and Hertzberg (1995) acknowledge that "these additivity procedures include a general assumption that interaction effects at low dose levels either

do not occur at all or are small enough to be insignificant to the risk estimate. Although risk assessors often assume that interactions do not occur at low doses, any known antagonism or synergism should be reflected in the risk assessment whenever possible.” Therefore, due to its potential implications in risk assessment, the interaction threshold model is an important development in statistical methodology.

In 2002, the Society of Toxicology commissioned an expert panel to suggest directions for future research in the risk assessment of chemical mixtures (Teuschler et al., 2002). A number of the alternative hypotheses presented by the panel dealt with interaction thresholds. Specifically, the panel stressed the importance of demonstrating additivity among the chemicals in a mixture at low exposure levels. The panel also emphasized the importance of being able to determine that the interaction threshold occurs at a higher dose than the individual chemical dose thresholds.

The ability to detect dose-dependent interactions has important implications for toxicology; evidence of an interaction should be incorporated into the risk assessment of a mixture. Among the chemicals present in a mixture, it may be that a synergism exists at high doses, such that the toxicity of the combination is greater than that predicted under additivity. In this case, considerable resources may be devoted to restricting human exposure to low dose levels for safety reasons. Conversely, if an antagonism exists at high doses, the toxicity of the combination at high doses is less than that predicted under additivity. As a result, the public health community may not express much interest in restricting exposure. Because public health importance depends, in part, on the nature and the location of the interaction, Teuschler and Hertzberg (1995) state

that “the study of toxicity thresholds and interaction thresholds for combinations of chemicals” is of interest in the EPA’s mixture research program.

The study of therapeutic thresholds and interaction thresholds for combinations of drugs and/or medications is likely to be of interest to the FDA as well. Among a mixture of medications, it may be that a synergism exists at high doses, such that the combination produces a better response than expected under additivity of the drugs. Knowledge of such a synergism has the potential to impact the course of treatment, since it might allow a physician to lower the prescribed dose of the combination while still achieving the desired response. Conversely, it may be that an antagonism exists at high doses, such that the medications do not work as well in combination as predicted under additivity.

1.6 Statistical Design and Methodology

The implementation of statistical models is an important tool in risk assessment. Threshold models allow the user to detect the presence of a dose threshold for a single agent or a complex mixture. The interaction threshold model allows the researcher to address the assumption of low-dose additivity. The statistical significance of the interaction threshold parameter provides evidence that there is a low-dose region of the mixture for which the relationship among the chemicals is additive. Because statistical models account for the variability in the data, the location of the threshold, if it exists, can be estimated with some degree of confidence. These models are significant advancements in the study of mixture toxicology; however, further developments in

statistical modeling are crucial for advancing the current state of science and addressing the continually changing needs of the field of toxicology.

While statistical interaction threshold models are capable of detecting dose-dependent interaction, the statement of such models is generally complicated. Since full factorial designs are often unfeasible for complex mixtures, care should be taken to ensure that mixture studies are designed to conserve resources while still providing sufficient information to meet research objectives. Simmons (1995) emphasized the need for improvement in the experimental design of mixtures studies, stating that “adoption of the use of statistically appropriate, innovative experimental designs that are efficient in their use of animals would increase the ability of toxicologists to determine the presence of nonadditive toxicity resulting from exposure to chemical mixtures.” In this regard, the ray design is an important advancement. In addition, however, the application of statistical optimality criteria to the development of experimental designs is crucial to the successful study of complex mixtures. Specifically, the implementation of a statistically optimal design allows the investigator to focus experimental efforts on the areas of most importance for the study under consideration. Since the optimal design depends on the model of interest, developments in statistical methodology should necessarily correspond to consideration of the experimental design characteristics necessary to implement them.

1.7 Prospectus

This dissertation is written in a distinct style. Chapters 2, 4 and 5 are preliminary versions of manuscripts in preparation for submission to statistical and toxicological

journals. These chapters are meant to stand alone; the literature review and notation development for each dissertation topic are completely contained within each chapter. As a result, there may be some overlap in the coming chapters.

As previously described, the expert panel commissioned by the Society of Toxicology indicated the need for statistical methodology capable of demonstrating low-dose additivity. The panel also stated the importance of methodology capable of determining that the interaction threshold of a mixture is higher than the dose threshold. In Chapter 2, we develop methodology that addresses both of these areas of research. The proposed dose and interaction thresholds model allows the simultaneous estimation of both the dose threshold and the interaction threshold along a fixed-ratio mixture ray. As a result, we can test the primary null hypothesis of additivity. If this hypothesis is rejected, then we can test for an active region of low-dose additivity by comparing the interaction threshold to the dose threshold. We also present the characteristics of the optimal design associated with this model. The method is demonstrated for a fixed-ratio mixture of nine single agents.

We have previously discussed the need for developments in design methodology to keep up with developments in statistical methodology. In Chapters 3 and 4, we address this issue as it pertains to the Flexible Single Chemical Required (FSCR) methodology developed by Gennings et al. (2004, 2006). We develop a method for constructing the associated parameter covariance matrix, which forms the basis of many alphabetic optimality criteria. The method is demonstrated for a fixed-ratio mixture of 18 single agents, and the optimal design is presented. In addition, we apply the penalized

optimality criterion developed by Parker and Gennings (2006) to improve the practicality of the optimal design.

Likelihood ratio-based procedures may be preferred when making inference regarding nonlinear dose-response relationships. In Chapter 5, we propose a procedure for using the likelihood ratio-based confidence interval as an optimality criterion, which can be used to find the optimal second stage design for improving the inference on a particular model parameter. The method is demonstrated for a single agent, as a means of improving the inference on the dose threshold. The resulting quasi-likelihood ratio-based design is presented and compared to the associated D- and Ds- optimal designs, and general design implications are discussed.

Finally, in Chapter 6, we summarize our contribution to the development of statistical methodology for the analysis of complex chemical mixtures and present areas for future research in this area.

Chapter 2

Detecting Departure from Additivity Along a Fixed-Ratio Mixture Ray with a Piecewise Model for Dose and Interaction Thresholds

2.1 Introduction and Motivation

As a result of complicated biotransformations, eliminations and/or prevention and repair processes, many biological systems exhibit some level of tolerance to a toxic insult (Cox, 1987). In such instances, in order for the chemical to cause an observable change in response, the body must be exposed to quantities sufficient to overwhelm these processes. It is often reasonable, therefore, to assume that a chemical must be administered above an unknown dose threshold if it is to produce an effect that is different from background. To account for this, Cox (1987) and Ulm (1991) described single chemical dose response relationships using models in which the dose threshold is a parameter to be estimated. Schwartz, Gennings and Chinchilli (1995) extended the concept and methodology of the dose threshold to the estimation of the response surface associated with a mixture of chemicals. Gennings et al. (1997) applied the dose threshold to the estimation of an additivity surface using single chemical data only. More recently, Casey et al. (2004) developed methodology for including the dose threshold in the analysis of a chemical mixture along a fixed-ratio ray.

It is also reasonable to consider the situation in which the relationship among the chemicals in a mixture is dose-dependent (Gennings et al., 2002; Könemann and Pieters, 1996). For example, Gennings et al. (2002) found that interaction in the mixture of arsenic, cadmium, chromium and lead was synergistic for some dose combinations and antagonistic for others, while no interaction was detected for still other dose combinations. To account for the dose-dependent nature of interactions, El-Masri, Tessari and Yang (1996) developed the concept of an interaction threshold using PBPK models. Hamm, Carter and Gennings (2005) incorporated the concept of an interaction threshold in the modeling of a dose response surface using an interaction boundary. Hamm (2004) also developed a statistical model for describing the dose response relationship of a chemical mixture along a fixed-ratio ray, in which the interaction threshold is a parameter to be estimated. The interaction threshold model allows the interaction among the chemicals to vary, depending on the location of the total dose of the mixture with respect to the unknown interaction threshold. In particular, Hamm's interaction threshold model describes additivity for total doses smaller than the interaction threshold, while the model allows the mixture to exhibit departure from additivity for larger total doses. The assumption of low-dose additivity is acknowledged by the U.S. EPA (2000) and supported by Carpy, Kobel and Doe (2000), whose literature review demonstrated that interactions are rare at low exposure levels. It may be that additivity truly prevails in the low dose reason. On the other hand, it may be that the study design was insufficient to detect an interaction, or that the interaction effect is of such small magnitude that it is difficult to detect or of no practical significance.

The interaction threshold model is an important development in statistical methodology. In 2002, the Society of Toxicology commissioned an expert panel to suggest directions for future research in the risk assessment of chemical mixtures (Teuschler et al., 2002). A number of the alternative hypotheses presented by the panel dealt with interaction thresholds. Specifically, the panel stressed the importance of demonstrating additivity among the chemicals in a mixture at low exposure levels. The panel also emphasized the importance of being able to determine that the interaction threshold is higher than the dose threshold. This suggests the need for statistical methodology capable of simultaneously estimating both a dose threshold and an interaction threshold.

To our knowledge, however, the concepts of the dose threshold and the interaction threshold have not been combined in a single model to describe the dose-response relationship between a response and a chemical mixture along a fixed-ratio ray. The situation in which both thresholds exist is considered as follows. For total doses smaller than the dose threshold, the response of the mixture is not different from background. For total doses between the dose threshold and the interaction threshold, the model describes an additive relationship among the chemicals in the mixture. For total doses larger than the interaction threshold, the model allows for interaction among the chemicals in the mixture. The development of such a model for describing the dose response relationship of a chemical mixture along a fixed-ratio ray is the primary objective of this chapter.

In section 2.2, we propose a model that includes both a dose threshold and an interaction threshold, and we describe the estimation process and related hypothesis tests. In section 2.3, we present an example in which the dose and interaction thresholds model is used to analyze the cell viability of a mixture of nine haloacetic acids (HAAs) along a fixed-ratio ray. In section 2.4, we discuss the design implications of this model and present optimal ray designs for the support of the dose and interaction thresholds model.

2.2 Methodology

2.2.1 Data Description

Consider the situation where we have single chemical data available for each of the c chemicals in the mixture, as well as mixture data taken along the fixed-ratio ray of interest, for a total of $c+1$ rays. For mixture doses below the interaction threshold, both the single chemical data and the mixture data are used to estimate the slopes associated with each of the single chemicals under the assumption of additivity. For mixture doses exceeding the interaction threshold, we use the mixture data alone to estimate the model parameters associated with departure from additivity. Therefore, the dose and interaction thresholds model is supported by both the single chemical data and the mixture data along the fixed-ratio ray.

To set notation, let the first c rays represent the single chemical data, and the last ray represents the mixture data along the fixed-ratio ray. The fixed-ratio ray for the mixture is defined by the mixing ratio, $\mathbf{a} = [a_1 \ a_2 \ \cdots \ a_c]$. The i^{th} single chemical is represented by a fixed-ratio ray where the mixing ratio is defined such that the i^{th} element

of \mathbf{a} is 1, and the remaining elements are 0. In either case, $x_i = a_i t$ is the dose of the i^{th} chemical associated with total dose t .

Let \mathbf{a}_h represent the mixing ratio for the h^{th} ray, for $h=1, \dots, c+1$. Let x_{hji} be the dose of the i^{th} chemical at the j^{th} dose on the h^{th} ray. Then $t_{hj} = \sum_{i=1}^c x_{hji}$ is the j^{th} total dose along the h^{th} ray, and $x_{hji} = a_{hi} t_{hj}$. Let y_{hjk} represent the k^{th} response from the j^{th} dose of the h^{th} ray, for $j=1, \dots, w_h$, and $k=1, \dots, n_{hj}$, where w_h represents the number of treatment levels for the h^{th} ray and n_{hj} represents the number of observations taken at the j^{th} treatment level on the h^{th} ray. Let $\boldsymbol{\beta}$ represent the vector of regression parameters, such that $E\{y_{hjk}\} = \mu_{hj}(\boldsymbol{\beta})$ and $Var\{y_{hjk}\} = \tau V(\mu_{hj}(\boldsymbol{\beta}))$. The total sample size for both the

single chemical and the mixture ray data is
$$N = \sum_{h=1}^{c+1} \sum_{j=1}^{w_h} n_{hj}.$$

Because the amount of the individual chemicals in the mixture can be determined, the use of the ray design allows the investigator to look at questions regarding both the interaction among the chemicals involved in the mixture and the main effects associated with the chemicals. When single chemical data are available for all of the chemicals involved in the mixture, the ray design is a useful tool for efficiently investigating the interaction among the chemicals along a given ray. Following the Single Chemical Required (Casey et al., 2004) methodology, the single chemical data can be used to predict additivity along the ray. To test for interaction, the coincidence of the additivity model and the mixture model is tested. When single chemical data are not available, the

Single Chemical Not Required (Casey et al., 2004) methodology, which associates the coefficient of the j^{th} degree term to j -factor interactions, can be applied.

2.2.2 The Dose and Interaction Thresholds Model

The goal of this section is to develop the piecewise model containing both a dose threshold and an interaction threshold. We will first describe a dose threshold additivity model for any mixture of chemicals; then we will demonstrate how the model is modified when we consider only a mixture along a fixed-ratio ray. Based on this two segment threshold additivity model for the ray, we develop a three-segment model that includes both a dose threshold and an interaction threshold. The third segment in the model allows the relationship among the chemicals to depart from additivity. Lastly, we constrain the model so that the response is continuous throughout the dose range.

Consider the dose threshold model in which the relationship among the c chemicals is that of additivity. If the dose-response curve is decreasing, we want this two segment model to specify that there is no change in the mean response prior to the dose threshold, and that the mean response decreases as a possibly nonlinear function of the predictor variables under the assumption of additivity. Let γ represent the vector of parameters necessary to describe the nonlinear dose-response relationship; for example, γ might consist of a minimum response and a range parameter. The basic dose threshold model, then, is of the following form.

$$g(\mu_{add}, \gamma) = \begin{cases} \beta_{01} & , \sum_{i=1}^c \beta_i x_i > \delta_{add} \\ \beta_{02} + \sum_{i=1}^c \beta_i x_i & , \sum_{i=1}^c \beta_i x_i \leq \delta_{add} \end{cases} \quad (2.1)$$

In the above dose threshold model, β_{01} and β_{02} are parameters related to the intercept of each segment, β_i represents the slope of the i^{th} chemical in the mixture, and δ_{add} is a parameter related to the unknown threshold for the mixture under the assumption of additivity.

Along the ray, however, we know that $x_i = a_i t$. Therefore, the dose threshold model given in equation (2.1) can be rewritten to describe the dose-response relationship expected under additivity along the ray, as follows.

$$\begin{aligned} g(\mu_{add}, \gamma) &= \begin{cases} \beta_{01} & , \sum_{i=1}^c \beta_i a_i t > \delta_{add} \\ \beta_{02} + \sum_{i=1}^c \beta_i a_i t & , \sum_{i=1}^c \beta_i a_i t \leq \delta_{add} \end{cases} \\ &= \begin{cases} \beta_{01} & , \theta_{add} t > \delta_{add} \\ \beta_{02} + \theta_{add} t & , \theta_{add} t \leq \delta_{add} \end{cases} \Rightarrow \\ g(\mu_{add}, \gamma) &= \begin{cases} \beta_{01} & , t < \delta_{add_t} \\ \beta_{02} + \theta_{add} t & , t \geq \delta_{add_t} \end{cases} \end{aligned} \quad (2.2)$$

Equation (2.2) defines the dose threshold additivity model along the ray. In this parameterization of the model, $\theta_{add} = \sum_{i=1}^c \beta_i a_i$ is the slope of total dose along the ray and

$\delta_{add_t} = \delta_{add} / \theta_{add}$ is the dose threshold in terms of total dose. The dose threshold for

the i^{th} single chemical in the mixture is represented by δ_{add} / β_i (Casey et al., 2004).

To incorporate an interaction threshold into our model, the model for the dose-response relationship along the ray consists of three segments. To address the assumption of low dose additivity, the additivity region is described according to the threshold additivity model for a fixed-ratio ray, and the interaction region includes parameters that allow for departure from additivity. The model specifies that there is no change in the mean response prior to the dose threshold and describes an additive relationship for mixture doses below the interaction threshold. For mixture doses beyond the interaction threshold, the model allows for departure from additivity.

Consider, then, the following model for the dose-response relationship of a mixture along a ray.

$$g(\mu_{mix}, \gamma) = \begin{cases} \beta_{01} & , t < \delta_{add_t} \\ \beta_{02} + \theta_{add}t & , \delta_{add_t} \leq t < \Delta \\ \beta_{03} + \theta_{add}t + \theta_{mix}t & , t \geq \Delta \end{cases} \quad (2.3)$$

As in equation (2.2), δ_{add_t} represents the dose threshold in terms of total dose, and θ_{add} represents the slope of total dose along the ray under the assumption of additivity. In the three segment model given in equation (2.3), Δ represents the interaction threshold, and θ_{mix} is the change in the slope for total doses administered above the interaction threshold.

To make the model given in equation (2.3) continuous at the dose threshold δ_{add_t} , we set $\beta_{02} + \theta_{add}\delta_{add_t} = \beta_{01}$. This continuity constraint requires that $\beta_{02} = \beta_{01} - \theta_{add}\delta_{add_t}$, and the model then becomes

$$g(\mu_{mix}, \gamma) = \begin{cases} \beta_{01} & , t < \delta_{add_t} \\ \beta_{01} + \theta_{add}t - \theta_{add}\delta_{add_t} & , \delta_{add_t} \leq t < \Delta \\ \beta_{03} + \theta_{add}t + \theta_{mix}t & , t \geq \Delta \end{cases} \quad (2.4)$$

The model given in equation (2.4) is continuous at the dose threshold. To make it continuous at the interaction threshold as well, we set

$\beta_{03} + \theta_{add} \Delta + \theta_{mix} \Delta = \beta_{01} + \theta_{add} \Delta - \theta_{add} \delta_{add_t}$. This continuity constraint requires that

$\beta_{03} = \beta_{01} - \theta_{add} \delta_{add_t} - \theta_{mix} \Delta$, and the model then becomes

$$g(\mu_{mix}, \gamma) = \begin{cases} \beta_{01} & , t < \delta_{add_t} \\ \beta_{01} + \theta_{add} t - \theta_{add} \delta_{add_t} & , \delta_{add_t} \leq t < \Delta \\ \beta_{01} - \theta_{add} \delta_{add_t} - \theta_{mix} \Delta + \theta_{add} t + \theta_{mix} t & , t \geq \Delta \end{cases} \Rightarrow$$

$$g(\mu_{mix}, \gamma) = \begin{cases} \beta_{01} & , t < \delta_{add_t} \\ \beta_{01} + \theta_{add} (t - \delta_{add_t}) & , \delta_{add_t} \leq t < \Delta \\ \beta_{01} + \theta_{add} (t - \delta_{add_t}) + \theta_{mix} (t - \Delta) & , t \geq \Delta \end{cases} \quad (2.5)$$

The model given in equation (2.5) is continuous at both the dose threshold and the interaction threshold.

Recall that both the single chemical data and the mixture data are used to fit the dose and interaction thresholds model, which is of the form

$$g(\mu_h, \gamma) = \begin{cases} \beta_0 & , t_{(h)} < \delta_{add_t(h)} \\ \beta_0 + \left(\sum_{i=1}^c \beta_i a_{i(h)} \right) (t_{(h)} - \delta_{add_t(h)}) & , \begin{cases} t_{(h)} \geq \delta_{add_t(h)} & , h \leq c \\ \delta_{add_t(h)} \leq t_{(h)} < \Delta & , h = c + 1 \end{cases} \\ \beta_0 + \left(\sum_{i=1}^c \beta_i a_{i(h)} \right) (t_{(h)} - \delta_{add_t(h)}) + \theta_{mix} (t_{(h)} - \Delta) & , t_{(h)} \geq \Delta \text{ and } h = c + 1 \end{cases} \quad (2.6)$$

Both the single chemical data and the mixture data are used to support the region of additivity, while only the mixture data are used to support the interaction region. Higher order terms can be added to the third segment of the model as necessary.

The dose and interaction thresholds model relates the total dose of the mixture along a fixed-ratio ray to the mean through a link function, $g(\mu_h, \gamma)$. For decreasing dose-response relationships with an approximate sigmoid shape, consider $\gamma = [\alpha \ \gamma]$, where α represents the minimum response and γ represents the range of the response. Then the mean response, μ , can be represented by any nonlinear function of the form $\alpha + \gamma F(g(\mu))$, where $F(\bullet)$ is any decreasing, sigmoidal function that takes on values between zero and one. For increasing dose-response relationships, $F(\bullet)$ is any increasing, sigmoidal function that takes on values between zero and one.

2.2.3 Quasi-likelihood Estimation

In the study of dose-response relationships, it is common to see that the variance of the response is not constant. Instead, the variance of the response often changes as a function of the mean response. Quasi-likelihood methods can be used to account for this in the data analysis.

Let $\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_c \ \delta_{add} \ \theta_{mix} \ \Delta]$ be the $p \times 1$ parameter vector represented in the dose and interaction thresholds model given in equation (2.6). Under the usual Gaussian framework, it is assumed that each of the observations comes from some normal distribution, that each observation is independent of the other observations, and that $Y_{hjk} \sim N(\mu_{hj}(\boldsymbol{\beta}), \sigma^2)$. Here, $\mu_{hj}(\boldsymbol{\beta})$ is the expected value of the response at the j^{th} dose on the h^{th} ray, which is related to the total dose administered through the $p \times 1$

parameter vector $\boldsymbol{\beta}$. The quasi-likelihood framework, however, allows some of these assumptions to be relaxed.

Under the quasi-likelihood framework, only the form of the first and second moments of the response, i.e. $E\{Y_{hjk}\} = \mu_{hj}(\boldsymbol{\beta})$ and $Var\{Y_{hjk}\} = \tau V(\mu_{hj}(\boldsymbol{\beta}))$, are assumed, without having to assume the distributional form of the response (McCullagh, 1983). For the $N \times 1$ response vector \mathbf{Y} , where the s^{th} element of \mathbf{Y} is the response of the s^{th} independent observation, we arrive at the assumptions $E\{\mathbf{Y}\} = \boldsymbol{\mu}(\boldsymbol{\beta})$ and $Var\{\mathbf{Y}\} = \tau \mathbf{V}$. Here \mathbf{V} is an $N \times N$ diagonal matrix where the $(s,s)^{\text{th}}$ element is $V(\boldsymbol{\mu}(\boldsymbol{\beta}))$ resulting from the s^{th} observation.

The quasi-likelihood method can be used to estimate model parameters following the estimation procedures described in McCullagh and Nelder (1989). The quasi-likelihood for an individual observation is the integral

$$Q(\boldsymbol{\mu}(\boldsymbol{\beta}); y) = \int_y^{\mu} \frac{y-t}{\tau V(t)} dt,$$

which behaves in the same way as a log-likelihood. The quasi-likelihood for the sample is the sum of the individual quasi-likelihoods

$$Q(\boldsymbol{\mu}(\boldsymbol{\beta}); \mathbf{y}) = \sum_{s=1}^N Q_s(\boldsymbol{\mu}_s(\boldsymbol{\beta}); y_s). \quad (2.7)$$

We can use numeric methods to find the maximum quasi-likelihood estimate $\hat{\boldsymbol{\beta}}$, the value of $\boldsymbol{\beta}$ that maximizes the quasi-likelihood given in equation (2.7).

Under the regularity conditions given in Seber and Wild (2003), the quasi-likelihood estimate $\hat{\boldsymbol{\beta}}$ has properties similar to maximum likelihood estimates. As set forth by

McCullagh and Nelder (1989), the quasi-likelihood estimates are asymptotically normal with mean $\boldsymbol{\beta}$. Let the matrix \mathbf{F} represent the matrix of derivatives of the mean response with respect to the model parameters, such that $\mathbf{F} = \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} \right]$ is an $n \times p$ matrix. Then the asymptotic covariance matrix is (McCullagh and Nelder, 1989)

$$\text{Var}\{\hat{\boldsymbol{\beta}}\} = \Sigma = \tau (\mathbf{F}' \mathbf{V}^{-1} \mathbf{F})^{-1} = \tau \boldsymbol{\Omega} \quad (2.8)$$

where $\boldsymbol{\Omega} = (\mathbf{F}' \mathbf{V}^{-1} \mathbf{F})^{-1}$. The moment estimate,

$$\hat{\tau} = \frac{1}{N-p} \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{V_i(\hat{\mu}_i)} = \frac{X^2}{N-p}, \quad (2.9)$$

is generally used to estimate τ (McCullagh and Nelder, 1989), where X^2 is the generalized Pearson statistic, which has an asymptotic χ_{N-p}^2 .

2.2.4 Inference

The quasi-likelihood ratio test can be used to compare two nested models in much the same way as the likelihood ratio test. Let Q_{full} be the maximum quasi-likelihood achieved under the full model, and let Q_{red} be the maximum quasi-likelihood achieved under the reduced model. Then the likelihood ratio test statistic is

$$LRT = \frac{-2\{Q_{red} - Q_{full}\}}{\tau},$$

which has an approximate χ^2 distribution with M degrees of freedom (McCullagh, 1983), where M is the number of parameters constrained by the reduced model and τ is known

under the full model. However, under the quasi-likelihood framework, τ is often unknown and estimated in the modeling process. Replacing τ with a consistent estimate $\hat{\tau}$, the quasi-likelihood ratio test statistic is

$$QLRT = \frac{-2\{Q_{red} - Q_{full}\}}{\hat{\tau}M}. \quad (2.10)$$

Conservatively, for large samples, $QLRT$ has an approximate F distribution with M and $N-p$ degrees of freedom.

The asymptotic properties of the maximum quasi-likelihood estimates can also be used to form hypothesis tests and confidence intervals on the parameter vector or some subset of the parameter vector. Based on the asymptotic normality of $\hat{\beta}$, the Wald test statistic for testing the null hypothesis $\beta = \beta_0$ is

$$\begin{aligned} W &= (\hat{\beta} - \beta_0)' \Sigma^{-1} (\hat{\beta} - \beta_0) \\ &= (\hat{\beta} - \beta_0)' (\tau \Omega)^{-1} (\hat{\beta} - \beta_0) \end{aligned} \quad (2.11)$$

where Σ is as defined in equation (2.8). Based on the asymptotic normality of $\hat{\beta}$, W has an approximate χ^2 distribution with M degrees of freedom. However, under the quasi-likelihood framework, both τ and Ω are unknown and estimated in the modeling process. Replacing these parameters in equation (2.11) with consistent estimates, the Wald test statistic is

$$W_{QL} = \frac{(\hat{\beta} - \beta_0)' (\hat{\Sigma})^{-1} (\hat{\beta} - \beta_0)}{M} \quad (2.12)$$

Conservatively, for large samples, W_{QL} has an approximate F distribution with M and $N-p$ degrees of freedom. It follows that an approximate 95% confidence interval around the b^{th} model parameter, $b=1, \dots, p$, is given by

$$\hat{\beta}_{(b)} \pm \sqrt{F_{1, N-p, 1-\frac{\alpha}{2}}} \sqrt{(\hat{\Sigma})_{(b,b)}}, \quad (2.13)$$

where $(\hat{\Sigma})_{(b,b)}$ is the $(b,b)^{\text{th}}$ element of the parameter covariance matrix $\hat{\Sigma}$, and $F_{1, N-p, 1-\frac{\alpha}{2}}$

is the $\left(1 - \frac{\alpha}{2}\right)^{\text{th}}$ percentile of the F distribution with 1 and $N-p$ degrees of freedom.

The Wald test can be extended for nonlinear hypotheses of model parameters using the Delta method as described by Seber and Wild (2003). Let λ denote a $q \times 1$ vector of functions of the parameter vector β , representing some nonlinear contrast, so that the null hypothesis can be written as $\lambda=0$. Letting $\mathbf{D} = \left[\frac{\partial \lambda}{\partial \beta} \right]$ represent the $q \times p$ matrix of derivatives, it follows from the Delta Method (Agresti, 2002) that $\hat{\lambda}$ is asymptotically distributed $N(\lambda, \mathbf{D}\Sigma \mathbf{D}')$. The Wald test statistic for the nonlinear contrast λ is of the form

$$W_{QL_\lambda} = \frac{(\hat{\lambda})' (\mathbf{D}\hat{\Sigma}\mathbf{D}')^{-1} (\hat{\lambda})}{q}. \quad (2.14)$$

Conservatively, for large samples, W_{QL_λ} follows an F distribution with q and $N-p$ degrees of freedom.

2.2.4.1 Test of Additivity

Often, the goal of a ray design is to detect departure from additivity among the chemicals present in the mixture. Following the Single Chemical Required (SCR) method of Casey et al. (2004) for dose threshold models and the extension of Hamm (2004) for interaction threshold models, the test of additivity is a test of coincidence. If the dose and interaction thresholds model does not coincide with the two segment additivity model, then we can conclude departure from additivity. Note that evidence of departure from additivity is captured in the parameter θ_{mix} . Consider the situation where θ_{mix} is equal to zero, indicating that there is no evidence of departure from additivity. In this case, only one threshold parameter is estimable – the dose threshold, δ_{add_t} . If the interaction threshold is equal to the dose threshold, and the parameter θ_{mix} is equal to zero, then the dose and interaction thresholds model collapses over the second segment, resulting in the threshold additivity model. For the dose and interaction thresholds model given in equation (2.6) to coincide with the two segment additivity model along the mixture ray, the null hypothesis of additivity is given by

$$H_0 : \begin{cases} \Delta = \delta_{add_t} \\ \theta_{mix} = 0 \end{cases} .$$

Recall, however, that $\delta_{add_t} = \delta_{add} / \theta_{add}$ and that $\theta_{add} = \sum_{i=1}^c \beta_i a_i$. The hypothesis of additivity, then, is a nonlinear contrast of the model parameters, where the null hypothesis can be written $\lambda=0$ for

$$\lambda = \begin{bmatrix} \Delta - \delta_{add_t} \\ \theta_{mix} \end{bmatrix}.$$

This hypothesis can be tested using the large sample properties of either the quasi-likelihood ratio test described in equation (2.10) or the Wald test described in equation (2.14). If the hypothesis test of additivity is significant, then we can conclude that there is departure from additivity somewhere along the ray.

2.2.4.2 Test for a Region of Additivity

The dose and interaction thresholds model is designed to allow dose-dependent interactions. If we can conclude that departure from additivity exists along the ray, a secondary goal is to determine where that departure exists. If the interaction threshold is significantly larger than the dose threshold for the mixture, then we can conclude that there is an active region of additivity, a region of additivity that is associated with a response other than background. Therefore, the one-sided test of $H_0 : \Delta - \delta_{add_t} \leq 0$ is a test for an active region of additivity along the ray. This hypothesis can be tested using a Wald test as described in equation (2.14), where $\lambda = \Delta - \delta_{add_t}$.

2.3 Chlorination Ray Example

2.3.1 Background

When drinking water is disinfected, chemical byproducts are put into the water. These chemicals are known as ‘disinfectant byproducts’, or DBPs. There are approximately 500 known DBPs, and a substantial proportion of these DBPs have yet to

be studied for adverse health effects (Richardson, Simmons and Rice, 2002). Some of the known DBPs fall into a class of chemicals known as haloacetic acids (HAAs). Nine such HAAs result from the disinfection process known as chlorination and are currently under investigation by the EPA. Table 2.1 contains a list of these HAAs and the relative proportion of each resulting from the chlorination process.

Table 2.1. Haloacetic Acid Byproducts of the Chlorination Process

<u>Chemical</u>	<u>Abbreviation</u>	<u>Ratio in the Mixture</u>
Chloroacetic acid	CA	0.03
Dichloroacetic acid	DCA	0.30
Trichloroacetic acid	TCA	0.26
Bromoacetic acid	BA	0.06
Dibromoacetic acid	DBA	0.03
Tribromoacetic acid	TBA	0.02
Bromodichloroacetic acid	BDCA	0.12
Dibromochloroacetic acid	DBCA	0.04
Bromochloroacetic acid	BCA	0.14

The primary research objective is to determine whether the nine HAAs given in Table 2.1 interact when combined according to the chlorination mixing ratio. To investigate, an *in vitro* cell assay was conducted to study the cytotoxic effects of the single chemicals. A similar study of the chlorination mixture was also conducted. The response is cell viability, defined as the percentage of the negative control, an agent known to have no impact on cell viability, surviving the administered concentration. A plot of cell viability versus concentration for each single chemical in the mixture is given in Appendix A1. Figure 2.1 is the plot of cell viability versus the total concentration (in μM) of the chlorination mixture administered.

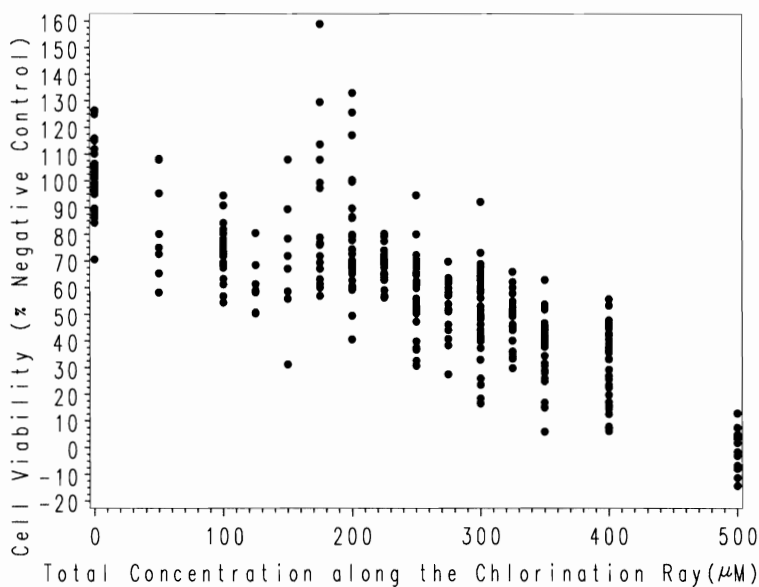


Figure 2.1. Cell Viability vs. Total Concentration of the Mixture Along the Chlorination Ray

2.3.2 Model Selection

As can be seen in Figure 2.1, the concentration-response relationship between the chlorination mixture and cell viability is decreasing overall. More specifically, cell viability appears to be strictly decreasing for mixture concentrations prior to approximately 100 μM . Beyond 100 μM , the cell viability appears to increase slightly and then decrease for the remaining concentrations within the experimental range. In other words, larger concentrations of the chlorination mixture produce smaller cell viability, while low concentrations of the mixture tend to produce less cytotoxicity.

As a preliminary analysis of the data, we used the single chemical data to predict additivity along the chlorination ray and compared the predicted additivity model to the observed mixture data. Because the response is a percentage of control, it is theoretically reasonable to restrict the mean response range to the interval between and including 0 and

100. In other words, at a total concentration of 0 μM , we expect a mean response of 100% of the cell viability of the control group. Additionally, as the total concentration increases, we expect that the cell viability will decrease until it reaches 0%. In equation (2.2), μ approaches α as t increases. Therefore, we set α equal to 0. Similarly, if we want to restrict our model so that μ equals 100 when t is equal to 0, and we let β_0 equal 0, then it can be shown that γ should be set to 200. We chose to model the logistic function of the mean; however, we could have worked with other link functions as well.

There are few concentration groups in the low concentration region, so that a concentration threshold is not visible in Figure 2.1. However, a number of the single chemical plots shown in Appendix A1 suggest the presence of a concentration threshold, suggesting that the preliminary analysis of the mixture should involve a threshold. The single chemical data were used to fit the threshold additivity model of the form

$$\mu_{(add)} = \begin{cases} 100 & , \sum_{i=1}^9 \beta_i a_i t > \delta_{add} \\ \frac{200}{1 + \exp\left\{-\left(\left(\sum_{i=1}^9 \beta_i a_i\right)\left(t - \delta_{add} / \sum_{i=1}^9 \beta_i a_i\right)\right)\right\}} & , \sum_{i=1}^9 \beta_i a_i t \leq \delta_{add} \end{cases} \quad (2.15)$$

to predict additivity along the chlorination ray. The threshold parameter under additivity, δ_{add} , was found to be statistically significant. Therefore, we decided to incorporate a concentration threshold in the mixture model to describe the relationship between cell viability and the concentration of the mixture along the chlorination ray. Figure 2.2A demonstrates the threshold additivity model based only on the single chemical data for total concentrations (in μM) along the chlorination ray. Figure 2.2B presents the fit of the threshold additivity model, based on the single chemical data only, to the mixture data

observed along the chlorination ray. The threshold additivity model predicts no change from a background response of 100% for mixture doses less than the concentration threshold; for mixture doses larger than the concentration threshold, the threshold additivity model predicts a decrease in cell viability until the response plateaus at 0%.

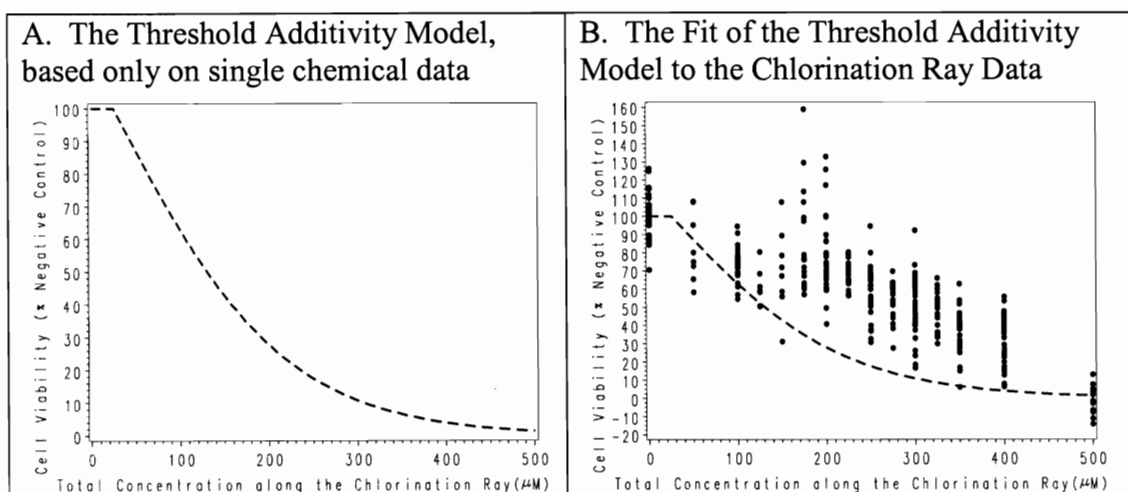


Figure 2.2. Predicted Cell Viability for the Chlorination Ray Under Additivity

As can be seen in Figure 2.2B, the predicted additivity model adequately fits the chlorination ray data for total concentrations less than approximately 100 μM . For larger total concentrations, the additivity model predicts a smaller cell viability than indicated by the observed mixture data. This region of disagreement between the chlorination ray data and the predicted additivity model may be indicative of an interaction threshold. Thus, the model chosen to describe this concentration-response relationship is of the form given in equation (2.6).

It does not appear, however, that the difference between the concentration-response relationship prior to the interaction threshold and that beyond the interaction threshold is simply a change in slope. Rather, the temporary increase in observed cell viability visible

in Figure 2.2B seems to indicate that a higher order term is necessary to describe the relationship in the third segment. Specifically, we added a quadratic term for added flexibility in the interaction region, so that the dose and interaction thresholds model is of the form

$$\mu_h = \begin{cases} 100 & , t < \delta_{add_1} \\ \frac{200}{1 + \exp\left\{-\left(\left(\sum_{i=1}^9 \beta_i a_{i(h)}\right)(t - \delta_{add_1})\right)\right\}} & \left. \begin{array}{l} t \geq \delta_{add_1}, h \leq 9 \\ \delta_{add_1} \leq t < \Delta, h = 10 \end{array} \right\} \\ \frac{200}{1 + \exp\left\{-\left(\left(\sum_{i=1}^9 \beta_i a_{i(h)}\right)(t - \delta_{add_1}) + \theta_{mix}(t - \Delta) + \theta_{mix2}(t - \Delta)^2\right)\right\}} & , t \geq \Delta \text{ and } h = 10 \end{cases} \quad (2.16)$$

It is common in concentration-response modeling for the variability of the response to change as a function of the mean response. If such non-constant variance exists, quasi-likelihood methods can be used to fit the model. Figure 2.3 contains plots of the variance versus the mean cell viability, which can be used to assess the need to use quasi-likelihood methods in our analysis. It is easily visible in Figure 2.3A that the variability associated with the response is not constant. In fact, the variability appears to increase and then decrease as the mean increases. Based on this observation, we found that the form of the variance function that most adequately fit the data was

$$Var\{Y\} = \tau \left(\mu + \frac{\mu^2}{k} \right), \text{ where } k = -124. \text{ This function resulted in an } R^2 \text{ value of } 0.65,$$

indicating that 65% of the variability present in the response variance can be accounted

for by this relationship. A plot of the predicted variance versus the mean cytotoxicity is given in Figure 2.3B.

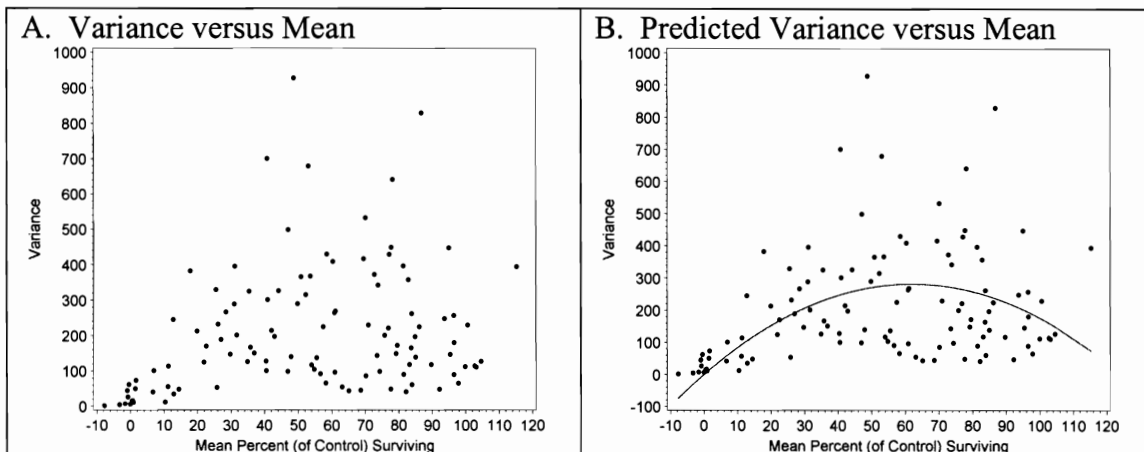


Figure 2.3. Change in Variance as a Function of the Mean Response

2.3.3 Results

Quasi-likelihood methods were used to fit the dose and interaction thresholds model, given in equation (2.16), to the single chemical data and the mixture data simultaneously. Figure 2.4 displays three plots which allow us to assess the fit of the dose and interaction thresholds model. Figure 2.4A demonstrates the fit of this model to the data observed from the chlorination mixture along the ray, as well as the fit of the additivity model given in equation (2.15). Prediction plots for the single agents are given in Appendix A2. Figure 2.4B demonstrates the relationship between the observed and predicted responses, while Figure 2.4C demonstrates the relationship between the predicted responses and their corresponding residuals.

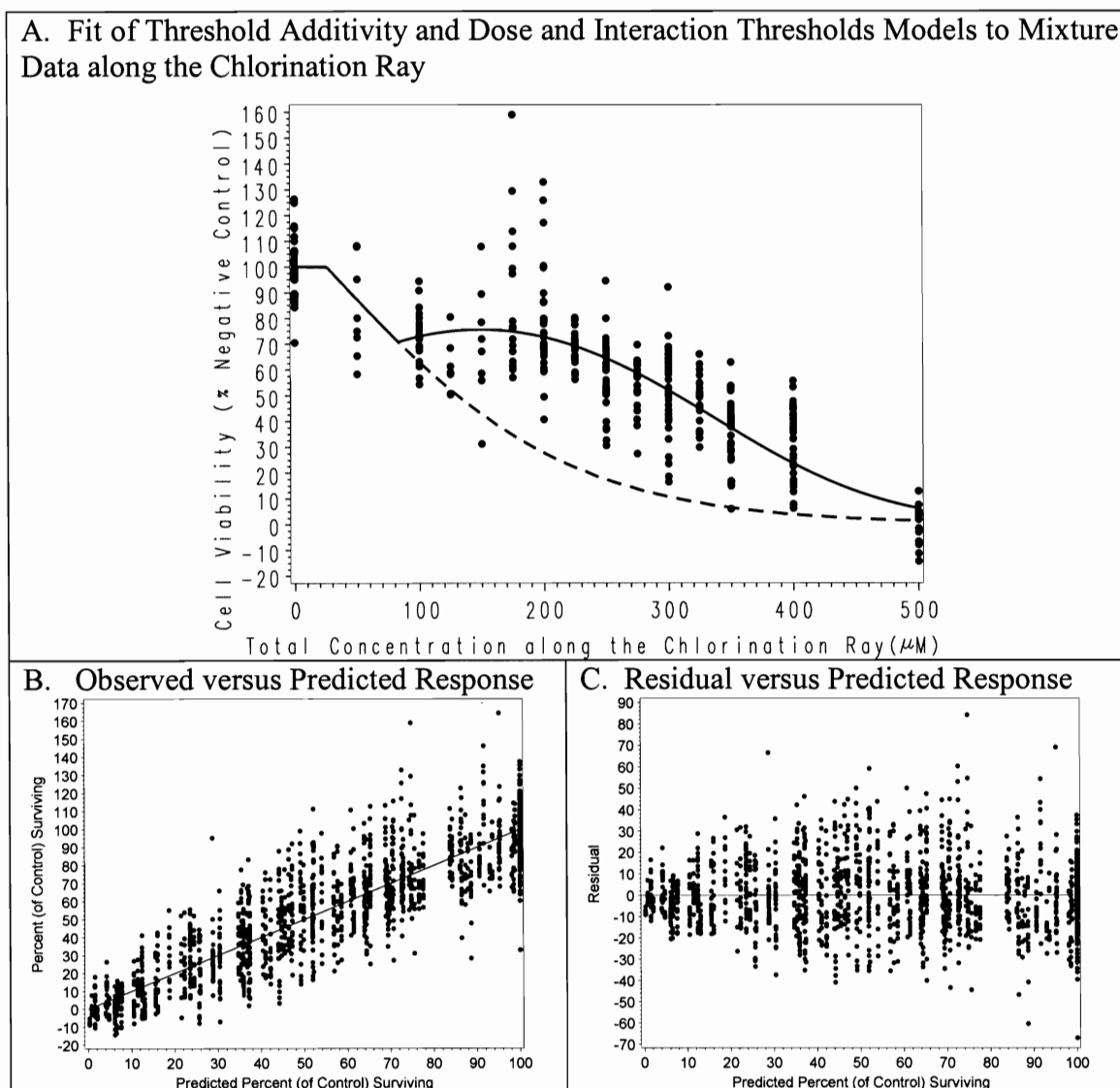


Figure 2.4. Assessment of Model Fit

As can be seen in Figure 2.4, the dose and interaction thresholds model more closely represents the relationship depicted by the points on the scatter plot than the additivity model. Additionally, the plots in Figures 2.4B and 2.4C do not indicate a problem with the fit of the dose and interaction thresholds model. The relationship between the observed and predicted responses is increasing, and there is no visible pattern in the variability of the points about the perfect prediction line. There is no visible pattern in

the variability of the points about the line where the residual is equal to zero. These graphical diagnostics did not give us cause to question the validity of the dose and interaction thresholds model, so we based our inference on it.

2.3.4 Inference

Recall that the primary research objective was to determine whether the nine HAAs involved in the mixture interact when combined according to the chlorination ray. This objective can be thought of as a test of the coincidence of the additivity and the mixture models. For the dose and interaction thresholds model given in equation (2.16) to coincide with the additivity model given in equation (2.15), restrictions must be placed on three of the model parameters. In particular, the null hypothesis of additivity can be written

$$H_0 : \begin{cases} \Delta = \delta_{add_t} \\ \theta_{mix} = 0 \\ \theta_{mix2} = 0 \end{cases} ,$$

which can be tested according to the methods described in section 2.2.4 as follows.

The hypothesis of additivity can be tested using a quasi-likelihood ratio test as described in equation (2.10). The additivity model described in equation (2.15) was fit to both the chlorination ray data and the single chemical data simultaneously. The achieved quasi-likelihood is that of the reduced model, Q_{red} , while the quasi-likelihood achieved under the dose and interaction thresholds model is that of the full model, Q_{full} . The

resulting quasi-likelihood ratio test was statistically significant ($F_{3,1765}=159.7$, $pvalue<0.0001$), indicating that there is departure from additivity along the ray.

Alternatively, the hypothesis of additivity can be tested using the Wald test for nonlinear hypotheses. Let $\lambda' = [\Delta - \delta_{add_t} \quad \theta_{mix} \quad \theta_{mix2}]$ be a vector of functions of the parameter vector β , representing the parameters involved in the null hypothesis. In particular, let

$$\begin{aligned} \lambda' &= \mathbf{h}(\beta) \\ &= [h_1(\beta) \quad h_2(\beta) \quad h_3(\beta)] \\ &= \begin{bmatrix} \Delta - \frac{\delta_{add}}{\sum_{i=1}^c \beta_i a_i} & \theta_{mix} & \theta_{mix2} \end{bmatrix}. \end{aligned}$$

Then $D = \left[\frac{\partial \lambda}{\partial \beta'} \right]$ is the corresponding matrix of derivatives, which is given in Appendix A3. From the Delta Method, we know that $\hat{\lambda}$ is asymptotically normally distributed with mean λ and covariance matrix $\mathbf{D}\Sigma\mathbf{D}'$. The Wald test statistic of this hypothesis is of the form given in equation (2.14),

$$W_{QL_\lambda} = \frac{(\hat{\lambda})' (\mathbf{D}\Sigma\mathbf{D}')^{-1} (\hat{\lambda})}{3} \Bigg|_{\beta=\hat{\beta}},$$

which has an approximate F distribution with 3 and 1765 degrees of freedom. The result of the hypothesis test of additivity is statistically significant ($F_{3,1765}=132.08$, $pvalue<0.0001$). Therefore, we again conclude that there is departure from additivity along the chlorination ray.

Table 2.2 lists the parameter estimates resulting from the fit of the dose and interaction thresholds model, given in equation (2.16). Each of the single agent slopes is statistically significant and negative. The parameter associated with the concentration threshold is statistically significant and negative, indicating a positive estimate for the concentration threshold of each single agent. The estimate of θ_{add} , the slope of total concentration under additivity, which is a function of the single chemical slopes and the proportions defining the chlorination ray, is -0.0105 (SE 0.0003). The interaction threshold is statistically significant and positive. The estimate of θ_{mix} is statistically significant and positive. Recall that θ_{mix} is the change in slope that occurs for total concentrations larger than the interaction threshold. Therefore, for total concentrations larger than $83.2 \mu\text{M}$, the slope of the total concentration of the mixture is less negative than the slope in the additivity region. The estimate of θ_{mix2} is statistically significant and positive.

Table 2.2. Parameter Estimates Resulting from Fit of the Dose and Interaction Thresholds Model to the Single Chemical and Mixture Data along the Ray

Parameter	Estimate	Standard Error	P-value
β_1 (CA)	-0.00183	0.000077	<0.0001
β_2 (DCA)	-0.00022	8.441E-6	<0.0001
β_3 (TCA)	-0.00051	0.000022	<0.0001
β_4 (BA)	-0.1492	0.00532	<0.0001
β_5 (DBA)	-0.00293	0.000108	<0.0001
β_6 (TBA)	-0.0199	0.000896	<0.0001
β_7 (BDCA)	-0.00199	0.000074	<0.0001
β_8 (DBCA)	-0.00725	0.000339	<0.0001
β_9 (BCA)	-0.00180	0.000066	<0.0001
δ_{add}	-0.2692	0.0268	<0.0001
θ_{mix}	0.0136	0.000952	<0.0001
θ_{mix2}	-0.00002	3.33E-6	<0.0001
Δ	83.1674	6.6572	<0.0001
τ	12.47		

The process for constructing confidence intervals around the threshold parameters is similar. Let $(\delta^*) = [\delta_1^* \ \delta_2^* \ \dots \ \delta_9^* \ \delta_{add_t} \ \Delta]$. Then $D = \left[\frac{\partial \delta^*}{\partial \beta} \right]$ is the corresponding matrix of derivatives, which is given in Appendix A4. The estimates of the threshold for each single chemical, as well as the concentration threshold and the interaction threshold for the mixture, are given in Table 2.3. The corresponding standard errors and confidence intervals are also given. As can be seen in the table, each confidence interval around the single chemical thresholds has a lower limit that exceeds zero. It is interesting to note that two chemicals, DCA and TCA, make up 56% of the mixture; the corresponding concentration threshold estimates for these agents are the largest in the mixture. Conversely, BA represents only 6% of the mixture, and the corresponding concentration threshold is the smallest in the mixture, estimated at 1.8 μM .

At the concentration threshold for the mixture, estimated at 25.7 μM , the amount of BA in the mixture (approximately 1.5 μM) is approaching its concentration threshold.

Table 2.3. Threshold Parameter Estimates

Agent	Threshold Estimate	Standard Error	Approximate 95% Lower Confidence Bound	Approximate 95% Upper Confidence Bound
CA	147.23	12.92	121.9	172.56
DCA	1232.07	111.13	1014.11	1450.02
TCA	525.85	49.01	429.72	621.98
BA	1.8	0.16	1.5	2.11
DBA	91.83	7.97	76.21	107.46
TBA	13.54	1.24	11.1	15.98
BDCA	135.53	12.07	111.86	159.2
DBCA	37.12	3.41	30.44	43.8
BCA	149.16	13.1	123.48	174.85
Chlorination Ray Mixture (Concentration Threshold)	25.7	2.21	21.37	30.03
Chlorination Ray Mixture (Interaction Threshold)	83.17	6.66	70.11	96.22

The lower confidence bound around the concentration threshold for the mixture along the chlorination ray also exceeds zero. We are approximately 95% confident that, for the chlorination ray mixture, the response will be that of background for concentrations below 21 μM , the lower confidence bound for the concentration threshold of the mixture along the chlorination ray. The Wald test for a region of additivity described in Section 2.2.4 was found to be statistically significant ($F_{1,1765}=76.1$, $p\text{value}<0.0001$). Therefore, we conclude that there is a significant region of additivity along the ray. The estimate of $\delta_{add,t}$, the concentration threshold in terms of total concentration along the chlorination ray, is 25.7 μM (SE 2.21). The estimate of the interaction threshold, Δ , is 83.2 μM (SE 6.66). We are approximately 95% confident that

the chemicals involved in the mixture behave as we would expect under additivity for concentrations below $70\mu\text{M}$, the lower confidence bound for the interaction threshold. Beyond $96\mu\text{M}$, the upper confidence bound for the interaction threshold, we have detected significant departure from additivity.

According to our model, then, there is no change in cell viability prior to the concentration threshold, a total concentration of approximately $21\mu\text{M}$, the lower confidence bound for δ_{add_t} . Conservatively, the agents interact as expected under additivity for total concentrations less than approximately $70\mu\text{M}$, the lower confidence bound for Δ . It may be that the agents interact as expected under additivity for total concentrations up to $96\mu\text{M}$, the upper confidence bound for Δ . In the region between the concentration threshold and the interaction threshold, we found that cell viability decreases as total concentration increases, and that this decrease is significantly different from zero. Beyond the interaction threshold, we found that both the parameter representing the change in the slope of total concentration, as well as the quadratic term, were significant. In this region, the dose and interaction thresholds model predicts higher cell viability, or less cell death, than that predicted by the additivity model. This indicates that an antagonistic relationship exists among the nine HAAs involved in the chlorination ray when the mixture is administered at a concentration larger than $96\mu\text{M}$.

2.3.5 Design Consideration

We have seen that the experiment performed along the chlorination ray was sufficient to support the dose and interaction thresholds model and to demonstrate

departure from additivity. The mixture experiment performed was fairly large, including 15 concentration groups, spaced throughout the entire concentration range. Each concentration group contained between eight and 32 observations, for a total of 327 observations in the mixture experiment. Studies this large require a tremendous amount of resources and are rarely seen in the literature.

A goal of mixtures research, then, should be to design mixture studies that are feasible, with respect to the available resources, but are sufficiently powerful to investigate the possible interactions among a mixture of a large number of chemicals along a fixed-ratio ray. Consider the dose and interaction thresholds model, for instance, which states that the relationship among the chemicals is that of additivity for doses smaller than the interaction threshold. While both single chemical data and data from the mixture ray are used to fit the dose and interaction thresholds model, the support of the additivity region comes mainly from the single chemical data. With preliminary information about the interaction region, the optimal design for the dose and interaction thresholds model should take into account the single chemical data that is available and focus on the interaction region.

Statistical methodology can be used to find the optimal experimental design for supporting the dose and interaction thresholds model, as well as for testing the hypothesis of additivity. Optimal design strategies require a prior hypothesis regarding the shape and the location of the interaction region. Therefore, it is of interest to determine which features of the dose-response curve form the basis for the optimal experimental design.

2.4 Optimal Experimental Designs

The experimental design used to generate the data must be sufficient to support the model intended to describe the dose-response relationship. In the chlorination ray example, the mixture data included 15 concentration groups, which were roughly equally spaced throughout the experimental range. The model we intended to fit, the dose and interaction thresholds model, was made up of three segments, two of which received most of their support from single chemical data that had been previously collected. The use of design criteria to determine the optimal experimental design, then, may have allowed the experimenter to collect the information necessary to support this model with fewer concentration groups and/or fewer observations.

Consider the situation described in Section 2.3.1, the chlorination ray example. In this example, the chlorination process results in a mixture of nine HAAs, which combine according to the chlorination mixing ratio, or chlorination ray. The primary research objective is to determine whether these HAAs interact when combined in this manner. Sufficient single chemical data are available, for each of the nine HAAs of interest, to model the relationship between concentration and cell viability under the assumption of additivity. The plots of the single chemical data are given in Appendix A1.

Cell viability decreases as the concentration increases, regardless of the single chemical administered. In addition, a number of the single chemical plots indicate the presence of a concentration threshold. Since the response is a percent of control, it is theoretically reasonable to restrict the range of the response to the interval between and including 0 and 100. Let

$$\mu_{(add)} = \begin{cases} 100, & \sum_{i=1}^9 \beta_i x_i > \delta_{add} \\ \frac{200}{1 + \exp\left\{-\left(\sum_{i=1}^9 \beta_i x_i - \delta_{add}\right)\right\}}, & \sum_{i=1}^9 \beta_i x_i \leq \delta_{add} \end{cases} \quad (2.17)$$

describe the concentration-response relationship expected under additivity. This model indicates that there is no change in cell viability prior to the concentration threshold. Beyond the concentration threshold, the agents behave under the assumption of additivity. The additivity model in equation (2.17) is equivalent to

$$\mu_{(add)} = \begin{cases} 100, & t < \delta_{add_t} \\ \frac{200}{1 + \exp\left\{-\left(\theta_{add} (t - \delta_{add_t})\right)\right\}}, & t \geq \delta_{add_t} \end{cases} \quad (2.18)$$

along the chlorination ray. The additivity parameters can be estimated from the single chemical data.

The stability of the parameter covariance matrix was initially of concern, because the determinant of the estimated parameter covariance matrix under the assumption of additivity was very close to zero. The fit of the additivity model resulted in variances less than 10^{-6} for all but two parameters. In many instances, these variances are as small or smaller than what the computer considers to be zero. Therefore, in order to stabilize the estimate of $Var\{\hat{\beta}\}$, it was necessary to change the units associated with the single chemical data. The single chemical data presented in Appendix A1 was expressed in micromoles (μM). To find the optimal experimental design, the data were expressed in 10^{-4} M (10^{-1} mM). Modifying the units associated with the single chemical data

improved the stability of the parameter covariance matrix, since it increased $|Var\{\hat{\beta}\}|$. Therefore, it was determined that the instability of the parameter covariance matrix was not due to over-parameterization of the model but to the small variances associated with the parameter estimates. However, while this change in metameter increased the stability of $Var\{\hat{\beta}\}$, it did not solve the problem entirely.

Changing the units increased the variances associated with the slope parameter estimates as expected; however, the change in units decreases the concentration range, which decreases the interaction threshold. This decrease in the interaction threshold is associated with a corresponding decrease in the variance associated with the estimate of the interaction threshold, leaving us with a small variance problem similar to that seen in the additivity analysis conducted in μM . Therefore, we reparameterized the mixture model by setting $\Delta = \Delta_{new}/1000$. Reparameterizing the model so that we estimate Δ_{new} , rather than Δ , reduces the range of the variability associated with the parameter estimates. These changes together give us a more stable parameter covariance matrix, where $|Var\{\hat{\beta}\}|$ is larger than 1, on which to base our design criterion.

Using the quasi-likelihood methods described in Section 2.2.3, we fit the additivity model given in equation (2.17) to the single chemical data. The dashed line in Figure 2.5 demonstrates the predicted relationship between cell viability and total concentration, under the assumption of additivity, along the chlorination ray. Suppose that it has been hypothesized that the nine HAAs involved in the chlorination ray interact when

administered in large concentrations. Specifically, it has been hypothesized that the true concentration-response relationship is better described by the solid line in Figure 2.5.

Dashed Line: Additivity Model, Solid Line: Hypothesized Mixture Model

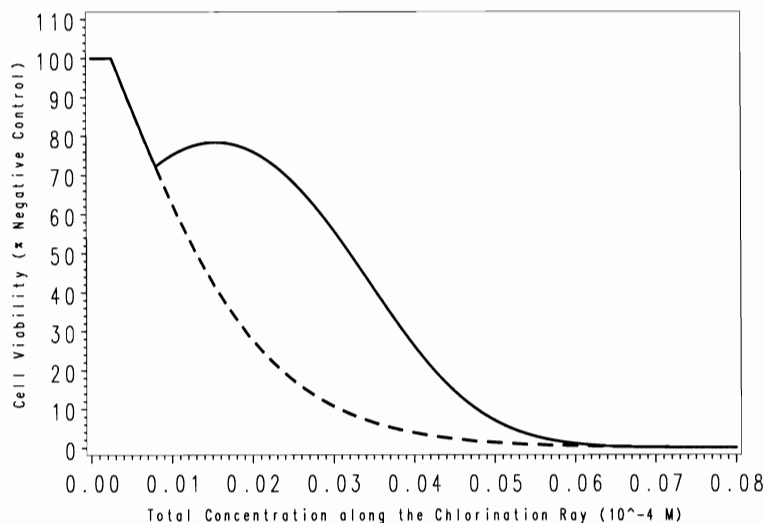


Figure 2.5. Hypothesized Relationship between Cell Viability and Total Concentration

It should be noted that we have single chemical data to support the additivity region of the hypothesized mixture model; we are using this single chemical data to design a mixture experiment. From this single chemical data, we have estimated the slope parameters for each of the nine chemicals under the assumption of additivity, and we have hypothesized parameter values for the interaction region. The parameter estimates resulting from the fit of the additivity model to the single chemical data are given in Appendix A5.

We restricted the concentration range of the mixture experiment. As can be seen in Figure 2.5, the hypothesized model plateaus at approximately zero for concentrations larger than $0.07 \cdot 10^{-4}$ M. We allowed the concentrations in our experiment to iterate as

high as $0.1 \cdot 10^{-4} \text{M}$. If we allowed the concentrations to iterate much larger than this value, the predicted response is essentially zero, which is problematic for the computation of the covariance matrix. Lastly, because we understand that a mixture experiment would not be conducted without a control group, we forced our design to include a control concentration. That is, the first concentration point in each design is always zero, though the proportion of subjects included at the zero concentration iterates within the design search.

Following Casey et al. (2005), we used the Nelder-Mead Simplex Algorithm to find the concentrations, $\underline{\mathbf{d}} = \{d_1 \quad d_2 \quad \cdots \quad d_m\}$, and the proportion of the sample allocated to each concentration, $\underline{\mathbf{q}} = \{q_1 \quad q_2 \quad \cdots \quad q_m\}$, that minimize the appropriate objective function. A SAS macro was written to find the optimal design for a number of starting values, under a number of design criteria. The corresponding SAS code is given in Appendix A6.

We considered the D-optimal design criterion as well as the Ds-optimal design criterion. We also considered a design criterion based on the hypothesis of additivity. Each design criterion is based on the estimate of the covariance matrix of the model parameters,

$$\text{Var}\{\hat{\boldsymbol{\beta}}\} = \Sigma = \tau(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}, \quad (2.19)$$

where the matrices \mathbf{F} and \mathbf{V} are as defined in Section 2.2.3.

Consider a mixture experiment with n_r observations taken at $r=1, \dots, m$ design points, and recall that there are n_s observations from the single chemical experiments.

The concentration values and the number of observations at each design point are represented in the \mathbf{F} matrix, which is defined as

$$\mathbf{F}_{N \times p} = \left[\begin{array}{c} \left. \begin{array}{c} \partial\mu_1(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \\ \vdots \\ \partial\mu_{n_s}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \end{array} \right\} n_s \times p \\ \text{-----} \\ \left. \begin{array}{c} \partial\mu_{11}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \\ \vdots \\ \partial\mu_{1n_1}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \end{array} \right\} n_1 \times p \\ \left. \begin{array}{c} \partial\mu_{21}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \\ \vdots \\ \partial\mu_{2n_2}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \end{array} \right\} n_2 \times p \\ \vdots \\ \left. \begin{array}{c} \partial\mu_{m1}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \\ \vdots \\ \partial\mu_{mn_m}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}' \end{array} \right\} n_m \times p \end{array} \right].$$

The total number of observations to be included in the mixture experiment, then, is

$n = \sum_{r=1}^m n_r$, and $N=n_s+n$ is the total number of observations to be used in the analysis. For

the dose and interaction thresholds model described in equation (2.16) and shown in Figure 2.5, p is 13. The derivatives contained in the \mathbf{F} matrix are given in Appendix A7.

The derivatives for the n_s observations associated with single chemical data are fixed. The derivatives for the n observations associated with the mixture experiment will iterate within the Nelder-Mead Simplex Algorithm, which will determine the concentration associated with each of the m design points, d_r , and the allocation of observations to each design point, q_r , $r=1, \dots, m$. Since the number of observations at the r^{th} design point

depends on the allocation of observations at the r^{th} design point, the number of observations at each design point will also iterate. The concentration and the allocation of observations that minimizes the design criterion is the optimal r point design.

For a given n , the D-optimal design is based on the proportion of observations at each concentration, rather than the number of observations at each concentration. In other words, the optimal design is not dependent on the number of observations available for the mixture experiment. This allows the experimenter to use the optimal design to determine the sample size needed for the mixture experiment to achieve a given power.

We began our search for the optimal designs with an eight point design. Recall that the results of Nelder-Mead can be somewhat dependent on starting values; to account for this, we used a grid of starting values to search for the optimal r point design. The starting values for the eight point design included two concentrations prior to the concentration threshold, two concentrations between the concentration threshold and the interaction threshold, and four concentrations beyond the interaction threshold. We considered 32 such sets of starting values. Each set of starting values was provided to the macro, and a candidate optimal design was reached in each case. The candidate optimal design with the minimum optimization criterion was considered to be the optimal r point design.

2.4.1 D-optimality

The D-optimal design is the design that minimizes the generalized variance, defined as the determinant of the covariance matrix of the model parameters, $|\text{Var}\{\hat{\boldsymbol{\beta}}\}|$. The D-

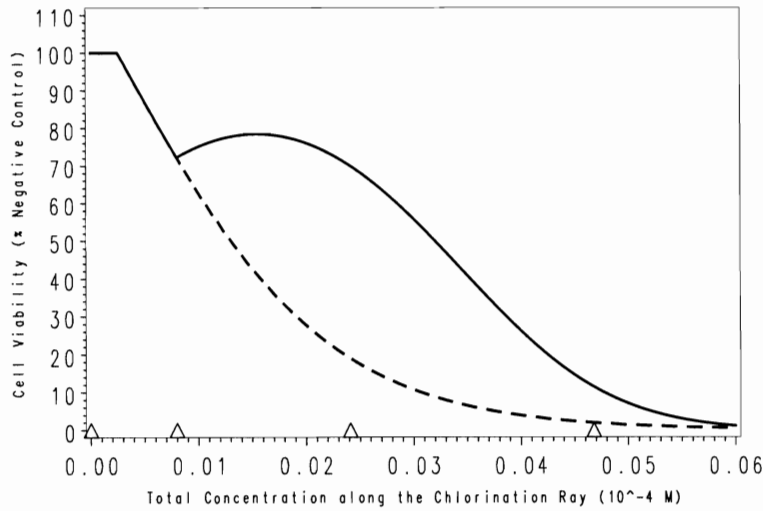
optimal designs are given in Table 2.4, where \mathbf{d} represents the vector of total concentration locations and \mathbf{q} represents the vector of corresponding proportions. As can be seen in Table 2.4, the D-optimal eight point design contains three design points that are associated with less than 0.5% of the available sample. This suggested that eight concentration groups were not necessary to support the model. Therefore, our starting values were reduced to a seven point design. Following the same procedure used to find the D-optimal eight point design, a D-optimal seven point design was found. Similar to the eight point design, the D-optimal seven point design indicated that seven concentration groups were not necessary to support the model. The D-optimal six point design, as well as the D-optimal five point and four point designs, led to the same conclusion. The D- optimal four point design, however, could not be reduced, because three design points are necessary to support the quadratic term in the interaction region. The design associated with the minimum determinant of the covariance matrix, among the optimal designs found in our research and presented in Table 2.4, is the D-optimal four point design, which results in a generalized variance of 3137.72.

Table 2.4. D-optimal Designs

Number of Points	Design: $\begin{Bmatrix} \mathbf{d} \\ \mathbf{q} \end{Bmatrix}$	Generalized Variance
8	$\begin{Bmatrix} 0 & 0.0002 & 0.0056 & 0.0065 & 0.0080 & 0.0248 & 0.0470 & 0.0475 \\ 0.001 & 0.003 & 0.068 & 0.000 & 0.224 & 0.317 & 0.290 & 0.097 \end{Bmatrix}$	4007.27
7	$\begin{Bmatrix} 0 & 0.0028 & 0.0080 & 0.0081 & 0.0237 & 0.0242 & 0.0468 \\ 0.000 & 0.019 & 0.297 & 0.017 & 0.187 & 0.212 & 0.269 \end{Bmatrix}$	3405.10
6	$\begin{Bmatrix} 0 & 0.0080 & 0.0080 & 0.0244 & 0.0470 & 0.0472 \\ 0.003 & 0.226 & 0.137 & 0.285 & 0.233 & 0.116 \end{Bmatrix}$	3191.29
5	$\begin{Bmatrix} 0 & 0.0080 & 0.0237 & 0.0467 & 0.0469 \\ 0.000 & 0.334 & 0.337 & 0.236 & 0.093 \end{Bmatrix}$	3139.65
4	$\begin{Bmatrix} 0 & 0.0080 & 0.0241 & 0.0468 \\ 0.001 & 0.338 & 0.325 & 0.337 \end{Bmatrix}$	3137.72

The D-optimal design, in relation to the hypothesized dose and interaction thresholds model, is presented in Figure 2.6. The concentration and the proportion of observations allocated to that concentration are given in the table next to the plot. The first design point represents the control group, though it is not required by the optimal design. It is interesting to note that the second design point is located on the hypothesized interaction threshold. All remaining design points are located in the interaction region, which is not surprising since the additivity region receives most of its support from the single chemical data. The Ds-optimal design, which minimizes the determinant of the subset of the covariance matrix related to the interaction parameters, yields a similar result.

Dashed Line: Additivity Model, Solid Line: Hypothesized Mixture Model
 Triangle: D-optimal Dose Location



D-Optimal Design	
Location (10^{-4} M)	Allocation (%)
0	0.1
0.0080	33.8
0.0241	32.5
0.0468	33.7

Figure 2.6. The D-Optimal Design for Supporting the Hypothesized Dose and Interaction Thresholds Model

2.4.2 Optimality with Regard to the Hypothesis Test of Additivity

The additivity optimal design is the design that minimizes the determinant of the covariance matrix associated with the hypothesis test of additivity. Recall from Section 2.3.4 that the hypothesis test of additivity for a model of the form given in equation (2.16) is

$$H_0 : \begin{cases} \Delta = \delta_{add_t} \\ \theta_{mix} = 0 \\ \theta_{mix2} = 0 \end{cases} .$$

Also recall, however, that we have reparameterized the model to allow $\Delta = \frac{\Delta_{new}}{1000}$.

Therefore, the hypothesis of additivity on which the additivity optimal design is based is

$$H_0 : \begin{cases} \Delta_{new}/1000 = \delta_{add_t} \\ \theta_{mix} = 0 \\ \theta_{mix2} = 0 \end{cases},$$

which can be tested according to the methods described in Section 2.2.4. Recall from section 2.3.4 that λ is the vector of functions of β that represents the null hypothesis of additivity. The above null hypothesis, rewritten in terms of λ , is $H_0 : \lambda = \mathbf{0}$. The additivity optimal design, then, is the design that minimizes the determinant of the covariance matrix of $\hat{\lambda}$, $\mathbf{D}\Sigma\mathbf{D}'$. The additivity optimal designs are given in Table 2.5, where \mathbf{d} represents the vector of total concentration locations and \mathbf{q} represents the vector of corresponding proportions.

The additivity optimal eight point design suggested that eight concentration groups were not necessary to support the model. Therefore, our starting values were reduced to a seven point design. Following the same procedure used to find the additivity optimal eight point design, an additivity optimal seven point design was found. Similar to the eight point design, the additivity optimal seven point design indicated that seven concentration groups were not necessary to support the model. The additivity optimal six point design, as well as the additivity optimal five and four point designs, led to the same conclusion. Again, however, the additivity optimal four point design cannot be reduced, because three design points are necessary to support the quadratic term in the interaction region.

Table 2.5. Additivity Optimal Designs for the Hypothesized Dose and Interaction Thresholds Model

Number of Points	Design(s): $\begin{Bmatrix} \mathbf{d} \\ \mathbf{q} \end{Bmatrix}$	$ \mathbf{D}\Sigma\mathbf{D}' $
8	$\begin{Bmatrix} 0 & 0.0026 & 0.0050 & 0.0080 & 0.0080 & 0.0241 & 0.0497 & 0.0505 \\ 0.008 & 0.030 & 0.003 & 0.147 & 0.219 & 0.224 & 0.237 & 0.132 \end{Bmatrix}$	0.33175
7	$\begin{Bmatrix} 0 & 0.0005 & 0.0008 & 0.0081 & 0.0248 & 0.0488 & 0.0492 \\ 0.003 & 0.046 & 0.110 & 0.203 & 0.230 & 0.291 & 0.117 \end{Bmatrix}$	0.53766
6	$\begin{Bmatrix} 0 & 0.0080 & 0.0081 & 0.0247 & 0.0500 & 0.0518 \\ 0.016 & 0.270 & 0.130 & 0.263 & 0.200 & 0.120 \end{Bmatrix}$	0.31714
5	$\begin{Bmatrix} 0 & 0.0080 & 0.0244 & 0.0284 & 0.0502 \\ 0.002 & 0.442 & 0.190 & 0.037 & 0.328 \end{Bmatrix}$	0.31127
4	$\begin{Bmatrix} 0 & 0.008 & 0.0261 & 0.0514 \\ 0.003 & 0.417 & 0.270 & 0.310 \end{Bmatrix}$	0.31941

The design associated with the minimum determinant of the covariance matrix associated with the nonlinear contrast representing additivity, among the optimal designs found in our research and presented in Table 2.5, is the additivity optimal five point design, which results in a determinant of 0.311. The additivity optimal five point design suggests that five points are not necessary to support the dose and interaction thresholds model specified under our hypothesis; in fact, the control concentration that we have required to be included in the design is not necessary, requiring 0.2% of the study sample size. However, when we simplified to the four point additivity optimal design, we saw an increase in the determinant. The additivity optimal design, in relation to the hypothesized dose and interaction thresholds model, is presented in Figure 2.7. The concentration and the proportion of observations allocated to that dose are given in the table next to the plot.

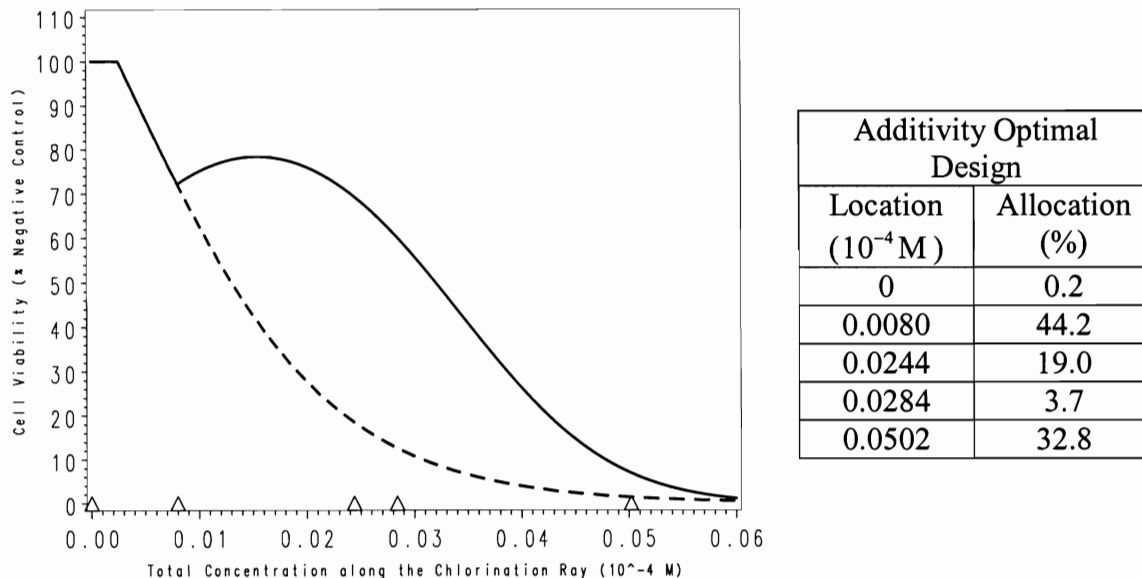


Figure 2.7. The Additivity Optimal Design for the Hypothesized Dose and Interaction Thresholds Model

The first design point represents the control group, though it is not required by the optimal design. It is interesting to note that the second design point is located on the hypothesized interaction threshold. All remaining design points are located in the interaction region, which is not surprising since the additivity region receives most of its support from the single chemical data.

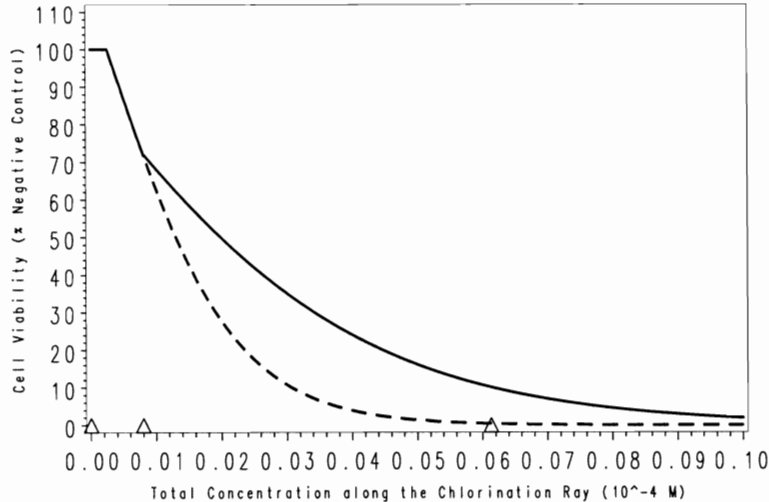
2.4.3 Additivity Optimal Designs for the General Case

Additivity optimal designs for the general cases of synergism and antagonism, where no quadratic term is included in the third segment of the model, were constructed as well. We have seen that the additivity optimal design includes only five design points; therefore, our search for the general additivity optimal designs began with a five point design as well. In both the synergistic case and the antagonistic case, the additivity

optimal five point design suggests that five points are not necessary to support the model. The additivity optimal five point designs are not much different from a three point design, with data allocated to doses at the interaction threshold and higher. The additivity optimal four point design is not much different from a two point design, with data allocated to doses at the interaction threshold and higher. The same is true for the additivity optimal three point design. While the three point design does not require the use of a control concentration, we cannot simplify the design further because two design points are necessary to support the linear form of the interaction region.

The additivity optimal design, in relation to the hypothesized dose and interaction thresholds model under the general antagonism assumption, is presented in Figure 2.8.

Dashed Line: Additivity Model, Solid Line: Hypothesized Mixture Model
 Triangle: Additivity Optimal Dose Location



Additivity Optimal Design	
Location (10 ⁻⁴ M)	Allocation (%)
0	0.2
0.0080	65.4
0.0613	34.4

Figure 2.8. The Additivity Optimal Design in the General Antagonism Case

The additivity optimal design, in relation to the hypothesized dose and interaction thresholds model under the general synergism assumption, is presented in Figure 2.9.

Dashed Line: Additivity Model, Solid Line: Hypothesized Mixture Model
 Triangle: Additivity Optimal Dose Location

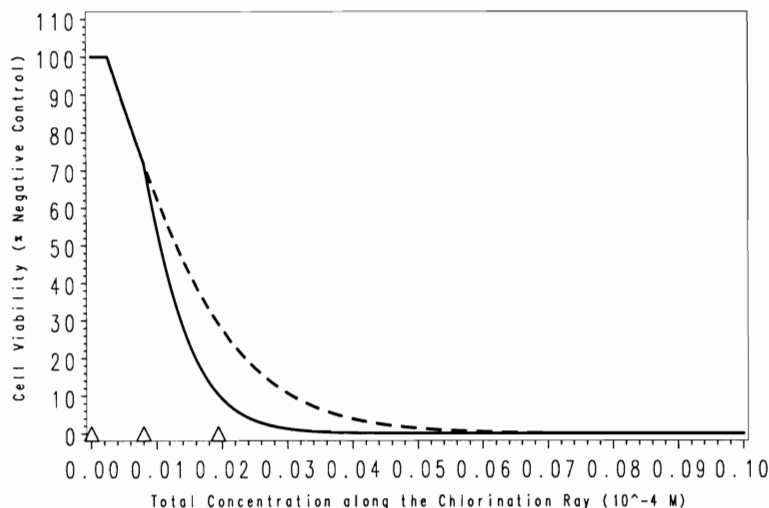


Figure 2.9. The Additivity Optimal Design in the General Synergism Case

In both cases, the first design point represents the control group, though it is not required by the optimal design. It is again interesting to note that the second design point is located at the hypothesized interaction threshold. The remaining design point is located in the interaction region, just before the response reaches a plateau at its minimum.

2.5 Conclusion

Ray designs are an efficient method for studying a mixture of a large number of chemicals that appear in a fixed mixing ratio. The dose and interaction thresholds model can be used to describe the relationship between the mixture, in terms of total dose, and the observed response. The hypothesis of additivity associated with this model is a nonlinear contrast of model parameters, which can be tested using either a Wald test or a quasi-likelihood ratio test. When the null hypothesis of additivity is rejected, the

investigator can make inference regarding the location of the departure from additivity. We can test the equality of the dose threshold and the interaction threshold; if the interaction threshold is significantly larger than the dose threshold, then a significant region of additivity can be claimed. The dose and interaction thresholds model is complicated in its parameterization, however, and it is important that the experimental design is sufficient to support the fit of the model. For example, if the hypothesized interaction region involves a quadratic term, the design must include at least three design points in that region in order to estimate the model parameters associated with interaction.

Statistical methodology can be used to find the optimal experimental design for supporting the dose and interaction thresholds model, as well as for testing the hypothesis of additivity. Because this model uses single chemical data to support the region of additivity, the optimal designs require support only in the area around the hypothesized interaction threshold and beyond. Regardless of the shape of the interaction region, the first required design point is located at the hypothesized interaction threshold. For quadratic interaction regions, a point just beyond the peak of the response within the interaction region is required, as well as a point just before the response achieves a plateau at its minimum. For linear antagonism regions, a point just before the response achieves a plateau at its minimum is all that is required by the optimal design. The same is true for linear synergism regions.

The optimal designs previously described included a control dose, though it was found to be unnecessary since it required a small proportion of the total sample size.

Investigators may prefer a design that includes more dose groups than those presented here, to allow for verification of the shape of the mixture dose-response curve. Augmenting the optimal designs, however, does not preserve optimal properties. Parker and Gennings (2006) developed a penalized optimal design criterion, which can be used to incorporate such verification groups into the design using desirability functions.

Chapter 3

The Flexible Single Chemical Required (FSCR) Method for Detecting Departure from Additivity Along a Fixed-Ratio Mixture Ray

3.1 Introduction

In the previous chapter, we found the optimal design for testing for departure from additivity using the dose and interaction thresholds model. In Chapter 4, we develop designs based on the Flexible Single Chemical Required (FSCR) method for detecting departure from additivity, which was developed by Gennings et al. (2004, 2006) and illustrated in Crofton et al. (2005). The example that serves as the motivation for this research was described in Gennings' et al. (2006). Gennings et al. found significant departure from additivity in a mixture of 18 polyhalogenated aromatic hydrocarbons (PHAHs). The FSCR interaction threshold model was fit to the fixed-ratio mixture ray, and the estimate of the interaction threshold was found to be positive. Gennings et al. verified the adequacy of the model fit by visually comparing the predicted model to the observed dose-response relationship. A quasi-likelihood ratio-based confidence interval was used to describe the location of the interaction threshold; this confidence interval included zero, so the authors were unable to conclude that a region of additivity exists along the ray.

Since Gennings et al. (2005) found no problem with the adequacy of the model fit, there are two possible explanations for the width of the confidence interval around the interaction threshold. The first explanation is that there is no interaction threshold along the fixed-ratio mixture ray. If this is the case, the authors' conclusion is satisfactory. The second, and perhaps more troublesome, explanation is that the study did not have sufficient power to detect a significant interaction threshold. If this is the case, the experimental data does not permit an adequate answer to the question at hand. In order to determine which explanation drives the inclusion of zero in the confidence interval around the interaction threshold, we can use statistical criteria to determine the optimal second stage design. These criteria will incorporate the data from, and the analysis of, the original experiment to find the optimal location and allocation of the second stage design points.

In order to understand the intricacies associated with developing optimal designs for the FSCR method of analysis, it is important to understand the FSCR method itself. The goals of Chapter 3 are to familiarize the reader with the FSCR method of detecting departure from additivity, developed by Gennings et al. (2004, 2006), and to describe in detail the analysis of the PHAH mixture presented in Gennings et al. (2006). For convenience, we will not continue to reference the FSCR papers of Gennings et al. (2004, 2006). However, it should be kept in mind that we present Chapter 3 not as new research, but as the basis for the new methodology presented in Chapter 4. The goal of Chapter 4 is to develop the optimal design for the support of the FSCR interaction threshold model described in Gennings et al. (2006).

3.2 Motivation

The usual methods for predicting additivity among a mixture of c agents rely on a single model that includes terms for each of the active single agents (e.g., Casey et al., 2004; Gennings et al., 2005). Such a model implies that certain features of the model are similar across all of the single agents. If a plateau parameter, for instance, is included in the model of the single agents, the usual additivity model requires that the true value of the plateau parameter is the same for each of the single agents. Often, however, this is not the case.

Our analysis of mixture data generally begins with a visual inspection of the data, by plotting the dose-response relationship exhibited by the raw data. This allows the researcher to make more informed decisions regarding the type of analysis to be performed – what form the model should take, whether additional nonlinear parameters should be included, etc. In the PHAH example provided in Section 3.3, it is evident from the graphs that a single plateau parameter is not sufficient to accurately describe the single agent data. The FSCR method allows the researcher to predict additivity while incorporating the different plateaus associated with the single agents.

3.3 Methodology Review

The goal of this section is to review the methodology associated with the FSCR method for detecting departure from additivity. In Section 3.3.1, we describe the model definitions and associated notation. In Section 3.3.1.1, the process for modeling the

single agent dose-response data and the mixture data, which constitutes the full model, is described. Based on the single agent models and Berenbaum's interaction index, the FSCR model for depicting additivity is described in Section 3.3.1.2, and the FSCR interaction threshold model is described in Section 3.3.1.3. Lastly, we discuss the complicated nature of the estimation process in Section 3.3.2 and describe the associated inference in Section 3.3.3.

3.3.1 Model Definitions

Consider the situation where the mean response can be represented by any nonlinear function of the form $\alpha + \gamma F(g(\mu))$, where $F(\bullet)$ is monotone, so that the inverse exists. For decreasing dose-response relationships, $F(\bullet)$ is any specified, decreasing, sigmoidal function that takes on values between zero and one. For increasing dose-response relationships, $F(\bullet)$ is any specified, increasing, sigmoidal function that takes on values between zero and one. As is common in dose-response modeling, we consider the quasi-likelihood approach, where it is assumed that $Var\{Y\} = \tau V(\mu)$.

3.3.1.1 The Full Model

The single agent models form the basis for the FSCR model. We begin by defining a model for each of the single agents and the fixed-ratio mixture ray, as follows. Let

$$\mu_i = \begin{cases} \alpha_i + \gamma_i, & x < \delta_i \\ \alpha_i + \gamma_i F(\beta_i(x - \delta_i)), & x \geq \delta_i \end{cases} \quad (3.1)$$

for $i=1, \dots, c+1$. Following this notation, $i=1, \dots, c$ represents the ray for each of the c single agents, and $i=c+1$ represents the fixed-ratio mixture ray under consideration. Along the i^{th} ray, the parameter β_i represents the slope, δ_i represents the unknown threshold, α_i represents the unknown minimum response, and γ_i represents the unknown range of the response.

3.3.1.2 The Dose Threshold Additivity Model

Recall that the basic dose threshold additivity model for a mixture of c agents consists of two segments. For decreasing dose-response relationships, the threshold additivity model is written as follows.

$$\mu_{add} = \begin{cases} \alpha + \gamma F(\beta_{01}), & \sum_{i=1}^c \beta_i x_i > \delta_{add} \\ \alpha + \gamma F\left(\beta_{02} + \sum_{i=1}^c \beta_i x_i\right), & \sum_{i=1}^c \beta_i x_i \leq \delta_{add} \end{cases}$$

To ensure that the additivity model is continuous at the dose threshold, δ_{add} , we set $\beta_{01} = \beta_{02} + \delta_{add}$. This continuity constraint requires that $\beta_{02} = \beta_{01} - \delta_{add}$, and the model then becomes

$$\mu_{add} = \begin{cases} \alpha + \gamma F(\beta_{01}), & \sum_{i=1}^c \beta_i x_i > \delta_{add} \\ \alpha + \gamma F\left(\beta_{01} + \sum_{i=1}^c \beta_i x_i - \delta_{add}\right), & \sum_{i=1}^c \beta_i x_i \leq \delta_{add} \end{cases}$$

This is the usual form of the threshold additivity model. An assumption that is implicit to this model, however, is that the nonlinear parameters required to describe the

dose-response relationship of the mixture, α and γ , are the same for each of the c single agents and the mixture. For instance, the above additivity model assumes that the minimum response is the same for all c agents and the mixture, ie $\alpha_1 = \alpha_2 = \dots = \alpha_c = \alpha$. If this assumption is not valid, then the above additivity model is not appropriate.

3.3.1.3 The FSCR Additivity Model

Recall, however, that the basic form of additivity is given by Berenbaum's interaction index. If equation (1.1) is true, if the interaction index equals one, then there is no interaction among the c agents in the mixture. Furthermore, based on the single agent models given in (3.1), we know the form of E_i , the concentration/dose of the i^{th} agent alone that yields a given response. Rewriting equation (3.1), we see that

$$E_i = \frac{F^{-1}\left(\frac{\mu_i - \alpha_i}{\gamma_i}\right)}{\beta_i} + \delta_i, \quad i = 1, \dots, c$$

is the concentration/dose of the i^{th} agent alone that yields response μ_i .

We can rewrite the interaction index in terms of the single agent model parameters, where μ is a given response of interest and is constant across the c single agents. Then

$$\sum_{i=1}^c \frac{x_i}{E_i} = \sum_{i=1}^c \frac{x_i}{\frac{F^{-1}\left(\frac{\mu - \alpha_i}{\gamma_i}\right)}{\beta_i} + \delta_i} = 1 \quad (3.2)$$

indicates an additive relationship among the c chemicals in the mixture. Recall, however, that we are interested in defining additivity along a fixed-ratio mixture ray; therefore, we

know that $x_i = a_i t$, where t is the total dose of the mixture along the fixed-ratio ray. As such, equation (3.2) can be rewritten

$$\sum_{i=1}^c \frac{a_i t}{\frac{F^{-1}\left(\frac{\mu - \alpha_i}{\gamma_i}\right) + \delta_i}{\beta_i}} = t \sum_{i=1}^c \frac{a_i}{\frac{F^{-1}\left(\frac{\mu - \alpha_i}{\gamma_i}\right) + \delta_i}{\beta_i}} = 1.$$

This equation represents the contour of constant response, which is planar under additivity. The above equation implies that the additive relationship between total dose and the mean response along a fixed-ratio mixture ray can be implicitly stated as follows.

For a specified value of the mean μ ,

$$t(\mu) = \left[\sum_{i=1}^c \frac{a_i}{\frac{F^{-1}\left(\frac{\mu - \alpha_i}{\gamma_i}\right) + \delta_i}{\beta_i}} \right]^{-1}$$

is the corresponding total dose under additivity. The general form of the FSCR additivity model, then, can be written in the following manner.

$$\mu_i = \begin{cases} \left. \begin{aligned} &\alpha_{i(add)} + \gamma_{i(add)}, & x_i < \delta_{i(add)} \\ &\alpha_{i(add)} + \gamma_{i(add)} F\left(\beta_{i(add)}(x - \delta_{i(add)})\right), & x_i \geq \delta_{i(add)} \end{aligned} \right\}, & i = 1, \dots, c \\ \mu_{(add)} \ni t_{(add)}(\mu_{(add)}) = \left[\frac{\sum_{i=1}^c \frac{a_i}{F^{-1}\left(\frac{\mu_{(add)} - \alpha_{i(add)}}{\gamma_{i(add)}}\right)} + \delta_{i(add)}}{\beta_{i(add)}} \right]^{-1}, & i = c + 1 \end{cases} \quad (3.3)$$

It is important to note that the subscripts on the slope and threshold parameters associated with the single agents have been modified from the full model. The subscript (*add*) indicates that these parameters are estimated under the constraint of additivity, using both single agent and fixed-ratio mixture ray data.

3.3.1.4 The FSCR Interaction Threshold Model

The FSCR interaction threshold model is an extension of the FSCR additivity model in equation (3.3), which divides the model into two segments. The model specifies that, for total doses smaller than the interaction threshold, the relationship among the agents in the mixture is that of additivity. This first segment is defined by the implicit function discussed in the previous section. Beyond the interaction threshold, the model allows for a departure from additivity. This segment is defined by an explicit function that describes the departure.

Along the fixed-ratio mixture ray, let Δ be the interaction threshold parameter, and let μ_Δ be the mean response at the interaction threshold as predicted under additivity.

Then the FSCR interaction threshold model is defined as follows.

$$\mu_i = \left\{ \begin{array}{l} \left. \begin{array}{l} \alpha_{i(add)} + \gamma_{i(add)}, \quad x_i < \delta_{i(add)} \\ \alpha_{i(add)} + \gamma_{i(add)} F\left(\beta_{i(add)}(x - \delta_{i(add)})\right), \quad x_i \geq \delta_{i(add)} \end{array} \right\} \quad i = 1, \dots, c \\ \left. \begin{array}{l} \left\{ \mu_{(add)} \ni t_{(add)}(\mu_{(add)}) = \left[\sum_{i=1}^c \frac{\alpha_i}{F^{-1}\left(\frac{\mu_{(add)} - \alpha_{i(add)}}{\gamma_{i(add)}}\right) + \delta_{i(add)}} \right]^{-1}, \quad t < \Delta \\ \alpha_i + (\mu_\Delta - \alpha_i) F(\beta_i(t - \Delta)), \quad t \geq \Delta \end{array} \right\} \quad i = c + 1 \end{array} \right\} \quad (3.4)$$

3.3.2 Estimation

Some discussion regarding the complexity of the implicit model may be helpful here, prior to the explanation of the estimation process. In an explicitly stated model, we can easily obtain the prediction for any total dose of interest. These predictions are required for parameter estimation, whether we use a least-squares approach or a quasi-likelihood approach. In an implicit model, such as the FSCR additivity model given in equation (3.3) and the additivity region of the FSCR interaction threshold model, however, we cannot arrive at the prediction for a given total dose simply by substituting into the model. Instead, the bisection algorithm is used to find the value of $\mu_{(add)}$ that is most closely associated with each observed mixture point.

To start the estimation process, the user must specify a set of starting values for the model parameters. A direct-search algorithm is used to search the parameter space for the parameter values that optimize the appropriate objective function. For each set of potential parameter values, mean predictions are required for each observed mixture point in order to calculate the objective function. Predictions for observed mixture points above the interaction threshold are obtained in the usual way, by substituting into equation (3.4). For observed mixture points below the interaction threshold, a bisection algorithm is used to find the value of $\mu_{(add)}$ that corresponds to each point. A more detailed description of the bisection algorithm is given in the following paragraph.

Let $t_{(add)}$ represent the total dose that is associated with $\mu_{(add)}$ under the assumption of additivity, and let t_{obs} represent the observed total dose along the fixed-ratio mixture ray. The bisection algorithm searches the values of $\mu_{(add)}$ in the response range and calculates the associated $t_{(add)}$. Each value $t_{(add)}$ is compared to the observed mixture point under consideration, t_{obs} . At such time that $|t_{obs} - t_{(add)}| < \varepsilon$, where ε is a small positive number, $\mu_{(add)}$ is considered to be the prediction for t_{obs} , the observed mixture point under consideration.

3.3.3 Inference

The primary goal of such an analysis is to detect departure from additivity along the fixed-ratio mixture ray. We can test the null hypothesis of additivity by comparing the

full model, defined in equation (3.1), to the FSCR additivity model, defined in equation (3.3), using a quasi-likelihood ratio test.

The properties of the quasi-likelihood ratio test were described in detail in Section 2.2.4. Let Q_{full} be the quasi-likelihood achieved under the full model, and let Q_{red} be the quasi-likelihood achieved under the FSCR additivity model. As before, M is the difference in the degrees of freedom between the FSCR full and additivity models. Recall that the quasi-likelihood ratio test statistic is

$$QLRT = \frac{-2\{Q_{red} - Q_{full}\}}{\hat{\tau}M}, \quad (3.5)$$

which has an approximate F distribution with M and $N-p$ degrees of freedom.

If the hypothesis test of additivity is significant, then the additivity model is not sufficient to describe the dose-response relationship along the fixed-ratio mixture ray. In fact, a significant test of additivity indicates that departure from additivity exists somewhere along the ray. When we conclude that there is an interaction among the agents, it is reasonable to question whether the interaction is dose-dependent. In this case, we fit the FSCR interaction threshold model, which allows for dose-dependent interaction among the agents in the mixture.

After the FSCR interaction threshold model has been fit to the data, it is of interest to determine where the departure from additivity exists. Recall that the interaction threshold is the dose beyond which the agents in the fixed-ratio mixture ray interact and is represented in the FSCR interaction threshold model by Δ . If the interaction threshold is significantly positive, then we can conclude that there is a significant region of

additivity. Therefore, the one-sided test of $H_0 : \Delta \leq 0$ is a test for a region of additivity along the mixture ray.

Alternatively, the test for a region of additivity can be conducted using the confidence interval associated with the interaction threshold parameter. Recall that the quasi-likelihood ratio test is statistically significant for

$$\frac{-2\{Q_{red} - Q_{full}\}}{\hat{t}M} \geq F_{M, N-p, 1-\alpha}.$$

All values of Δ_0 which do not result in a significant quasi-likelihood ratio test are plausible values of the interaction threshold.

Let $\boldsymbol{\theta}$ represent the $(p-1) \times 1$ vector of model parameters, excluding Δ , the interaction threshold parameter. Let $\hat{\boldsymbol{\theta}}$ and $\hat{\Delta}$ represent the maximum quasi-likelihood estimate of $\boldsymbol{\theta}$ and Δ , respectively, found by maximizing the quasi-likelihood with respect to all model parameters simultaneously. Let $\hat{\boldsymbol{\theta}}_{\Delta_0}$ represent the maximum quasi-likelihood estimate of $\boldsymbol{\theta}$ associated with a fixed value of the interaction threshold, Δ_0 . Using this notation, the quasi-likelihood ratio confidence interval around Δ is given by all Δ_0 such that

$$\left\{ \frac{-2\{Q_{red(\Delta_0, \hat{\boldsymbol{\theta}}_{\Delta_0})} - Q_{full(\hat{\Delta}, \hat{\boldsymbol{\theta}})}\}}{\hat{t}} < F_{1, N-p, 1-\alpha} \right\}, \quad (3.6)$$

since $M=1$. If the confidence interval around the interaction threshold does not include zero, then we can conclude that a significant region of additivity exists. Further, we can

conclude that the region of additivity is exceeded at the upper confidence bound of the interaction threshold.

Gennings et al. used quasi-likelihood ratio-based inference due to the complexity of the FSCR additivity and interaction threshold models. Because additivity is defined implicitly, the derivatives necessary to estimate a parameter covariance matrix, which is the basis for Wald-type inference, are not immediately available. An objective of Chapter 4 is to construct the parameter covariance matrix associated with the FSCR interaction threshold model.

3.4 Example

3.4.1 Background

Thyroid disrupting chemicals are “xenobiotics that alter the structure or function of the thyroid gland, alter regulatory enzymes associated with thyroid hormone homeostasis, or change circulating or tissue concentrations of thyroid hormones” (Crofton et al., 2005). One such class of thyroid disrupting chemicals, polyhalogenated aromatic hydrocarbons (PHAHs), induces enzymes which increase the elimination of thyroid hormones such as thyroxine (T4) from the body. Crofton et al. studied 18 PHAHs individually, as well as a fixed-ratio mixture based on environmental concentrations, to test for interaction among the chemicals with respect to the resulting decrease in serum T4 concentrations. The concentration of each chemical in the mixture was below the concentration that was found to have significant biological activity, with the exception of PCB126. The

concentration of PCB126 associated with the highest mixture dose is associated with an approximate 16% decrease in T4 (Crofton et al., 2005).

This example was illustrated by Gennings et al. (2006), and the following sections describe the analysis performed by Gennings et al. These sections are included to confirm those results and to familiarize the reader with the methods used. The primary research objective is to determine whether the response associated with the fixed-ratio mixture of 18 PHAHs departs from what is expected under additivity among the chemicals. If evidence of an interaction is found, a secondary objective is to determine whether the interaction is dose-dependent. The response of interest is thyroid function, as measured by serum total T4 as a percentage of control. The observed dose-response relationship for each of the single chemicals in the mixture is given in Appendix B1. Figure 3.1 is the plot of the observed dose-response relationship for the fixed-ratio mixture ray under study.

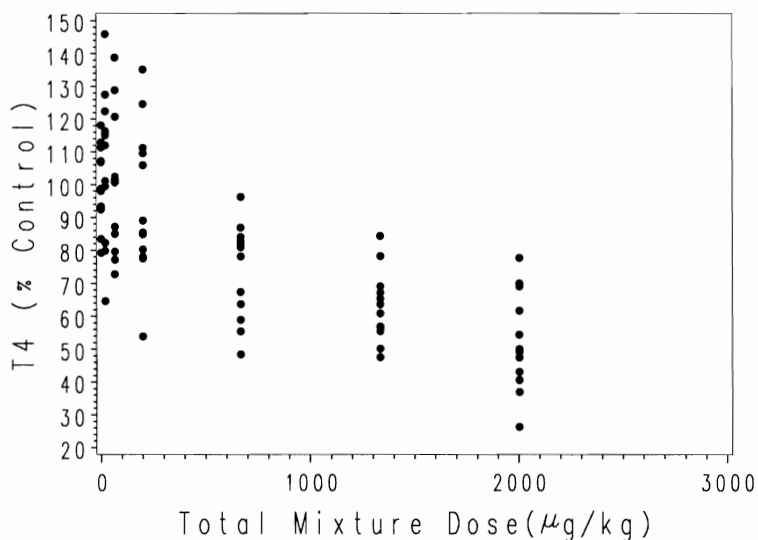


Figure 3.1. T4 vs. Total Dose of the Mixture Along the Ray

3.4.2 Model Selection

We chose to describe the convexity of the dose-response relationship evident in the plots using the exponential function, so that $F(\bullet) = \exp(\bullet)$. Additionally, because the T4 response was measured as a percent of control, it is reasonable to restrict the response to values less than 100. As such, we let $\gamma_i = 100 - \alpha_i$ represent the range of the response. As is often the case in dose-response modeling, the variance of the response changes as a function of the mean. To account for this non-constant variance in our response, quasi-likelihood methods were used to estimate parameters and establish model fit. We found that the form of the variance function that most adequately fit the data was $Var\{Y\} = \tau\mu$. This function resulted in an R^2 value of 0.68, indicating that 68% of the variability present in the response variance can be accounted for by this relationship. A plot of the observed and predicted variance versus the mean T4 is given in Figure 3.2.

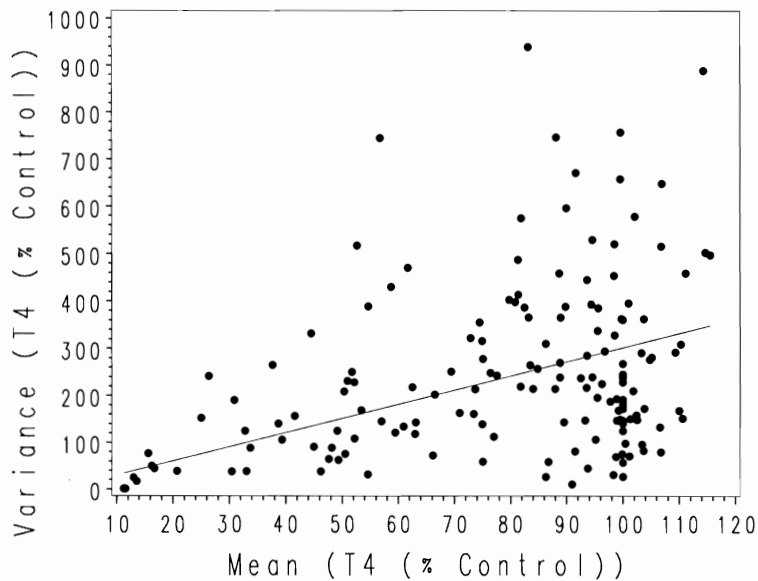


Figure 3.2 Predicted Variance vs. Mean

Following Gennings et al., the model given in equation (3.1) was fit simultaneously to all of the single chemical data and the mixture data in terms of total dose, with one exception. Slope, threshold, and minimum effect parameters for OCDF were not included in the full model, since OCDF was not dose-responsive (Gennings et al., 2006). The resulting parameter estimates are given in Appendix B2. Upon inspection, it can be seen that the agents can be grouped in terms of their minimum response. Agents 1, 2, 6, 12, 13, 16 and 17 appear to plateau in the same region of the response range, as do agents 3, 10, 11, 14 and 15, and agents 4, 5, 7, 8 and 9. Therefore, the minimum response parameters were combined in this manner in an effort to arrive at a more parsimonious model. The resulting parameter estimates are given in Appendix B3.

It is important to think through the meaning associated with the threshold parameter in the dose threshold model of the single agents. If there is no threshold ($\delta = 0$), then the model consists of a single segment which has intercept $\alpha + \gamma$. If the estimate of the threshold is positive, then the model consists of two segments. Both segments are contained in the positive dose range, and the intercept is $\alpha + \gamma$. If the estimate of the threshold is negative, however, only one segment of the model is represented in the positive dose range. The intercept in this case is something less than $\alpha + \gamma$, so that a negative threshold estimate corresponds to a downward shift in the intercept. Because the response is measured as a percent of control, it makes sense that we should require a 100% response at the control dose. Therefore, if a single agent results in a negative threshold estimate, the threshold is removed from the model to preserve a mean control response of 100.

A number of the dose thresholds have negative estimates. Note that the confidence interval on the dose threshold for agent 15, PCDD, is entirely negative, indicating that δ_{15} is negative. This threshold was removed, and the model was refit. At this stage in the model-building process, a number of the threshold estimates were again negative. However, the upper bound of each of the thresholds in question was in the experimental range, so they were not immediately removed from the model. Instead, we conducted an individual search of a grid of starting values for each threshold outside of the experimental range. The threshold that was most notably negative was removed from the model. We continued this process until all remaining thresholds had estimates that were within the experimental region. After the last grid search, two of the remaining threshold estimates remained negative, although the corresponding confidence intervals both included zero. As such, these estimates were also removed from the model.

The resulting model is the “full” model ($Q_{full} = 371577, df = 1308, \hat{\tau} = 3.33$), and the associated parameter estimates are given in Table 3.1. These parameter estimates agree with those described in Gennings et al. (2006). Plots demonstrating the fit of the full model to the single chemical data and the mixture data are given in Appendix B4. It is interesting to note that the confidence intervals around the minimum response parameters for the single chemicals do not overlap, which supports our decision to group the agents in this way. All of the slope estimates are statistically significant and negative, with the exception of the mixture slope and the slope for agent 17, TCDF. However, both the mixture and TCDF had positive estimates for the dose threshold. Therefore, these parameters were left in the model. All of the dose threshold estimates are positive,

though most are not statistically significant. We are able to conclude significant positive thresholds for agents 4 and 5, PCB105 and PCB118 respectively.

Table 3.1. Parameter Estimates Resulting from the Fit of the Full FSCR Model

Parameter	Estimate	Standard Error	p-value	Approximate 95% Confidence Limits	
$\alpha_{1_2_6_12_13_16_17}$	50.2556	1.4495	<0.0001	47.4120	53.0992
$\alpha_{3_10_11_14_15}$	30.7384	1.9501	<0.0001	26.9126	34.5642
$\alpha_{4_5_7_8_9}$	14.3276	1.0261	<0.0001	12.3146	16.3406
α_{mix}	42.2947	14.7759	0.0043	13.3070	71.2823
β_1	-0.0608	0.0164	0.0002	-0.0929	-0.0286
β_2	-0.0378	0.0142	0.008	-0.0657	-0.00989
β_3	-0.00012	0.000020	<0.0001	-0.00016	-0.00008
β_4	-0.00083	0.000171	<0.0001	-0.00117	-0.00050
β_5	-0.00070	0.000109	<0.0001	-0.00091	-0.00048
β_6	-0.7186	0.2249	0.0014	-1.1598	-0.2774
β_7	-0.00005	5.541E-6	<0.0001	-0.00006	-0.00004
β_8	-0.00003	3.444E-6	<0.0001	-0.00004	-0.00003
β_9	-0.00057	0.000061	<0.0001	-0.00069	-0.00045
β_{10}	-0.00262	0.000549	<0.0001	-0.00369	-0.00154
β_{11}	-0.00002	4.815E-6	<0.0001	-0.00003	-0.00001
β_{12}	-0.00001	2.525E-6	<0.0001	-0.00002	-7.19E-6
β_{13}	-0.00003	6.331E-6	<0.0001	-0.00004	-0.00002
β_{14}	-0.00067	0.000102	<0.0001	-0.00087	-0.00047
β_{15}	-0.3749	0.0639	<0.0001	-0.5003	-0.2494
β_{16}	-6.5054	2.4139	0.0071	-11.2411	-1.7697
β_{17}	-0.3312	0.2744	0.2277	-0.8695	0.2072
β_{mix}	-0.00087	0.000467	0.0622	-0.00179	0.000044
δ_1	0.3891	1.4995	0.7953	-2.5527	3.3309
δ_2	3.0937	3.9265	0.4309	-4.6094	10.7969
δ_3	76.4097	398.4	0.848	-705.3	858.1
δ_4	513.1	147.9	0.0005	223.0	803.3
δ_5	669.5	147.5	<0.0001	380.2	958.8

Parameter	Estimate	Standard Error	p-value	Approximate 95% Confidence Limits	
δ_6	0.0427	0.1295	0.7417	-0.2114	0.2968
δ_{10}	9.8125	25.1041	0.696	-39.4372	59.0621
δ_{11}	3227.3	5581.7	0.5632	-7723.0	14177.5
δ_{16}	0.00361	0.0197	0.8546	-0.0350	0.0423
δ_{17}	1.8587	0.9770	0.0573	-0.0580	3.7755
δ_{mix}	49.4723	74.3290	0.5058	-96.3473	195.3

The FSCR additivity model given in equation (3.3) was fit to both the single chemical and the mixture data simultaneously ($Q_{full} = 371547.1, df = 1311, \hat{\tau} = 3.34$). Figure 3.3 compares the fit of the FSCR additivity model to the mixture data with the fit of the FSCR full model to the mixture data. The associated parameter estimates are given in Table 3.2, and plots demonstrating the fit of the additivity model to the single chemical data and the mixture data are given in Appendix B5.

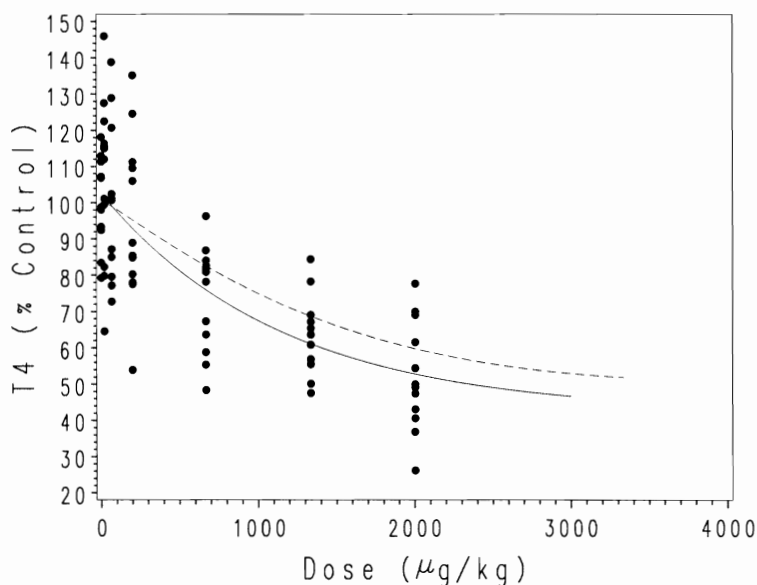


Figure 3.3. Fit of FSCR Full Model (solid line) and FSCR Additivity Model (dashed line) to Mixture Data Along the Fixed-Ratio Mixture Ray

Table 3.2. Parameter Estimates Resulting from the Fit of the FSCR Additivity Model

Parameter	Estimate
$\alpha_{1_2_6_12_13_16_17}$	50.8
$\alpha_{3_10_11_14_15}$	30.9
$\alpha_{4_5_7_8_9}$	14.3
β_1	-0.06027
β_2	-0.04032
β_3	-0.00012
β_4	-0.00083
β_5	-0.00069
β_6	-1.51304
β_7	-0.00005
β_8	-0.00003
β_9	-0.00056
β_{10}	-0.00257
β_{11}	-0.00002
β_{12}	-0.00001
β_{13}	-0.00003
β_{14}	-0.00068
β_{15}	-0.37213
β_{16}	-7.41317
β_{17}	-0.4366
δ_1	0.2
δ_2	3.0
δ_3	40.4
δ_4	498.9
δ_5	622.9
δ_6	0.02
δ_{10}	3.3
δ_{11}	3099.3
δ_{16}	0.003
δ_{17}	2.0

At this point in the analysis, it is reasonable to test the null hypothesis of additivity. If the test of additivity is significant, then the single chemical parameters under the assumption of additivity are not sufficient to describe the mixture response. In other words, a significant test of additivity indicates that there is significant departure from additivity along the ray. We construct the quasi-likelihood ratio test statistic according to equation (3.5), where $M=3$. The quasi-likelihood ratio test of additivity is statistically significant ($F_{3,1308} = 5.99$, p-value 0.00047). Therefore, we conclude that there is an interaction among the chemicals along the fixed-ratio mixture ray. Furthermore, Figure 3.3 indicates that the predictions in the low-dose region are similar for the FSCR full model and the FSCR additivity model. This could indicate the presence of an interaction threshold.

To determine whether the interaction among the chemicals is dose-dependent, the FSCR interaction threshold model, given in equation (3.4), was fit to the single chemical and the mixture data simultaneously ($Q_{full} = 371577.4$, $df = 1308$, $\hat{\tau} = 3.33$). The associated parameter estimates are given in Table 3.3, and plots demonstrating the fit of the FSCR interaction threshold model to the single chemical data are given in Appendix B6. Figure 3.4 compares the fit of the FSCR interaction threshold model to the mixture data with the fit of the FSCR full model to the mixture data. Figure 3.5 compares the fit of the FSCR interaction threshold model to the mixture data with what is predicted under additivity using the FSCR interaction threshold model.

Table 3.3. Parameter Estimates Resulting from the Fit of the FSCR Interaction Threshold Model

Parameter	Estimate
$\alpha_{1_2_6_12_13_16_17}$	50.4
$\alpha_{3_10_11_14_15}$	30.9
$\alpha_{4_5_7_8_9}$	14.2
β_1	-0.05927
β_2	-0.04015
β_3	-0.00012
β_4	-0.00081
β_5	-0.00069
β_6	-0.70073
β_7	-0.00005
β_8	-0.00003
β_9	-0.00056
β_{10}	-0.00254
β_{11}	-0.00002
β_{12}	-0.00001
β_{13}	-0.00003
β_{14}	-0.00068
β_{15}	-0.37296
β_{16}	-6.39152
β_{17}	-0.49315
δ_1	0.2
δ_2	3.0
δ_3	41.7
δ_4	495.4
δ_5	665.1
δ_6	0.02
δ_{10}	3.4
δ_{11}	3015.0
δ_{16}	0.004
δ_{17}	2.2

Parameter	Estimate
α_{mix}	44.9
β_{mix}	-0.00099
Δ	105.3

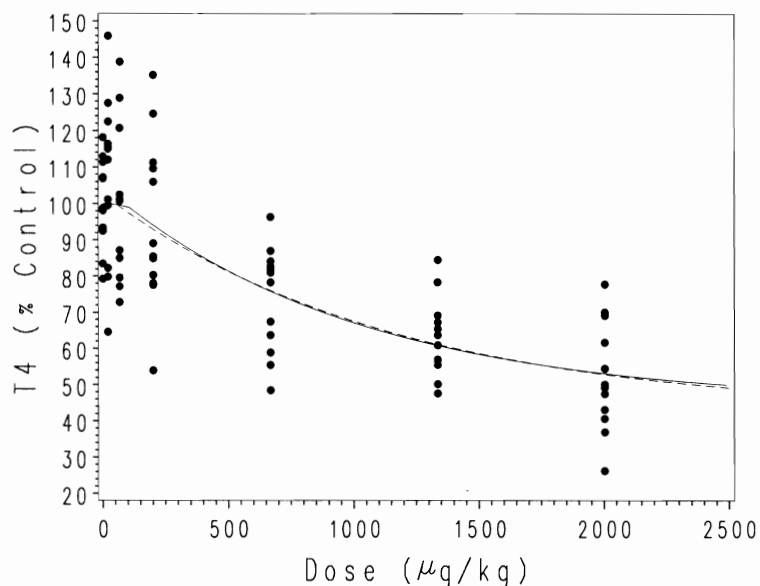


Figure 3.4. Fit of FSCR Full Model (dashed line) and FSCR Interaction Threshold Model (solid line) to Mixture Data Along the Fixed-Ratio Mixture Ray

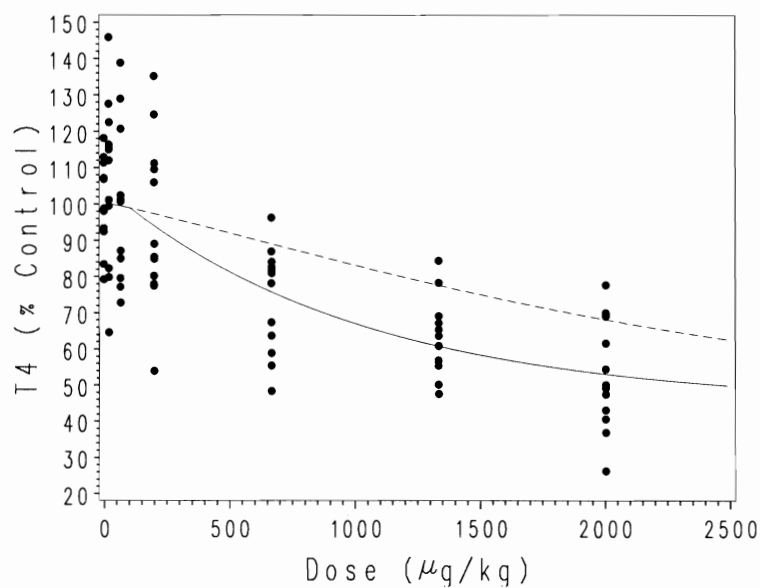


Figure 3.5. Fit of FSCR Interaction Threshold Model (solid line) and FSCR Additivity Model (dashed line) to Mixture Data Along the Fixed-Ratio Mixture Ray

Agents 1, 2, 6, 12, 13, 16 and 17 plateau at a minimum response of 50.4 % of control. Agents 3, 10, 11, 14 and 15 plateau at a minimum response of 30.9 % of control. Agents 4, 5, 7, 8 and 9 plateau at a minimum response of 14.2 % of control. The mixture ray plateaus at a minimum response of 44.9 % of control. Each of the slope parameters is negative, as we would expect from the single agent and mixture scatter plots. The estimate of Δ , the dose beyond which the chemicals in the mixture interact, is 105.3 $\mu\text{g}/\text{kg}$.

3.4.3 Inference

Recall that the primary research objective was to determine whether the 18 chemicals in the mixture interact when combined along the fixed-ratio mixture ray. The quasi-likelihood ratio test of additivity was conducted according to equation (3.5) and was statistically significant ($F_{3,1308} = 5.99$, p-value 0.00047). Therefore, we can conclude that there is significant departure from additivity along the fixed-ratio mixture ray.

Figure 3.5 demonstrates that the FSCR interaction threshold model falls below the FSCR additivity model. That is, the mixture results in a greater decrease in T4 than is predicted using only the single chemicals. Therefore, we describe the interaction among the chemicals along the fixed-ratio mixture ray as a synergism; the mixture of chemicals works more efficiently than the individual chemicals predict to decrease T4 as a percent of control. However, the estimate of the interaction threshold is positive, which is some evidence that this interaction may be dose-dependent. In other words, there may be a region of additivity along the ray.

The test for a region of additivity is based on the quasi-likelihood ratio confidence interval around Δ , the interaction threshold. If the confidence interval does not include zero, then we can conclude that a significant region of additivity exists along the fixed-ratio mixture ray. The quasi-likelihood ratio confidence interval around the interaction threshold was constructed using Nelder-Mead to estimate model parameters and the bisection algorithm to find the endpoints of the interval. The approximate 95% confidence interval is (0,544) $\mu\text{g}/\text{kg}$. Since the confidence interval includes zero, there is insufficient evidence for us to conclude that a region of additivity exists along the fixed-ratio mixture ray.

3.4.4 Design Consideration

We have seen that the mixture experiment performed was sufficient to support the complex FSCR interaction threshold model and to demonstrate departure from additivity. However, the data did not support a significant region of additivity, as evidenced by the inclusion of zero in the quasi-likelihood ratio confidence interval around the interaction threshold. Because the interaction threshold was not statistically significant, our secondary goal has not been met. There are two possible reasons for the lack of significance of the interaction threshold. The first possibility, of course, is that there is no interaction threshold, so that the departure from additivity exists everywhere along the fixed-ratio mixture ray. The second, and perhaps more troublesome, possibility is that the design of the mixture experiment was insufficient to precisely estimate the interaction threshold, and hence to demonstrate a region of additivity.

If it is not sufficient to report that departure from additivity exists somewhere along the ray, if it is also of importance to know whether a region of additivity exists along the ray, what options are available to the researcher? Certainly the information obtained from the FSCR interaction threshold analysis can be used to design and conduct a new mixture experiment. It is probably not feasible, however, to conduct an entirely new mixture experiment, without making use of the information available from the original experiment. Instead, we are interested in supplementing the original mixture experiment with additional design points, strategically placed to support the precise estimation of the interaction threshold. Model derivatives are not easily obtained, so a D- or Ds- optimal design is a complex, but possible, solution.

The following chapter is primarily concerned with overcoming the problems associated with finding the parameter covariance matrix associated with the FSCR interaction threshold model. The parameter covariance matrix will be used to develop optimal designs to support the FSCR interaction threshold model. We will place additional design points to supplement the original mixture experiment, in an effort to obtain the most precise estimate of the interaction threshold possible. Designs based on a Ds-optimality criterion will be considered. When necessary, the desirability of a given design may be considered to ensure that the resulting designs are of practical use.

Chapter 4

Optimal Design for the Flexible Single Chemical Required (FSCR)

Interaction Threshold Model

4.1 Introduction and Motivation

Recall that the usual methods for predicting additivity rely on a single model that includes terms for each of the active single agents. The usual additivity model is assumed to adequately represent the single agent data, as it is often used to describe the dose-response relationships of the single agents. Such a model implies that certain features of the model are similar across all of the single agents. If a plateau parameter, for instance, is included in the model of the single agents, the usual additivity model requires that the true value of the plateau parameter is the same for each of the single agents. Often, however, this is not the case.

To allow more flexibility in adequately modeling the dose-response relationships of the single agents, Gennings et al. (2004) developed the Flexible Single Chemical Required (FSCR) method. The basis of the FSCR method is the adequate representation of the single agent dose-response relationships. Additivity is then defined according to Berenbaum's interaction index, which can be represented using the parameters associated with the single agent dose-response relationships. The definition of additivity in the

FSCR methods is equivalent to that in the Single Chemical Required methods; however, the FSCR methods allow for the estimation of different range and threshold parameters for each of the single agents and the fixed-ratio mixture ray. Gennings et al. demonstrated the FSCR additivity model for a mixture of six chemicals (2004) and demonstrated the FSCR interaction threshold model for a mixture of eighteen PHAHs (2006).

As shown in Section 3.2, the statement of the FSCR interaction threshold model, as well as the FSCR additivity model, is complex. As a result, we want to consider employing statistical optimality criteria to help us determine the optimal experimental design for supporting these models. We can use a number of statistical design criteria to find the optimal design, in terms of both the location of the design points and the proportion of subjects allocated to each design point. A commonly used design criterion is the D-optimality criterion, which is based on the parameter covariance matrix. The calculation of the estimated parameter covariance matrix uses the derivatives of the model with respect to the model parameters. Recall, however, that the active region of additivity in the FSCR models is implicitly stated. These models cannot be stated such that the mean is an explicit function of the data; therefore, we cannot take derivatives with respect to the model parameters in the usual way. It follows that the construction of the parameter covariance matrix for the FSCR models is an obstacle to employing an alphabetic optimality criterion.

The goal of this chapter is to develop a method for constructing the parameter covariance matrix, so that we can determine the optimal experimental design for

supporting the FSCR interaction threshold model. In Section 4.2, we review the mixture of 18 PHAHs described in Chapter 3, which will serve to illustrate our design work, and address the stability of the parameter covariance matrix associated with the full model. In section 4.3, we develop the process for constructing model derivatives. In section 4.4, we present the experimental design that is optimal for improving the precision of the interaction threshold estimate. The practicality of the optimal experimental design is addressed in section 4.5, where we consider the penalized optimal design criterion developed by Parker and Gennings (2006).

4.2 Motivating Example: The Estimation of an Interaction Threshold in a Mixture of Eighteen PHAHs

In the previous chapter, we described an analysis aimed at detecting departure from additivity among 18 PHAHs combined according to a fixed-ratio mixture ray. A secondary goal of the analysis was to determine whether the departure from additivity, if significant, was dose-dependent. The design of the experiment was sufficient to support the complicated FSCR interaction threshold model, and we were able to show that significant departure from additivity exists along the ray. However, although the estimate of the interaction threshold was positive, the quasi-likelihood ratio confidence interval includes zero. Therefore, there is insufficient evidence for us to conclude that the presence of the interaction is dose-dependent.

Since the model adequately describes the dose-response relationship of the mixture, there are two possible explanations for the width of the confidence interval. One

explanation is that the true interaction threshold is zero, which means that the interaction along the fixed-ratio mixture ray is not dose-dependent. The other explanation is more problematic. It is possible that a positive interaction threshold does exist along the ray, but the experiment performed was not sufficient to precisely estimate it. The lack of significance of the interaction threshold may be due to a lack of precision, rather than a lack of significance. If this is the case, then our second goal has not been met; the data do not permit an adequate answer to the question at hand.

In the example presented in Chapter 3, we are without a true answer to the question of the dose-dependent nature of the interaction. We can say that we don't have enough evidence to conclude that the interaction is dose-dependent. There is some evidence, however, since the estimate of the interaction threshold is positive. Therefore, we are interested in increasing the precision associated with our estimate of the interaction threshold by adding design points to the original experiment. We want to put these additional design points in a location such that they optimize the improvement to our precision.

For ease of discussion, consider a nonlinear exponential threshold model, as described in Gennings et al. (2006). The full model used to describe each of the 18 single agents and the fixed-ratio mixture ray is as described in Chapter 3,

$$\mu_i = \begin{cases} 100, & x < \delta_i \\ \alpha_i + (100 - \alpha_i) \exp\{\beta_i(x - \delta_i)\} & x \geq \delta_i \end{cases} \quad (4.1)$$

where $i=1, \dots, c+1$. Following this notation, $i=1, \dots, c$ represents the ray for each of the c single agents, and $i=c+1$ represents the fixed-ratio mixture ray under consideration.

This model was fit to the single chemical and the mixture data simultaneously. Let θ represent the vector of full model parameters, as described in Chapter 3. Since the full model is stated explicitly, the parameter covariance matrix, $Var\{\hat{\theta}\}$, can be easily obtained. We found, however, that the parameter covariance matrix resulting from the fit of the full model was unstable, in that $|Var\{\hat{\theta}\}|$ was approximately equal to zero.

In Chapter 2, we addressed this issue by modifying the units associated with the single chemical data, which improves the stability of the parameter covariance matrix here as well. Therefore, it was determined that the instability of the parameter covariance matrix was not due to over-parameterization of the full model but to the small variances associated with the parameter estimates. The fit of the FSCR full model resulted in standard errors less than 10^{-4} for nine slope parameters of the 17 active agents. In this example, however, the dose scale varies widely across the 18 single chemicals. Experimental doses for PCB105, for instance, range from 0 to 100,000 $\mu\text{g}/\text{kg}$, while experimental doses for PCDD range only from 0 to 10 $\mu\text{g}/\text{kg}$. Changing the dose scale increases the variances associated with the slope parameters but decreases the variances associated with the threshold parameters. Therefore, in order to stabilize the estimate of $Var\{\hat{\theta}\}$, we found it useful to reparameterize the model, rather than change the dose scale. The scaled FSCR full model includes scaled slope parameters for the single agents and the fixed-ratio mixture ray as necessary.

4.3 Construction of the Parameter Covariance Matrix for the Scaled FSCR Interaction Threshold Model

Without loss of generality, let the scaled FSCR full model be described as follows,

$$\mu_i = \begin{cases} 100, & x < \delta_i \\ \alpha_i + (100 - \alpha_i) \exp\left\{\frac{\beta_i}{SF_i}(x - \delta_i)\right\}, & x \geq \delta_i \end{cases} \quad (4.2)$$

for $i=1, \dots, c+1$. The scaling factor is represented by SF_i , which is determined separately for each of the c single agents and the mixture ray. Note that, under the scaled model, β_i is not the slope of the i^{th} agent but is associated with the slope. If the dose scales are similar across the c chemicals, such that the scaling factor is unnecessary, let SF_i equal one for all i . In this case, the scaled FSCR full model given in equation (4.2) reduces to the FSCR full model given in equation (4.1), and β_i again represents the slope of the i^{th} agent. Scaling the slope parameter does not affect the prediction of the model; it reduces the range of the variance associated with the slope parameters.

The scaled FSCR additivity model can be written as follows,

$$\mu = \begin{cases} \begin{cases} 100, & x < \delta_i \\ \alpha_i + (100 - \alpha_i) \exp\left\{\frac{\beta_i}{SF_i}(x - \delta_i)\right\}, & x \geq \delta_i \end{cases}, & i = 1, \dots, c \\ \mu_{(add)} \ni t_{(add)}(\mu_{(add)}) = \left[\sum_{i=1}^c \frac{\alpha_i}{\frac{SF_i}{\beta_{i(add)}} \log\left(\frac{\mu_{(add)} - \alpha_{i(add)}}{\gamma_{i(add)}}\right) + \delta_{i(add)}} \right]^{-1}, & i = c + 1 \end{cases}$$

where SF_i is as defined above. Then the scaled FSCR interaction threshold model is written as follows.

$$\mu = \begin{cases} \left. \begin{aligned} &100, & x < \delta_i \\ &\alpha_i + (100 - \alpha_i) \exp \left\{ \frac{\beta_i}{SF_i} (x - \delta_i) \right\}, & x \geq \delta_i \end{aligned} \right\}, & i = 1, \dots, c \\ \left. \begin{aligned} &\mu_{(add)} \ni t_{(add)}(\mu_{(add)}) = \left[\sum_{i=1}^c \frac{\alpha_i}{\frac{SF_i}{\beta_{i(add)}} \log \left(\frac{\mu_{(add)} - \alpha_{i(add)}}{\gamma_{i(add)}} \right) + \delta_{i(add)}} \right]^{-1}, & t < \Delta \\ &\alpha_{mix} + (\mu_{\Delta} - \alpha_{mix}) \exp \{ \beta_{mix} (t - \Delta) \}, & t \geq \Delta \end{aligned} \right\}, & i = c + 1 \end{cases} \quad (4.3)$$

Let $\boldsymbol{\varphi} = [\boldsymbol{\theta} \ \Delta]$ represent the $p \times 1$ vector of model parameters, where $\boldsymbol{\theta}$ represents the $(p-1) \times 1$ vector of model parameters excluding Δ , as described in Chapter 3. Let the matrix \mathbf{F} represent the matrix of derivatives of the mean response with respect to the model parameters, such that $\mathbf{F} = \left[\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\varphi}} \right]$ is an $n \times p$ matrix. Then the asymptotic covariance matrix (McCullagh and Nelder, 1989) is

$$\text{Var} \{ \hat{\boldsymbol{\varphi}} \} = \tau (\mathbf{F}' \mathbf{V}^{-1} \mathbf{F})^{-1}. \quad (4.4)$$

The variance associated with the interaction threshold is given by the $(p,p)^{\text{th}}$ element of the parameter covariance matrix, $\text{Var} \{ \hat{\boldsymbol{\varphi}} \}$.

To compute the covariance matrix for the parameter estimates, then, we must be able to take first derivatives of the model with respect to the model parameters. In the case of the single agents and mixture doses beyond the interaction threshold, for which the model

is explicitly stated, we can find these derivatives in the usual way. The additivity region of the model, however, is an implicit function, so we have to take implicit derivatives of the model with respect to the model parameters.

In section 4.3.1, we develop these implicit derivatives for a mixture of two agents. In section 4.3.2, we describe how the derivatives are affected by simplification of the minimum response parameters. Complete model derivatives are given in Appendix C1.

4.3.1 Differentiation of the Scaled FSCR Additivity Model for Two Agents

Without loss of generality, consider the situation where $c=2$, so that, for $t < \Delta$, the additivity region of the model is written as follows.

$$\mu_{add} = \left\{ \mu \ni t(\mu) = \left[\sum_{i=1}^2 \frac{a_i}{\frac{SF_i}{\beta_i} \ln \left(\frac{\mu - \alpha_i}{(100 - \alpha_i)} \right) + \delta_i} \right]^{-1} \right\}$$

$$= \left\{ \mu \ni t(\mu) = \left[\frac{a_1}{\frac{SF_1}{\beta_1} \ln \left(\frac{\mu - \alpha_1}{(100 - \alpha_1)} \right) + \delta_1} + \frac{a_2}{\frac{SF_2}{\beta_2} \ln \left(\frac{\mu - \alpha_2}{(100 - \alpha_2)} \right) + \delta_2} \right]^{-1} \right\}$$

Let $\xi_1 = \frac{SF_1}{\beta_1} \ln \left(\frac{\mu - \alpha_1}{(100 - \alpha_1)} \right) + \delta_1$ and $\xi_2 = \frac{SF_2}{\beta_2} \ln \left(\frac{\mu - \alpha_2}{(100 - \alpha_2)} \right) + \delta_2$. Now we can rewrite

the additivity region of the model in the following manner.

$$\begin{aligned}\mu_{add} &= \left\{ \mu \ni t(\mu) = \left[\frac{a_1}{\xi_1} + \frac{a_2}{\xi_2} \right]^{-1} \right\} \\ &= \left\{ \mu \ni t(\mu) = \left[a_1 \xi_1^{-1} + a_2 \xi_2^{-1} \right]^{-1} \right\}\end{aligned}\quad (4.5)$$

Following Finney (1994), we can take the derivative of both sides of the implicit function with respect to each of the model parameters. The result of this implicit differentiation is that we obtain equations that can be solved for the derivative of the mean with respect to the model parameters.

To find the derivative of the mean with respect to α_1 , we implicitly differentiate both sides of equation (4.5) with respect to α_1 , as follows.

$$\begin{aligned}\frac{\partial t(\mu)}{\partial \alpha_1} &= \frac{\partial \left[a_1 \xi_1^{-1} + a_2 \xi_2^{-1} \right]^{-1}}{\partial \alpha_1} \\ 0 &= - \left[a_1 \xi_1^{-1} + a_2 \xi_2^{-1} \right]^{-2} \left[\left(-a_1 \xi_1^{-2} \right) \frac{\partial \xi_1}{\partial \alpha_1} - a_2 \xi_2^{-2} \frac{\partial \xi_2}{\partial \alpha_1} \right]\end{aligned}\quad (4.6)$$

To complete this equation, we can differentiate ξ_1 and ξ_2 with respect to α_1 in the following manner.

$$\frac{\partial \xi_1}{\partial \alpha_1} = \frac{SF_1}{\beta_1} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) * \left[(\mu - \alpha_1) \left((100 - \alpha_1)^{-2} \right) + \left((100 - \alpha_1)^{-1} \right) \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \right]$$

As defined, ξ_i is a function of only α_i , β_i , and δ_i . Because ξ_i represents the $ED_i(\mu)$, the μ contained within ξ_i is associated only with the i^{th} chemical. Therefore $\frac{\partial \xi_2}{\partial \alpha_1} = 0$.

We can substitute these derivatives into (4.6) and solve the resulting equation for $\frac{\partial \mu_{add}}{\partial \alpha_1}$

as follows.

$$\begin{aligned}
0 &= -[a_1\xi_1^{-1} + a_2\xi_2^{-1}]^{-2} \left[(-a_1\xi_1^{-2}) \frac{\partial \xi_1}{\partial \alpha_1} - a_2\xi_2^{-2} \underbrace{\frac{\partial \xi_2}{\partial \alpha_1}}_0 \right] \\
0 &= -[a_1\xi_1^{-1} + a_2\xi_2^{-1}]^{-2} \left[(-a_1\xi_1^{-2}) \left\{ \frac{SF_1}{\beta_1} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) \left[(\mu - \alpha_1)((100 - \alpha_1)^{-2}) + ((100 - \alpha_1)^{-1}) \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \right] \right\} \right] \\
0 &= \left[(-a_1\xi_1^{-2}) \left\{ \frac{SF_1}{\beta_1} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) \left[(\mu - \alpha_1)((100 - \alpha_1)^{-2}) + ((100 - \alpha_1)^{-1}) \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \right] \right\} \right] \\
0 &= \left[(\mu - \alpha_1)((100 - \alpha_1)^{-2}) + ((100 - \alpha_1)^{-1}) \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \right] \Rightarrow \\
&\quad -(\mu - \alpha_1)((100 - \alpha_1)^{-2}) = ((100 - \alpha_1)^{-1}) \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \\
-(\mu - \alpha_1)((100 - \alpha_1)^{-2})((100 - \alpha_1)^{+1}) &= \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \Rightarrow \\
1 - (\mu - \alpha_1)((100 - \alpha_1)^{-1}) &= \frac{\partial \mu_{add}}{\partial \alpha_1} \tag{4.7}
\end{aligned}$$

A similar result can be shown for $\frac{\partial \mu_{add}}{\partial \alpha_2}$.

To find the derivative of the mean with respect to β_1 , we can implicitly differentiate both sides of the implicit additivity function with respect to β_1 , as follows.

$$\begin{aligned} \frac{\partial t(\mu)}{\partial \beta_1} &= \frac{\partial [a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-1}}{\partial \beta_1} \\ 0 &= -[a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-2} \left[(-a_1 \xi_1^{-2}) \frac{\partial \xi_1}{\partial \beta_1} - a_2 \xi_2^{-2} \frac{\partial \xi_2}{\partial \beta_1} \right] \end{aligned} \quad (4.8)$$

To complete this equation, we can differentiate ξ_1 and ξ_2 with respect to β_1 in the following manner.

$$\begin{aligned} \frac{\partial \xi_1}{\partial \beta_1} &= \frac{SF_1}{\beta_1} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) \left(\frac{1}{100 - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \beta_1} + (-SF_1 \beta_1^{-2}) \ln \left(\frac{\mu - \alpha_1}{100 - \alpha_1} \right) \\ &= \frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \beta_1} + (-SF_1 \beta_1^{-2}) \ln \left(\frac{\mu - \alpha_1}{100 - \alpha_1} \right) \\ \frac{\partial \xi_2}{\partial \beta_1} &= 0 \end{aligned}$$

We can substitute these derivatives into equation (4.8) and solve the resulting equation for $\frac{\partial \mu_{add}}{\partial \beta_1}$ as follows.

$$\begin{aligned} 0 &= -[a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-2} \left[(-a_1 \xi_1^{-2}) \frac{\partial \xi_1}{\partial \beta_1} - a_2 \xi_2^{-2} \frac{\partial \xi_2}{\partial \beta_1} \right] \\ 0 &= \left[(-a_1 \xi_1^{-2}) \left(\frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \beta_1} + (-SF_1 \beta_1^{-2}) \ln \left(\frac{\mu - \alpha_1}{100 - \alpha_1} \right) \right) \right] \\ 0 &= \frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \beta_1} + (-SF_1 \beta_1^{-2}) \ln \left(\frac{\mu - \alpha_1}{100 - \alpha_1} \right) \Rightarrow \end{aligned}$$

$$\begin{aligned}
(SF_1\beta_1^{-2})\ln\left(\frac{\mu-\alpha_1}{100-\alpha_1}\right) &= \frac{SF_1}{\beta_1}\left(\frac{1}{\mu-\alpha_1}\right)\frac{\partial\mu_{add}}{\partial\beta_1} \\
\frac{\cancel{SF_1}}{\beta_1^2}\ln\left(\frac{\mu-\alpha_1}{100-\alpha_1}\right)\frac{(\mu-\alpha_1)\cancel{\beta_1}}{\cancel{SF_1}} &= \frac{\partial\mu_{add}}{\partial\beta_1} \\
\frac{(\mu-\alpha_1)}{\beta_1}\ln\left(\frac{\mu-\alpha_1}{100-\alpha_1}\right) &= \frac{\partial\mu_{add}}{\partial\beta_1} \tag{4.9}
\end{aligned}$$

A similar result can be shown for $\frac{\partial\mu_{add}}{\partial\beta_2}$.

To find the derivative of the mean with respect to δ_1 , we can implicitly differentiate both sides of the implicit additivity function with respect to δ_1 as follows.

$$\begin{aligned}
\frac{\partial t(\mu)}{\partial\delta_1} &= \frac{\partial[a_1\xi_1^{-1} + a_2\xi_2^{-1}]^{-1}}{\partial\delta_1} \\
0 &= -[a_1\xi_1^{-1} + a_2\xi_2^{-1}]^{-2} \left[(-a_1\xi_1^{-2})\frac{\partial\xi_1}{\partial\delta_1} - a_2\xi_2^{-2}\frac{\partial\xi_2}{\partial\delta_1} \right] \tag{4.10}
\end{aligned}$$

To complete this equation, we can differentiate ξ_1 and ξ_2 with respect to δ_1 in the following manner.

$$\begin{aligned}
\frac{\partial\xi_1}{\partial\delta_1} &= \frac{SF_1}{\beta_1}\left(\frac{\cancel{100-\alpha_1}}{\mu-\alpha_1}\right)\left(\frac{1}{\cancel{100-\alpha_1}}\right)\frac{\partial\mu_{add}}{\partial\delta_1} + 1 \\
&= \frac{SF_1}{\beta_1}\left(\frac{1}{\mu-\alpha_1}\right)\frac{\partial\mu_{add}}{\partial\delta_1} + 1 \\
\frac{\partial\xi_2}{\partial\delta_1} &= 0
\end{aligned}$$

We can substitute these derivatives into equation (4.10) and solve the resulting equation

for $\frac{\partial\mu_{add}}{\partial\delta_1}$.

$$\begin{aligned}
0 &= -[a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-2} \left[(-a_1 \xi_1^{-2}) \frac{\partial \xi_1}{\partial \delta_1} - a_2 \xi_2^{-2} \underbrace{\frac{\partial \xi_2}{\partial \delta_1}}_0 \right] \\
0 &= -[a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-2} \left[(-a_1 \xi_1^{-2}) \left(\frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \delta_1} + 1 \right) \right] \\
0 &= (-a_1 \xi_1^{-2}) \left(\frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \delta_1} + 1 \right) \\
0 &= \frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \delta_1} + 1 \\
-1 &= \frac{SF_1}{\beta_1} \left(\frac{1}{\mu - \alpha_1} \right) \frac{\partial \mu_{add}}{\partial \delta_1} \\
-\frac{(\mu - \alpha_1) \beta_1}{SF_1} &= \frac{\partial \mu_{add}}{\partial \delta_1} \tag{4.11}
\end{aligned}$$

A similar result can be shown for $\frac{\partial \mu_{add}}{\partial \delta_2}$.

4.3.2 Differentiation of the Scaled FSCR Interaction Threshold Model with Combined Minimum Response Parameters

The derivatives specified in the previous section stem from the case where each single agent has its own α , β , and δ parameters. If suggested by the data, however, the minimum response parameters can be combined to arrive at a more parsimonious model.

In the case where $c=2$, consider that $\alpha_2 = \alpha_1$. Then we can rewrite

$$\xi_2 = \frac{SF_2}{\beta_2} \ln \left(\frac{\mu - \alpha_1}{(100 - \alpha_1)} \right) + \delta_2,$$

so that when we implicitly differentiate ξ_1 and ξ_2 with respect to α_1 , we get the following. As before,

$$\begin{aligned} \frac{\partial \xi_1}{\partial \alpha_1} &= \frac{SF_1}{\beta_1} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) \left[(\mu - \alpha_1) \left((100 - \alpha_1)^{-2} \right) + \left((100 - \alpha_1)^{-1} \right) \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right) \right] \\ &= \frac{SF_1}{\beta_1} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) \left[\frac{\mu - \alpha_1}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right)}{(100 - \alpha_1)} \right], \end{aligned}$$

and since ξ_2 is also a function of α_1 ,

$$\frac{\partial \xi_2}{\partial \alpha_1} = \frac{SF_2}{\beta_2} \left(\frac{100 - \alpha_1}{\mu - \alpha_1} \right) \left[\frac{\mu - \alpha_1}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1 \right)}{(100 - \alpha_1)} \right].$$

We can substitute these derivatives into equation (4.6) and solve the resulting equation

for $\frac{\partial \mu_{add}}{\partial \alpha_1}$ as follows.

$$\begin{aligned}
0 &= -[a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-2} \left[(-a_1 \xi_1^{-2}) \frac{\partial \xi_1}{\partial \alpha_1} - a_2 \xi_2^{-2} \frac{\partial \xi_2}{\partial \alpha_1} \right] \\
0 &= -[a_1 \xi_1^{-1} + a_2 \xi_2^{-1}]^{-2} \left[\begin{aligned} & \left(-a_1 \xi_1^{-2} \right) \left\{ \frac{SF_1 (100 - \alpha_1)}{\beta_1 (\mu - \alpha_1)} \left[\frac{(\mu - \alpha_1)}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add} - 1}{\partial \alpha_1} \right)}{(100 - \alpha_1)} \right] \right\} \\ & - a_2 \xi_2^{-2} \left\{ \frac{SF_2 (100 - \alpha_1)}{\beta_2 (\mu - \alpha_1)} \left[\frac{\mu - \alpha_1}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add} - 1}{\partial \alpha_1} \right)}{(100 - \alpha_1)} \right] \right\} \end{aligned} \right] \\
0 &= \left[\begin{aligned} & \left(-a_1 \xi_1^{-2} \right) \left\{ \frac{SF_1 (100 - \alpha_1)}{\beta_1 (\mu - \alpha_1)} \left[\frac{(\mu - \alpha_1)}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add} - 1}{\partial \alpha_1} \right)}{(100 - \alpha_1)} \right] \right\} - a_2 \xi_2^{-2} \left\{ \frac{SF_2 (100 - \alpha_1)}{\beta_2 (\mu - \alpha_1)} \left[\frac{\mu - \alpha_1}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add} - 1}{\partial \alpha_1} \right)}{(100 - \alpha_1)} \right] \right\} \end{aligned} \right] \\
0 &= \left[\frac{(\mu - \alpha_1)}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add} - 1}{\partial \alpha_1} \right)}{(100 - \alpha_1)} \right] \left[\begin{aligned} & \left(-a_1 \xi_1^{-2} \right) \left(\frac{SF_1 (100 - \alpha_1)}{\beta_1 (\mu - \alpha_1)} \right) - a_2 \xi_2^{-2} \left(\frac{SF_2 (100 - \alpha_1)}{\beta_2 (\mu - \alpha_1)} \right) \end{aligned} \right] \\
0 &= \left[\frac{(\mu - \alpha_1)}{(100 - \alpha_1)^2} + \frac{\left(\frac{\partial \mu_{add} - 1}{\partial \alpha_1} \right)}{(100 - \alpha_1)} \right] \Rightarrow
\end{aligned}$$

$$\begin{aligned}
 -\frac{(\mu - \alpha_1)}{(100 - \alpha_1)^2} &= \frac{\left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1\right)}{(100 - \alpha_1)} \\
 -\frac{(\mu - \alpha_1)}{(100 - \alpha_1)^2} \cancel{(100 - \alpha_1)} &= \left(\frac{\partial \mu_{add}}{\partial \alpha_1} - 1\right) \\
 1 - \frac{(\mu - \alpha_1)}{(100 - \alpha_1)} &= \frac{\partial \mu_{add}}{\partial \alpha_1}
 \end{aligned}$$

In this case, $\frac{\partial \mu_{add}}{\partial \alpha_1}$ will be nonzero for all agents which include α_1 in the model, and zero otherwise.

4.4 Optimal Experimental Design for Improving the Precision of the Interaction Threshold Estimate in the Mixture of 18 PHAHs

A thorough analysis of the mixture of 18 PHAHs was presented in Chapter 3. Let us consider the data presented in Chapter 3 to represent the first stage data. In this section, our attention is focused on the procedure for determining the optimal second stage design for improving the precision associated with the estimate of the interaction threshold. Recall that the original analysis of this data involved likelihood-ratio based inference. We have since developed a method for constructing the parameter covariance matrix. Since we intend to base our design work on the parameter covariance matrix, we have to be concerned with its stability. As a result, the analysis on which our design work is based differs slightly from that presented in Chapter 3.

As discussed in Section 4.2, the dose scale varies widely across the 18 single chemicals present in the mixture. As a result, the parameter covariance matrix associated

with the fit of the FSCR full model, given in equation (4.1), was unstable, in that the determinant was approximately equal to zero. In order to stabilize the parameter covariance matrix, the scaled FSCR full model, given in equation (4.2), was fit to the single agents and the fixed-ratio mixture ray simultaneously. Non-unity scaling factors were used for each single agent with a dose range larger than 1000 $\mu\text{g}/\text{kg}$. After scaling, the determinant of the parameter covariance matrix associated with the full model is 88.255, which demonstrates considerable improvement in the stability of the matrix. In addition, as described in Chapter 3, the minimum response parameters were combined to arrive at a more parsimonious model. The resulting parameter estimates are given in Appendix C2. Parameter estimates resulting from the fit of the scaled FSCR additivity model are given in Appendix C3. The parameter estimates resulting from the fit of the scaled FSCR interaction threshold model are given in Appendix C4.

The D-optimality design criterion is based on the estimated parameter covariance matrix, which is constructed using hypothesized values of the model parameters. Because we have estimates of the model parameters from the analysis of the first stage data, those estimates serve as our hypothesized values. The hypothesized dose-response relationship is described in Figure 4.1.

Dashed Line: Additivity Model, Solid Line: Hypothesized Mixture Model

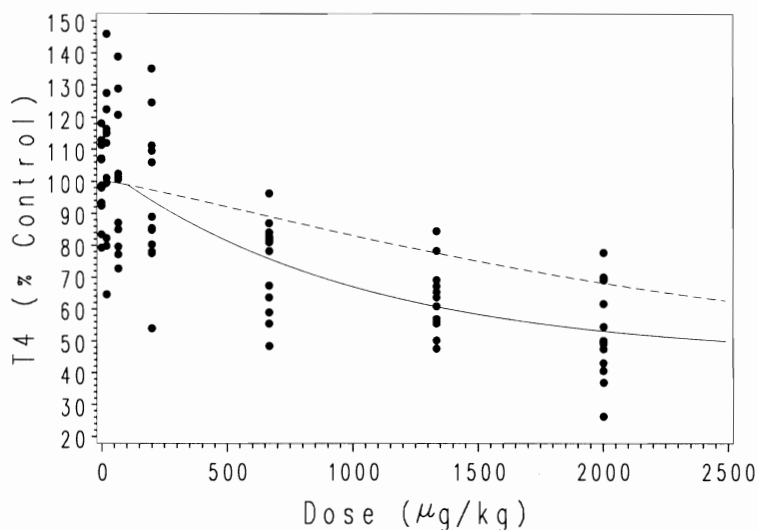


Figure 4.1. Predicted T4 vs. Total Dose of the Mixture

It should be noted that we have both single chemical and mixture data to support the hypothesized model parameters. Neither the single chemical data nor the mixture data from the first stage experiment will be affected by our design work. We are interested in constructing an optimal design to supplement the data already available.

We required that the covariance matrix of the model parameters be positive definite. In addition, because we understand that a mixture experiment would not be conducted without a control group, we forced our design to include a control dose. That is, the first dose point in each design is always zero, though the proportion of subjects included at the zero dose iterates within the search algorithm.

Consider a second stage mixture experiment with n_{mix_design} observations taken from m design points, and recall that there are n_s observations from the single chemical experiments and $n_{mix_original}$ observations from the first stage mixture experiment. The

dose values and the number of observations at each design point are represented in the \mathbf{F} matrix, which is defined as

$$\mathbf{F} = \left[\begin{array}{l} \left. \begin{array}{l} \partial\mu_1 / \partial\phi' \\ \vdots \\ \partial\mu_{n_s} / \partial\phi' \end{array} \right\} n_s \times p \\ \left. \begin{array}{l} \partial\mu_{n_s+1} / \partial\phi' \\ \vdots \\ \partial\mu_{n_s+n_{mix_original}} / \partial\phi' \end{array} \right\} n_{mix_original} \times p \\ \hline \left. \begin{array}{l} \partial\mu_{1n_1} / \partial\phi' \\ \vdots \\ \partial\mu_{2n_2} / \partial\phi' \\ \vdots \\ \partial\mu_{mn_m} / \partial\phi' \end{array} \right\} \begin{array}{l} n_1 \times p \\ n_2 \times p \\ \vdots \\ n_m \times p \end{array} \end{array} \right]$$

$n_s + n_{mix_original}$ observations
from the first stage experiments

n_{mix_design} observations to be included
in the second stage experiment

The total number of observations to be included in the second stage mixture experiment, then, is $n_{mix_design} = \sum_{r=1}^m n_m$, and $N = n_s + n_{mix_original} + n_{mix_design}$ is the total

number of observations to be used in the final analysis. For the scaled FSCR interaction threshold model described in equation (4.3) and shown in Figure 4.1, p is 33. The derivatives contained in the \mathbf{F} matrix are calculated as described in Section 4.3.

The derivatives for the n_s observations associated with the single chemical data are fixed, as are the derivatives for the $n_{mix_original}$ observations associated with the first stage mixture experiment. The derivatives for the n_{mix_design} observations associated with the second stage mixture experiment will iterate within the Nelder-Mead Simplex Algorithm, which will determine the dose associated with each of the m design points, t_r , and the allocation of observations to each design point, q_r , for $r=1, \dots, m$. Since the number of observations at the r^{th} design point depends on the allocation of observations at the r^{th} design point, the number of observations at each design point will also iterate.

A SAS macro was written to find the optimal design for a number of starting values, under a Ds-optimal design criterion. Our Ds-optimal design criterion seeks to minimize the variance associated with the hypothesized interaction threshold, Δ . The algorithm searches to find the doses, as well as the proportion of observations allocated to each dose, that minimize the variance of the interaction threshold. The dose and allocation of observations that minimizes the variance of the interaction threshold is the optimal m point design.

For a given n_{mix_design} , the Ds-optimal design is based on the proportion of observations at each dose, rather than the number of observations at each dose. In other words, while the variance achieved by the optimal design is based on the number of observations available for the second-stage mixture experiment, the design itself is not.

This allows the experimenter to use the optimal design to determine the sample size needed to achieve a given power.

We began the search for the optimal design with an eight point design. The starting values included four doses prior to the interaction threshold and four doses larger than the interaction threshold. We considered 32 such sets of starting values. Each set of starting values was provided to the macro, with an equal allocation of the available sample to the m design points, and a candidate optimal design was reached in each case. In some instances, due to the rounding of the sample size allocated to each design point, the number of subjects required by the design was larger than n_{mix_design} . When this was the case for the candidate optimal m point design, observations were removed from the design point with the largest sample size in order to preserve the sample size associated with the design. Let us call the variance of the interaction threshold based on the adjusted samples sizes the modified variance; if no sample size adjustment was required, the modified variance is equal to the actual variance. The candidate optimal design with the smallest modified variance is the optimal m point design.

In the case where the optimal m point design required fewer than m dose groups, as a result of overlapping dose locations or insufficient allocation, we reduced our grid of starting values to an $m-1$ point design. The number of design points was reduced until the design could be reduced no further, either because all remaining design points were required or because a smaller design was considered unacceptable. The Ds-optimal designs are given in Table 4.1, where \mathbf{t} represents the vector of total dose locations and \mathbf{q} represents the vector of corresponding proportions.

Table 4.1. Ds-Optimal Second Stage Designs for the Estimation of the Interaction Threshold

The “Actual” design is the actual result of the Nelder-Mead Simplex Algorithm over a grid of starting values. In the “Practical” design, design points have been collapsed and/or eliminated as indicated by the “Actual” design. The asterisk (*) denotes that design points have been collapsed in keeping with their location relative to the interaction threshold.

Number of Points		Second Stage Design: $\begin{Bmatrix} \mathbf{t} \\ \mathbf{q} \end{Bmatrix}$	Variance of the Interaction Threshold
8	Actual	$\begin{Bmatrix} 0 & 38.9 & 105.3_{<\Delta} & 105.3_{>\Delta} & 105.3_{>\Delta} & 105.9 & 116.6 & 578.0 \\ 0.093 & 0.125 & 0.113 & 0.112 & 0.185 & 0.222 & 0.105 & 0.045 \end{Bmatrix}$	3413.8
	Practical	$\begin{Bmatrix} 0 & 39 & 105 & 106 & 117 & 578 \\ 0.093 & 0.125 & 0.410 & 0.222 & 0.105 & 0.045 \end{Bmatrix}$	4113.0
		* $\begin{Bmatrix} 0 & 39 & 105 & 106 & 117 & 578 \\ 0.093 & 0.125 & 0.113 & 0.519 & 0.105 & 0.045 \end{Bmatrix}$	3438.0
7	Actual	$\begin{Bmatrix} 0 & 37.0 & 105.0 & 105.3_{>\Delta} & 105.3_{>\Delta} & 105.3_{>\Delta} & 113.4 \\ 0.001 & 0.156 & 0.155 & 0.250 & 0.165 & 0.157 & 0.116 \end{Bmatrix}$	3281.5
	Practical	$\begin{Bmatrix} 0 & 37 & 105 & 113 \\ 0.001 & 0.156 & 0.727 & 0.116 \end{Bmatrix}$	7041.9
		* $\begin{Bmatrix} 0 & 37 & 106 & 113 \\ 0.001 & 0.156 & 0.727 & 0.116 \end{Bmatrix}$	3587.8
6	Actual	$\begin{Bmatrix} 0 & 4.0 & 104.5 & 105.3_{>\Delta} & 105.4 & 109.7 \\ 0.020 & 0.001 & 0.291 & 0.095 & 0.319 & 0.274 \end{Bmatrix}$	3228.7
	Practical	$\begin{Bmatrix} 0 & 105 & 110 \\ 0.020 & 0.705 & 0.274 \end{Bmatrix}$	4458.6
		* $\begin{Bmatrix} 0 & 105 & 106 & 110 \\ 0.020 & 0.291 & 0.414 & 0.274 \end{Bmatrix}$	3232.3

5	Actual	$\left\{ \begin{array}{ccccc} 0 & 96.7 & 99.9 & 105.5 & 106.8 \\ <0.001 & 0.205 & 0.164 & 0.326 & 0.305 \end{array} \right\}$	3213.9
	Practical	$\left\{ \begin{array}{ccccc} 0 & 97 & 100 & 106 & 107 \\ <0.001 & 0.205 & 0.164 & 0.326 & 0.305 \end{array} \right\}$	3216.6
4	Actual	$\left\{ \begin{array}{cccc} 0 & 98.3 & 104.5 & 105.5 \\ <0.001 & 0.009 & 0.335 & 0.656 \end{array} \right\}$	3195.6
	Practical	$\left\{ \begin{array}{ccc} 0 & 105 & 106 \\ <0.001 & 0.335 & 0.656 \end{array} \right\}$	3200.1
3	Actual	$\left\{ \begin{array}{ccc} 0 & 105.2 & 105.3_{>\Delta} \\ 0.120 & 0.255 & 0.625 \end{array} \right\}$	3238.5
	Practical	$\left\{ \begin{array}{cc} 0 & 105 \\ 0.120 & 0.880 \end{array} \right\}$	16593.8
		$* \left\{ \begin{array}{ccc} 0 & 105 & 106 \\ 0.12 & 0.255 & 0.625 \end{array} \right\}$	3245.0
2	Actual	$\left\{ \begin{array}{cc} 0 & 105.3_{>\Delta} \\ 0.329 & 0.671 \end{array} \right\}$	3428.72
	Practical	$\left\{ \begin{array}{cc} 0 & 105 \\ 0.329 & 0.671 \end{array} \right\}$	42304.7
		$* \left\{ \begin{array}{cc} 0 & 106 \\ 0.329 & 0.671 \end{array} \right\}$	3434.6

The Ds-optimal eight point design presented in Table 4.1 appears to be, in reality, a six point design. There are three dose groups located at the hypothesized interaction threshold, 105.3 $\mu\text{g}/\text{kg}$. These three design points can be collapsed into one dose group at 105 $\mu\text{g}/\text{kg}$, as described in the table. If we look carefully at these doses, however, we see that one dose is just below the interaction threshold, and the other two are just above. In keeping with this distinction, when we round the design points to the nearest integer, the total dose just below the interaction threshold can be rounded down, whereas the total dose just above the interaction threshold can be rounded up. This design is also described in the table and is denoted with an asterisk (*). This allows us to investigate the importance of their location relative to the interaction threshold, which is an important distinction. Combining dose groups without regard to their location relative to the interaction threshold results in a dramatic increase in the variance of the interaction threshold. Thus, the Ds-optimal second stage design requires data both above and below the hypothesized interaction threshold in order to precisely estimate its location.

Similarly, the Ds-optimal seven point design presented in Table 4.1 is, practically speaking, a three point design. The allocation to the control dose is only 0.1% of the sample, indicating that the control dose is not required by the optimal design. In addition, there are three dose groups located at the hypothesized interaction threshold, 105.3 $\mu\text{g}/\text{kg}$. Again, the design can be reduced in one of two ways – we can round the resulting total doses in the usual fashion, or we can round according to the placement of the design point with regard to the interaction threshold. As we saw previously, the variance of the

interaction threshold increases dramatically when we combine dose groups without regard to their location relative to the interaction threshold.

Similarly, the Ds-optimal six point design presented in Table 4.1 can be reduced by removing unnecessary design points and collapsing dose groups that are sufficiently close together. The design allocates only 0.1% of the sample to a total dose of 4 $\mu\text{g}/\text{kg}$, making this design point unnecessary. In addition, there are three design points situated in the vicinity of the interaction threshold. When these groups are collapsed to form one dose group, the variance of the interaction threshold increases dramatically. If we combine these groups according to their location relative to the interaction threshold, however, there is only a slight increase in variance over the Ds-optimal six point design.

The Ds-optimal five point design presented in Table 4.1 is actually a four point design, as the control dose is not required by the optimal design. The Ds-optimal four point design is, in reality, a two point design. The control dose is not necessary, and the design allocates less than 1% of the sample to a dose of 98 $\mu\text{g}/\text{kg}$. The Ds-optimal three point design contains two design points in the vicinity of the interaction threshold. As previously described, the variance of the interaction threshold increases dramatically if we combine these groups without considering their location relative to the hypothesized interaction threshold. In keeping with our observation that the Ds-optimal design requires data both below and above the interaction threshold, the Ds-optimal two point design includes a control group and a dose just above the interaction threshold.

The Ds-optimal second stage design is the design that minimizes the variance of the interaction threshold. We can see from Table 4.1, then, that the Ds-optimal design is the

Ds-optimal four point design, which is described in Figure 4.2. The dashed line represents the additivity model, and the hypothesized mixture model is represented by the solid line. The dots indicate data available from the first stage of the experiment, and the triangles represent the location of the design points indicated by the Ds-optimal second stage design.

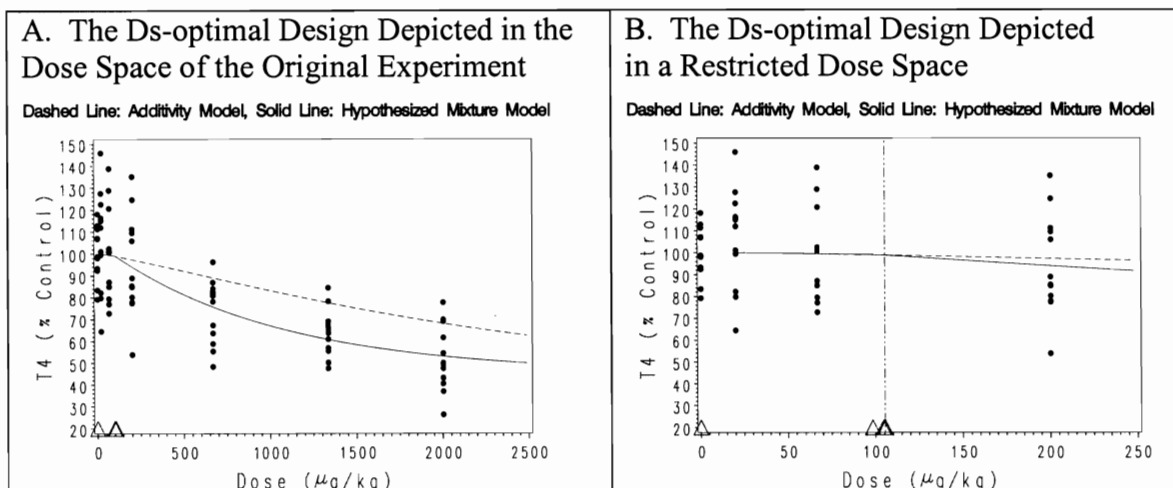


Figure 4.2. The Ds-optimal Second Stage Design for the Estimation of the Interaction Threshold

The vertical dotted line in Figure 4.2B represents the location of the hypothesized interaction threshold.

Though this design may be statistically optimal, there are some practical problems with its implementation. It does not require the use of a control group, which may be a problem for some scientists. A toxicologist may not be willing to conduct an experiment that does not include a control group. A potential solution is to include a control group in the second stage experiment but to exclude it from the analysis. In other words, we can compare the response of the controls from each stage for quality control purposes, but we can analyze the results of the complete experiment without the second stage control

group. This alone, however, does not completely solve the problems with the practicality of the Ds-optimal design shown in Figure 4.2. The remaining design points are fairly close together, so that the design reduces to a single design point. It is unlikely, for instance, that a toxicologist will be able to distinguish a dose of 104.5 $\mu\text{g}/\text{kg}$ from a dose of 105.5 $\mu\text{g}/\text{kg}$. Many of the designs given in Table 4.1 share these practical problems. In the next section, we apply the penalized optimality criterion, developed by Parker and Gennings (2006), to arrive at a second stage design that is more likely to be carried out in practice.

4.5 Penalized Optimal Design Methodology

Parker and Gennings (2006) developed the penalized optimal design criterion, which uses desirability functions to penalize impractical experimental designs. Recall that $\boldsymbol{\varphi} = [\boldsymbol{\theta} \ \Delta]$ represents the $p \times 1$ vector of model parameters, where $\boldsymbol{\theta}$ represents the $(p-1) \times 1$ vector of model parameters excluding Δ , as described in Chapter 3. Let $h(\mathbf{t}, \mathbf{q}, f(\boldsymbol{\varphi}), m, n)$ be the value of an alphabetic optimality criterion, such as the Ds-optimal design criterion used in Section 4.4. As previously discussed, \mathbf{t} is the vector of design points associated with the optimal design, \mathbf{q} is the vector of corresponding sample allocations, m is the number of design points, and $n_{\text{mix_design}}$ is the sample size available.

Let $j=1, \dots, k$ represent the k experimental design preferences under consideration. Each characteristic, p_j , is transformed into its desirability, d_j , where $0 \leq d_j \leq 1$. In this setting, $d_j = 0$ indicates that characteristic p_j is highly undesirable, and $d_j = 1$ indicates

that characteristic p_j is highly desirable. The overall desirability of the k combined characteristics is represented by

$$D(\mathbf{t}, \mathbf{q}) = (d_1 d_2 \cdots d_k)^{1/k}, \quad (4.12)$$

and the penalty function is represented by $1 - D(\mathbf{t}, \mathbf{q})$.

As described in Parker and Gennings (2006), the penalty function is added to the design criterion to penalize experimental designs, based on the investigator's pre-defined experimental design preferences. The penalty function can take on any value between and including 0 and 1; therefore, a user-defined scaling constant, Λ , is used to control the weight of the penalty function, relative to the minimum optimality criterion. For a given Λ , the penalized optimality criterion, as a function of \mathbf{t} and \mathbf{q} , is written

$$h(\mathbf{t}, \mathbf{q}, f(\boldsymbol{\varphi}), m, n) + \Lambda(1 - D(\mathbf{t}, \mathbf{q})). \quad (4.13)$$

The penalized optimal design is the design that jointly minimizes $h(\mathbf{t}, \mathbf{q}, f(\boldsymbol{\varphi}), m, n)$ and $1 - D(\mathbf{t}, \mathbf{q})$ for a given value of Λ .

Parker recommends choosing a value of Λ in the range where there is stability in the desirability function and the alphabetic optimality criterion. Parker considers multiples of $\min(h(\mathbf{t}, \mathbf{q}, f(\boldsymbol{\varphi}), m, n))$, the minimum value of the alphabetic optimality criterion, as potential values for the scaling constant Λ . The penalized optimal design is found for possible values of Λ , where $\Lambda = \lambda \left[\min(h(\mathbf{t}, \mathbf{q}, f(\boldsymbol{\varphi}), m, n)) \right]$ and λ is any positive number. To choose λ , the alphabetic optimality criterion and the desirability function are plotted against values of λ . Parker and Gennings (2006) choose λ such that, compared to

the optimal design resulting from the alphabetic optimality criterion, there is a sizable improvement in the desirability function, and the increase in the alphabetic optimality criterion is minimal. Parker admits that “although absolute optimality properties are lost, the penalized optimal design is optimal in accordance with design preferences set by the investigator and the alphabetic optimality criterion” (2005). This tradeoff is reasonable in situations such as that described in Section 4.4, where certain characteristics of the statistically optimal design may not be considered reasonable by scientific investigators.

4.5.1 Experimental Design Preferences for the Second Stage Fixed-Ratio Ray Mixture Experiment of 18 PHAHs

Recall from Section 4.4 that we had a number of practical concerns with the Ds-optimal second stage designs described in Table 4.1. The Ds-optimal design is the Ds-optimal four point design, which is described as follows

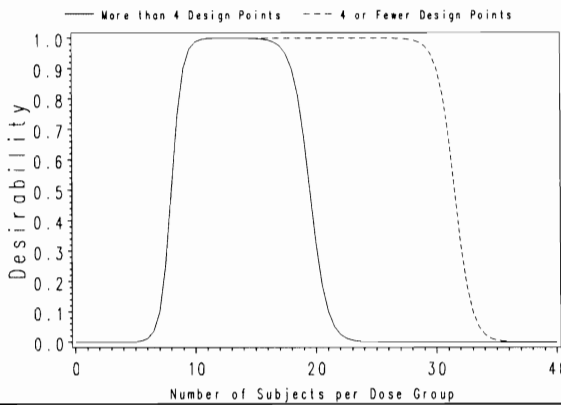
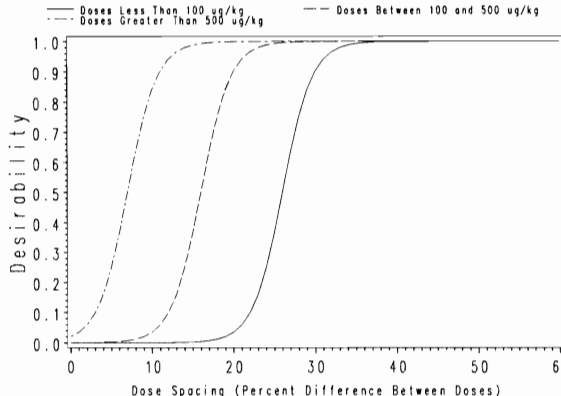
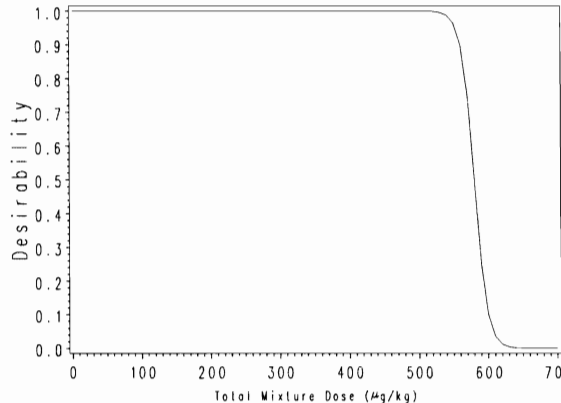
$$\begin{Bmatrix} \mathbf{t} \\ \mathbf{q} \end{Bmatrix} = \begin{Bmatrix} 0 & 98.3 & 104.5 & 105.5 \\ < 0.001 & 0.009 & 0.335 & 0.656 \end{Bmatrix},$$

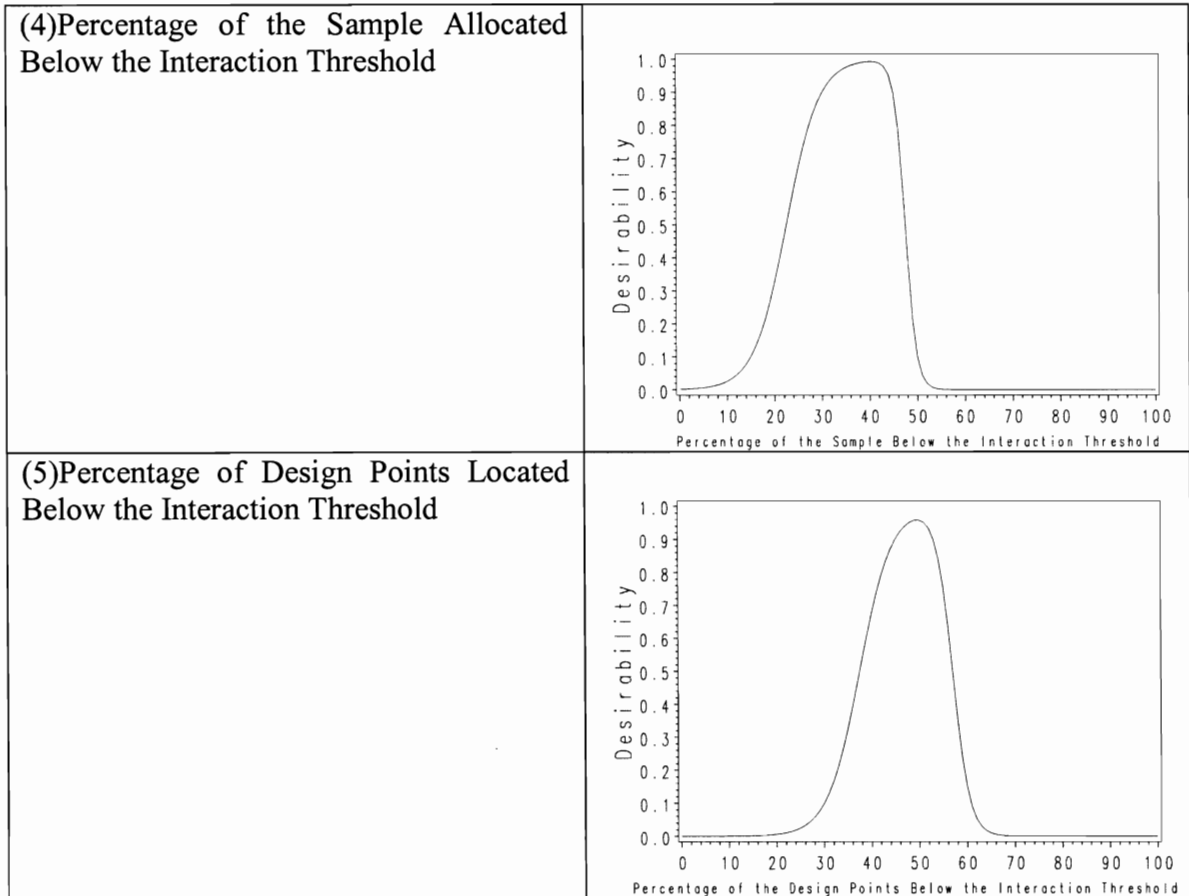
where total dose is given in $\mu\text{g}/\text{kg}$. First, it should be noted that less than 0.1% of the sample is allocated to the control dose. Per our personal communication with Dr. Kevin Crofton, a neurotoxicologist with the Neurobehavioral Toxicology Branch of the National Health and Environmental Effects Research Laboratory of the U.S. EPA, and the scientist who conducted the original study, a scientist is not likely to conduct an experiment that does not include a control group. In addition, the remaining three design points are located with 10 $\mu\text{g}/\text{kg}$ of each other. We are particularly concerned with the

placement of the third and fourth design points, which are located only 1 $\mu\text{g}/\text{kg}$ apart. Depending on the sensitivity of the equipment used in the experiment, a scientist may not be able to reliably distinguish 104.5 $\mu\text{g}/\text{kg}$ from 105.5 $\mu\text{g}/\text{kg}$.

To address these issues, the penalized optimality criterion (Parker and Gennings, 2006) was used to find the second stage design that minimizes the variance of the hypothesized interaction threshold while maximizing the desirability of the associated design. This criterion is based on desirability functions, which capture the investigator's experimental design preferences. These desirability functions were determined through collaboration with Dr. Crofton via personal communication. The logistic function was used to create our desirability functions, though other functions can be used to achieve the appropriate shape. Table 4.2 contains plots of these desirability functions. The experimental design preferences used, and their corresponding desirability functions, are described below.

Table 4.2. Desirability Functions for the Penalized Optimal Second Stage Design

Characteristic	Desirability Curve
(1) Number of Subjects per Dose Group	 <p>— More than 4 Design Points - - - 4 or Fewer Design Points</p>
(2) Dose Spacing, Described as a Percent Difference Between Doses	 <p>— Doses Less Than 100 ug/kg - - - Doses Between 100 and 500 ug/kg - - - Doses Greater Than 500 ug/kg</p>
(3) Dose Region	



(1) Number of Subjects per Dose Group

This design preference stems directly from the lack of a control group in the Ds-optimal second stage design. The number of subjects per dose group can be thought of as a target approach; the allocation of between 10 and 18 subjects per dose group is fairly standard in the toxicological literature. This target region is captured in the desirability function described by the solid curve. For experimental designs with more than four design points, this target range is appropriate, since our design work is based on a second stage sample size of 100 subjects. For experimental designs with four or fewer design

points, it is acceptable to have more than 18 subjects per dose group. This expanded target region is captured in the desirability function described by the dashed curve. In addition, it is important to note that the target region may depend on the importance of the particular agent or mixture under study. For important compounds, the overall sample size may be increased, so it may be acceptable to increase the number of subjects allowed at each design point.

(2) Dose Spacing (Percent Difference Between Doses)

Due to the accuracy of the serial dilution method of obtaining doses, the acceptable spacing between dose groups depends on dose region. We describe dose spacing as the percent difference between two consecutive dose groups. A smaller percentage is acceptable for doses in the high dose region than in the low dose region. For instance, a 10% difference is more reliably detected between doses of 500 and 550 $\mu\text{g}/\text{kg}$ than between doses of 50 and 55 $\mu\text{g}/\text{kg}$. After careful consideration, the dose space was divided into three dose regions, and a desirability function was constructed for each region. The low dose region requires a large percentage difference between doses, and the percentage difference required decreases as the dose increases. The desirability function for doses less than 100 $\mu\text{g}/\text{kg}$ specifies that each dose in this region must be at least 30% larger than the dose below, as shown in the the solid curve. The desirability function for doses between 100 and 500 $\mu\text{g}/\text{kg}$ is a parallel shift to the left, and the desirability function for doses larger than 500 $\mu\text{g}/\text{kg}$ is a further shift to the left.

(3) Dose Region

Teuschler et al. (2002) commented on the importance of low dose studies. As such, it may be of interest to restrict the design search to doses within a particular region of interest. In general, this desirability function represents a constraint on the dose region. In the mixture of 18 PHAHs, recall that we have information from the first stage experiment. The FSCR analysis of the first stage experiment indicated that the interaction threshold is no larger than 544 $\mu\text{g}/\text{kg}$. Therefore, the desirability function representing dose region restricts the placement of dose groups larger than 550 $\mu\text{g}/\text{kg}$.

(4) Percentage of the Sample Below the Interaction Threshold

If the nonlinear model of interest contains one or more threshold parameters, used to define two or more regions of the model, it may be of interest to place dose groups in each region. In the mixture of 18 PHAHs, we would like to be able to use the data from the second stage to validate our analysis of the first stage experiment. After the second stage experiment has been conducted, we are able to compare the data to the model resulting from our analysis of the first stage data. Ideally, we are looking for the predicted model to adequately describe the data from both stages. It is important, then, that the second stage design allocates enough of the sample below the threshold for us to inspect model fit in the additivity region. The same is true for the interaction region. The desirability function for the percentage of the sample below the interaction threshold describes a target region of 30 to 45%.

As the number of subjects included in the second stage experiment increases, this desirability curve may shift to the left. For experiments with a large sample size, we may require a smaller percentage of the sample allocated to a particular dose region.

(5) Percentage of Design Points Below the Interaction Threshold

This design preference is somewhat related to that described in (4) above, the percentage of the sample below the interaction threshold. If we are interested in allocating a certain percentage of the sample below the interaction region, for validation purposes, we may want to ensure that a certain percentage of the second stage design points are located below the interaction threshold. This is to guard against a situation in which 35% of the second stage sample is allocated below the interaction threshold but to the same design point. In an experiment with less than ten groups, a roughly equal distribution of design points above and below the interaction threshold is appropriate. This gives us a number of validation points in the additivity region. The desirability function for the percentage of design points below the interaction threshold describes a target region of 40 to 55% percent. As the number of dose groups increases, however, this curve will likely shift to the left. For experiments with a large number of design points, we would require a smaller percentage of those design points to be allocated in a particular dose region.

Preferences for sample size and the number of design points were also considered initially; however, these cannot be controlled through desirability functions. The minimum alphabetic optimality criterion depends on both the sample size available and the number of design points specified. Because these characteristics are pre-specified and fixed for a given design, we cannot allow them to contribute to the desirability function.

Five design preferences have been identified for our second stage experimental design work. The number of desirability functions involved in the design search, however, is much larger. The target region for the number of subjects allocated to each dose group involves the use of m desirability functions, one for each design point in the study. There are $\binom{m-1}{2}$ desirability functions associated with the dose spacing requirement, since each nonzero dose must be compared with each other nonzero dose. There are $m-1$ desirability functions associated with the constraint on the dose region, since all nonzero doses have the potential to iterate outside of the region. In addition, there is one desirability function associated with the preference for the percentage of the sample allocated below the interaction threshold, and one desirability function associated with the preference for the percentage of the design located below the interaction threshold. The total number of desirability functions for the second stage design involving the mixture of 18 PHAHs is $k = 2m + 1 + \binom{m-1}{2} (m > 2)$.

4.5.2. Penalized Optimal Design for Improving the Precision of the Interaction Threshold Estimate

Following Parker and Gennings (2006), a SAS macro was written to find the optimal design for a number of starting values, under a penalized Ds-optimal design criteria, using the Nelder-Mead Simplex Algorithm. The penalized Ds-optimal design criteria seeks to minimize simultaneously the variance associated with the hypothesized interaction threshold and the penalty associated with the design. For a given number of design points m , the value of λ is set to one, and a grid of starting values is searched for the penalized optimal m point design. The value of λ is updated, and the same grid of starting values is searched for the penalized optimal m point design. For each value of λ , both the variance of the hypothesized interaction threshold and the desirability associated with the penalized optimal m point design are recorded. The desirability and variance are plotted against values of λ . We choose λ to be such that we get a sizable increase in the desirability of the design, and that the associated increase in the variance of the hypothesized interaction threshold is considered both reasonable and tolerable to the investigator. We considered designs consisting of four, six and eight design points.

We began our search for the penalized optimal design with an eight point design. The starting values for the eight point design included four doses prior to the interaction threshold and four doses larger than the interaction threshold. We considered eight such sets of starting values. For a given λ , each set of starting values was provided to the macro, with an equal allocation of the available sample to the m design points, and a candidate penalized optimal design was reached in each case. As before, in some

instances, due to the rounding of the sample size allocated to each design point, the number of subjects required by the design was larger than n_{mix_design} . When this was the case for the candidate penalized optimal m point design, observations were removed from the design point with the largest sample size in order to preserve the sample size available to the design. Let us call the variance of the interaction threshold based on the adjusted sample sizes the modified variance, the desirability based on the adjusted sample sizes the modified desirability, and the objective based on the adjusted sample sizes the modified objective. If no sample size adjustment was required, then the modified variance is equal to the actual variance, the modified desirability is equal to the actual desirability, and the modified objective is equal to the actual objective. The candidate penalized optimal design with the smallest modified objective was considered to be the penalized optimal m point design for a given λ . The corresponding SAS code is given in Appendix C5.

The plot of the achieved variance and desirability versus λ for the eight point penalized optimal second stage designs is given in Figure 4.3. Recall that the variance associated with the Ds-optimal eight point design is 3413.8, and the corresponding desirability is 0.13. It can be seen in the plot that the penalized optimal design with $\lambda = 0.25$ results in a desirability of approximately 0.90, which is a substantial increase over the Ds-optimal eight point design. The associated variance is approximately 3800, which is a slight increase over the Ds-optimal eight point design.

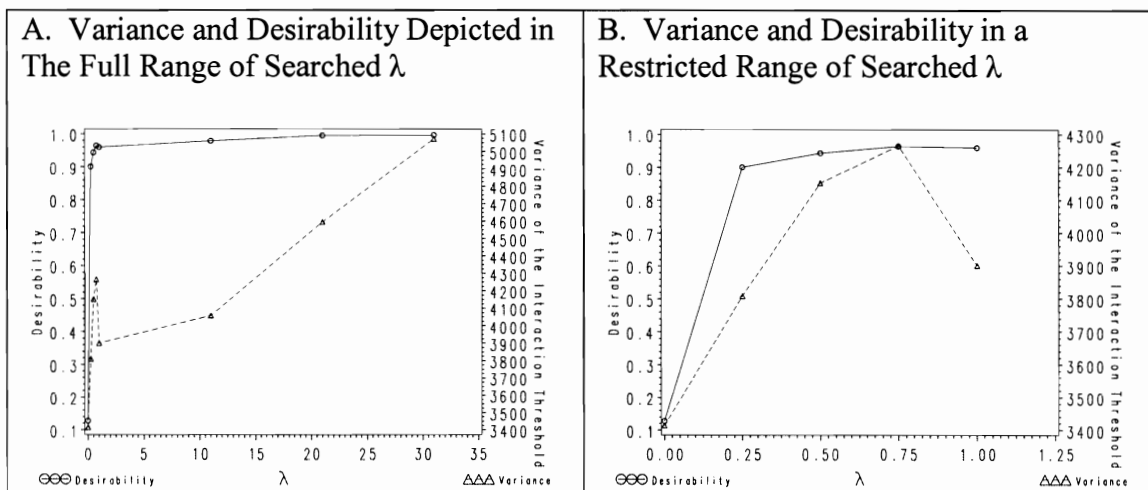


Figure 4.3. Variance of the Hypothesized Interaction Threshold and Desirability vs. λ for the Eight Point Penalized Optimal Second Stage Designs

The plot of the achieved variance and desirability versus λ for the six point penalized optimal second stage designs is given in Figure 4.4. Recall that the variance associated with the Ds-optimal six point design is 3228.74, and the corresponding desirability is 0.003. It can be seen in the plot that the penalized optimal design with $\lambda = 0.75$ results in a desirability larger than 0.90, which is a substantial increase over the Ds-optimal six point design. The associated variance is approximately 3800, which is a slight increase over the Ds-optimal six point design.

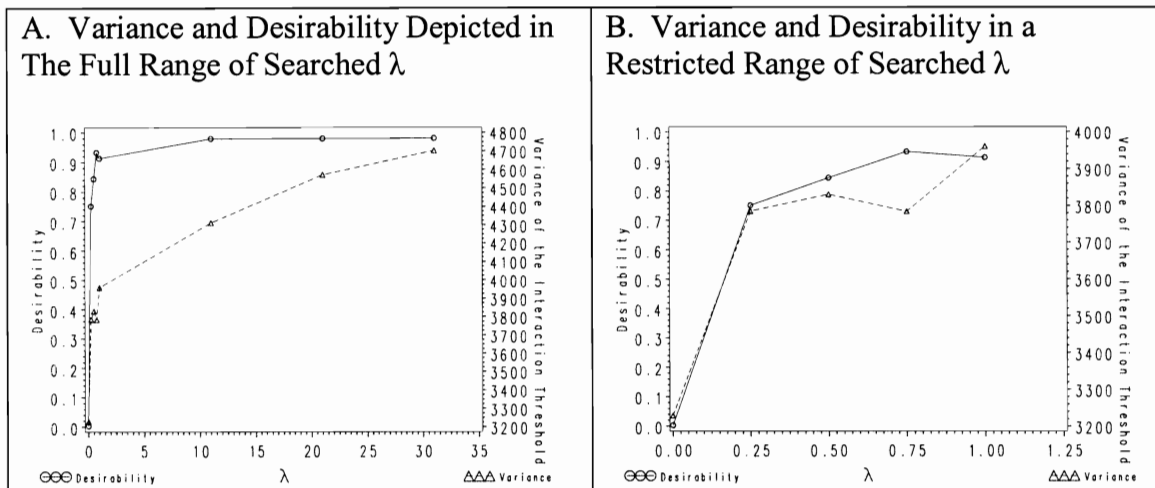


Figure 4.4. Variance of the Hypothesized Interaction Threshold and Desirability vs. λ for the Six Point Penalized Optimal Second Stage Designs

The plot of the achieved variance and desirability versus λ for the four point penalized optimal second stage designs is given in Figure 4.5. Recall that the variance associated with the Ds-optimal four point design is 3195.6, and the corresponding desirability is 0.00003. It can be seen in the plot that the penalized optimal design with $\lambda = 0.75$ results in a desirability larger than 0.95, which is a substantial increase over the Ds-optimal four point design. The associated variance is approximately 3400, which is a slight increase over the Ds-optimal four point design.

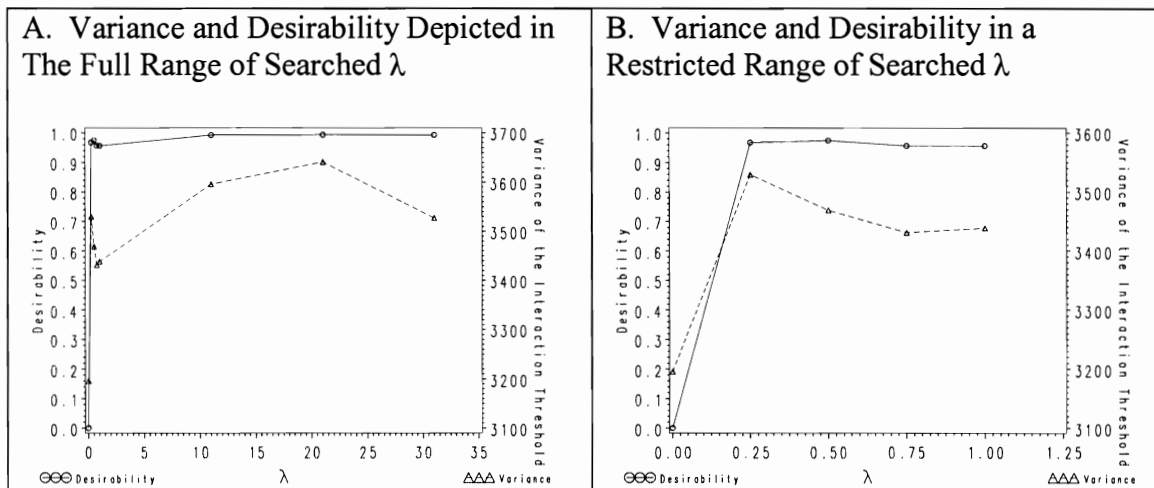


Figure 4.5. Variance of the Hypothesized Interaction Threshold and Desirability vs. λ for the Four Point Penalized Optimal Second Stage Designs

Regardless of the value of m , the number of experimental design points, we saw that increasing λ beyond its chosen value results in further increases in desirability. However, the increases are slight and are associated with further increases in the variance of the hypothesized interaction threshold. The penalized optimal second stage designs, based on the selected λ values, are given in Table 4.3, where \mathbf{t} represents the vector of total dose locations and \mathbf{q} represents the vector of corresponding proportions. The Ds-optimal second stage designs are also provided for comparison.

Table 4.3. Optimal Second Stage Designs for the Estimation of the Interaction Threshold

Number of Points	Optimality Criterion	Second Stage Design: $\begin{Bmatrix} \mathbf{d} \\ \mathbf{q} \end{Bmatrix}$	Variance of the Interaction Threshold	Desirability
8	Ds ($\lambda = 0$)	$\begin{Bmatrix} 0 & 38.9 & 105.3_{<\Delta} & 105.3_{>\Delta} & 105.3_{>\Delta} & 105.9 & 116.6 & 578.0 \\ 0.093 & 0.125 & 0.113 & 0.112 & 0.185 & 0.222 & 0.105 & 0.045 \end{Bmatrix}$	3413.8	0.128
	Penalized ($\lambda = 0.25$)	$\begin{Bmatrix} 0 & 65.9 & 85.6 & 106.7 & 123.5 & 152.9 & 181.4 & 503.5 \\ 0.076 & 0.115 & 0.145 & 0.166 & 0.184 & 0.145 & 0.095 & 0.075 \end{Bmatrix}$	3809.1	0.901
6	Ds ($\lambda = 0$)	$\begin{Bmatrix} 0 & 4.0 & 104.5 & 105.3_{>\Delta} & 105.4 & 109.7 \\ 0.020 & 0.001 & 0.291 & 0.095 & 0.319 & 0.274 \end{Bmatrix}$	3228.7	0.003
	Penalized ($\lambda = 0.75$)	$\begin{Bmatrix} 0 & 64.7 & 84.7 & 107.5 & 142.4 & 173.3 \\ 0.110 & 0.175 & 0.167 & 0.195 & 0.195 & 0.157 \end{Bmatrix}$	3784.3	0.933
4	Ds ($\lambda = 0$)	$\begin{Bmatrix} 0 & 98.3 & 104.5 & 105.5 \\ <0.001 & 0.009 & 0.335 & 0.656 \end{Bmatrix}$	3195.6	<0.001
	Penalized ($\lambda = 0.75$)	$\begin{Bmatrix} 0 & 78.7 & 105.3 & 127.8 \\ 0.153 & 0.250 & 0.313 & 0.285 \end{Bmatrix}$	3431.3	0.957

As can be seen in Table 4.3, the penalized optimal four point design results in the highest desirability and the lowest variance of the penalized optimal designs. The increase in desirability of the penalized optimal four point design over the Ds-optimal four point design is substantial, while the corresponding increase in the variance of the hypothesized interaction threshold is a mere 7.4%, which is a fairly minimal increase. Therefore, the penalized optimal four point design, which is described in Figure 4.6, is the penalized optimal second stage design. The dashed line represents the additivity model, and the hypothesized mixture model is represented by the solid line. The dots indicate data available from the first stage of the experiment, and the triangles represent the location of the design points indicated by the penalized optimal second stage design.

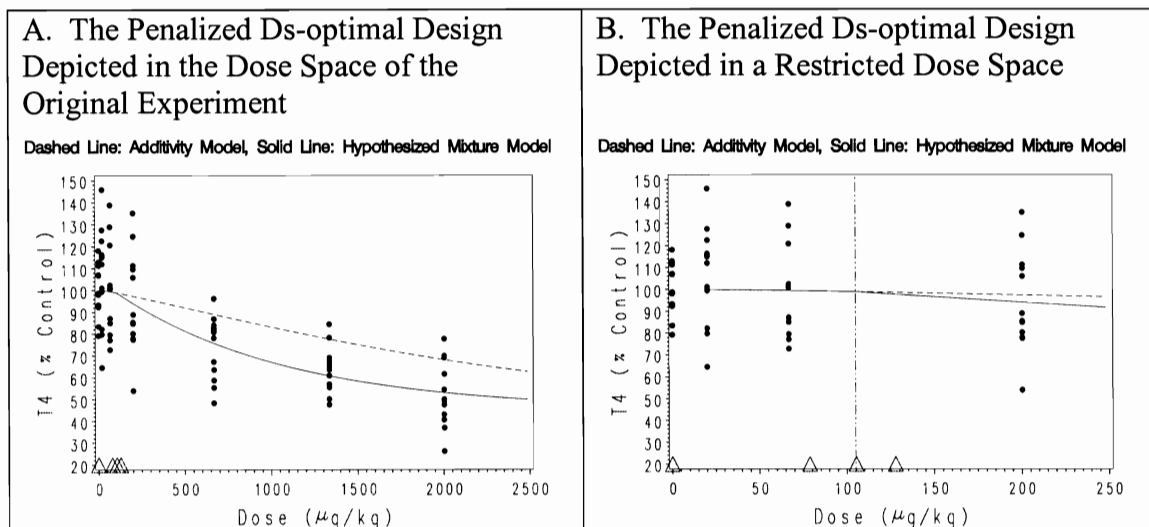


Figure 4.6. The Penalized Optimal Second Stage Design for the Estimation of the Interaction Threshold

The vertical dotted line in Figure 4.2B represents the location of the hypothesized interaction threshold.

4.6. Discussion

While Ds-optimal designs may satisfy statistical optimality criterion, there may be practical issues associated with their implementation. The Ds-optimal second stage design presented in Section 4.4 is a worthwhile example of these issues. The Ds-optimal second stage design is a four point design which, in reality, can be described by two design points. The allocation of the Ds-optimal second stage design is such that the entire sample is divided between 104.5 and 105.5 $\mu\text{g}/\text{kg}$, the third and fourth design points. The lack of a control group is likely to be a concern for the investigator. Additionally, the close proximity of the third design point to the fourth is likely to be an issue in the implementation of this design. The investigator may not be able to reliably distinguish between these doses. If the implementation of the Ds-optimal design is a problem because it conflicts with the design preferences of the investigator, then the design preferences of the investigator can and should be taken into account when determining the optimal design.

The penalized optimality criterion developed by Parker and Gennings (2006) uses both an alphabetic optimality criterion and the design preferences specified by the investigator to determine the optimal design. To apply this criterion in the example presented in this chapter, we collaborated with toxicologist Dr. Kevin Crofton at the U.S. EPA regarding design preferences for this scenario. We characterized these preferences, described in Section 4.5.1, using the desirability functions presented in Table 4.2. The overall desirability of the design, as well as the variance of the hypothesized interaction

threshold, were taken into account in the search for the penalized optimal second stage design.

Whereas the Ds- optimality criterion applied in Section 4.4 seeks to minimize the variance of the hypothesized interaction threshold, the penalized optimality criterion seeks to minimize both the variance of the hypothesized interaction threshold and the penalty function associated with the design. The penalized optimal second stage design, presented in Figure 4.6, clearly addresses the practical issues of the Ds-optimal design. Approximately 15% of the sample is allocated to the control dose. The remaining three design points are scattered around the hypothesized interaction threshold. Their proximity to each other is controlled by the desirability functions, as is the number of subjects per design point, the location of the design points, the percentage of the sample below the interaction threshold, and the percentage of the design points below the interaction threshold. Thus, the resulting penalized optimal design has the characteristics deemed important in practice, as quantified by the desirability functions, while allowing only a minimal increase in the variance of the interaction threshold.

Chapter 5

Determining Optimal Experimental Designs for Nonlinear Models Using Likelihood Ratio-Based Inference

5.1 Introduction and Motivation

In some cases where nonlinear models are concerned, (quasi-) likelihood ratio-based testing procedures may be preferred over Wald-type procedures. One reason for this preference, O'Brien (2002) argued, is that the coverage properties of the (quasi-) likelihood ratio-based confidence interval are more reasonable than the Wald-type intervals. To demonstrate, O'Brien (2002) provided a number of simple examples where Wald-type procedures proved inadequate. Wald-type procedures are based on the asymptotic normality of the vector of model parameter estimates, which comes from a linear Taylor-series approximation to the nonlinear functional form of the mean. If the linear approximation is poor, inference based on this assumption of asymptotic normality may be misleading (Seber and Wild, 2003). Whereas Wald-type procedures are based on a linear approximation to the nonlinear function, (quasi-) likelihood ratio-based procedures are based on the nonlinear function itself.

First-order, such as D- and Ds-, optimality criteria are also based on a linear approximation to the nonlinear function (Seber and Wild, 2003). Locally D- optimal

designs minimize the volume of the linearized confidence region for the unknown parameters β (Seber and Wild, 2003; O'Brien, 1992). As a result, these criteria have fundamental limitations if the linear approximation is poor. Quadratic design criteria, based on a quadratic approximation to the nonlinear function, minimize the volume of the second-order approximation to the confidence region (O'Brien, 1992). Both the first-order and quadratic optimality criteria are based on approximations to the nonlinear function. Recall, however, that (quasi-) likelihood ratio-based procedures do not involve approximations to the nonlinear function; instead, the (quasi-) likelihood is a function of the nonlinear function itself. If the planned analysis involves (quasi-) likelihood ratio-based inference, then, it may be more appropriate to use a (quasi-) likelihood ratio-based criterion to develop the design.

The primary obstacle to the development of a (quasi-) likelihood ratio-based optimality criterion is that data are required in order to construct the interval. In order to construct the (quasi-) likelihood ratio-based confidence interval for a potential design, the response at each design point must be known. Minkin and Kundhal (1999) developed an optimality criterion based on the width of the likelihood ratio-based confidence interval. In particular, Minkin and Kundhal considered the situation in which logistic regression was to be used to analyze binomial data, in order to obtain an estimate of the ED_{50} . Minkin and Kundhal restricted their attention to a symmetric, two-point design and addressed the need for data by replacing each unknown data point with its expected value. By doing so, Minkin and Kundhal demonstrated that the optimal design can be found numerically by solving a system of two equations.

In the scenario of Minkin and Kundhal, the logistic model based on the hypothesized parameter values predicted the proportion responding at a given dose. As a result, they are not concerned with estimating the variability of the response. In the described situation, the variability of the response at a given dose is a known function of the corresponding mean. Replacing the unknown data points with their expected values, then, seems reasonable. For continuous data, however, using the mean is representative of the best case scenario and leads us to underestimate the variance of the response.

We are interested in developing a more general procedure for using the (quasi-) likelihood ratio-based confidence interval as an optimality criterion. In particular, we want to find the design that maximizes the lower bound of the (quasi-) likelihood ratio-based confidence interval around a given parameter of interest. Let us consider the situation in which we have dose-response data available; these data constitute the first stage data. If these data do not permit an adequate answer to the question of interest, we may want to supplement the first stage data with additional design points, which we refer to as the second stage design.

The goal of this chapter is to propose a procedure for using a (quasi-) likelihood ratio-based confidence interval as an optimality criterion in the development of an optimal second stage design. In Section 5.2, we describe our method for constructing the quasi-likelihood ratio-based optimality criterion. The method and resulting designs will be illustrated using a nonlinear threshold model for single agent dose-response data. The analysis of the first stage data is presented in Section 5.3. In Section 5.4, we demonstrate the method and resulting second stage design and compare the resulting quasi-likelihood

ratio-based optimal second stage design to the corresponding D- and Ds- optimal second stage designs. Lastly, we demonstrate and compare the designs under a different set of hypothesized parameter values in Section 5.5, and we conclude the chapter by discussing overall design implications.

5.2 Methods

To take into account the variability inherent in dose-response data where the response of interest is a continuous variable, we propose a bootstrap resampling procedure, which makes use of information available from the first stage data to predict responses for a given second stage design. Let y_{sjk} represent the k^{th} response observed at the j^{th} dose in the s^{th} stage of the experiment, for $s = 1, 2$, $j = 1, \dots, m_s$, and $k = 1, \dots, n_{sj}$.

Here m_s represents the number of treatment levels in the s^{th} stage of the experiment, and n_{sj} represents the number of observations taken at the j^{th} treatment level of the s^{th} stage.

The total sample size for the s^{th} stage is $n_s = \sum_{j=1}^{m_s} n_{sj}$, where n_2 is fixed prior to beginning the design search.

Let x_{sj} represent the dose associated with the j^{th} treatment level of the s^{th} stage. Let $\boldsymbol{\beta}$ represent the $p \times 1$ vector of regression parameters, such that $E\{y_{sjk}\} = \mu_{sj}(\boldsymbol{\beta}) = f(x_{sj}, \boldsymbol{\beta})$ and $Var\{y_{sjk}\} = \tau V(\mu_{sj}(\boldsymbol{\beta}))$, where $f(\bullet)$ is the nonlinear functional form of the mean response. As described in Section 2.2.3, numeric methods are used to find the maximum (quasi-) likelihood estimate $\hat{\boldsymbol{\beta}}$, and the moment estimate is

used to estimate τ . The properties of the quasi-likelihood ratio-test were described in detail in Section 2.2.4. Here we provide a general discussion of the procedure for finding the maximum quasi-likelihood ratio-based lower bound around parameter of interest ζ , where n represents the number of observations available for analysis. This procedure will be used numerous times during the design search. Prior to beginning the search, the parameter estimates and quasi-likelihood ratio-based lower bound are found using only the first stage data, so that $n = n_1$. During the design search, the parameter estimates and quasi-likelihood ratio-based lower bound are found using information from both the first and second stages, so that $n = n_1 + n_2$.

Based on the available data, let Q_{full} be the quasi-likelihood achieved under the full model, and let Q_{red} be the quasi-likelihood achieved under the reduced model. As before, M is the difference in the degrees of freedom between the full and reduced models. Recall that the quasi-likelihood ratio test statistic is statistically significant for

$$\frac{-2\{Q_{red} - Q_{full}\}}{\hat{t}M} \geq F_{M, n-p, 1-\alpha}. \quad (5.1)$$

Consider $\boldsymbol{\beta} = [\zeta \quad \boldsymbol{\theta}]$, where $\boldsymbol{\theta}$ represents the $(p-1) \times 1$ vector of model parameters, excluding the parameter of interest ζ . Let $\hat{\boldsymbol{\theta}}$ and $\hat{\zeta}$ represent the maximum quasi-likelihood estimate of $\boldsymbol{\theta}$ and ζ , respectively, found by maximizing the quasi-likelihood with respect to all model parameters simultaneously. Let $\hat{\boldsymbol{\theta}}_{\zeta}$ represent the maximum quasi-likelihood estimate of $\boldsymbol{\theta}$ associated with a fixed value of the parameter of interest,

ζ_0 . The quasi-likelihood ratio-based confidence interval around ζ is given by all ζ_0 such that

$$\left\{ \frac{-2 \left\{ Q_{red}(\zeta_0, \hat{\theta}) - Q_{full}(\zeta, \hat{\theta}) \right\}}{\hat{\tau}} < F_{1, n-p, 1-\alpha} \right\},$$

since $M=1$. Equivalently, the quasi-likelihood ratio-based confidence interval is given by all ζ_0 such that

$$\left\{ Q_{red}(\zeta_0, \hat{\theta}) > Q_{full}(\zeta, \hat{\theta}) - \frac{1}{2} \hat{\tau} F_{1, n-p, 1-\alpha} \right\}. \quad (5.2)$$

This indicates that any ζ_0 for which the associated quasi-likelihood is larger than the comparison value, given by the right-hand side of equation (5.2), is included in the confidence interval around ζ . Then the quasi-likelihood ratio-based lower bound, ζ_L , is defined as

$$\zeta_L = \left\{ \min_{\zeta_0} \ni \left[Q_{red}(\zeta_0, \hat{\theta}) = Q_{full}(\zeta, \hat{\theta}) - \frac{1}{2} \hat{\tau} F_{1, n-p, 1-\alpha} \right] \right\} \quad (5.3)$$

The objective is to maximize the lower bound of this confidence interval, ζ_L , which is our optimality criterion.

Consider a second stage design of n_2 observations taken at m_2 design points, represented by x_{2j} , with the proportion of observations taken at the j^{th} second stage design point represented by q_{2j} . Then $n_{2j} = n_2 q_{2j}$ represents the sample size allocated

to the j^{th} second stage design point. For nonlinear models, optimality criteria depend not only on the location of the design points x_{21}, \dots, x_{2m_2} , but also on the unknown parameters $\boldsymbol{\beta}$. The goal of this method is to select x_{2j} and q_{2j} in order to maximize ζ_L , which is found according to equation (5.3).

Based on hypothesized parameter values, the mean can be predicted at each potential design point. In this regard, the analysis of the first stage data plays an important role in designing the second stage experiment. Let the estimates of $\boldsymbol{\beta}$ and τ resulting from the analysis of the first stage data define the hypothesized model. Based on these parameter estimates, we can compute the predicted mean for each design point in the first stage experiment, $\hat{\mu}_{1j}(\hat{\boldsymbol{\beta}}) = f(x_{1j}, \hat{\boldsymbol{\beta}})$, as well as the standardized residual, e_{1jk} , associated with each observation. In the quasi-likelihood case, where it is assumed that $\text{Var}(y_{sjk}) = \tau V(\mu_{sj}(\boldsymbol{\beta}))$, the standardized residual of the k^{th} observation on the j^{th} dose of the first stage is defined as

$$e_{1jk} = \frac{y_{1jk} - \hat{\mu}_{1j}(\hat{\boldsymbol{\beta}})}{\sqrt{\hat{\tau} V(\hat{\mu}_{1j}(\hat{\boldsymbol{\beta}}))}}. \quad (5.4)$$

The observed standardized residuals provide important information regarding the variability of the first stage observations about the mean response.

We can randomly sample from the observed standardized residuals to create a bootstrap sample of size n_2 , and this bootstrap sample can be used to predict the response associated with each second stage observation as follows. Let the prediction for the k^{th}

observation on the j^{th} design point in the second stage experiment be represented by \tilde{y}_{2jk} , for $k = 1, \dots, n_{2j}$. The predicted mean for the j^{th} design point, x_{2j} , is $\hat{\mu}_{2j}(\hat{\boldsymbol{\beta}}) = f(x_{2j}, \hat{\boldsymbol{\beta}})$. We generate n_2 random samples of size one from the standardized residuals, with replacement, such that the predicted standardized residual associated with response \tilde{y}_{2jk} is \tilde{e}_{2jk} . Solving equation (5.4) for y , we see that the predicted response for the k^{th} observation on the j^{th} design point is

$$\tilde{y}_{2jk} = \tilde{e}_{2jk} \sqrt{\hat{t}V(\hat{\mu}_{2j}(\hat{\boldsymbol{\beta}}))} + \hat{\mu}_{2j}(\hat{\boldsymbol{\beta}}), \quad (5.5)$$

a function of the j^{th} dose level, its predicted mean and its predicted standardized residual.

Using this technique to predict responses for a given second stage design, each bootstrap sample drawn will result in a different estimate of the lower bound. We can obtain an empirical distribution for the lower bound of the confidence interval by drawing a large number, B , of bootstrap samples. We considered the median of the bootstrap distribution to represent the location of the lower bound, although other measures of central tendency may also be considered. For a given second stage design, the median of the bootstrap distribution of the lower bound is constructed using all available data, according to the following procedure.

Bootstrap Algorithm

1. Select a bootstrap sample of size n_2 from the standardized residuals.
2. Predict responses for the second stage design using equation (5.5). The complete design is represented by the $(n_1 + n_2) \times 1$ vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots & \mathbf{x}_{1m_1} & \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots & \mathbf{x}_{2m_2} \end{bmatrix}',$$

where \mathbf{x}_{1j} is the $n_{1j} \times 1$ vector representing the j^{th} dose, x_{1j} , and \mathbf{x}_{2j} is the $n_{2j} \times 1$ vector representing the j^{th} design point, x_{2j} . The corresponding response vector is

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{11} & \mathbf{y}_{12} & \cdots & \mathbf{y}_{1m_1} & \tilde{\mathbf{y}}_{21} & \tilde{\mathbf{y}}_{22} & \cdots & \tilde{\mathbf{y}}_{2m_2} \end{bmatrix}',$$

where \mathbf{y}_{1j} is the $n_{1j} \times 1$ vector of responses observed at the j^{th} dose of the first stage experiment, and $\tilde{\mathbf{y}}_{2j}$ is the $n_{2j} \times 1$ vector of predicted responses for the j^{th} design point in the second stage experiment.

3. Using the Nelder-Mead Simplex Algorithm, determine the maximum likelihood estimate, $\hat{\boldsymbol{\beta}}$, of the vector of model parameters $\boldsymbol{\beta}$, using the complete design and response vectors defined in step (2).
4. Find the lower bound of the confidence interval, ζ_L , according to equation (5.3), using the complete design and response vectors defined in step (2).
5. Repeat steps (1) through (4) a total of B times, noting the lower bound ζ_L .
6. The location of the lower bound is represented by the median of the B lower bounds, $Mdn \zeta_L$.

Use of the median of the bootstrap distribution of the lower bound as the objective function stabilizes the design search; however, there is still a fair amount of variability in the median. Because of this variability, the design search is not a simple numerical optimization problem.

The design search requires use of a direct search algorithm; we chose to use the Nelder-Mead Simplex Algorithm. This procedure involves two optimization problems – the maximization of the median of the bootstrap distribution of the lower bound and the maximization of the quasi-likelihood, which is used to find the maximum quasi-likelihood estimate of the vector of model parameters. As a result, nested Nelder-Mead Simplex Algorithms are used to find the optimal second stage design based on the quasi-likelihood ratio-based confidence interval. Because this algorithm can be somewhat dependent on starting values, a grid search of dose locations is used to find the optimal second stage design. The candidate optimal design is the design that maximizes the median of the bootstrap distribution for a particular starting design. For each potential starting design, the objective is to find the location and corresponding allocation that maximize the median. The design that maximizes the median of the bootstrap distribution of the lower bound across all candidate optimal designs is the optimal design. The procedure for finding the optimal design is defined as follows.

Design Algorithm

1. Based on the first stage data, use the Nelder-Mead Simplex Algorithm to determine the maximum likelihood estimate, $\hat{\beta}$, of the vector of model parameters β .
2. Construct the quasi-likelihood ratio-based confidence interval around ζ associated with the first stage data, according to equation (5.2).
3. Calculate the standardized residual associated with each of the first stage observations, according to equation (5.4).

4. Using the Nelder-Mead Simplex Algorithm, find the second stage design that maximizes the median of the bootstrap distribution of the lower bound, which is found according to the bootstrap algorithm described previously.
5. Record the candidate optimal design and the median ${}_{Mdn}\zeta_L$.
6. Repeat steps (4) and (5) for each potential starting design in the grid search.
7. Compare the medians across candidate optimal designs. The optimal design is the candidate optimal design that is associated with the maximum of the median of the bootstrap distribution.

The SAS/IML code used to run this procedure is given in Appendix D.

In the Simplex Algorithm, the default stopping rule associated with the convergence of the response is 10^{-8} ; this means that the algorithm will stop when, for a candidate simplex, the maximum response is no more than $10^{-6}\%$ larger than the minimum response. This stopping rule is appropriate when the objective function is a constant function of the design. Consider, for instance, the D_s - optimality criterion, which uses the variance of the dose threshold as the objective function. The variance of the threshold associated with a particular design is the same each time the algorithm encounters that design; therefore, it is reasonable to require that the change in the objective function be sufficiently small to convince the user that the algorithm has converged to its minimum.

This stopping rule is too strict, however, when the objective function is a variable function of the design. Consider the lower bound of the (quasi-) likelihood ratio-based confidence interval around parameter of interest ζ as an optimality criterion. The objective is to maximize the lower bound. This situation is complicated because the

value of the lower bound varies as a result of the bootstrap resampling procedure used to find it. As a result, the lower bound associated with a particular design, ζ_L , will not necessarily be the same each time the algorithm encounters that design. Using the median of the bootstrap distribution of the lower bound, ${}_{Mdn}\zeta_L$, smoothes the search, because the median represents the location of the lower bound for a given design. However, there is still variability in the median for a given design. As a result, it is necessary to adjust the stopping rule associated with the Nelder-Mead Simplex Algorithm. To this end, a study of the distribution of ${}_{Mdn}\zeta_L$ for a number of candidate designs may be helpful.

The default stopping rule associated with the convergence of the design is also 10^{-8} . This means that the algorithm will stop when, for a candidate simplex, the maximum value of a design parameter is no more than $10^{-6}\%$ larger than the minimum value of that parameter. Changes in the design parameters that occur beyond the third decimal place do not appreciably change the design. As a result of the variability in the median of the bootstrap distribution, however, the same design will result in a different value of the objective function. Therefore, this stopping rule is too strict as well.

In the following section, we present a dose-response analysis of the deltamethrin data described in Wolansky, Gennings and Crofton (2005). This analysis provides motivation for the method that we have proposed. In addition, we demonstrate the method using the deltamethrin data and present the second stage design resulting from the use of the quasi-likelihood ratio-based lower confidence bound as an optimality criterion.

5.3 Example: Analysis of Dose-Response Data for Deltamethrin

Deltamethrin, a pyrethroid, is a pesticide that acts primarily on the nervous system (Wolansky, Gennings and Crofton, 2005). Wolansky, Gennings and Crofton (2005) conducted a study to assess the individual dose response functions for deltamethrin and other pyrethroids. Each of the 55 rats included in the experiment was randomly allocated to one of seven deltamethrin dose groups. The response of interest was motor activity, recorded as a percentage of the control response. The observed dose-response relationship is given in Figure 5.1.

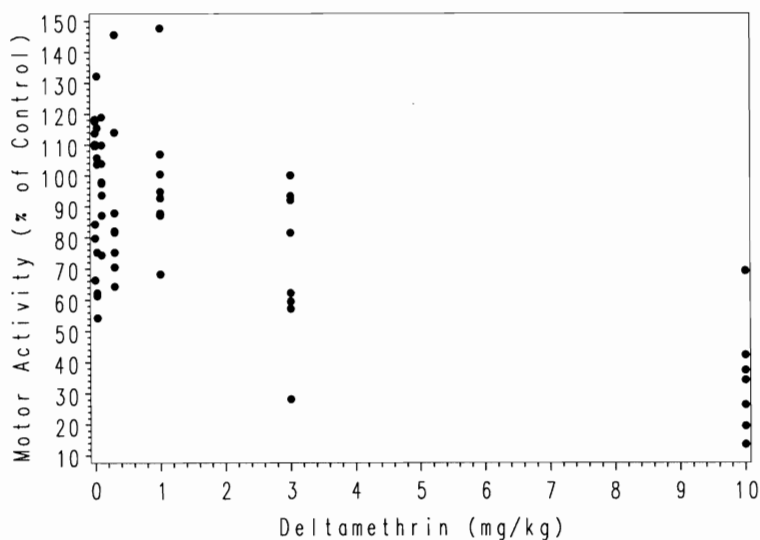


Figure 5.1. Motor Activity vs. Administered Dose of Deltamethrin

The data were analyzed using a nonlinear exponential threshold model of the form

$$\mu = \begin{cases} \alpha + \gamma, & x \leq \delta \\ \alpha + \gamma \exp\{\beta(x - \delta)\}, & x > \delta \end{cases}$$

The response variable is a percentage of the response at the control dose. An hormetic effect could be considered to model mean responses larger than 100. Using the threshold model, however, it is theoretically reasonable to restrict the mean response to percentages no larger than 100. The response does not appear to reach a minimum at 0, but at some higher percentage of the control response. Therefore, we allowed the minimum response to be estimated by the data, rather than forcing a fixed minimum response. We fit the model

$$\mu = \begin{cases} 100, & x \leq \delta \\ \alpha + (100 - \alpha) \exp\{\beta(x - \delta)\}, & x > \delta \end{cases} \quad (5.6)$$

to the deltamethrin data pictured in Figure 5.1.

In this example, let $\boldsymbol{\beta} = [\delta \quad \alpha \quad \beta]$ be the vector of model parameters. We are particularly interested in determining whether a threshold exists for this agent. Quasi-likelihood methods were used to find the maximum quasi-likelihood estimate of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}$. The resulting parameter estimates are given in Table 5.1, and the resulting parameter fit is shown in Figure 5.2.

Table 5.1. Parameter Estimates Resulting from the Fit of the Model

Parameter	Estimate	Approximate 95% Wald-type Confidence Interval
α	24.99	(-12.3, 62.2)
β	-0.22	(-0.55, 0.10)
δ	0.89	(-0.24, 2.0)

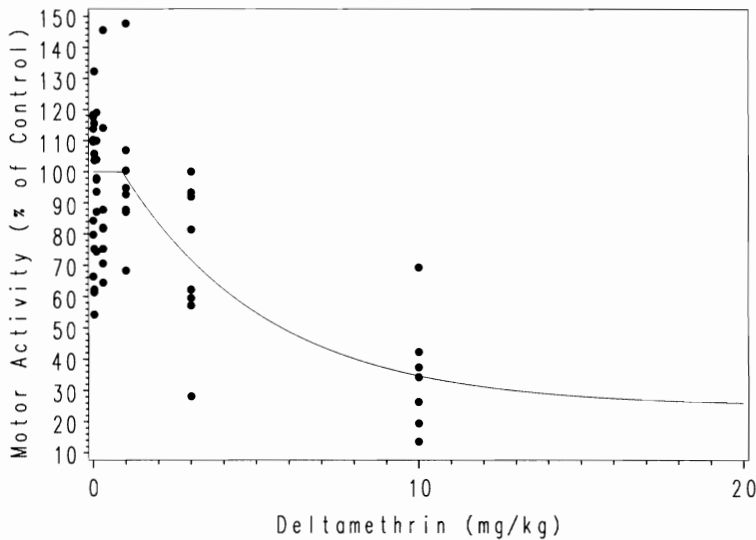


Figure 5.2. The Fit of the Exponential Threshold Model to the Deltamethrin Data

It can be seen from the plot that the model adequately represents the observed dose-response relationship. The existence of a dose threshold appears to be reasonable. The parameter estimates, however, lead to a different conclusion. As the threshold is of primary concern, consider $\beta = [\delta \ \theta]$. Then the quasi-likelihood ratio-based confidence interval can be found according to equation (5.2), where $\zeta = \delta$. The estimate of the threshold is 0.89 mg/kg; the approximate 95% quasi-likelihood ratio-based confidence interval around the threshold is $[-1.37, 3)$. The corresponding Wald-type confidence interval differs markedly, though both intervals lead us to conclude that the threshold is not statistically significant. The resulting estimate of the slope is -0.22, while the resulting estimate of the minimum response parameter is 24.99. According to our parameter estimates, the minimum response of approximately 25% will be reached at a dose larger than 30 mg/kg. The largest dose for which we have data available, however,

is 10 mg/kg. In other words, lack of data support in the region of minimum response is an obstacle to the precise estimation of these model parameters.

This problem with parameter estimation can be evidenced in our search of potential threshold values to include in the quasi-likelihood ratio-based confidence interval. To search for the approximate location of the confidence interval, we estimated the remaining parameters while holding the threshold fixed at ζ_0 . The resulting quasi-likelihood value was plotted against its corresponding threshold value, which is given in Figure 5.3. This plot is a graphical technique for finding threshold values, ζ_0 , for which equation (5.2) is satisfied. The solid line represents the confidence interval cutoff, as defined by the right side of equation (5.2). Any potential threshold value, ζ_0 , that has its corresponding quasi-likelihood above the solid line is included in the confidence interval around the threshold.

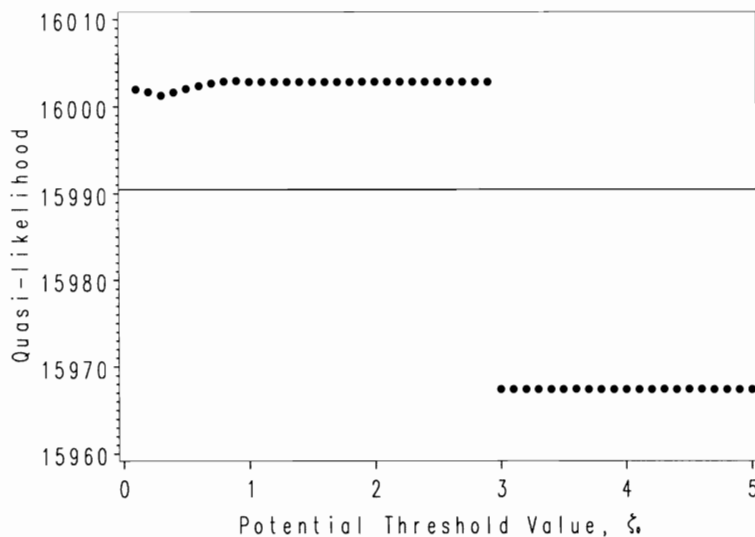


Figure 5.3. Quasi-likelihood Associated with a Fixed Value of the Threshold

As can be seen in Figure 5.3, the quasi-likelihoods associated with potential threshold values up to, but not including, three, are approximately equal. In fact, these quasi-likelihoods are equivalent to the third decimal place. Recall that the maximum quasi-likelihood estimate of the threshold is 0.89 mg/kg, which corresponds to a quasi-likelihood value of 16002.934; however, the above plot indicates that larger thresholds are approximately just as likely (the associated decrease in the quasi-likelihood is only 0.13). Again, this uncertainty in the estimate of the threshold likely stems from the fact that there is insufficient data support for the estimation of the minimum response.

Fixing the minimum response at 25%, which is the maximum quasi-likelihood estimate of the minimum response, the lower confidence bound on the dose threshold increases from -1.34 to -0.88 mg/kg. This increase is a demonstration of the impact that insufficient data support has on parameter estimation. In addition, further demonstration of the impact of insufficient data support is provided in Figure 5.4, where we give the plot of the quasi-likelihood associated with potential threshold values when the minimum response is held fixed. Again, this plot is a graphical technique for finding threshold values, ζ_0 , for which equation (5.2) is satisfied. The behavior of the quasi-likelihood in Figure 5.4 is certainly more typical than in Figure 5.3; the quasi-likelihood is largest for the maximum quasi-likelihood estimate of the dose threshold and decreases steadily as the potential threshold values increase.

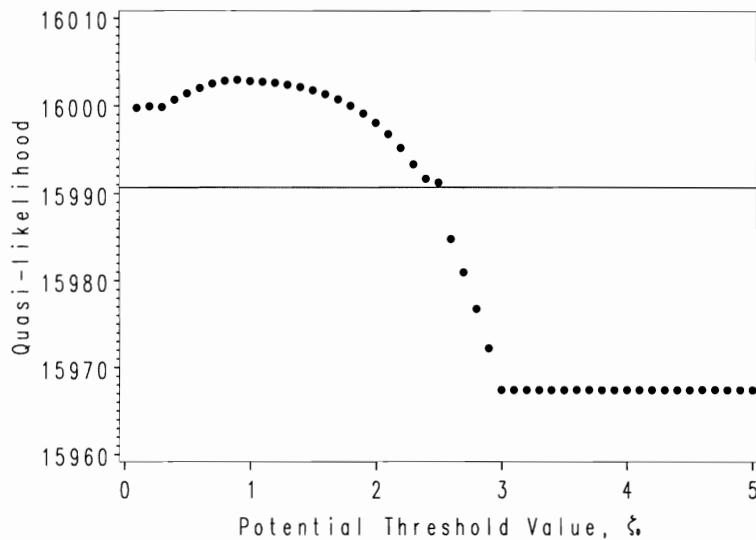


Figure 5.4. Quasi-likelihood Associated with a Fixed Value of the Threshold for Fixed Minimum Response

In order to more precisely estimate the dose threshold, we are interested in using statistical optimality criteria to design a second stage experiment. The first stage data will remain unchanged; we will use the Nelder-Mead Simplex Algorithm to determine the location and allocation of the second stage sample that maximizes the median of the bootstrap distribution the lower bound of the quasi-likelihood ratio-based confidence interval around the dose threshold. We present this design in Section 5.4 and compare it to the D- and Ds- optimal designs. In Section 5.5, we investigate the change in these designs that results from a change in the hypothesized model.

5.4 Comparison of the Second Stage Designs Associated with the D-, Ds- and Quasi-Likelihood Ratio-Based Lower Confidence Bound Optimality Criteria

We set the stopping rule associated with the convergence of the design at 10^{-2} ; this means that the algorithm will stop when, for a candidate simplex, the maximum value of a designparameter is no more than 1% larger than the minimum value of that parameter. This stopping rule is meant to ensure that, if the changes in the design parameters are such that the design is not appreciably altered, the algorithm will recognize that the design has converged.

Recall that each bootstrap sample results in a different set of predicted responses, and each set of predicted responses results in a different estimate of the lower confidence bound around the dose threshold, ζ_L . As a result, the objective function is the median of the bootstrap distribution of the lower bound. To study the empirical distribution of the median, we considered each of three candidate designs. For each candidate design, we simulated 1000 bootstrap distributions of the lower bound and recorded the median of each bootstrap distribution, which is defined as the median lower bound resulting from 100 randomly selected bootstrap samples. This study provided insight regarding the variability associated with the median, which was taken into consideration when setting the stopping rule associated with convergence of the response.

The empirical distribution of the median lower bound indicated that the magnitude of the first and third quartiles represents, on average, a 20% shift from the median of the median lower bounds. This means that we might expect that, for a candidate design, at least 50% of bootstrap distributions will result in a median lower bound that falls within

20% of the median of the median lower bounds. Due to chance, the median of the bootstrap distribution of the lower bound may vary from the median of the median lower bounds by as much as 20% in either direction. Therefore, we set the stopping rule associated with the convergence of the response at 0.2; the algorithm will stop when, for a candidate simplex, the maximum response is no more than 20% larger than the minimum response.

Because the results of Nelder-Mead can be somewhat dependent on starting values, we conducted a search over 27 sets of starting values. The second stage design consists of $n_2 = 100$ subjects. We required that the resulting design contain a control dose; that is, the first design point is always the control, though the proportion of the second stage sample allocated to the control dose is allowed to iterate within the search algorithm.

The second stage design resulting from the use of the quasi-likelihood ratio-based lower bound on the dose threshold as an optimality criterion is presented in Table 5.2. For comparison purposes, the D- and Ds- optimal second stage designs are also given in the table. In our notation, \mathbf{d} represents the second stage vector of dose locations, and the vector \mathbf{q} represents the proportion of the second stage sample allocated to each dose.

Table 5.2. Optimal Second Stage Designs for the Estimation of the Dose Threshold

Criterion	Second Stage Design: $\begin{Bmatrix} \mathbf{d} \\ \mathbf{q} \end{Bmatrix}$
Quasi-likelihood Ratio-based	$\begin{Bmatrix} 0 & 1.0 & 6.9 & 20.3 \\ .12 & .45 & .32 & .12 \end{Bmatrix}$
D- optimal	$\begin{Bmatrix} 0 & 0.9 & 6.7 & 37.6 \\ <.001 & .29 & .24 & .47 \end{Bmatrix}$
Ds- optimal	$\begin{Bmatrix} 0 & .891 & .890 & .890 \\ .005 & .41 & .36 & .24 \end{Bmatrix}$

The quasi-likelihood ratio-based second stage design allocates 12% of the second stage sample to the control group. Approximately 45% of the sample is allocated to the second design point, located just above the hypothesized dose threshold. The remaining design points are placed so as to capture information regarding the slope and the minimum response. These design characteristics are similar to those seen in the D-optimal design. Both designs place the second design point just above the hypothesized dose threshold and place remaining design points to improve the estimation of the slope and minimum response parameters. The Ds-optimal second stage design, however, places design points only in the vicinity of the hypothesized dose threshold. In Figure 5.4, we compare the location of the second stage design points, with regard to the hypothesized model and the first stage data, for the quasi-likelihood ratio-based design to the D-optimal design.

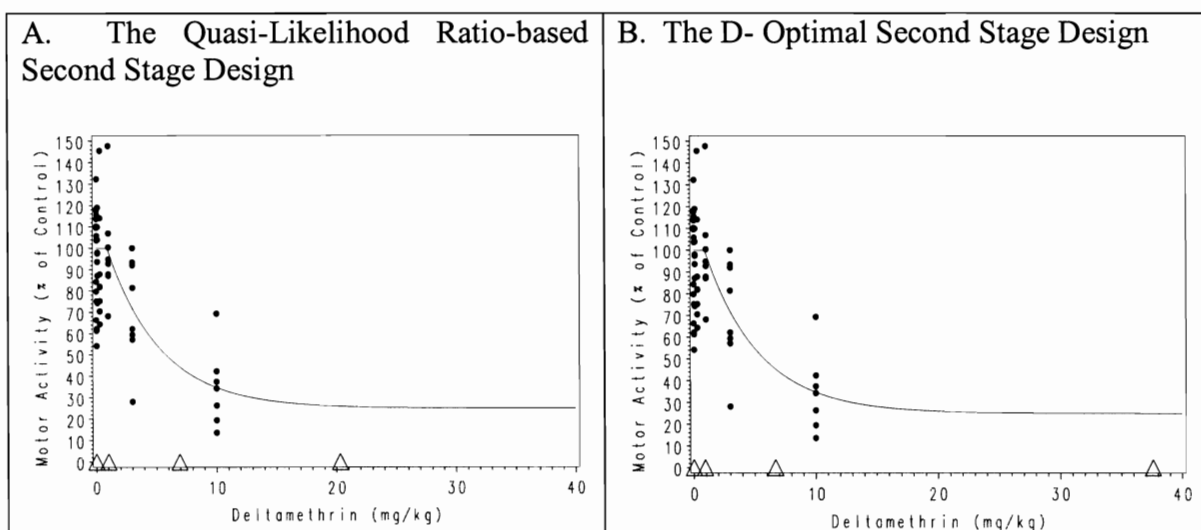


Figure 5.5. The Second Stage Designs for the Estimation of the Dose Threshold

A visual comparison of the design plots in Figure 5.5 makes evident the similarities in the quasi-likelihood ratio-based design and the D-optimal design. The primary

distinction is the location of the largest dose group. The quasi-likelihood ratio-based design places the largest dose group at 20 mg/kg, which is approximately one-half the dose associated with the largest dose group in the D- optimal design. It is important to keep in mind, however, that the D- optimal design is based on the generalized inverse of the parameter covariance matrix, while the quasi-likelihood ratio-based design uses the median of the bootstrap distribution of the lower bound on the dose threshold as the optimality criterion.

Since the preferred method of analysis is quasi-likelihood ratio-based inference, we are interested in the performance of the quasi-likelihood ratio-based confidence interval associated with these designs. To investigate, we created 1000 bootstrap samples of the standardized residuals. Each of the bootstrap samples was used to predict responses for the quasi-likelihood ratio-based design, and the lower confidence bound on the dose threshold, ζ_L , was recorded for each bootstrap sample. The process was repeated for both the D- and the Ds- optimal designs. We compared the distribution of lower bounds by creating side-by-side boxplots and histograms, which are presented in Figure 5.6.

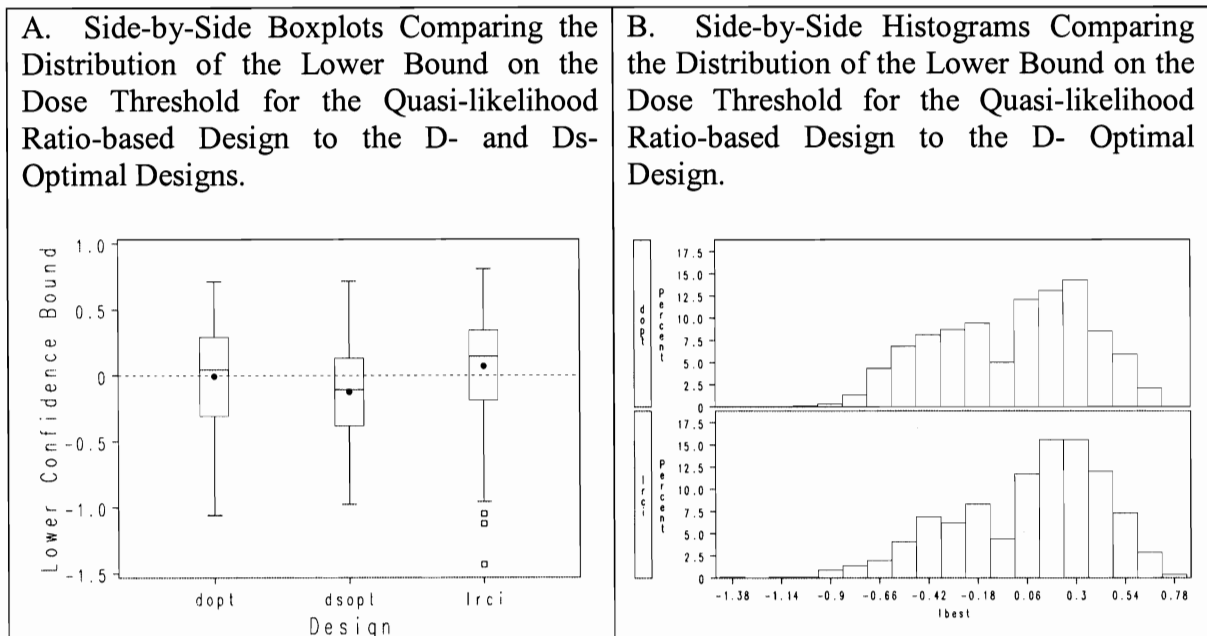


Figure 5.6. The Distribution of Lower Confidence Bounds on the Dose Threshold by Optimality Criterion.

The D- optimal design results in a median quasi-likelihood ratio-based lower confidence bound on the dose threshold of 0.04 mg/kg, whereas the quasi-likelihood ratio-based design results in a median of 0.14 mg/kg. In Figure 5.6A, it is evident that the median lower bound associated with the quasi-likelihood ratio-based design is higher than the median associated with the D- optimal design. The variability in the distributions, however, appears similar, as indicated by a visual comparison of the box width associated with these two designs. Based on this plot, one might think that the distribution resulting from the quasi-likelihood ratio-based design is simply a location shift of the distribution resulting from the D- optimal design. In Figure 5.6B, however, it is evident that the distributions are remarkably similar in location; instead, it appears that the tails of the distribution are heavier for the D- optimal design than for the quasi-likelihood ratio-based design, which results in a lower median for the D- optimal design.

Recall that insufficient data support in the region of minimum response is an obstacle to the estimation of the deltamethrin model parameters. To assess the resulting impact on the design, we found the quasi-likelihood ratio-based design where the minimum response is fixed at 25%, the maximum quasi-likelihood estimate of the minimum response. The resulting design includes a control group, which receives 3% of the sample. The design also includes dose groups at 1.3, 10.6 and 12.0 mg/kg with proportions 0.38, 0.31, and 0.28 respectively. This is a slight change from the quasi-likelihood ratio-based design given in Table 5.2, which is based on the estimation of the minimum response parameter, though it exhibits similar characteristics. In either case, the second stage design requires data just beyond the hypothesized dose threshold; the remaining design points are placed so as to obtain information regarding the other model parameters. In the case where the minimum response parameter is fixed rather than estimated, it is not necessary to have data support in the dose region of the minimum response. Instead, the design places two dose groups in the active region of the dose-response curve to support the estimation of the slope parameter.

5.5 Comparison of the Second Stage Designs Associated with the D-, Ds- and Quasi-Likelihood Ratio-Based Lower Confidence Bound Optimality Criteria under an Alternate Set of Hypothesized Parameter Values

Recall that Figure 5.3 is a plot of the quasi-likelihood achieved versus the corresponding value of the dose threshold. In terms of the quasi-likelihood, threshold values up to, but not including, 3 mg/kg are approximately as likely as the maximum

quasi-likelihood estimate, 0.89 mg/kg. This uncertainty in our estimate of the dose threshold may be the result of insufficient data support in the region of minimum response. In Figure 5.2, which demonstrates the fit of the model to the deltamethrin data, it can be seen that only two groups received doses larger than 2 mg/kg. With so few observations in the active region of the dose-response curve, it is difficult to estimate the manner in which the response changes beyond the dose threshold. This uncertainty is also problematic in terms of determining the optimal second stage design, since the design will depend on the hypothesized model.

Based on the minimal decrease in the quasi-likelihood described in Section 5.3 and shown in Figure 5.3, a dose threshold estimate of 1.8 mg/kg is approximately as likely as the maximum quasi-likelihood estimate, 0.89 mg/kg. Therefore, we let the parameter estimates corresponding to a fixed dose threshold of 1.8 mg/kg define an alternate hypothesized model. In Figure 5.7, we demonstrate the fit of this alternative model with regard to the first stage data; a visual assessment does not indicate problems with the fit of the alternative model. Under this alternative model, we found the optimal quasi-likelihood ratio-based second stage design, which we compared to the corresponding D- and Ds- optimal second stage designs. The procedure and discussion are similar to that described in Section 5.4.

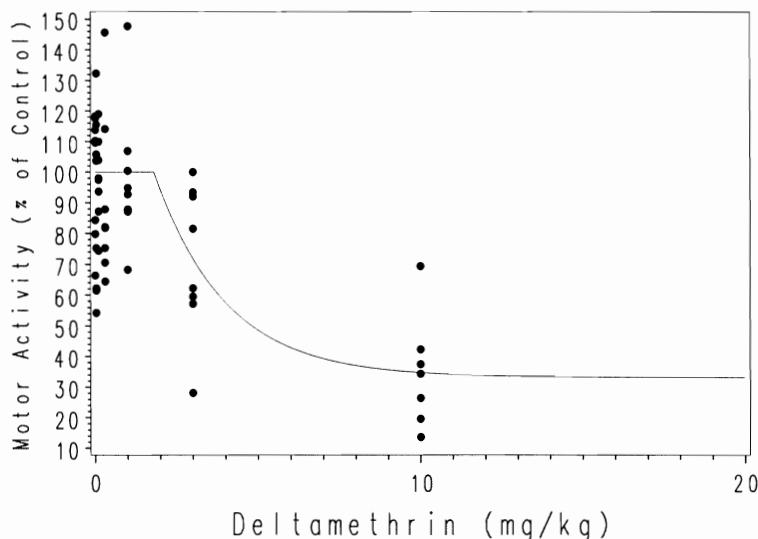


Figure 5.7. The Fit of the Alternative Model to the Deltamethrin Data

We set the stopping rule associated with the convergence of the design at 10^{-2} ; this means that the algorithm will stop when, for a candidate simplex, the maximum value of a design parameter is no more than 1% larger than the minimum value of that parameter. We conducted a study of 1000 simulated bootstrap distributions and recorded the corresponding estimates of the median lower bound to investigate the associated variability. This study indicated that the magnitude of the first and third quartiles represents, on average, a 5% shift from the median of the median lower bounds. Therefore, we set the stopping rule associated with the convergence of the response at 0.05; this means that the algorithm will stop when, for a candidate simplex, the maximum response is no more than 5% larger than the minimum response.

The second stage design resulting from the use of the quasi-likelihood ratio-based lower bound on the dose threshold as an optimality criterion is presented in Table 5.3.

For comparison purposes, the D- optimal and the Ds- optimal second stage designs are also given in the table.

Table 5.3. Optimal Second Stage Designs for the Estimation of the Dose Threshold Under the Alternative model

Criterion	Second Stage Design: $\begin{Bmatrix} \mathbf{d} \\ \mathbf{q} \end{Bmatrix}$
Quasi-likelihood Ratio-based	$\begin{Bmatrix} 0 & 1.9 & 5.1 & 17.1 \\ .08 & .50 & .35 & .07 \end{Bmatrix}$
D- Optimal	$\begin{Bmatrix} 0 & 1.8 & 4.5 & 23.8 \\ .002 & .36 & .36 & .28 \end{Bmatrix}$
Ds- Optimal	$\begin{Bmatrix} 0 & 1.61 & 1.80 & 1.81 \\ .11 & .01 & .57 & .31 \end{Bmatrix}$

The quasi-likelihood ratio-based design allocated 8% of the second stage sample to the control group. Approximately 50% of the sample is allocated to the second design point, located just above the hypothesized dose threshold. The remaining design points are placed so as to capture information regarding the slope and the minimum response. These design characteristics are similar to those seen in the D- optimal design. Both designs place the second design point just above the hypothesized dose threshold and place remaining design points to improve the estimation of the slope and minimum response parameters. The Ds- optimal second stage design, however, places design points only in the vicinity of the hypothesized dose threshold. In Figure 5.8, we compare the location of the second stage design points for the quasi-likelihood ratio-based design to the D- optimal design, with regard to the alternative model and the data from the first stage.

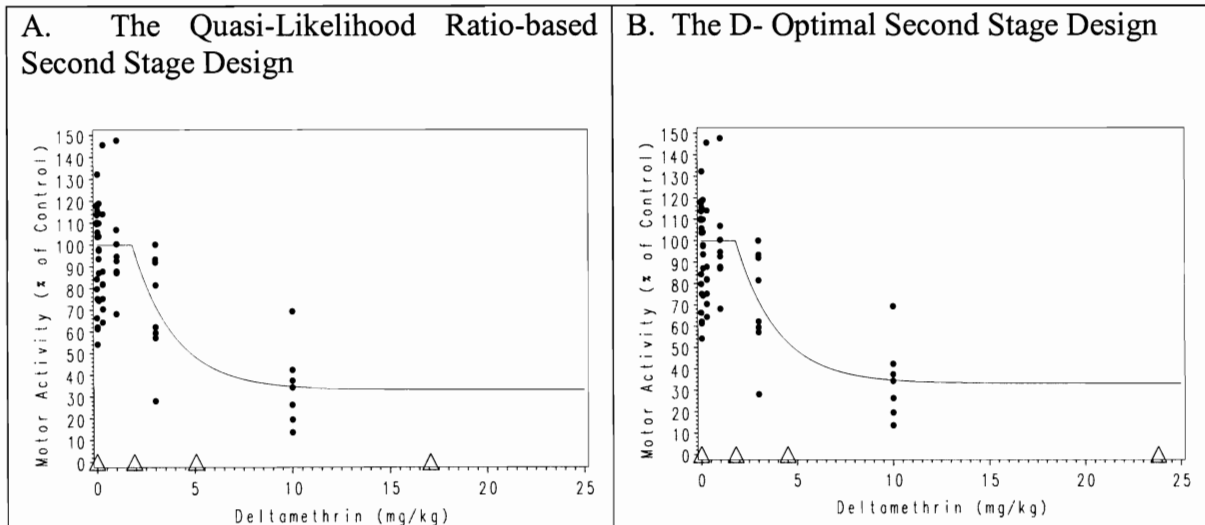


Figure 5.8. The Second Stage Designs for the Estimation of the Dose Threshold Under the Alternative Model

A visual comparison of the design plots in Figure 5.8 makes evident the similarities in the quasi-likelihood ratio-based design and the D- optimal design. As we saw in the previous section, the primary distinction is the location of the largest dose group. The quasi-likelihood ratio-based design places the largest dose group at 17 mg/kg; the largest dose group in the D- optimal design is placed slightly higher, at 24 mg/kg. Under this alternative model, the minimum response of 33% is predicted for doses larger than 16 mg/kg, a considerable shift from the previous model, which predicted a minimum response for doses larger than 30mg/kg. This explains the effect of the hypothesized model on the location of the last design point.

To investigate the performance of the quasi-likelihood ratio-based lower confidence bound associated with these designs, we created 1000 bootstrap samples of responses predicted by the alternative model and the appropriate design. We compared the

distributions of the lower confidence bound by creating side-by-side boxplots and histograms, which are presented in Figure 5.9.

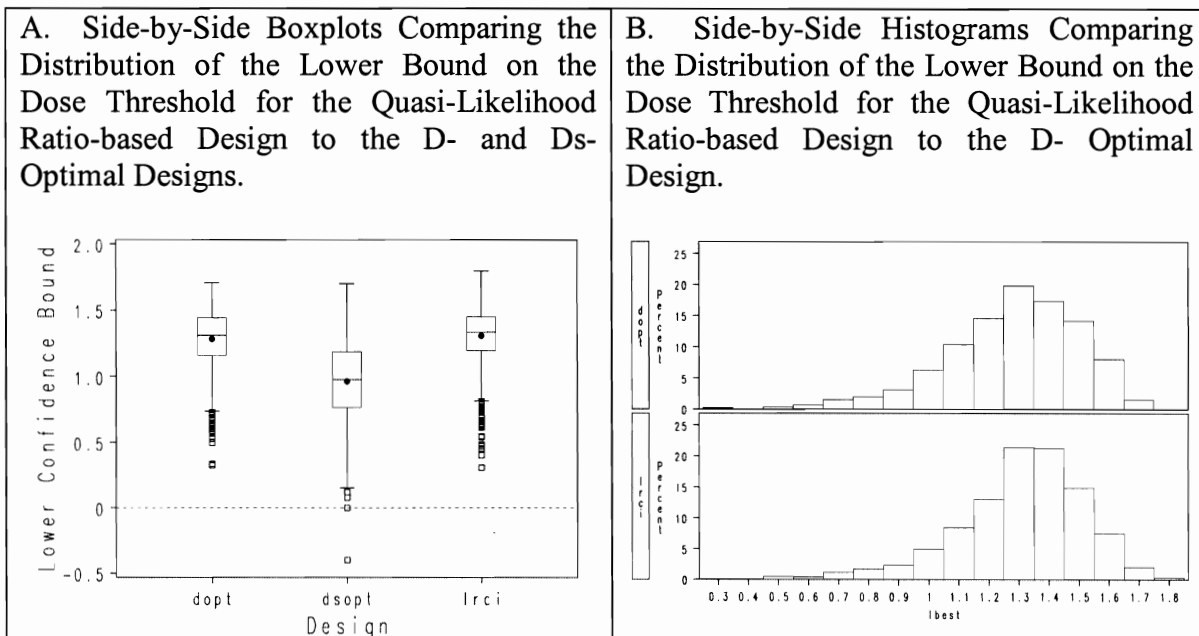


Figure 5.9. The Distribution of Lower Confidence Bounds on the Dose Threshold by Optimality Criterion.

The D- optimal design results in a median quasi-likelihood ratio-based lower confidence bound on the dose threshold of 1.31 mg/kg, whereas the quasi-likelihood ratio-based design results in a median of 1.33 mg/kg. In Figure 5.9A, it appears that the distribution of the lower bound for the D- optimal design is very similar to that of the quasi-likelihood ratio-based design; in fact, the similarity in the distributions is confirmed in Figure 5.9B.

5.6 Discussion

For nonlinear models, Wald-type testing procedures are based on a linear approximation to the mean. If the linear approximation is poor, Seber and Wild (2003) state that the associated inference may be misleading. As a result, (quasi-) likelihood ratio-based testing procedures, which are based on the nonlinear function itself, rather than a linear approximation to the nonlinear function, may be preferred. Similarly, first-order, such as D- and Ds-, optimality criteria are based on a linear approximation to the nonlinear function; if the approximation is poor, designs based on these criteria may also be misleading. Quadratic optimality criteria (O'Brien, 1992) are an improvement over first-order criteria, since they are based on a quadratic approximation. If (quasi-) likelihood ratio-based procedures are preferred, however, it is of interest to develop an optimality criterion based on these procedures.

A major obstacle to the implementation of a (quasi-) likelihood ratio-based optimality criterion is that data are required in order to construct the confidence interval. Minkin and Kundhal (1999) developed an optimality criterion based on the width of the likelihood ratio-based confidence interval. Minkin and Kundhal considered the situation in which logistic regression was to be used to analyze binomial data, in order to obtain an estimate of the ED_{50} . They restricted their attention to a symmetric, two-point design and addressed the need for data by replacing each unknown data point with its expected value. In this situation, where the response is the proportion responding, replacing the unknown data points with their expected values seems reasonable. For analyses

involving a continuous response, however, this solution is representative of the best-case scenario.

We have developed a procedure, loosely based on the method of Minkin and Kundhal described above, for using a (quasi-) likelihood ratio-based confidence interval to find the optimal second stage design. In particular, we are interested in maximizing the lower bound of the (quasi-) likelihood ratio-based confidence interval. We developed a bootstrap resampling procedure for predicting responses at a given design point; the resulting variability in the lower bound is reduced by using the median of the bootstrap distribution of the lower bound as the objective function. We have demonstrated the proposed method for improving the inference of the dose threshold in the single agent deltamethrin.

In the preceding sections, we have described a number of difficulties associated with the use of the quasi-likelihood ratio-based confidence interval as an optimality criterion. Perhaps the primary difficulty is that, in order to construct the confidence interval, it is necessary to have data at the candidate design points. The proposed method involves a bootstrap resampling technique that allows us to predict responses at the candidate design points. While this addresses the need for data, it brings about another set of obstacles.

The implementation of this method is computationally intensive. The method involves the use of nested direct search algorithms, which are handled in the following manner. The design search is conducted using the Nelder-Mead Simplex Algorithm; the objective function is the median of the bootstrap distribution of the lower bound on the dose threshold, which is based on 100 bootstrap samples. This means that, for each

candidate design, the lower confidence bound must be constructed 100 times. Each lower bound is found using another Nelder-Mead Simplex Algorithm to estimate model parameters. For large samples, complicated models, or data sets with insufficient support, significant computational time may be required to find the median of the bootstrap distribution of the lower bound for a given design. In this case, completion of the design search may be impractical, or even unfeasible.

The choice of stopping rules for the Nelder-Mead Simplex Algorithm handling the design search is somewhat subjective. We based the stopping rule associated with the convergence of the response on our study of the distribution of the median of the bootstrap distribution. Care should be taken when setting this value. If the stopping rule is set too high, then the design search will terminate without providing sufficient evidence that the response has reached its approximate maximum. If the stopping rule is set too low, the design search will continue even though the response has reached its approximate maximum.

We have demonstrated that the characteristics of the quasi-likelihood ratio-based second stage design are similar to those of the D- optimal second stage design. Both designs place the first dose just above the hypothesized dose threshold. The remaining design points are placed so as to improve estimation of the slope and minimum response parameters. The last design point is placed at a dose large enough to validate the hypothesized minimum response parameter; the remaining dose is placed in the active region of the dose-response relationship to improve estimation of the slope parameter.

Perhaps more importantly, we have shown that the distributions of the quasi-likelihood ratio-based lower bound associated with these designs are similar. As such, the D- optimal design is adequately suited for analysis by quasi-likelihood ratio-based inference methods. Considering the complexities of the implementation of the quasi-likelihood ratio-based confidence interval as an optimality criterion, it appears that the resources required to determine the quasi-likelihood ratio-based design outweigh the benefits of using the quasi-likelihood ratio-based design over the D- optimal design.

Chapter 6

Summary and Extensions

6.1 Summary

The implementation of a full factorial design, which is traditionally used to study chemical mixtures, can be very costly in terms of time and resources when the mixture of interest is complex. As a result, focus has been shifted toward the implementation of more efficient experimental designs that are still capable of showing interaction among the chemicals in a complex mixture. In addition, because the number of potential mixtures for study is astronomical, it has been suggested that research should focus on relevant mixtures, mixtures in which the components and relative proportions reflect those encountered in the environment (Teuschler et al., 2002). In these regards, the ray design is an important development in the study of complex chemical mixtures. The implementation of statistical optimality criteria when developing experimental designs will also help the investigator to focus resources in areas that are most important for achieving research objectives.

As a result of complicated biotransformation, elimination and/or repair processes, many biological systems exhibit some level of tolerance to a toxic insult (Cox, 1987). As such, it is often of interest to estimate the dose threshold, the boundary beyond which the

agent or mixture causes a change in response. In addition, there are numerous examples of dose-dependent interaction in the peer-reviewed literature. For the purposes of risk assessment, the U.S. EPA often assumes that additivity prevails in the low-dose region, whereas departure from additivity may exist for large doses. This assumption implies the presence of an interaction threshold, a boundary that separates the additivity region from the interaction region. It is often of interest to estimate the interaction threshold using toxicity data from the mixture of interest. The detection of these thresholds is an important objective in mixtures research, since the dose-response relationship evident at high doses may differ markedly from the low-dose relationship (Teuschler et al., 2002; Teuschler and Hertzberg, 1995).

This dissertation is primarily focused on the development of statistical methodology and experimental designs for the estimation of thresholds. In Chapter 2, we presented methodology for simultaneously estimating a dose threshold and an interaction threshold along a fixed-ratio ray. The novel model presented consists of three segments. Prior to the dose threshold, there is no change from the background response; between the dose threshold and the interaction threshold, the model describes an active region of additivity; beyond the interaction threshold, the model allows for departure from additivity. Inference associated with the dose and interaction thresholds model begins with a test of additivity, or a test for interaction. If the test of additivity is significant, then we conclude that there is interaction along the ray. It is then of interest to test for a region of additivity. If the interaction threshold is significantly larger than the dose threshold, we conclude that there is an active region of additivity along the ray, the location of which

can be described using the confidence intervals around the threshold parameters. The methodology was demonstrated for a mixture of nine HAAs along a fixed-ratio ray. In addition, we presented optimal design work for the dose and interaction thresholds model. The additivity region of the model receives a great deal of support from the single chemical data; as a result, it was shown that the optimal design requires mixture data only in the interaction region.

The objective of Chapter 3 was to review FSCR methodology and the analysis presented by Gennings et al. (2006) of a mixture of 18 PHAHs along a fixed-ratio ray. This mixture serves as the motivating example for the design work presented in Chapter 4. The FSCR additivity and interaction threshold models are described in Chapter 3; the statement of the additivity region in these models is implicit, which makes the parameter covariance matrix difficult to construct. As a result, Gennings et al. used quasi-likelihood ratio-based testing procedures in the analysis. In Chapter 4, we developed a method for constructing the parameter covariance matrix of the FSCR models using implicit differentiation. The methodology is demonstrated using the mixture of 18 PHAHs. We used the parameter covariance matrix to find the optimal second stage design for improving the estimation of the interaction threshold parameter. In addition, we collaborated with Dr. Kevin Crofton, a toxicologist with the U.S. EPA and the scientist who conducted the first stage experiment, to find the penalized optimal design (Parker and Gennings, 2006).

The methodology presented in Chapters 2, 3, and 4 involved the use of nonlinear models and, often, Wald-type testing procedures. However, Wald-type testing

procedures are based on a linear approximation to the nonlinear model; as a result, (quasi-) likelihood ratio-based procedures may be preferred. In Chapter 5, we developed a procedure for constructing an optimal second stage design using the lower bound of the (quasi-) likelihood ratio-based confidence interval as the optimality criterion. We proposed a bootstrap resampling technique to predict responses for a given design point; the objective function is the median of the bootstrap distribution of the lower confidence bound on the model parameter of interest. The methodology is demonstrated using a nonlinear threshold model for a single agent, where the parameter of interest is the dose threshold. We compared the resulting quasi-likelihood ratio-based design to the usual D- and Ds- optimal designs under two hypothesized models. It was found that the quasi-likelihood ratio-based design and the D- optimal designs are similar, and the quasi-likelihood ratio-based lower bound on the threshold behaves similarly for these designs as well.

6.2 Extensions

In our design work, we considered only the situation in which the design for a given stage is completely defined prior to beginning the corresponding experiment. For instance, in Chapters 4 and 5, we considered only the situation in which the second stage design is completely defined prior to beginning the second stage experiment. Alternatively, we can consider the continual reassessment method (O'Quigley, Pepe and Fisher, 1990; O'Quigley and Reiner, 1998). The continual reassessment method was proposed for use in Phase I clinical trials and is aimed toward the estimation of the

maximum tolerated dose; this sequential method uses response information available on the first $j-1$ subjects entered into the study to determine the dose for the j^{th} subject entered. In the risk assessment of chemical mixtures, the continual reassessment method may be an efficient design for the estimation of the interaction threshold.

For the fixed design scenarios considered in this work, we considered D- and Ds-optimality criteria, which are commonly seen in the literature. There are other alphabetic optimality criteria that may also be useful. In addition, we may wish to consider optimality criteria based on the variance associated with linear or nonlinear combinations of model parameters, as in the additivity optimal design discussed in Chapter 2. For nonlinear models, optimality criteria depend not only on the design but on the value of the parameter vector. Since the value of the parameter vector is unknown, one may wish to consider Bayesian techniques for determining the optimal experimental design (Chaloner and Verdinelli, 1995).

Teuschler et al. (2002) encourage the development of methods for detecting thresholds. According to the threshold model used throughout this dissertation, the dose threshold divides the dose range into two regions – prior to the dose threshold, there is no change from the background response; beyond the dose threshold, the response changes as a possibly nonlinear function of the dose. However, Calabrese (2004) noted that the past several years have shown considerable interest in hormesis, “a dose-response phenomenon characterized by a low-dose stimulation and a high-dose inhibition.” If the data indicate potential stimulation of the response in the low-dose region (ie, for growth effects) or a potential reduction of the response in the low-dose region (ie, for toxicity

effects), one might consider the implementation of a hormetic threshold model. It might also be of interest to extend the dose and interaction thresholds model to allow for a hormetic response in the low-dose region.

The U.S. EPA regularly assumes that the components of a mixture behave additively in the low-dose region, and that interactions, if they exist, are strictly a high-dose phenomenon. As a result, the interaction threshold is generally thought to separate a low-dose region of additivity from a higher dose interaction region. Further, these models allow for a single region of additivity and a single region of interaction, throughout which the direction of the interaction remains constant. However, Gennings et al. (2002) demonstrated a low-dose region of additivity, followed by alternating regions of synergism, additivity and antagonism. Since the dose-dependent nature of interactions is generally accepted, it may be of interest to develop a model with sufficient flexibility to allow for multiple regions of additivity and interaction. The data requirements for the support of such a complicated model demonstrate the usefulness of statistical optimality criteria in developing experimental designs.

As we discussed in Chapter 3, the FSCR methodology is based on the adequate modeling of the single agent dose-response data and allows the user to detect interaction among the agents in the mixture. Gennings et al. (2004, 2006) developed and demonstrated the additivity and interaction threshold models for continuous data. However, certain histopathology and developmental endpoints may be measured as ordinal responses. As a result, it may be of interest to extend the FSCR methodology to

models, such as cumulative logit or proportional odds, that are capable of handling ordinal responses and/or correlated cluster data.

The methods for detecting interaction described in this dissertation involve an underlying assumption that the single agent dose-response relationships behave in the same way; that is, it is assumed that the active single agent dose-response relationships are either all increasing or all decreasing. In microarray studies, this is often not the case; one agent may increase the expression of a particular gene, while another agent decreases the expression of the same gene. Therefore, it is of interest to develop methodology sufficiently flexible to handle single agents whose dose-response relationships reveal conflicting directions. This necessitates a more basic definition of additivity than that provided by Berenbaum's interaction index. Methodology based on the fundamental concept of an interaction, that the slope of a chemical's dose-response curve changes in the presence of another chemical, may be useful (Gennings et al., 2005).

List of References

List of References

- Agresti, A. (2002). *Categorical Data Analysis* (2nd ed). New Jersey: John Wiley & Sons, Inc.
- Berenbaum, M.C. (1985). The Expected Effect of a Combination of Agents: the General Solution. *Journal of Theoretical Biology*, 114, 413-431.
- Brunden, M.N. and Vidmar, T.J. (1989). Optimal 3 x 3 Factorial and 3-Ray Designs for Drug Interaction Experiments with Dichotomous Responses. *Communications in Statistics*, 18, 2835-2859.
- Calabrese, E.J. (2004). Hormesis: From Marginalization to Mainstream - A Case for Hormesis as the Default Dose-Response Model in Risk Assessment. *Toxicology and Applied Pharmacology*, 197, 125-136.
- Carpy, S.A., Kobel, W. and Doe, J. (2000). Health Risk of Low-Dose Pesticides Mixtures: A Review of the 1985-1998 Literature on Combination Toxicology and Health Risk Assessment. *Journal of Toxicology and Environmental Health, Part B*, 3, 1-25.
- Carter, Jr., W.H. (1995). Relating Isobolograms to Response Surfaces. *Toxicology*, 105, 181-188.
- Casey, M., Gennings, C., Carter, Jr., W.H., Moser, V.C., and Simmons, J.E. (2004). Detecting Interaction(s) and Assessing the Impact of Component Subsets in a Chemical Mixture Using Fixed-Ratio Mixture Ray Designs. *Journal of Agricultural, Biological, and Environmental Statistics*, 9, 339-361.
- (2005). D_s-Optimal Designs for Studying Combinations of Chemicals Using Multiple Fixed-Ratio Ray Experiments. *Environmetrics*, 16, 129-147.
- Chaloner, K. and Verdinelli, I. (1995). Bayesian Experimental Design: A Review. *Statistical Science*, 10, 273-304.
- Cox, C. (1987). Threshold Dose-Response Models in Toxicology. *Biometrics*, 43, 511-523.

- Crofton, K.M., Craft, E.S., Hedge, J.M., Gennings, C., Simmons, J.E., Carchman, R.A., Carter, Jr., W.H., and DeVito, M.J. (2005). Thyroid Hormone Disrupting Chemicals: Evidence for Dose-Dependent Additivity or Synergism. *Environmental Health Perspectives*, 113, 1549-1554.
- El-Masri, H.A., Tessari, J.D. and Yang, R.S.H. (1996). Exploration of an Interaction Threshold for the Joint Toxicity of Trichloroethylene and 1,1-Dichloroethylene: Utilization of a PBPK Model. *Archives of Toxicology*, 70, 527-539.
- Finney, T. (1994). *Calculus* (2nd ed). Massachusetts: Addison-Wesley Publishing Co.
- Gennings, C., Carter, Jr., W.H., Campaign, J.A., Bae, D., and Yang, R.S.H. (2002). Statistical Analysis of Interactive Cytotoxicity in Human Epidermal Keratinocytes Following Exposure to a Mixture of Four Metals. *Journal of Agricultural, Biological and Environmental Statistics*, 7, 58-73.
- Gennings, C., Carter, Jr., W.H., Carchman, R.A., DeVito, M.J., Simmons, J.E., and Crofton, K.M. (2006). The Impact of Exposure to a Mixture of Eighteen PHAHs on Thyroid Function: Estimation of an Interaction Threshold. *Journal of Agricultural, Biological and Environmental Statistics*, under review.
- Gennings, C., Carter, Jr., W.H., Carchman, R.A., Teuschler, L.K., Simmons, J.E., and Carney, E.W. (2005). A Unifying Concept for Assessing Toxicological Interactions: Changes in Slope. *Toxicological Sciences*, 88, 287-297.
- Gennings, C., Carter, Jr., W.H., Carney, E.W., Charles, G.D., Gollapudi, B.B., and Carchman, R.A. (2004). A Novel Flexible Approach for Evaluating Fixed Ratio Mixtures of Full and Partial Agonists. *Toxicological Sciences*, 80, 134-150.
- Gennings, C., Schwartz, P., Carter, Jr., W.H., and Simmons, J.E. (1997). Detection of Departures from Additivity in Mixtures of Many Chemicals with a Threshold Model. *Journal of Agricultural, Biological, and Environmental Statistics*, 2, 198-211.
- Gennings, C., Schwartz, P., Carter, Jr., W.H., and Simmons, J.E. (2000). Erratum: Detection of Departures from Additivity in Mixtures of Many Chemicals with a Threshold Model. *Journal of Agricultural, Biological, and Environmental Statistics*, 5, 257-259.
- Gessner, P.K. and Cabana, B.E. (1970). A Study of the Interaction of the Hypnotic Effects and of the Toxic Effects of Chloral Hydrate and Ethanol. *The Journal of Pharmacology and Experimental Therapeutics*, 174, 247-259.
- Hamm, A.K., Carter, Jr., W.H., and Gennings, C. (2005). Analysis of an Interaction Threshold in a Mixture of Drugs and/or Chemicals. *Statistics in Medicine*, 24, 2493-2507.

- Hamm, A.K. (2004). Analysis of an Interaction Threshold in Drug/Chemical Mixtures. Unpublished Ph.D. Dissertation.
- Könemann, W.H. and Pieters, M.N. (1996). Confusion of Concepts in Mixture Toxicology. *Food and Chemical Toxicology*, 34, 1025-1031.
- Mantel, N. (1958). An Experimental Design in Combination Chemotherapy. *Annals of the New York Academy of Sciences*, 76, 909-914.
- McCullagh, P. (1983). Quasi-likelihood Functions. *The Annals of Statistics*, 11, 59-67.
- McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models* (2nd ed). New York: Chapman and Hall.
- Minkin, S. and Kundhal, K. (1999). Likelihood-Based Experimental Design for Estimation of ED₅₀. *Biometrics*, 55, 1030-1037.
- Monosson, E. (2005). Chemical Mixtures: Considering the Evolution of Toxicology and Chemical Assessment. *Environmental Health Perspectives*, 113, 383-390.
- O'Brien, T.E. (July 2002), "Hypothesis Tests, Confidence Intervals and Common Sense," *International Conference on Teaching Statistics 6*, Cape Town, South Africa.
- O'Brien, T.E. (1992). A Note on Quadratic Designs for Nonlinear Regression Models. *Biometrika*, 79, 847-849.
- O'Quigley, J., Pepe, M., and Fisher, L. (1990). Continual Reassessment Method: A Practical Design for Phase 1 Clinical Trials in Cancer. *Biometrics*, 46, 33-48.
- O'Quigley, J. and Reiner, E. (1998). Miscellanea: A Stopping Rule for the Continual Reassessment Method. *Biometrika*, 85, 741-748.
- Parker, S.M. (2005). Solutions to Reduce Problems Associated with Experimental Designs for Nonlinear Models: Conditional Analyses and Penalized Optimal Designs. Unpublished Ph.D. Dissertation.
- Parker, S.M. and Gennings, C. (2006). Penalized Optimal Experimental Designs for Nonlinear Models. *Under review*.
- Richardson, S.D., Simmons, J.E., and Rice, G. (2002). Disinfection Byproducts: The Next Generation. *Environmental Science and Technology*, 36, 198A-205A

Schwartz, P.F., Gennings, C. and Chinchilli, V.M. (1995). Threshold Models for Combination Data from Reproductive and Developmental Experiments. *Journal of the American Statistical Association*, 90, 862-870.

Seber, G.A.F. and Wild, C.J. (2003). *Nonlinear Regression*. New York: John Wiley & Sons, Inc.

Simmons, J.E. (1995). Chemical Mixtures: Challenge for Toxicology and Risk Assessment. *Toxicology*, 105, 111-119.

Teuschler, L.K. and Hertzberg, R.C. (1995). Current and Future Risk Assessment Guidelines, Policy, and Methods Development for Chemical Mixtures. *Toxicology*, 105, 137-144.

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.

Ulm, K. (1991). A Statistical Method for Assessing a Threshold in Epidemiological Studies. *Statistics in Medicine*, 10, 341-349.

U.S. Environmental Protection Agency (2000). "Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures." USEPA EPA/630/R-00/002. National Center for Environmental Assessment, Office of Research and Development, Washington, D.C. 51 FR 34014. Available:
http://www.epa.gov/NCEA/raf/chem_mix.htm

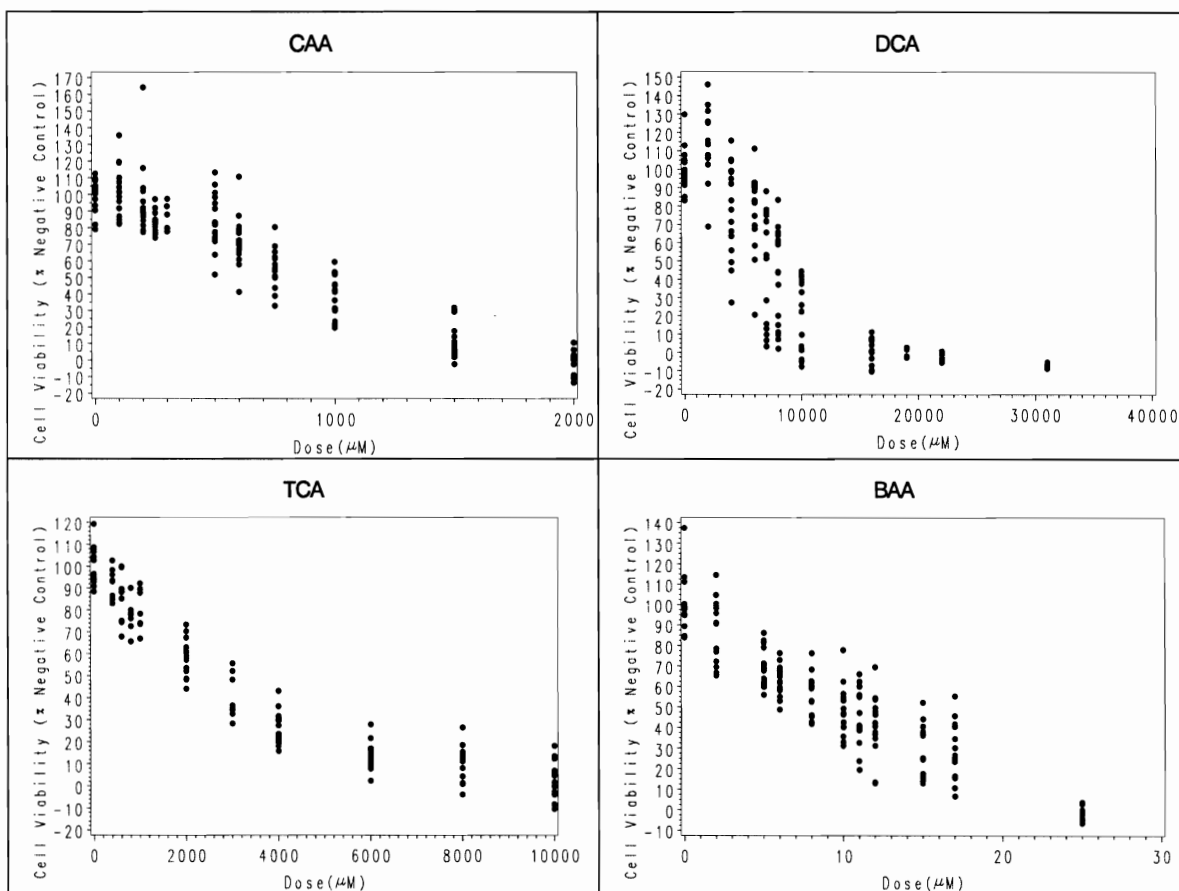
Wolansky, M.J., Gennings, C. and Crofton, K.M. (2006). Relative Potencies for Acute Effects of Pyrethroids on Motor Function in Rats. *Toxicological Sciences*, 89, 271-277.

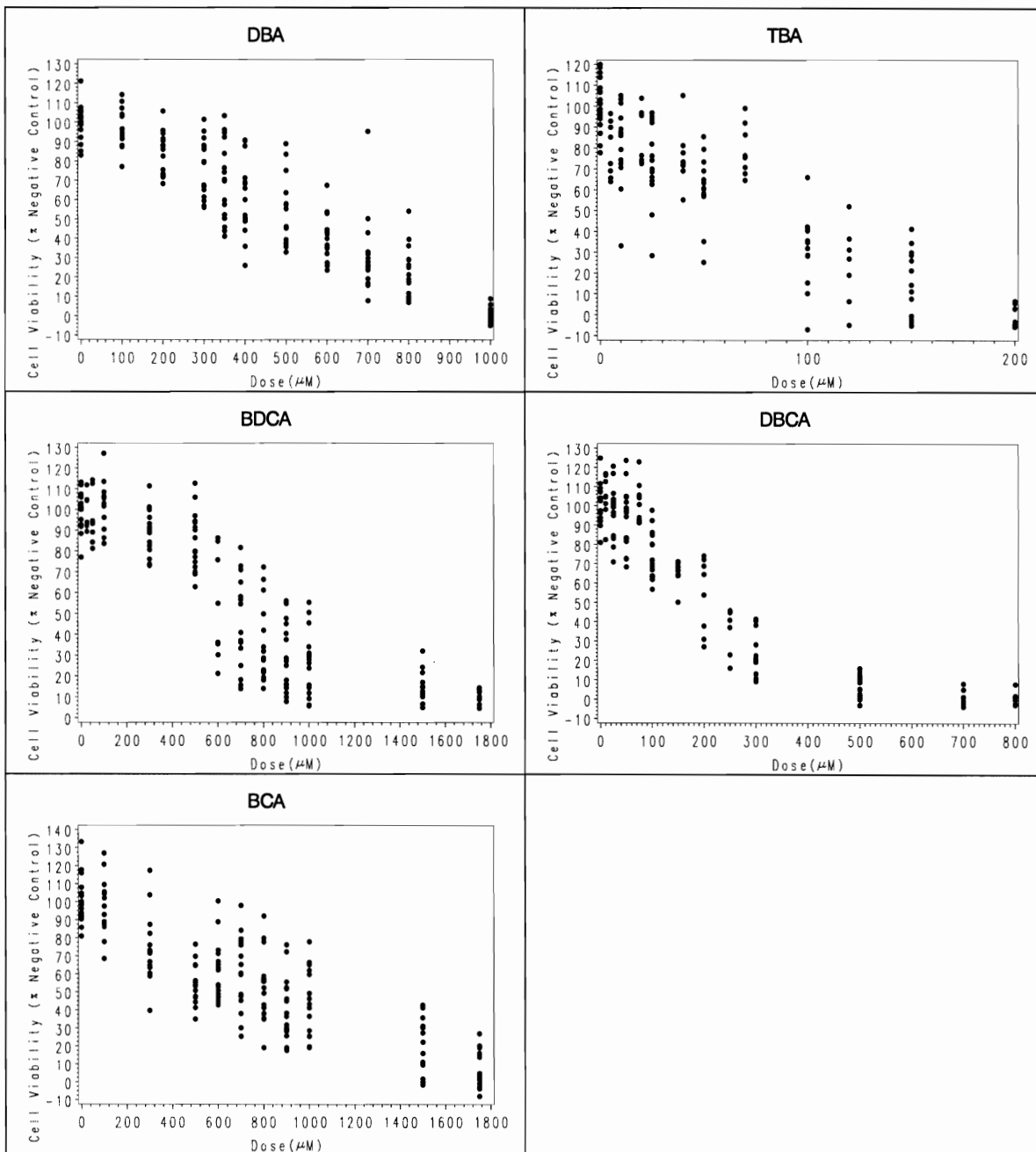
Appendices

Appendix A

Appendix A1

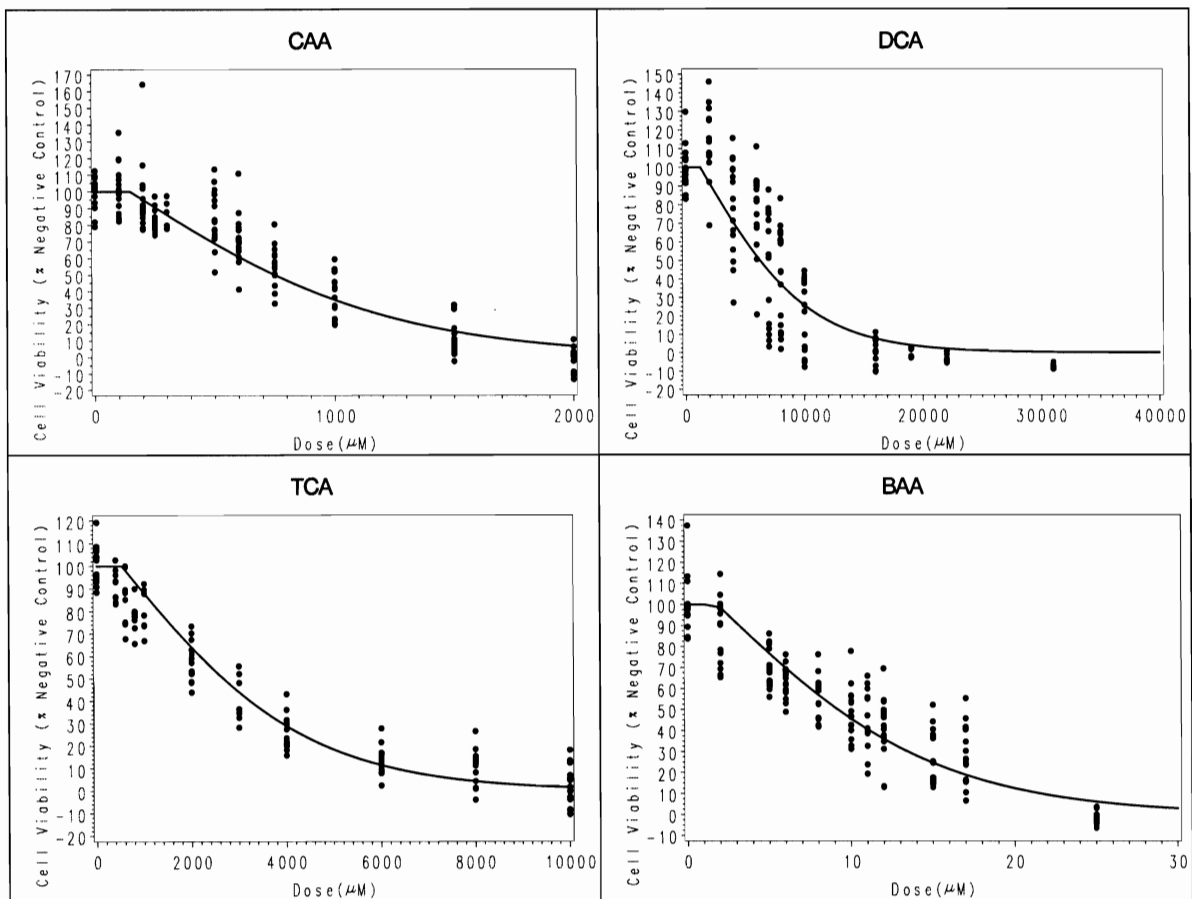
Plots of Cell Viability versus Dose for each Single Chemical in the Chlorination Ray Mixture

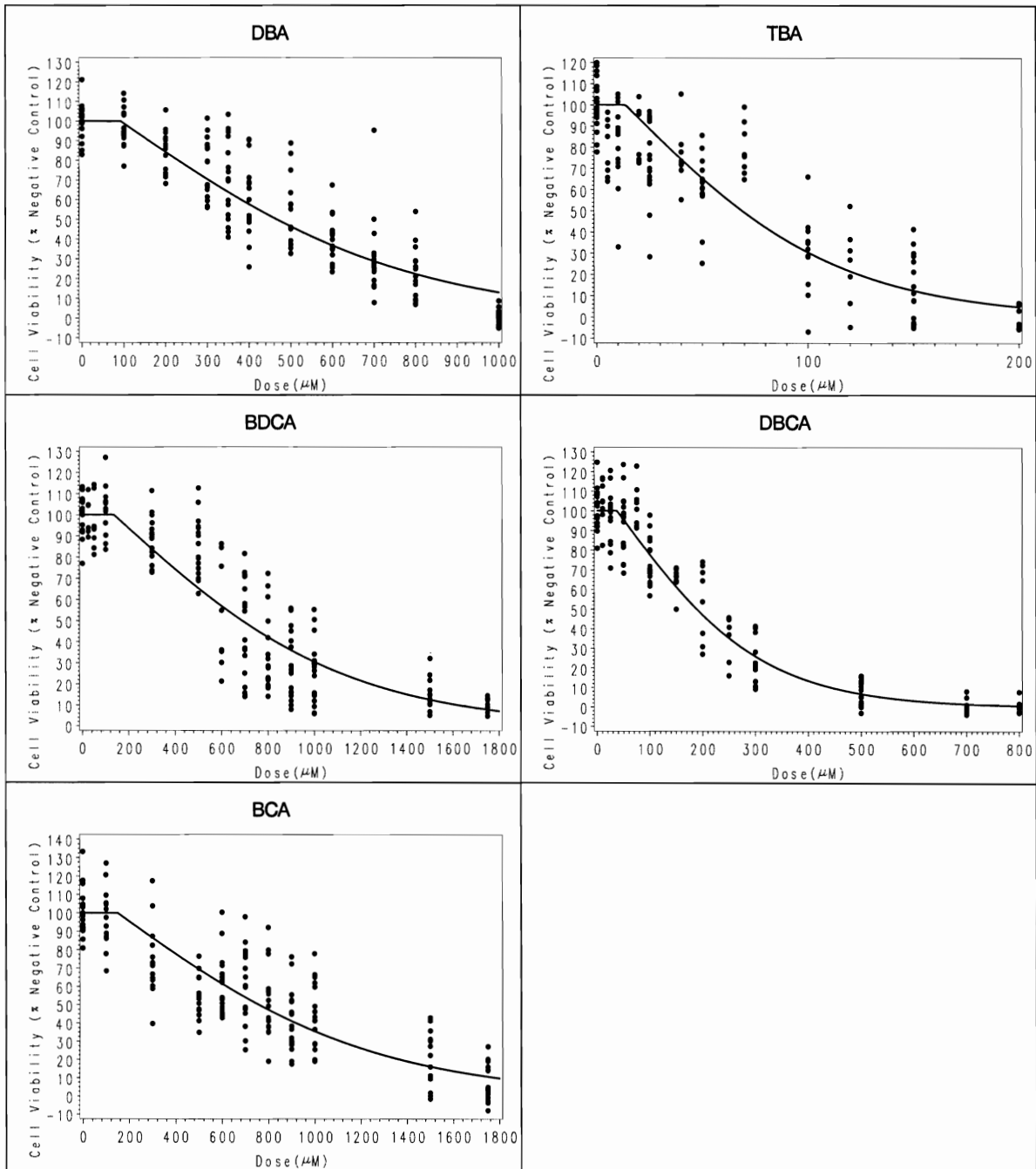




Appendix A2

Plots of Predicted Cell Viability versus Dose for each Single Chemical in the Chlorination Ray Mixture





Appendix A3

Derivatives Corresponding to the Hypothesis of Additivity

$$\mathbf{D} = \left[\frac{\partial \lambda}{\partial \boldsymbol{\beta}} \right]_{3 \times 13}$$

$$= \begin{bmatrix}
 a_1 \delta_{add} / \theta_{add}^2 & a_2 \delta_{add} / \theta_{add}^2 & \dots & a_9 \delta_{add} / \theta_{add}^2 & -1 / \theta_{add} & 0 & 0 & 1 \\
 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0
 \end{bmatrix}$$

Appendix A5**Parameter Estimates Resulting from the Fit of the Additivity Model to the Single Chemical Data**

Parameter	Estimate	Standard Error	P-value
β_1 (CA)	-18.1725	0.7918	<0.0001
β_2 (DCA)	-2.1601	0.0870	<0.0001
β_3 (TCA)	-5.1272	0.2328	<0.0001
β_4 (BA)	-1486.5	55.0268	<0.0001
β_5 (DBA)	-29.1572	1.1154	<0.0001
β_6 (TBA)	-198.6	9.3174	<0.0001
β_7 (BDCA)	-19.7544	0.7696	<0.0001
β_8 (DBCA)	-72.1209	3.5082	<0.0001
β_9 (BCA)	-18.0092	0.6873	<0.0001
δ_{add}	-0.2633	0.0271	<0.0001
τ	13.81		

Appendix A6

SAS Optimal Design Macro

```

%macro design (theta_mix=,theta_mix2=,cptl_delta=,points=,start_vals=,bign=,method=);
/*      method=1 ==> minimize determinant of entire covariance matrix.
      method=2 ==> minimize determinant of interaction-subset of covariance matrix.
      method=3 ==> minimize determinant of covariance matrix for nonlinear hypothesis of additivity.*/
start initial;

/*Specify assumptions.*/
a={0.03, 0.30, 0.26, 0.06, 0.03, 0.02, 0.12, 0.04, 0.14};
beta={      -18.1725, -2.1601, -5.1272, -1486.5, -29.1572, -198.6,
          -19.7544, -72.1209, -18.0092};
dadd=-0.2633;

/*Estimate of betas and dadd come from the fit of the additivity model to the single chemical data.*/
theta_mix=&theta_mix;
theta_mix2=&theta_mix2;
cptl_delta=&cptl_delta;
cptl_delta_new=cptl_delta*1000;
tau=14;      /*Estimate of tau comes from the fit of the additivity model to the single chemical
              data.*/
theta_add=a`*beta;
dadd_t=dadd/theta_add;
print '*****Assumed Parameter Estimates*****';
print 'Slopes for the Individual Agents' beta;
print 'Dose Threshold ' dadd_t;

```

```

print 'Theta under Additivity ' theta_add;
print 'Interaction Threshold ' cptl_delta;
print 'Theta under Mixture ' theta_mix;
print 'Theta under Mixture - quadratic term ' theta_mix2;
t_vals=j(&points,1);
q_vals=j(&points,1);
%do i=1 %to &points;
    q&i=1/&points;
    t&i=&start_vals[1,&i];
    t_vals[&i,1]=t&i;
    q_vals[&i,1]=q&i;
%end;
flag1=0; flag2=0;
print '*****Starting Values*****';
print 'Total Doses' t_vals;
print 'Proportion of Sample' q_vals;
print 'Total Sample Size Assumed ' &bign;

/*Define fixed, single chemical data.*/
use temp;
read all var {x1 x2 x3 x4 x5 x6 x7 x8 x9} into fixed_x;
singlen=nrow(fixed_x);
fixed_x1=fixed_x[,1]; fixed_x2=fixed_x[,2]; fixed_x3=fixed_x[,3];
fixed_x4=fixed_x[,4]; fixed_x5=fixed_x[,5]; fixed_x6=fixed_x[,6];
fixed_x7=fixed_x[,7]; fixed_x8=fixed_x[,8]; fixed_x9=fixed_x[,9];
fixed_xbeta=fixed_x*beta;

/*Define derivatives for fixed, single chemical data.*/
/*Model definition.*/
add=exp(-(fixed_xbeta-dadd));
muadd=100*(fixed_xbeta>dadd)
    +(200/(1+add))#(fixed_xbeta<dadd);
/*Derivatives*/
dmuadd_dbeta1=(-200)#((1+add)##(-2))#add#(-fixed_x1)#(fixed_xbeta<dadd);
dmuadd_dbeta2=(-200)#((1+add)##(-2))#add#(-fixed_x2)#(fixed_xbeta<dadd);
dmuadd_dbeta3=(-200)#((1+add)##(-2))#add#(-fixed_x3)#(fixed_xbeta<dadd);
dmuadd_dbeta4=(-200)#((1+add)##(-2))#add#(-fixed_x4)#(fixed_xbeta<dadd);
dmuadd_dbeta5=(-200)#((1+add)##(-2))#add#(-fixed_x5)#(fixed_xbeta<dadd);

```

```

dmuadd_dbeta6=(-200)#((1+add)##(-2))#add#(-fixed_x6)#(fixed_xbeta<dadd);
dmuadd_dbeta7=(-200)#((1+add)##(-2))#add#(-fixed_x7)#(fixed_xbeta<dadd);
dmuadd_dbeta8=(-200)#((1+add)##(-2))#add#(-fixed_x8)#(fixed_xbeta<dadd);
dmuadd_dbeta9=(-200)#((1+add)##(-2))#add#(-fixed_x9)#(fixed_xbeta<dadd);
dmuadd_ddadd=(-200)#((1+add)##(-2))#add#(fixed_xbeta<dadd);
/*Fixed subset of F matrix*/
fixed_f=dmuadd_dbeta1 || dmuadd_dbeta2 || dmuadd_dbeta3 ||
        dmuadd_dbeta4 || dmuadd_dbeta5 || dmuadd_dbeta6 ||
        dmuadd_dbeta7 || dmuadd_dbeta8 || dmuadd_dbeta9 ||
        dmuadd_ddadd || j(singlen,1,0) || j(singlen,1,0) ||
        j(singlen,1,0);
vmuadd=muadd+(muadd##2)/(-119);
fixed_v=diag(vmuadd);

/*Define derivatives for nonlinear hypothesis test of additivity.*/
lambda=j(4,1,0);
lambda[1]=cpt1_delta;
lambda[2]=dadd_t;
lambda[3]=theta_mix;
lambda[4]=theta_mix2;
dlambda_dbeta=j(4,13,0);
dlambda_dbeta[1,13]=1/1000;
dlambda_dbeta[2,1]=-dadd*a[1]/(theta_add**2);
dlambda_dbeta[2,2]=-dadd*a[2]/(theta_add**2);
dlambda_dbeta[2,3]=-dadd*a[3]/(theta_add**2);
dlambda_dbeta[2,4]=-dadd*a[4]/(theta_add**2);
dlambda_dbeta[2,5]=-dadd*a[5]/(theta_add**2);
dlambda_dbeta[2,6]=-dadd*a[6]/(theta_add**2);
dlambda_dbeta[2,7]=-dadd*a[7]/(theta_add**2);
dlambda_dbeta[2,8]=-dadd*a[8]/(theta_add**2);
dlambda_dbeta[2,9]=-dadd*a[9]/(theta_add**2);
dlambda_dbeta[2,10]=1/theta_add;
dlambda_dbeta[3,11]=1;
dlambda_dbeta[4,12]=1;
c=j(3,4,0); c[1,1]=1; c[1,2]=-1; c[2,3]=1; c[3,4]=1;

finish initial;
run initial;

```

```

start varcov;
  %do i=1 %to &points;
    n&i=round(&bign*q&i);
  %end;

  /*Sample Size constraints.*/
  sumn=0;
  %let temp_points=%eval(&points-1);
  %do i=1 %to &temp_points;
    sumn=sumn+n&i;
  %end;
  n&points=&bign-sumn;

  /*Define Ts.*/
  t=0;
  %do i=1 %to &points;
    if n&i>0 then t=t// j(n&i,1,t&i);
  %end;;          /*Initializes t to contain ni rows of ti
                  (in one statement)using vertical concatenation.*/
  t=t[2:&bign+1];

  /*Define Xs based on Ts.*/
  mixed_x=t*a`;
  x1=mixed_x[,1]; x2=mixed_x[,2]; x3=mixed_x[,3]; x4=mixed_x[,4];
  x5=mixed_x[,5]; x6=mixed_x[,6]; x7=mixed_x[,7]; x8=mixed_x[,8];
  x9=mixed_x[,9];
  xbeta=mixed_x*beta;

  /*Model definition.*/
  add_term=exp(-(xbeta-dadd));
  mix_term=exp(-(xbeta-dadd)-theta_mix#(t-cptl_delta)
               -theta_mix2#((t-cptl_delta)##2));
  mu=100*(t<(dadd/theta_add)
          +(200/(1+add_term))#((dadd/theta_add)<=t)#(t<cptl_delta)
          +(200/(1+mix_term))#(t>=cptl_delta));
  /*Derivatives.*/
  /*Additivity Derivatives.*/
  dadd_dbeta1=add_term#(-x1); dadd_dbeta2=add_term#(-x2);

```

```

dadd_dbeta3=add_term#(-x3); dadd_dbeta4=add_term#(-x4);
dadd_dbeta5=add_term#(-x5); dadd_dbeta6=add_term#(-x6);
dadd_dbeta7=add_term#(-x7); dadd_dbeta8=add_term#(-x8);
dadd_dbeta9=add_term#(-x9);
dadd_ddadd=add_term;
/*Mixture Derivatives.*/
dmix_dbeta1=mix_term#(-x1); dmix_dbeta2=mix_term#(-x2);
dmix_dbeta3=mix_term#(-x3); dmix_dbeta4=mix_term#(-x4);
dmix_dbeta5=mix_term#(-x5); dmix_dbeta6=mix_term#(-x6);
dmix_dbeta7=mix_term#(-x7); dmix_dbeta8=mix_term#(-x8);
dmix_dbeta9=mix_term#(-x9);
dmix_ddadd=mix_term;
dmix_dtheta_mix=mix_term#(-t+cptl_delta);
dmix_dtheta_mix2=mix_term#(-(t##2)+2#t#cptl_delta-(cptl_delta##2));
dmix_dcptl_delta_new=mix_term#
  ((theta_mix/1000) + (2#t#theta_mix2/1000) - (2#cptl_delta_new#theta_mix2/(1000##2)));
dmu_dbeta1=(-200)#((1+add_term)##(-2))#dadd_dbeta1#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta1#(t>=cptl_delta);
dmu_dbeta2=(-200)#((1+add_term)##(-2))#dadd_dbeta2#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta2#(t>=cptl_delta);
dmu_dbeta3=(-200)#((1+add_term)##(-2))#dadd_dbeta3#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta3#(t>=cptl_delta);
dmu_dbeta4=(-200)#((1+add_term)##(-2))#dadd_dbeta4#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta4#(t>=cptl_delta);
dmu_dbeta5=(-200)#((1+add_term)##(-2))#dadd_dbeta5#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta5#(t>=cptl_delta);
dmu_dbeta6=(-200)#((1+add_term)##(-2))#dadd_dbeta6#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta6#(t>=cptl_delta);
dmu_dbeta7=(-200)#((1+add_term)##(-2))#dadd_dbeta7#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta7#(t>=cptl_delta);
dmu_dbeta8=(-200)#((1+add_term)##(-2))#dadd_dbeta8#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta8#(t>=cptl_delta);
dmu_dbeta9=(-200)#((1+add_term)##(-2))#dadd_dbeta9#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))#dmix_dbeta9#(t>=cptl_delta);
dmu_ddadd=(-200)#((1+add_term)##(-2))
  #dadd_ddadd#(dadd_t<=t)#(t<cptl_delta)
  + (-200)#((1+mix_term)##(-2))
  #dmix_ddadd#(t>=cptl_delta);

```



```

dmu_dtheta_mix=(-200)#((1+mix_term)##(-2))
                                #dmix_dtheta_mix#(t>=cptl_delta);
dmu_dtheta_mix2=(-200)#((1+mix_term)##(-2))
                                #dmix_dtheta_mix2#(t>=cptl_delta);
dmu_dcptl_delta_new=(-200)#((1+mix_term)##(-2))
                                #dmix_dcptl_delta_new#(t>=cptl_delta);
iterating_f=(dmu_dbeta1 || dmu_dbeta2 || dmu_dbeta3 || dmu_dbeta4 ||
dmu_dbeta5 || dmu_dbeta6 || dmu_dbeta7 || dmu_dbeta8 ||
dmu_dbeta9 || dmu_ddadd ||
dmu_dtheta_mix || dmu_dtheta_mix2 ||dmu_dcptl_delta_new );
bigf=fixed_f // iterating_f;
vmu=mu+(mu##2)/(-119));
iterating_v=diag(vmu);
bigv=block(fixed_v,iterating_v);
p1=10;      p2=3; p=p1+p2;
inf=bigf`*inv(bigv)*bigf;
varmat=tau*ginv(inf);
detvar=det(varmat);
sub=varmat[p1+1:p,p1+1:p];
detsub=det(sub);
finish varcov;

start var_ts;
var_lambda=dlambda_dbeta*varmat*dlambda_dbeta`;
var_contrast=c*var_lambda*c`;
detcontrast=det(var_contrast);
finish var_ts;

%include simplex;
start function;
%do i=1 %to &temp_points;
%let j=%eval(&i+1);
t&j=parms[&i];
q&i=parms[&points+(&i-1)];
%end;
nparms=2*(&points)-2;
sumq=0;

```

```

%do i=1 %to &temp_points;
    sumq=sumq + q&i;
%end;
q&points=1-sumq;
test=0;
%do i=1 %to &points;
    if (t&i<0) then test=test+1;
    else if (t&i>0.1) then test=test+1;
    if (q&i<0) then test=test+1;
    else if (q&i>1) then test=test+1;
%end;
run varcov;
run var_ts;
if test>0 then do;
    fn_value=10**30;
    flag1=1;
end;
else if detvar<=0 then do;
    flag2=1;
    fn_value=10**30;
end;
else do;
    if &method=1 then fn_value=detvar;
    else if &method=2 then fn_value=detsub;
    else if &method=3 then fn_value=detcontrast;
end;
finish function;

start optima;
doses=t1
    %do i=2 %to &points;
        //t&i
    %end;;
proportions=q1
    %do i=2 %to &points;
        //q&i
    %end;;
in_parms=(doses[2:&points,1] // proportions[1:&temp_points,1]);

```

```
in_steps=in_parms#0.1;
run varcov; run var_ts;
print 'initial evaluation' varmat, detvar detsub detcontrast;

run simplex;
run varcov;
run var_ts;
print '*****The Final Design***** ';
print 'After ' count ' iterations: ';
print 'Total Doses '
      %do i=1 %to &points;
          t&i
      %end;;
print 'Proportion of Sample '
      %do i=1 %to &points;
          q&i
      %end;;
print 'Sample Size'
      %do i=1 %to &points;
          n&i
      %end;;
print 'with ' fn_value flag1 flag2, detvar detsub detcontrast;
finish optima;
run optima;
%mend design;
```

Appendix A7

Derivatives of the Mean with Respect to the Model Parameters for Use in the Optimization of the Experimental Design

$$F = \left[\frac{\partial \mu(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right]$$

$$= \left[\frac{\frac{\partial \mu(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \Big|_{\text{single chemical data } (n_s * 13)}}{\frac{\partial \mu(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \Big|_{\text{mixture data } (n * 13)}} \right]$$

Let

$$\mu = 100(t < \delta_{add_t}) + 200 \left[1 + \exp \left\{ - \left(\sum_{i=1}^c \beta_i x_i - \delta_{add} \right) \right\} \right]^{-1} (\delta_{add_t} \leq t) \left(t < \frac{\Delta_{new}}{1000} \right)$$

$$+ 200 \left[1 + \exp \left\{ - \left(\sum_{i=1}^c \beta_i x_i - \delta_{add} \right) - \theta_{mix} \left(t - \frac{\Delta_{new}}{1000} \right) \right\} - \theta_{mix2} \left(t - \frac{\Delta_{new}}{1000} \right)^2 \right]^{-1} \left(t \geq \frac{\Delta_{new}}{1000} \right)$$

$$= 100(t < \delta_{add_t}) + 200 [1 + add_term]^{-1} (\delta_{add_t} \leq t) \left(t < \frac{\Delta_{new}}{1000} \right)$$

$$+ 200 [1 + mix_term]^{-1} \left(t \geq \frac{\Delta_{new}}{1000} \right)$$

where

$$add_term = \exp \left\{ - \left(\sum_{i=1}^c \beta_i x_i - \delta_{add} \right) \right\},$$

$$mix_term = \exp \left\{ - \left(\sum_{i=1}^c \beta_i x_i - \delta_{add} \right) - \theta_{mix} \left(t - \frac{\Delta_{new}}{1000} \right) - \theta_{mix2} \left(t - \frac{\Delta_{new}}{1000} \right)^2 \right\}.$$

Then the derivatives of the mean with respect to the model parameters are as follows.

$$\frac{\partial \mu}{\partial \beta_i} = -200(1_add_term)^{-2}(add_term)(-x_i)(t \geq \delta_{add_t})(t < \Delta) \\ - 200(1 + mix_term)^{-2}(mix_term)(-x_i)$$

$$\frac{\partial \mu}{\partial \delta_{add}} = -200(1_add_term)^{-2}(add_term)(+1)(t \geq \delta_{add_t})(t < \Delta) \\ - 200(1 + mix_term)^{-2}(mix_term)(+1)$$

$$\frac{\partial \mu}{\partial \theta_{mix}} = -200(1 + mix_term)^{-2}(mix_term)\left(-t + \frac{\Delta_{new}}{1000}\right) \\ = -200(1 + mix_term)^{-2}(mix_term)(-t + \Delta)$$

$$\frac{\partial \mu}{\partial \theta_{mix2}} = -200(1 + mix_term)^{-2}(mix_term)\left(-t^2 + 2t \frac{\Delta_{new}}{1000} - \frac{\Delta_{new}^2}{1000^2}\right) \\ = -200(1 + mix_term)^{-2}(mix_term)(-t^2 + 2t\Delta - \Delta^2)$$

$$\frac{\partial \mu}{\partial \Delta_{new}} = -200(1 + mix_term)^{-2}(mix_term)\left(\frac{\theta_{mix}}{1000} + \frac{2t\theta_{mix2}}{1000} - \frac{2\theta_{mix2}\Delta_{new}}{1000^2}\right)$$

Appendix A8

Derivatives Corresponding to the Hypothesis of Additivity in the Reparameterized Model

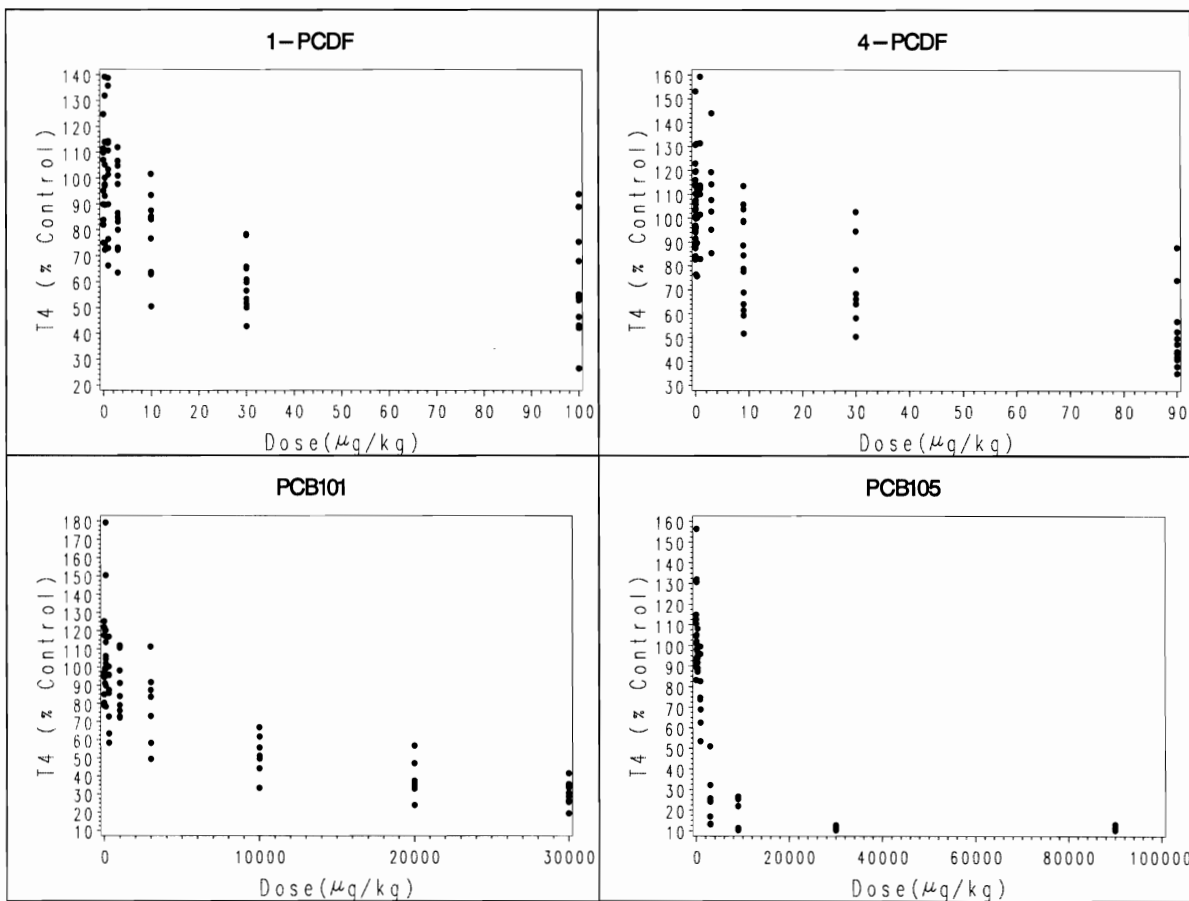
$$\mathbf{D} = \left[\frac{\partial \lambda}{\partial \boldsymbol{\beta}} \right]_{3 \times 13}$$

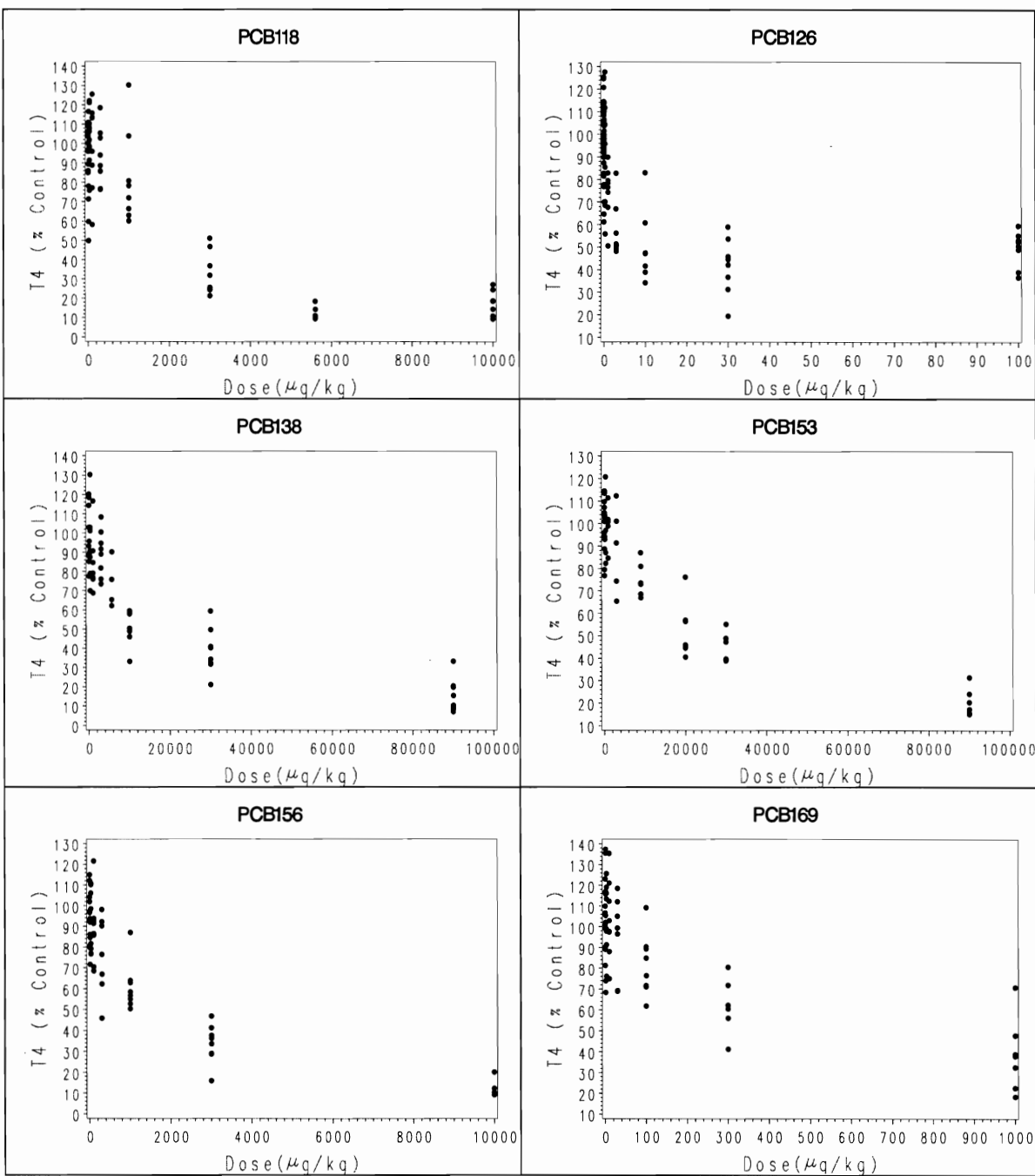
$$= \begin{bmatrix}
 a_1 \delta_{add} / \theta_{add}^2 & a_2 \delta_{add} / \theta_{add}^2 & \dots & a_9 \delta_{add} / \theta_{add}^2 & -1 / \theta_{add} & 0 & 0 & 1 / 1000 \\
 0 & 0 & \dots & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0
 \end{bmatrix}$$

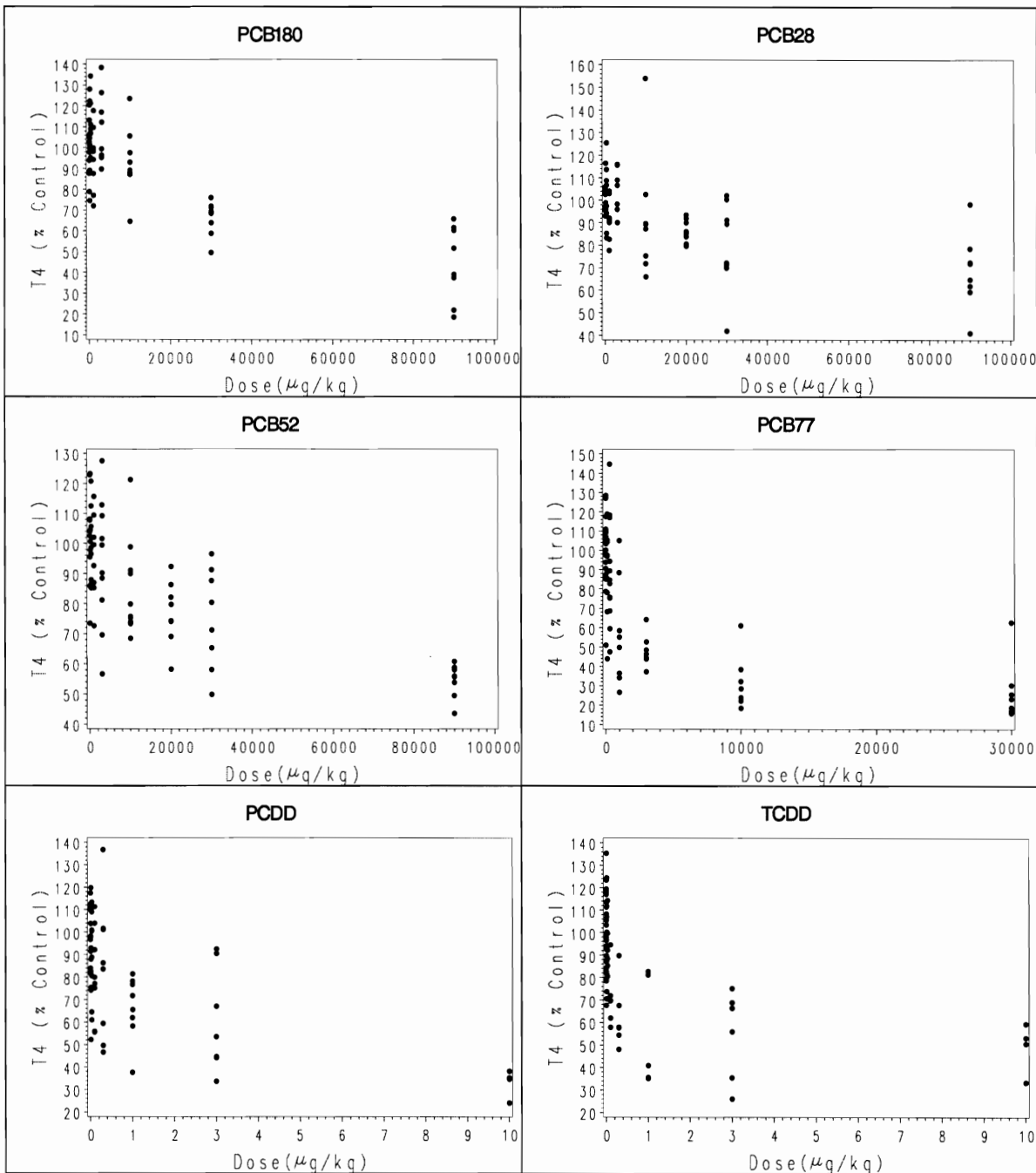
Appendix B

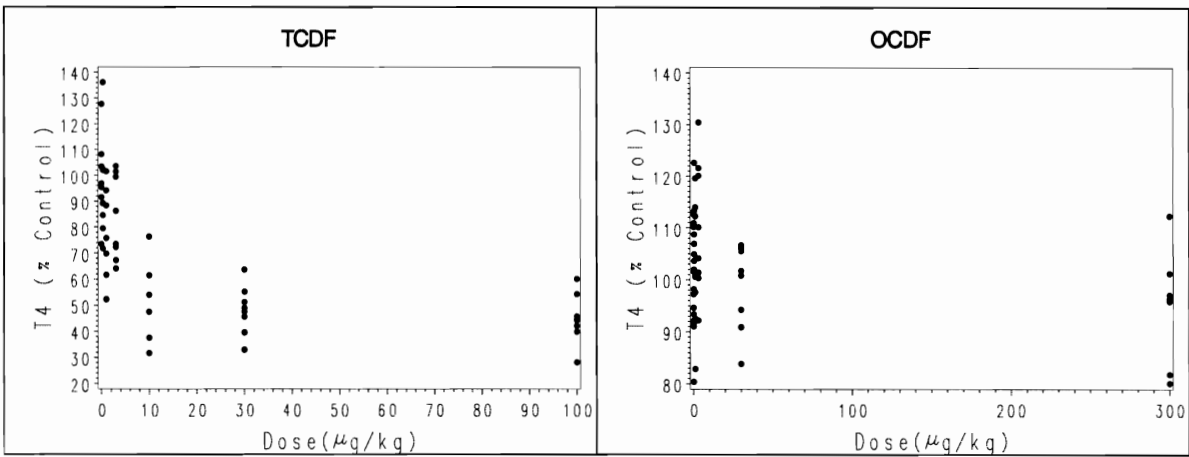
Appendix B1

Plots of T4 versus Dose for each Single Chemical in the Mixture









Appendix B2

Parameter Estimates Resulting from the full FSCR model

Parameter	Estimate	StdErr	pvalue	LowerCL	UpperCL
α_1	58.2612	1.8390	<0.0001	54.6534	61.8690
α_2	45.4477	26.7469	0.0895	-7.0256	97.9210
α_3	29.1118	4.3513	<0.0001	20.5752	37.6484
α_4	12.6959	1.3446	<0.0001	10.0580	15.3338
α_5	14.8527	2.2491	<0.0001	10.4403	19.2651
α_6	47.3391	2.0981	<0.0001	43.2230	51.4553
α_7	15.7056	2.9150	<0.0001	9.9868	21.4244
α_8	18.4523	4.2413	<0.0001	10.1315	26.7730
α_9	12.6602	3.6134	0.0005	5.5714	19.7491
α_{10}	34.1507	6.1429	<0.0001	22.0993	46.2022
α_{11}	42.1307	5.9563	<0.0001	30.4454	53.8160
α_{12}	68.1106	6.8620	<0.0001	54.6484	81.5728
α_{13}	51.3040	7.4068	<0.0001	36.7729	65.8350
α_{14}	29.3864	0.0987	<0.0001	29.1928	29.5800
α_{15}	26.9772	13.3387	0.0433	0.8089	53.1456
α_{16}	53.4973	0.0400	<0.0001	53.4189	53.5758
α_{17}	46.5578	0.00134	<0.0001	46.5552	46.5604
α_{mix}	42.6365	0.0279	<0.0001	42.5818	42.6911
β_1	-0.0968	0.0353	0.0062	-0.1661	-0.0276
β_2	-0.0232	0.0252	0.3577	-0.0726	0.0262
β_3	-0.00011	0.000028	<0.0001	-0.00016	-0.00006
β_4	-0.00065	0.000099	<0.0001	-0.00084	-0.00045
β_5	-0.00071	0.000147	<0.0001	-0.00100	-0.00042
β_6	-0.5719	0.2039	0.0051	-0.9719	-0.1720

Parameter	Estimate	StdErr	pvalue	LowerCL	UpperCL
β_7	-0.00005	8.151E-6	<0.0001	-0.00007	-0.00003
β_8	-0.00004	7.736E-6	<0.0001	-0.00005	-0.00002
β_9	-0.00046	0.000077	<0.0001	-0.00061	-0.00031
β_{10}	-0.00288	0.00115	0.0123	-0.00513	-0.00063
β_{11}	-0.00004	0.000018	0.0293	-0.00007	-3.93E-6
β_{12}	-0.00004	0.000034	0.2966	-0.00010	0.000031
β_{13}	-0.00003	0.000012	0.0199	-0.00005	-4.57E-6
β_{14}	-0.00056	0.000104	<0.0001	-0.00077	-0.00036
β_{15}	-0.2204	0.1111	0.0474	-0.4384	-0.00251
β_{16}	-8.0362	3.8720	0.0381	-15.6325	-0.4398
β_{17}	-0.3030	0.1727	0.0795	-0.6418	0.0357
β_{mix}	-0.00088	0.000142	<0.0001	-0.00116	-0.00060
δ_1	0.8573	1.3489	0.5252	-1.7890	3.5036
δ_2	-1.6113	4.8317	0.7388	-11.0904	7.8678
δ_3	41.2927	441.8	0.9256	-825.5	908.1
δ_4	262.8	103.8	0.0115	59.0520	466.5
δ_5	674.8	150.0	<0.0001	380.6	969.0
δ_6	-0.00249	0.2454	0.9919	-0.4840	0.4790
δ_7	-1066.0	867.3	0.2192	-2767.4	635.4
δ_8	-140.2	1350.3	0.9173	-2789.2	2508.8
δ_9	-127.3	96.6988	0.1882	-317.0	62.3874
δ_{10}	3.5751	35.0098	0.9187	-65.1086	72.2589
δ_{11}	6936.9	3646.4	0.0573	-216.8	14090.5
δ_{12}	2042.4	9966.8	0.8377	-17510.9	21595.7
δ_{13}	-494.2	2457.4	0.8407	-5315.2	4326.9
δ_{14}	-71.5264	98.4166	0.4675	-264.6	121.6
δ_{15}	-0.5156	0.4838	0.2868	-1.4648	0.4336
δ_{16}	0.00670	0.0270	0.8043	-0.0463	0.0597
δ_{17}	1.7817	0.8979	0.0474	0.0203	3.5432
δ_{mix}	50.2372	80.8661	0.5346	-108.4	208.9

Appendix B3

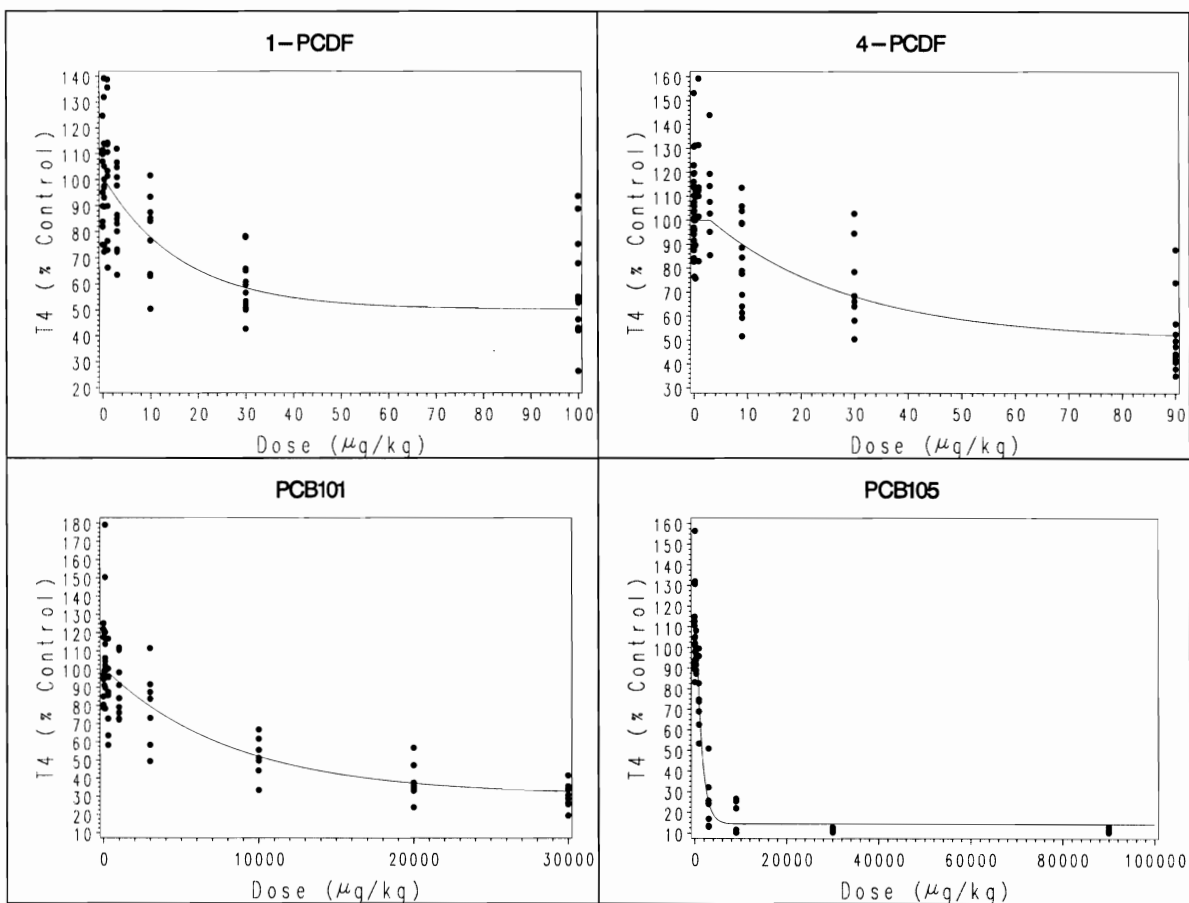
Parameter Estimates Resulting from the full FSCR model with combined minimum response parameters

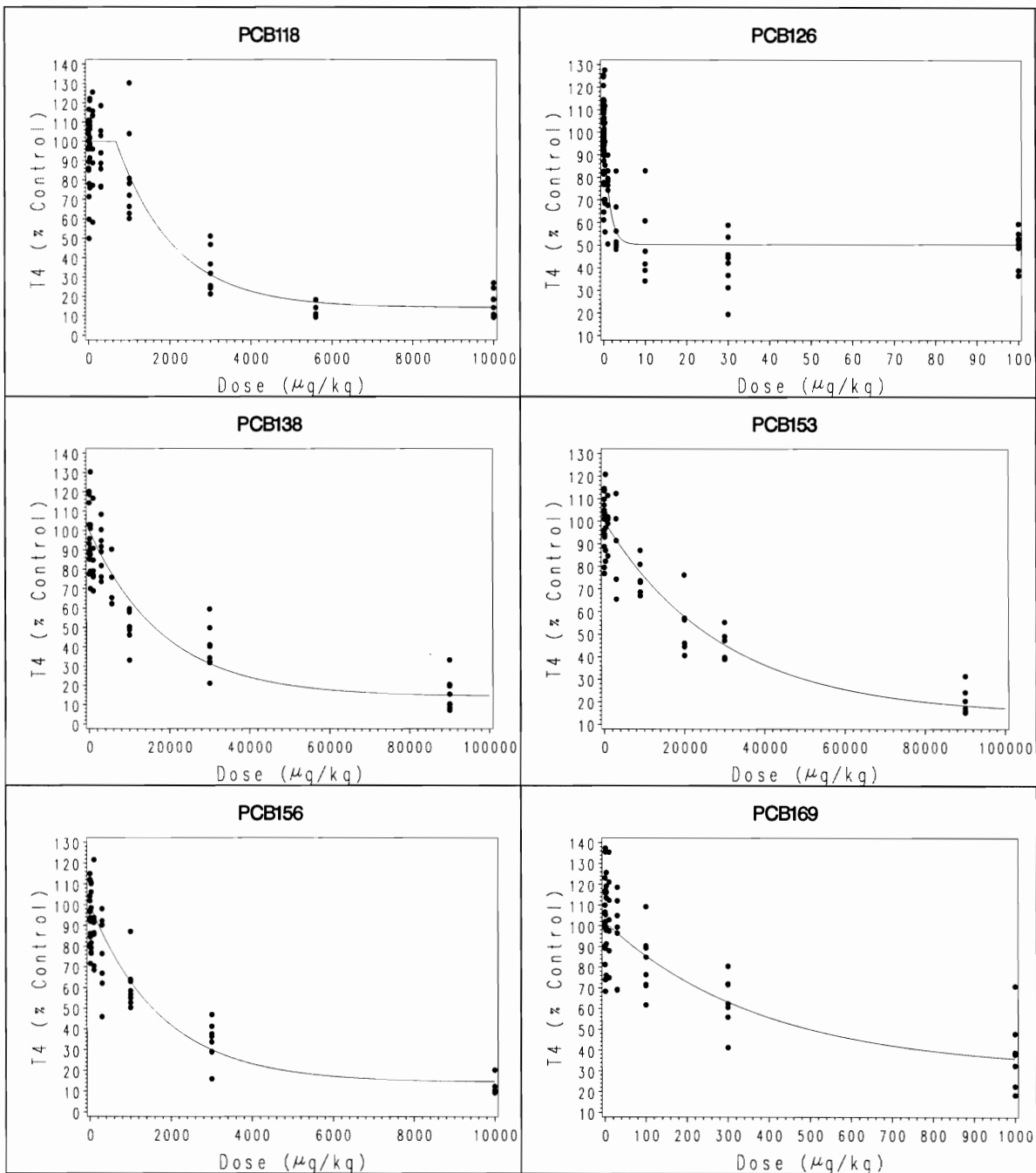
Parameter	Estimate	StdErr	pvalue	LowerCL	UpperCL
$\alpha_{1_2_6_12_13_16_17}$	51.5270	0.8855	<0.0001	49.7898	53.2643
$\alpha_{3_10_11_14_15}$	28.9070	2.0848	<0.0001	24.8171	32.9970
$\alpha_{4_5_7_8_9}$	14.0382	0.9949	<0.0001	12.0863	15.9901
α_{mix}	42.6607	0.8201	<0.0001	41.0519	44.2695
β_1	-0.0809	0.0203	<0.0001	-0.1207	-0.0411
β_2	-0.0352	0.0115	0.0023	-0.0578	-0.0126
β_3	-0.00011	0.000018	<0.0001	-0.00015	-0.00008
β_4	-0.00067	0.000067	<0.0001	-0.00080	-0.00053
β_5	-0.00066	0.000064	<0.0001	-0.00079	-0.00054
β_6	-0.5593	0.4234	0.1867	-1.3899	0.2713
β_7	-0.00005	5.248E-6	<0.0001	-0.00006	-0.00004
β_8	-0.00003	3.633E-6	<0.0001	-0.00004	-0.00003
β_9	-0.00049	0.000059	<0.0001	-0.00060	-0.00037
β_{10}	-0.00230	0.000402	<0.0001	-0.00309	-0.00151
β_{11}	-0.00002	3.923E-6	<0.0001	-0.00003	-9.99E-6
β_{12}	-0.00001	4.827E-6	0.0022	-0.00002	-5.34E-6
β_{13}	-0.00003	6.207E-6	<0.0001	-0.00004	-0.00001
β_{14}	-0.00044	0.000107	<0.0001	-0.00065	-0.00023
β_{15}	-0.2092	0.0442	<0.0001	-0.2960	-0.1225
β_{16}	-7.1871	0.7165	<0.0001	-8.5927	-5.7815
β_{17}	-0.2359	0.2170	0.2773	-0.6617	0.1899
β_{mix}	-0.00088	0.000026	<0.0001	-0.00093	-0.00083
δ_1	1.3862	1.2492	0.2674	-1.0646	3.8369

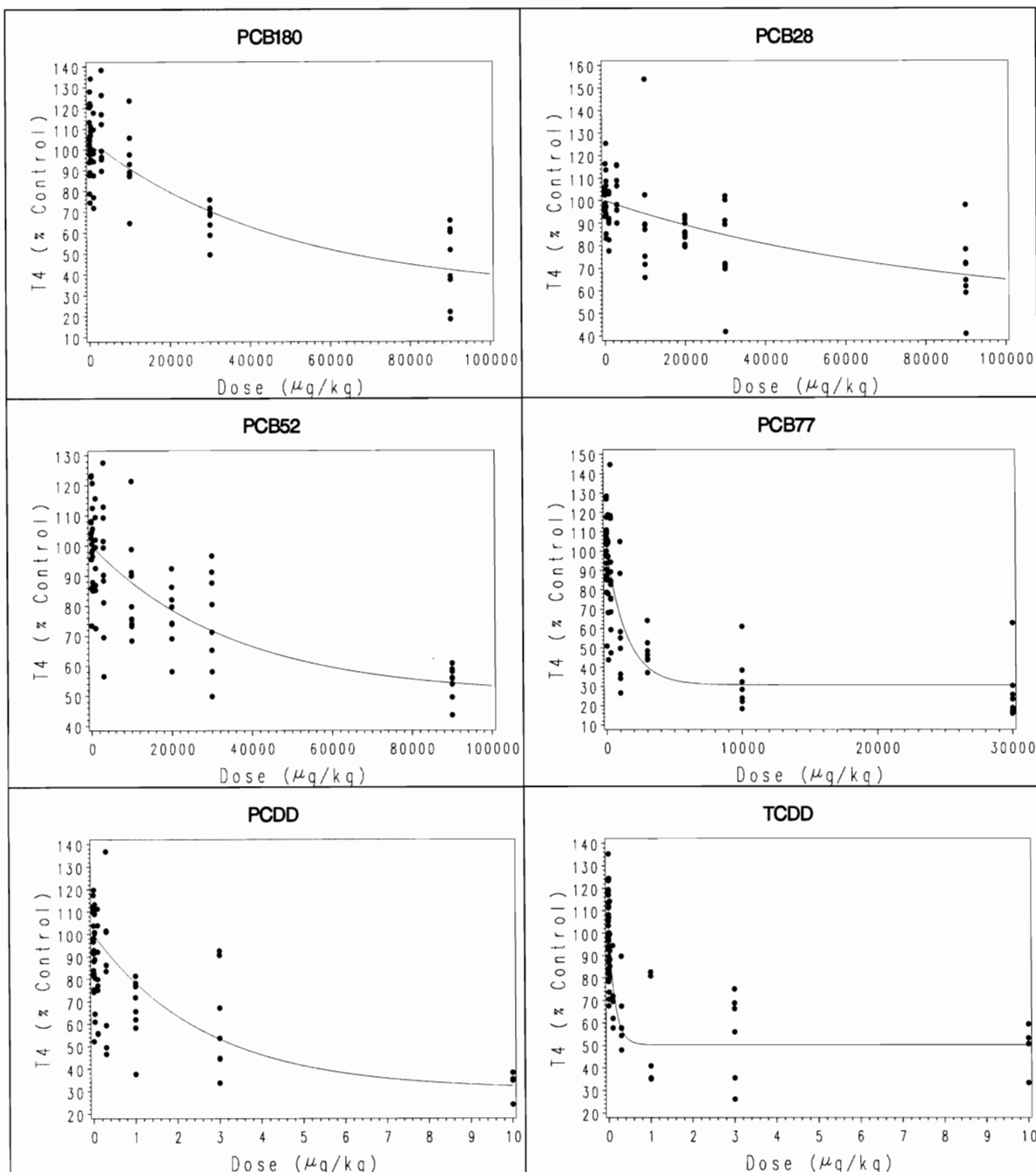
Parameter	Estimate	StdErr	pvalue	LowerCL	UpperCL
δ_2	2.1342	3.3684	0.5265	-4.4741	8.7424
δ_3	110.3	137.4	0.4224	-159.3	379.8
δ_4	286.9	84.1765	0.0007	121.7	452.0
δ_5	641.1	58.8314	<0.0001	525.7	756.5
δ_6	-0.0220	1.0843	0.9838	-2.1493	2.1053
δ_7	-1858.9	848.4	0.0286	-3523.4	-194.5
δ_8	-549.4	1106.2	0.6195	-2719.6	1620.8
δ_9	-232.3	94.0804	0.0137	-416.9	-47.7519
δ_{10}	7.5555	22.7472	0.7398	-37.0706	52.1815
δ_{11}	3597.7	5086.4	0.4795	-6380.8	13576.2
δ_{12}	1952.2	5714.8	0.7327	-9259.1	13163.5
δ_{13}	412.1	2812.8	0.8835	-5106.2	5930.4
δ_{14}	-101.1	767.5	0.8952	-1606.8	1404.6
δ_{15}	-0.8081	0.3200	0.0117	-1.4358	-0.1804
δ_{16}	0.0146	0.0160	0.3604	-0.0167	0.0460
δ_{17}	-0.1080	1.0836	0.9206	-2.2338	2.0178
δ_{mix}	49.4944	3.0264	<0.0001	43.5571	55.4317

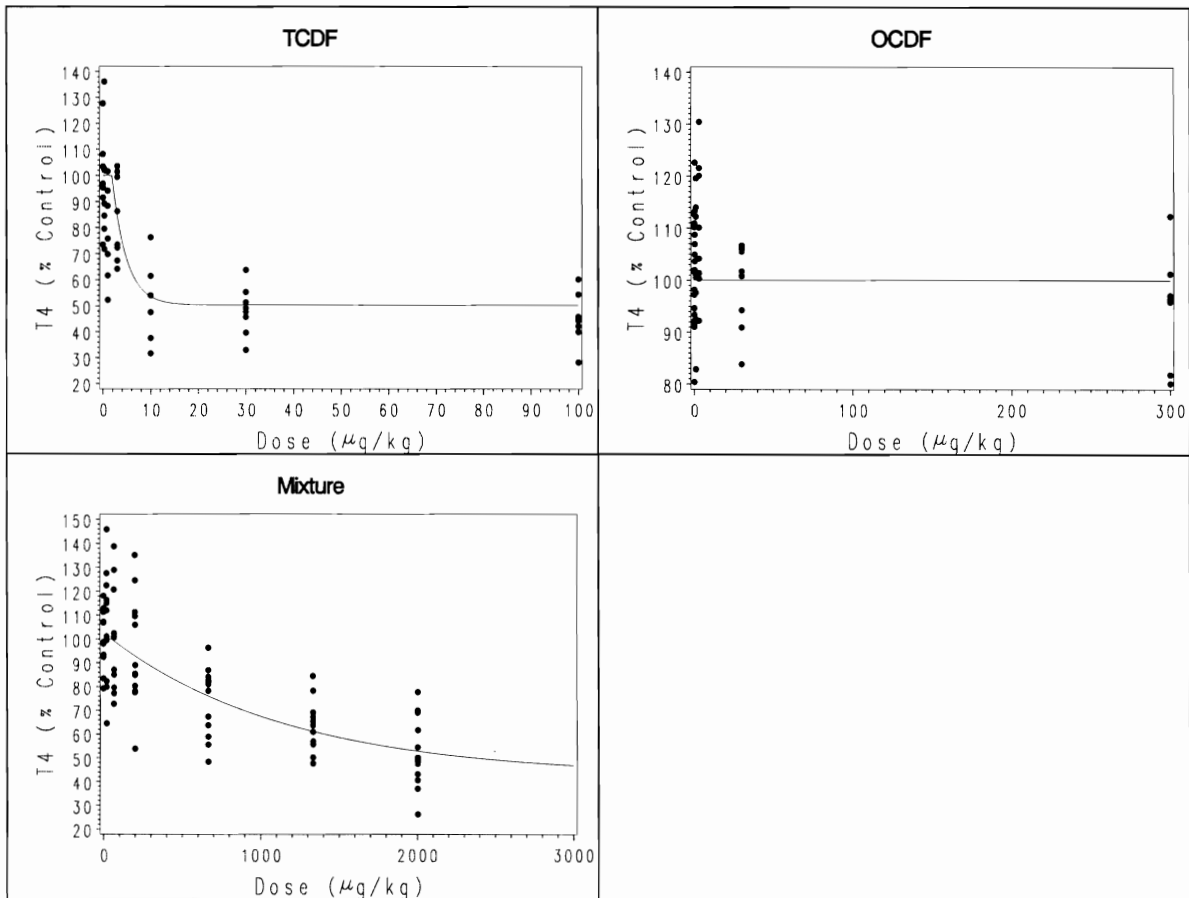
Appendix B4

Observed versus Predicted Plots Associated with the FSCR Full Model



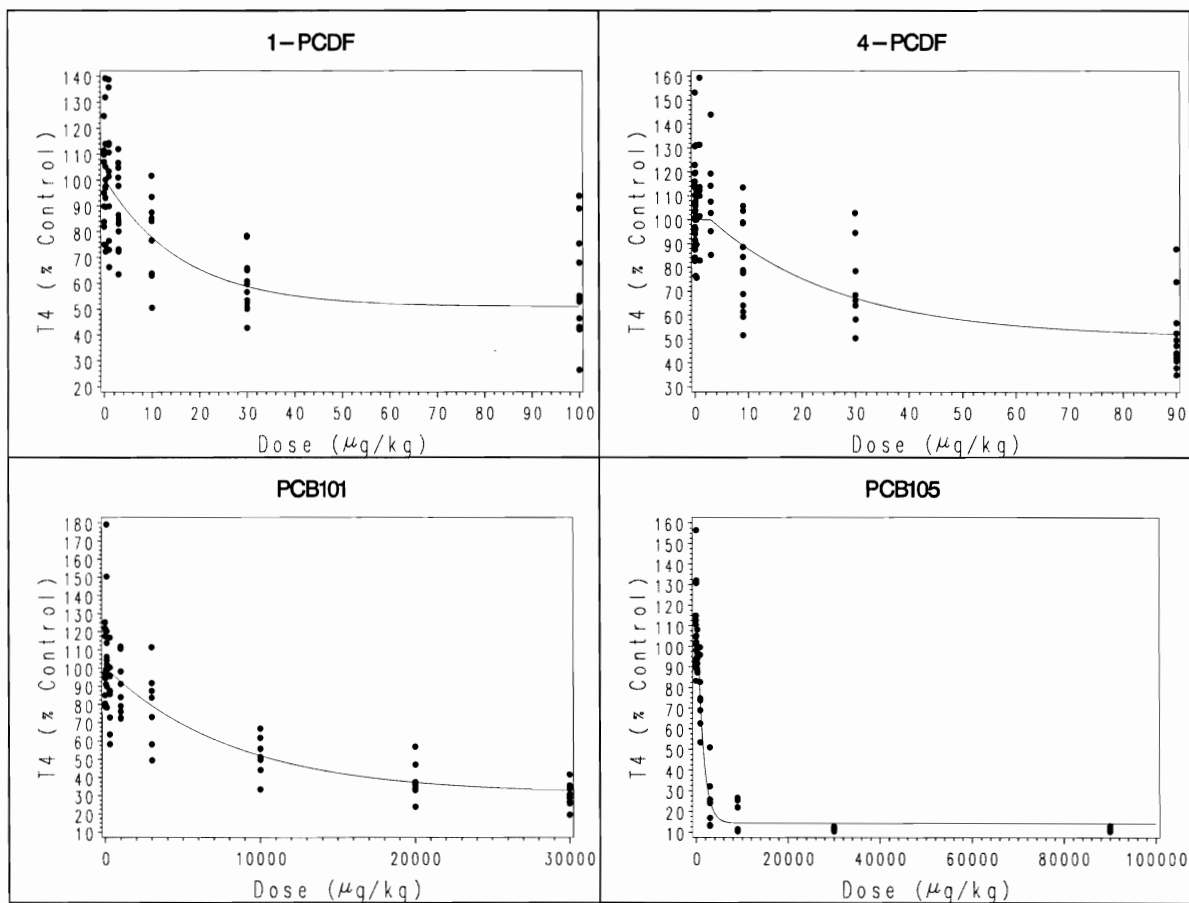


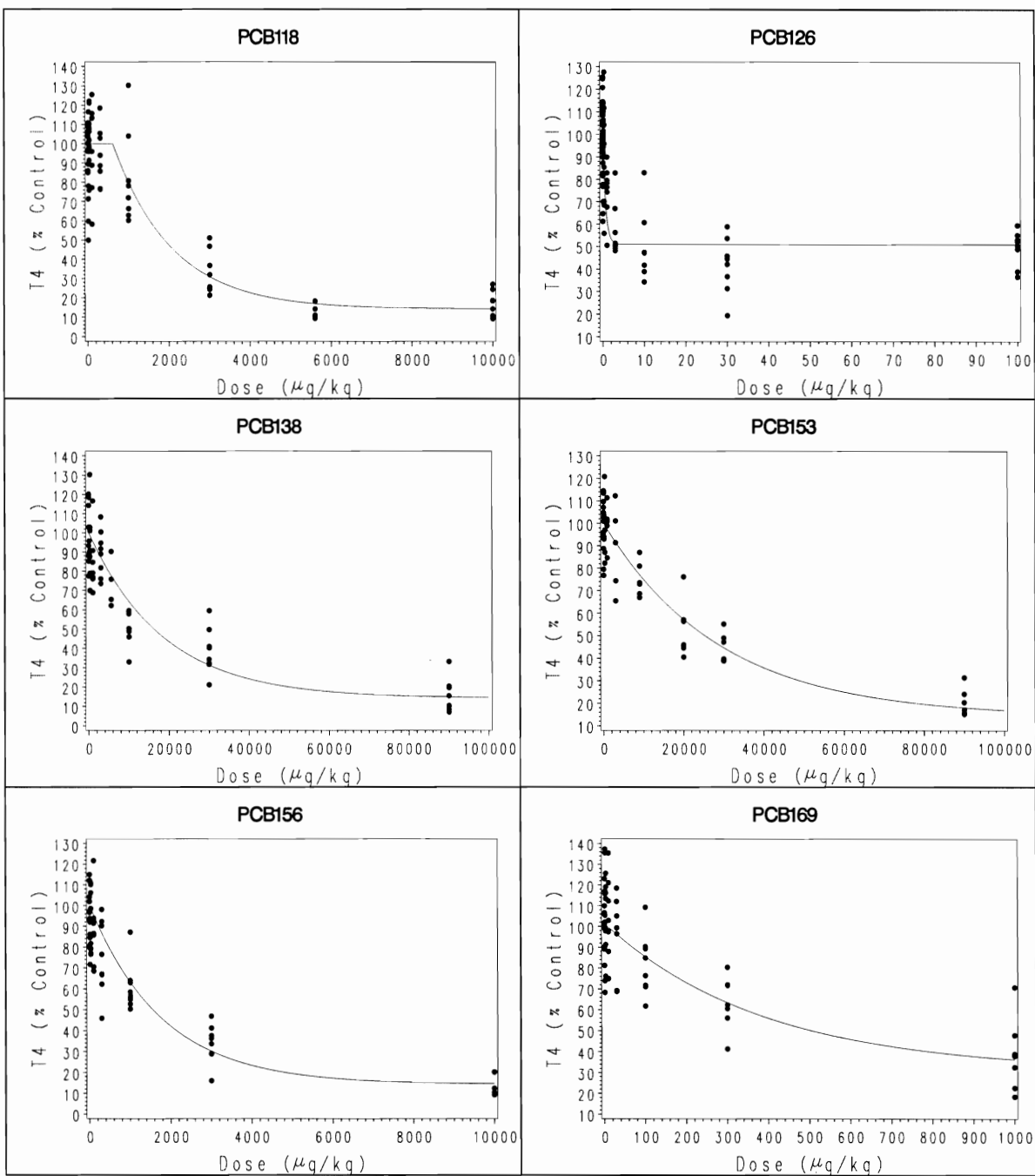


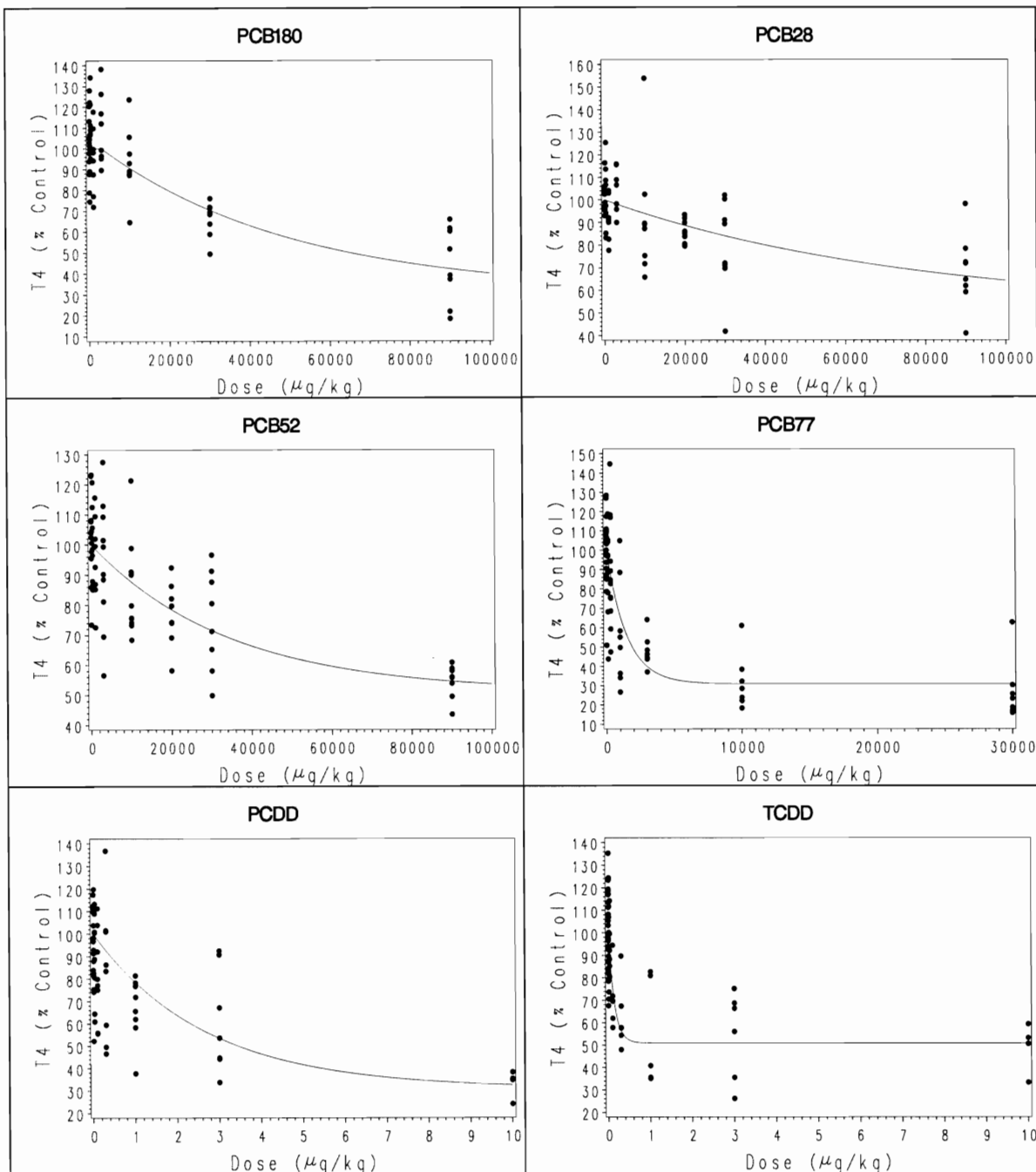


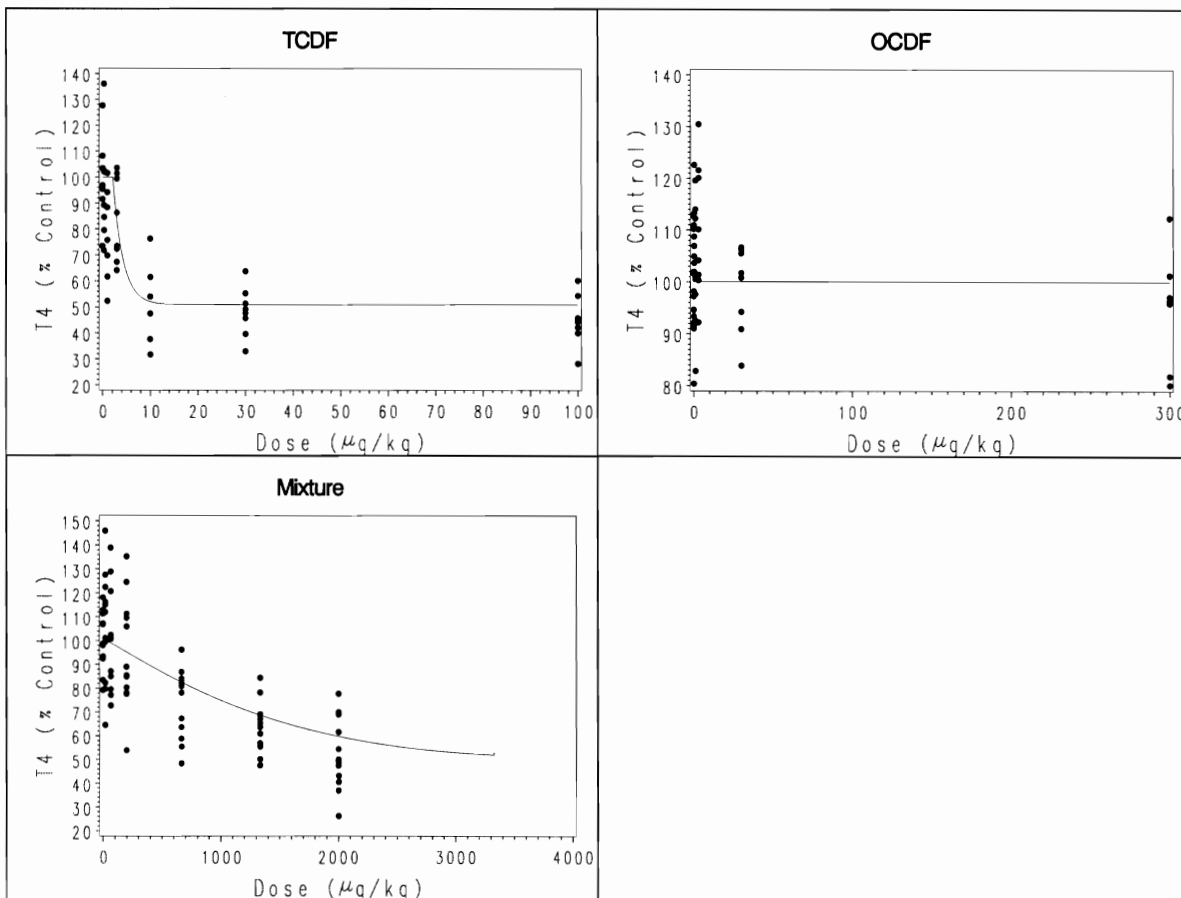
Appendix B5

Observed versus Predicted Plots Associated with the FSCR Additivity Model



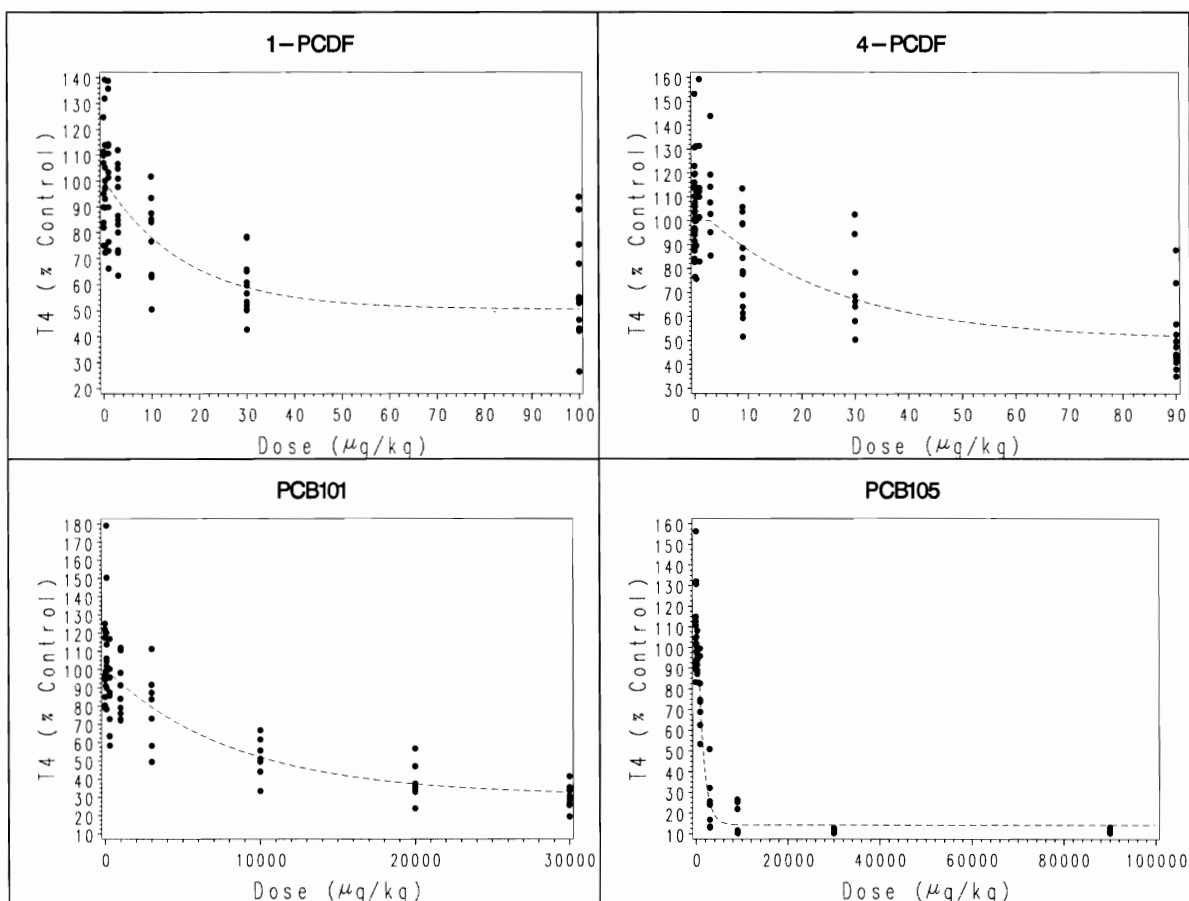


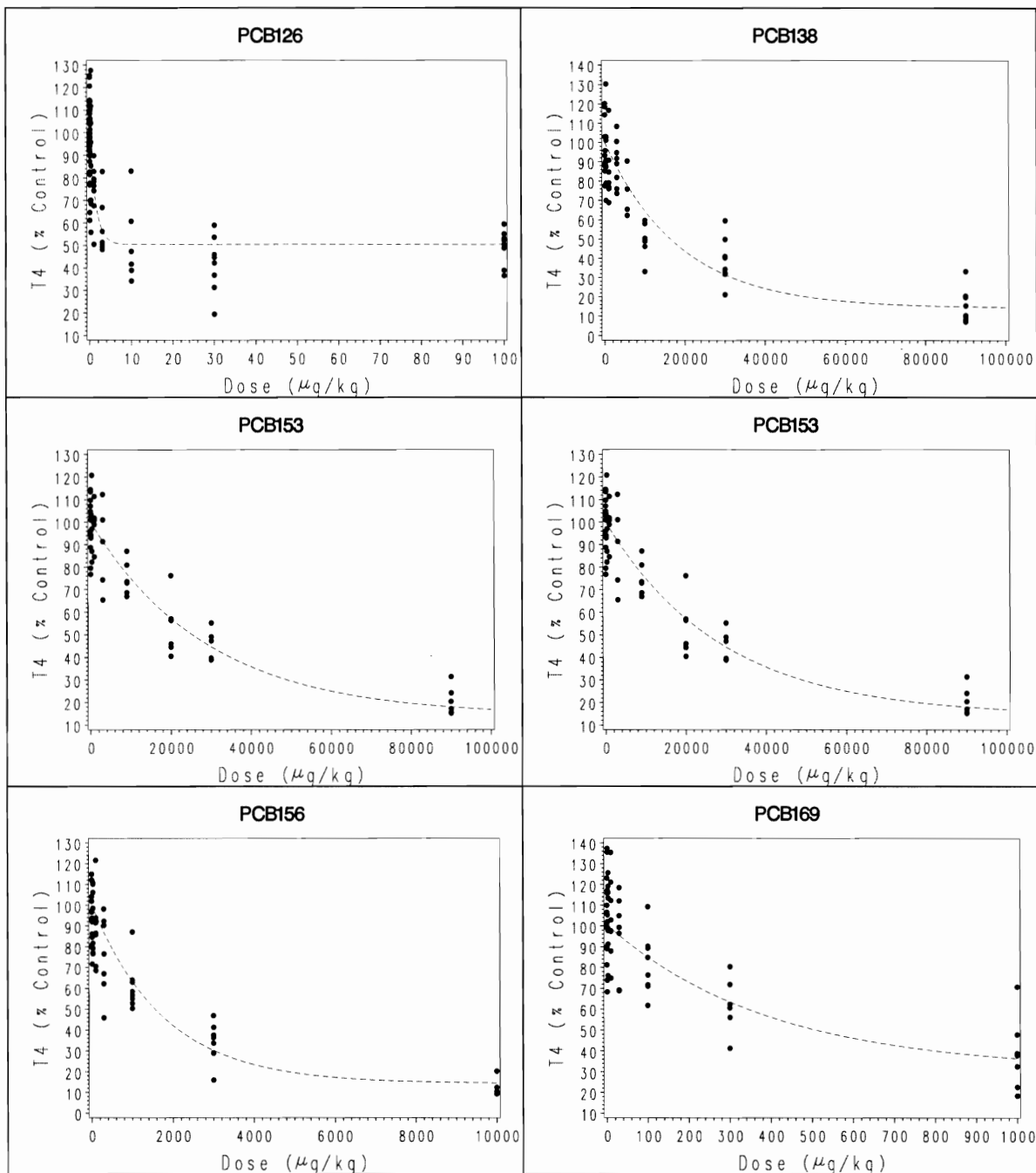


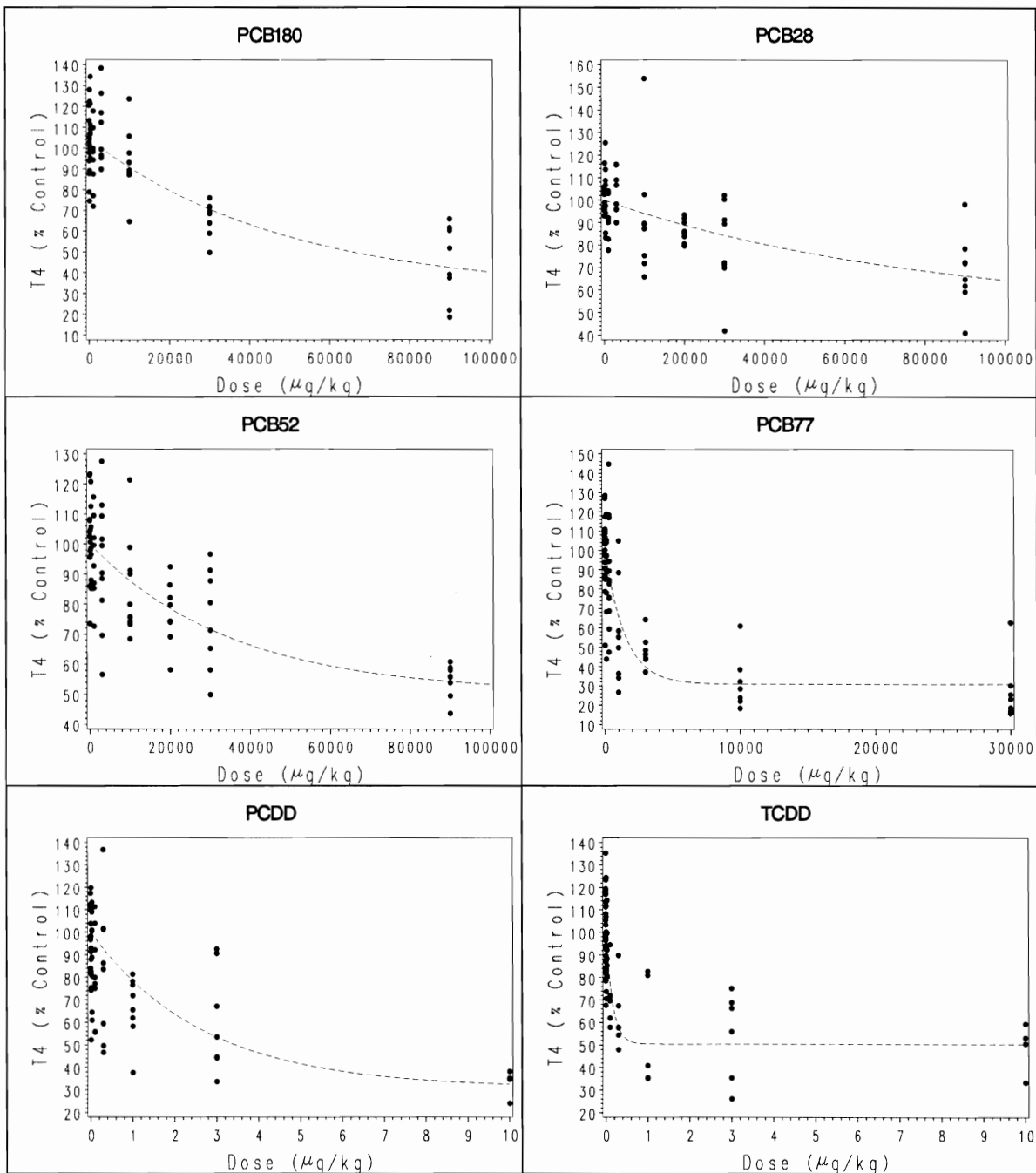


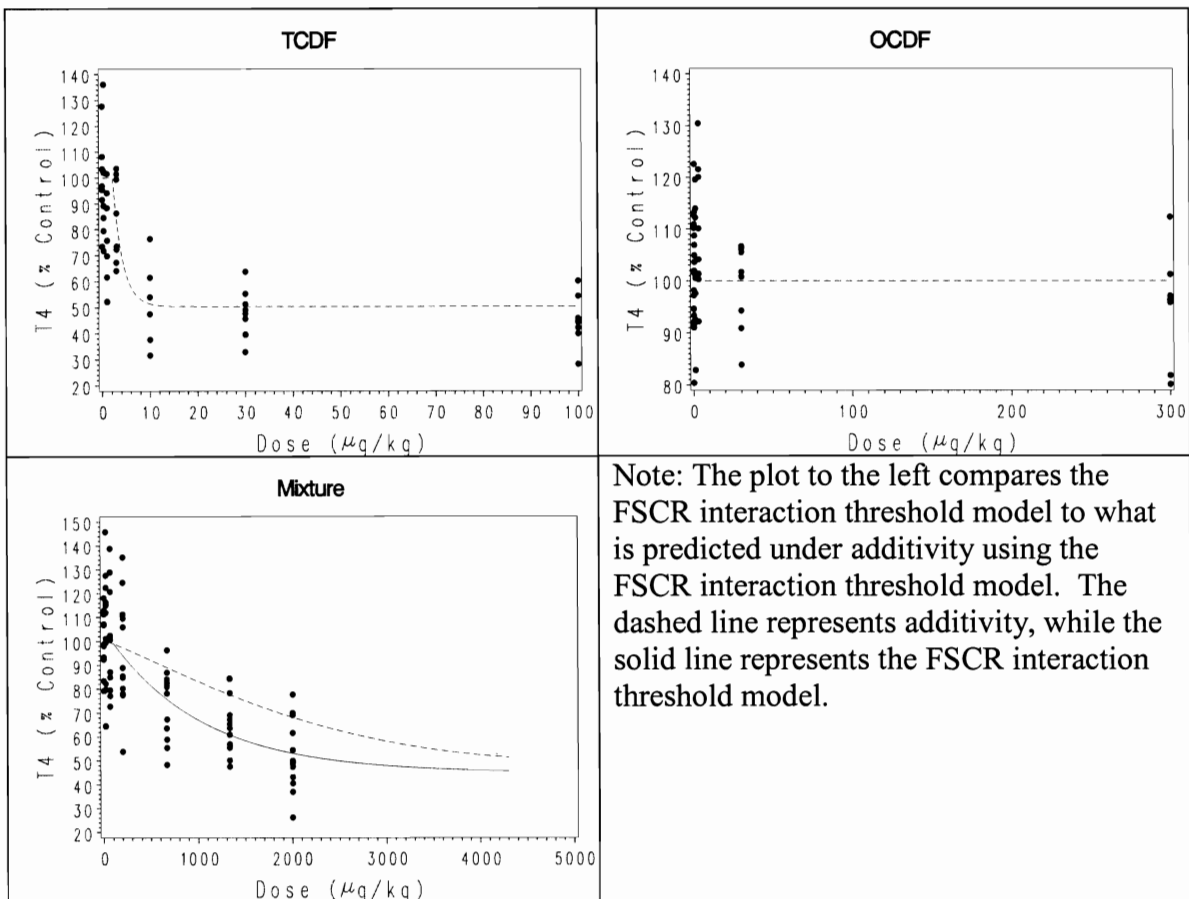
Appendix B6

Observed versus Predicted Plots Associated with the FSCR Interaction Threshold Model









Appendix C

Appendix C1

Derivatives of the FSCR Scaled Interaction Threshold Model with Respect to the Model Parameters

$$\begin{aligned} \frac{\partial \mu}{\delta \alpha_{1_2_6_12_13_16_17}} &= \left[1 - \exp \left\{ \frac{\beta_i}{SF_i} (x - \delta_i) \right\} \right] (x \geq \delta_i) (i = 1, 2, 6, 12, 13, 16, 17) \\ &+ \left[1 - \frac{(\mu - \alpha_{1_2_6_12_13_16_17})}{(100 - \alpha_{1_2_6_12_13_16_17})} \right] (i = c + 1) (t < \Delta) \\ &+ \left[\exp \{ \beta_{mix} (t - \Delta) \} \left[1 - \frac{(\mu_\Delta - \alpha_{1_2_6_12_13_16_17})}{(100 - \alpha_{1_2_6_12_13_16_17})} \right] \right] (i = c + 1) (t \geq \Delta) \end{aligned}$$

$$\begin{aligned} \frac{\partial \mu}{\delta \alpha_{3_10_11_14_15}} &= \left[1 - \exp \left\{ \frac{\beta_i}{SF_i} (x - \delta_i) \right\} \right] (x \geq \delta_i) (i = 3, 10, 11, 14, 15) \\ &+ \left[1 - \frac{(\mu - \alpha_{3_10_11_14_15})}{(100 - \alpha_{3_10_11_14_15})} \right] (i = c + 1) (t < \Delta) \\ &+ \left[\exp \{ \beta_{mix} (t - \Delta) \} \left[1 - \frac{(\mu_\Delta - \alpha_{3_10_11_14_15})}{(100 - \alpha_{3_10_11_14_15})} \right] \right] (i = c + 1) (t \geq \Delta) \end{aligned}$$

$$\begin{aligned} \frac{\partial \mu}{\delta \alpha_{4_5_7_8_9}} &= \left[1 - \exp \left\{ \frac{\beta_i}{SF_i} (x - \delta_i) \right\} \right] (x \geq \delta_i) (i = 4, 5, 7, 8, 9) \\ &+ \left[1 - \frac{(\mu - \alpha_{4_5_7_8_9})}{(100 - \alpha_{4_5_7_8_9})} \right] (i = c + 1) (t < \Delta) \\ &+ \left[\exp \{ \beta_{mix} (t - \Delta) \} \left[1 - \frac{(\mu_\Delta - \alpha_{4_5_7_8_9})}{(100 - \alpha_{4_5_7_8_9})} \right] \right] (i = c + 1) (t \geq \Delta) \end{aligned}$$

$$\begin{aligned} \frac{\partial \mu}{\partial \beta_i} = & \left[(100 - \alpha_i) \exp \left\{ \frac{\beta_i}{SF_i} (x - \delta_i) \right\} \frac{(x - \delta_i)}{SF_i} (x \geq \delta_i) \right] (i = 1, \dots, c) \\ & + \left[\frac{(\mu - \alpha_1)}{\beta_1} \ln \left(\frac{\mu - \alpha_1}{100 - \alpha_1} \right) \right] (i = c + 1) (t < \Delta) \\ & + \left[\exp \{ \beta_{mix} (t - \Delta) \} \left[\frac{(\mu_\Delta - \alpha_1)}{\beta_1} \ln \left(\frac{\mu_\Delta - \alpha_1}{100 - \alpha_1} \right) \right] \right] (i = c + 1) (t \geq \Delta) \end{aligned}$$

$$\begin{aligned} \frac{\partial \mu}{\partial \delta_i} = & \left[(100 - \alpha_i) \exp \left\{ \frac{\beta_i}{SF_i} (x - \delta_i) \right\} \left(-\frac{\beta_i}{SF_i} \right) (x \geq \delta_i) \right] (i = 1, \dots, 6, 10, 11, 16, 17) \\ & + \left[-\frac{(\mu - \alpha_1) \beta_1}{SF_1} \right] (i = c + 1) (t < \Delta) \\ & + \left[\exp \{ \beta_{mix} (t - \Delta) \} \left[-\frac{(\mu_\Delta - \alpha_1) \beta_1}{SF_1} \right] \right] (i = c + 1) (t \geq \Delta) \end{aligned}$$

$$\frac{\partial \mu}{\partial \alpha_{mix}} = \left[1 - \exp \{ \beta_{mix} (t - \Delta) \} \right] (i = c + 1) (t \geq \Delta)$$

$$\frac{\partial \mu}{\partial \beta_{mix}} = \left[(\mu_\Delta - \alpha_{mix}) \exp \{ \beta_{mix} (t - \Delta) \} (t - \Delta) \right] (i = c + 1) (t \geq \Delta)$$

$$\frac{\partial \mu}{\partial \Delta} = \left[(\mu_\Delta - \alpha_{mix}) \exp \{ \beta_{mix} (t - \Delta) \} (-\beta_{mix}) \right] (i = c + 1) (t \geq \Delta)$$

Appendix C2

Parameter Estimates Resulting from the FSCR Scaled Full Model with Combined Minimum Response Parameters

Parameter	Estimate	Standard Error	p-value	Approximate 95% Confidence Limits	
$\alpha_{1_2_6_12_13_16_17}$	50.2556	1.4495	<0.0001	47.4120	53.0992
$\alpha_{3_10_11_14_15}$	30.7384	1.9501	<0.0001	26.9126	34.5642
$\alpha_{4_5_7_8_9}$	14.3276	1.0261	<0.0001	12.3146	16.3406
β_1	-0.0608	0.0164	0.0002	-0.0929	-0.0286
β_2	-0.0378	0.0142	0.008	-0.0657	-0.00989
β_3	-11.9346	2.0221	<0.0001	-15.9015	-7.9677
β_4	-8.3300	1.7098	<0.0001	-11.6843	-4.9757
β_5	-6.9610	1.0886	<0.0001	-9.0966	-4.8253
β_6	-0.7186	0.2249	0.0014	-1.1598	-0.2774
β_7	-53.8744	5.5415	<0.0001	-64.7458	-43.0031
β_8	-33.9523	3.4438	<0.0001	-40.7084	-27.1962
β_9	-56.7315	6.0522	<0.0001	-68.6048	-44.8583
β_{10}	-0.00262	0.000549	<0.0001	-0.00369	-0.00154
β_{11}	-20.7825	4.8148	<0.0001	-30.2283	-11.3368
β_{12}	-12.1418	2.5247	<0.0001	-17.0948	-7.1889
β_{13}	-27.9797	6.3307	<0.0001	-40.3993	-15.5601
β_{14}	-6.6610	1.0155	<0.0001	-8.6533	-4.6687
β_{15}	-0.3749	0.0639	<0.0001	-0.5003	-0.2494
β_{16}	-6.5054	2.4139	0.0071	-11.2411	-1.7697
β_{17}	-0.3312	0.2744	0.2277	-0.8695	0.2072
δ_1	0.3891	1.4995	0.7953	-2.5527	3.3309
δ_2	3.0937	3.9265	0.4309	-4.6094	10.7969

Parameter	Estimate	Standard Error	p-value	Approximate 95% Confidence Limits	
δ_3	76.4097	398.4	0.848	-705.3	858.1
δ_4	513.1	147.9	0.0005	223.0	803.3
δ_5	669.5	147.5	<0.0001	380.2	958.8
δ_6	0.0427	0.1295	0.7417	-0.2114	0.2968
δ_{10}	9.8125	25.1041	0.696	-39.4372	59.0621
δ_{11}	3227.3	5581.7	0.5632	-7723.0	14177.5
δ_{16}	0.00361	0.0197	0.8546	-0.0350	0.0423
δ_{17}	1.8587	0.9770	0.0573	-0.0580	3.7755
α_{mix}	42.2947	14.7759	0.0043	13.3070	71.2823
β_{mix}	-0.00087	0.000467	0.0622	-0.00179	0.000044
δ_{mix}	49.4723	74.3290	0.5058	-96.3473	195.3

Appendix C3

Parameter Estimates Resulting from the Fit of the FSCR Scaled Additivity Model

Parameter	Estimate
$\alpha_{1_2_6_12_13_16_17}$	50.83547
$\alpha_{3_10_11_14_15}$	30.93872
$\alpha_{4_5_7_8_9}$	14.29674
β_1	-0.06027
β_2	-0.04032
β_3	-12.0046
β_4	-8.2648
β_5	-6.93473
β_6	-1.51304
β_7	-54.1238
β_8	-34.5465
β_9	-56.2282
β_{10}	-0.00257
β_{11}	-20.8711
β_{12}	-13.0153
β_{13}	-28.5603
β_{14}	-6.75876
β_{15}	-0.37213
β_{16}	-7.41317
β_{17}	-0.4366
δ_1	0.218393
δ_2	2.999922
δ_3	40.43141

Parameter	Estimate
δ_4	498.8989
δ_5	622.8992
δ_6	0.019764
δ_{10}	3.250986
δ_{11}	3099.349
δ_{16}	0.002821
δ_{17}	2.046182

Appendix C4

Parameter Estimates Resulting from the Fit of the FSCR Scaled Interaction Threshold Model

Parameter	Estimate
$\alpha_{1_2_6_12_13_16_17}$	50.43233
$\alpha_{3_10_11_14_15}$	30.90785
$\alpha_{4_5_7_8_9}$	14.23039
β_1	-0.05927
β_2	-0.04015
β_3	-11.9207
β_4	-8.13799
β_5	-6.9295
β_6	-0.70073
β_7	-53.433
β_8	-34.5557
β_9	-55.9946
β_{10}	-0.00254
β_{11}	-20.5839
β_{12}	-12.6627
β_{13}	-28.4047
β_{14}	-6.80725
β_{15}	-0.37296
β_{16}	-6.39152
β_{17}	-0.49315
δ_1	0.201199
δ_2	2.998452
δ_3	41.69784

Parameter	Estimate
δ_4	495.3648
δ_5	665.1093
δ_6	0.020615
δ_{10}	3.443008
δ_{11}	3014.988
δ_{16}	0.004053
δ_{17}	2.174192
α_{mix}	44.86722
β_{mix}	-0.00099
Δ	105.298

Appendix C5

SAS Penalized Optimal Second Stage Design Code for the FSCR Interaction Threshold Model and the Fixed-Ratio Mixture of 18 Polyhalogenated Aromatic Hydrocarbons

```
*Create design columns;
%macro vartemp;
  %do i=1 %to &ncol;
    bestdesign&i=bestdesign[1,&i];
  %end;
%mend;

*String together design columns;
%macro varlabel;
  %do i=1 %to &ncol;
    bestdesign&i
  %end;
%mend;

*Create list of design parameters;
%macro varlist;
  t1 || q1 || n1 ||
  %do i=2 %to &m;
    t&i || q&i || n&i ||
  %end;
%mend varlist;
```

```

*Create loop to allow lambda to iterate;
%macro lambdaloop(m=,lambda_low=,lambda_high=,lambda_by=,integer=1,divisor=1,sortcol=,ncol=);
  *m: number of design points;
  *integer=1 ==> lambda loops through integer values;
  *integer=0 ==> lambda loops through noninteger values;
  *multiplier: multiply lambdatemp by multiplier to get noninteger values;
  %do lambdatemp=&lambda_low %to &lambda_high %by &lambda_by;

    %if &integer=0 %then %do;
      %let lambdaval=%sysevalf(&lambdatemp/&divisor);
      %let lambdaindex=fraction&lambdatemp;
    %end;
    %else %do;
      %let lambdaval=&lambdatemp;
      %let lambdaindex=&lambdaval;
    %end;
    *&lambdaval is the lambda used in the penalization process;
    *&lambdaindex is the identifier used to keep track of &lambdaval;
    *these are different only when &lambdaval is a noninteger value;

    print 'lambda value ' &lambdaval '-----';

    use doses&m;
    read all into start_doses;
    results&m._&lambdaindex=j(1,&ncol,0);

    *Loop through starting designs;
    %do index=1 %to 8;
      tempd=start_doses[&index,];
      %penalized_design(m=&m,start_vals=tempd,lambda=&lambdaval);
      results&m._&lambdaindex=results&m._&lambdaindex //
        (count || &lambdaval || &index || %varlist
         det || actual_var || practical_var || actual_d || practical_d ||
         fn_value || practical_fn_value || flag1 || flag2 || flag3);

      %put 'Design for lambda ' &lambdaval ' and starting value set ' &index '
        completed.';
    %end;
  %end;

```

```

results&m._&lambdaindex=results&m._&lambdaindex[2:nrow(results&m._&lambdaindex),];
print results&m._&lambdaindex;

*Sort candidate designs;
do index1=0 to (nrow(results&m._&lambdaindex)-1);
  do index2=1 to (nrow(results&m._&lambdaindex)-1-index1);
    if
  (results&m._&lambdaindex[index2+1,&sortcol]<results&m._&lambdaindex[index2,&sortcol])
  then do;
    temp=results&m._&lambdaindex[index2,];
    results&m._&lambdaindex[index2,]=results&m._&lambdaindex[index2+1,];
    results&m._&lambdaindex[index2+1,]=temp;
    end;
  end;
end;

bestdesign=results&m._&lambdaindex[1,];
%vartemp;

*Insert optimal design for a given lambda into design data set;
edit penalized_ds&m var {%varlabel};
append var {%varlabel};
%put 'Design for lambda ' &lambdaval ' appended to data set.';

%end;
%mend lambdaloop;

*title 'Assumed Parameter Values';
data assumptions;
  set scaled.DrGstart_intest;
run;

data for_anal;
  set fscr.complete;
  where pert4 ne .;
run;

```

```
proc sort data=for_anal;
  by t_ind;
run;

proc freq data=for_anal;
  tables t / out=obst;
run;

/*Run macro for a grid of starting values (8 dose points) */
/*
data doses8;
  do t1=0;
    do t2=20, 40;
      do t3=60, 90;
        do t4=100;
          do t5=110;
            do t6=120, 250;
              do t7=500;
                do t8=700;
                  output;
                end;
              end;
            end;
          end;
        end;
      end;
    end;
  end;
end;
run;
```



```

data penalized_ds8;
  input bestdesign1 bestdesign2 bestdesign3 bestdesign4 bestdesign5 bestdesign6 bestdesign7
        bestdesign8 bestdesign9 bestdesign10 bestdesign11 bestdesign12 bestdesign13 bestdesign14
        bestdesign15 bestdesign16 bestdesign17 bestdesign18 bestdesign19 bestdesign20
        bestdesign21 bestdesign22 bestdesign23 bestdesign24 bestdesign25 bestdesign26
        bestdesign27 bestdesign28 bestdesign29 bestdesign30 bestdesign31 bestdesign32
        bestdesign33 bestdesign34 bestdesign35 bestdesign36 bestdesign37;

datalines;

;
run;
*/

proc iml;
*Define the dose-response relationship using parameter values;
  start model_definition;
  use assumptions;
  read all var {intest} into parm_assumptions;
  p=nrow(parm_assumptions);
  alphas=parm_assumptions[1:3];
  alpha1_2_6_12_13_16_17=alphas[1];
  alpha1=alpha1_2_6_12_13_16_17; alpha2=alpha1_2_6_12_13_16_17;
  alpha6=alpha1_2_6_12_13_16_17; alpha12=alpha1_2_6_12_13_16_17;
  alpha13=alpha1_2_6_12_13_16_17; alpha16=alpha1_2_6_12_13_16_17;
  alpha17=alpha1_2_6_12_13_16_17;
  alpha3_10_11_14_15=alphas[2];
  alpha3=alpha3_10_11_14_15; alpha10=alpha3_10_11_14_15;
  alpha11=alpha3_10_11_14_15; alpha14=alpha3_10_11_14_15;
  alpha15=alpha3_10_11_14_15;
  alpha4_5_7_8_9=alphas[3];
  alpha4=alpha4_5_7_8_9; alpha5=alpha4_5_7_8_9; alpha7=alpha4_5_7_8_9;
  alpha8=alpha4_5_7_8_9; alpha9=alpha4_5_7_8_9;
  betas=parm_assumptions[4:20];
  %separate_betas;
  deltas=parm_assumptions[21:30];
  delta1=deltas[1]; delta2=deltas[2]; delta3=deltas[3]; delta4=deltas[4]; delta5=deltas[5];
  delta6=deltas[6]; delta10=deltas[7]; delta11=deltas[8]; delta16=deltas[9];
  delta17=deltas[10];

```

```

    delta7=0; delta8=0; delta9=0; delta12=0; delta13=0; delta14=0; delta15=0;
alpha_mix=parm_assumptions[31];
beta_mix=parm_assumptions[32];
cpt1_delta=parm_assumptions[33];
tauhat=3.33;
print 'Assumed Form of the Dose-Response Relationship-----';
print 'alpha1_2_6_12_13_16_17' alpha1_2_6_12_13_16_17;
print 'alpha3_10_11_14_15' alpha3_10_11_14_15;
print 'alpha4_5_7_8_9' alpha4_5_7_8_9;
print 'Single Chemical Slopes' betas;
print 'Threshold, Agent 1' delta1;
print 'Threshold, Agent 2' delta2;
print 'Threshold, Agent 3' delta3;
print 'Threshold, Agent 4' delta4;
print 'Threshold, Agent 5' delta5;
print 'Threshold, Agent 6' delta6;
print 'Threshold, Agent 10' delta10;
print 'Threshold, Agent 11' delta11;
print 'Threshold, Agent 16' delta16;
print 'Threshold, Agent 17' delta17;
print 'Agents 7, 8, 9, 12, 13, 14, 15 do not have thresholds.';
print 'alpha_mix' alpha_mix;
print 'Mixture Slope' beta_mix;
print 'Interaction Threshold' cpt1_delta;
a1=0.000003; a2=0.000013; a3=0.076814; a4=0.038282; a5=0.190302; a6=0.000302; a7=0.190181;
a8=0.190861; a9=0.006541; a10=0.000197; a11=0.188700; a12=0.039237; a13=0.077523;
a14=0.000988; a15=0.000007; a16=0.000007; a17=0.000010; a18=0.000032;
sf1=1; sf2=1; sf3=100000; sf4=10000; sf5=10000; sf6=1; sf7=1000000; sf8=1000000;
sf9=100000; sf10=1; sf11=1000000; sf12=1000000; sf13=1000000; sf14=10000; sf15=1;
sf16=1; sf17=1; sft=1;
finish;

*Read in data;
start read_data;
use for_anal;
read all var {x1} into x1; read all var {x2} into x2;
read all var {x3} into x3; read all var {x4} into x4;
read all var {x5} into x5; read all var {x6} into x6;

```

```

read all var {x7} into x7; read all var {x8} into x8;
read all var {x9} into x9; read all var {x10} into x10;
read all var {x11} into x11; read all var {x12} into x12;
read all var {x13} into x13; read all var {x14} into x14;
read all var {x15} into x15; read all var {x16} into x16;
read all var {x17} into x17; read all var {x18} into x18;
read all var {t} into t;
read all var {x1_ind} into x1_ind; read all var {x2_ind} into x2_ind;
read all var {x3_ind} into x3_ind; read all var {x4_ind} into x4_ind;
read all var {x5_ind} into x5_ind; read all var {x6_ind} into x6_ind;
read all var {x7_ind} into x7_ind; read all var {x8_ind} into x8_ind;
read all var {x9_ind} into x9_ind; read all var {x10_ind} into x10_ind;
read all var {x11_ind} into x11_ind; read all var {x12_ind} into x12_ind;
read all var {x13_ind} into x13_ind; read all var {x14_ind} into x14_ind;
read all var {x15_ind} into x15_ind; read all var {x16_ind} into x16_ind;
read all var {x17_ind} into x17_ind; read all var {x18_ind} into x18_ind;
read all var {t_ind} into t_ind;
read all var {pert4} into y;
n=nrow(t); nmix=t_ind[+]; nsingle=n-nmix;
print 'Of the ' n ' observations contained in the original experiment,';
print '      ' nsingle ' observations were associated with single chemical data, and';
print '      ' nmix ' observations were associated with the mixture.';
finish;

*Define bisection algorithm;
start bisection;
epsilon=0.0001;
criterion=1;
mu_high=100; tmu_high=0;
mu_low=max(alpha1_2_6_12_13_16_17,alpha3_10_11_14_15,alpha4_5_7_8_9,alpha_mix)+epsilon;
%findt(mu=mu_low); tmu_low=t_est;
do until (criterion<epsilon);
    mu_new=(mu_high+mu_low)/2;
    %findt(mu=mu_new);
    tmu_new=t_est;
    criterion=abs(t_obs-tmu_new);
    if tmu_new>t_obs then mu_low=mu_new;
    else if tmu_new<t_obs then mu_high=mu_new;

```

```

end;
*print criterion t_obs tmu_new mu_new;
finish;

*Find additivity predictions;
start additivity;
  *print obst;
  alpha1=alpha1_2_6_12_13_16_17; alpha2=alpha1_2_6_12_13_16_17;
  alpha6=alpha1_2_6_12_13_16_17; alpha12=alpha1_2_6_12_13_16_17;
  alpha13=alpha1_2_6_12_13_16_17; alpha16=alpha1_2_6_12_13_16_17;
  alpha17=alpha1_2_6_12_13_16_17;
  alpha3=alpha3_10_11_14_15; alpha10=alpha3_10_11_14_15; alpha11=alpha3_10_11_14_15;
  alpha14=alpha3_10_11_14_15; alpha15=alpha3_10_11_14_15;
  alpha4=alpha4_5_7_8_9; alpha5=alpha4_5_7_8_9; alpha7=alpha4_5_7_8_9;
  alpha8=alpha4_5_7_8_9; alpha9=alpha4_5_7_8_9;
  addt=j(nrow(obst),1,0);
  mixadd=j(nrow(obst),1,0);
  mixadd[1]=100;
  do i=2 to nrow(obst);
    t_obs=obst[i];
    run bisection;
    addt[i]=tmu_new; mixadd[i]=mu_new;
  end;
  *print mixadd;
finish;

start addmix;
  mix_add=j(n,1,0);
  do i=1 to n;
    sum=0;
    do j=1 to nrow(obst);
      if t_ind[i]=0 then sum=0;
      else do;
        if _t_[i]=obst[j] then temp=mixadd[j];
        else temp=0;
        sum=sum+temp;
      end;
    end;
  end;
end;

```

```

        mix_add[i]=sum;
    end;
finish;

start interaction;
    t_obs=cptl_delta;
    run bisection;
    mu_cptl_delta=mu_new;
finish;

*Compute parameter covariance matrix associated with the original experiment;
start original_covmat;
    *%createplot;
    use obst; read all var {t} into obst;
    %calculatemu(original=1, _t_=t);
    single=t[1:nsingle];
    %singleagent_derivatives;
    mixture=t[nsingle+1:n];
    %mixture_derivatives(_t_=mixture);
    mixture_fixedf=alpha_derivs || beta_derivs || delta_derivs || alpha_mix_deriv ||
        beta_mix_deriv || cptl_delta_deriv;
    fixed_f=singleagent_fixedf // mixture_fixedf;
    fixed_vmat=diag(mu_original);
    original_varmat=tauhat*inv((fixed_f`*inv(fixed_vmat)*fixed_f));
    det_original=det(original_varmat); *print det_original;
    det11=det(original_varmat[1:30,1:30]); det22=det(original_varmat[31:33,31:33]);
    stderr=j(p,1,0);
    do i=1 to p;
        stderr[i]=sqrt(original_varmat[i,i]);
    end;
    var_cptldelta=original_varmat[p,p];
    print 'The initial experiment resulted in variance of the interaction threshold '
    var_cptldelta;
finish;

```

```

start minimumvariances;
    minh_xq8=3413.8;
    minh_xq6=3228.7;
    minh_xq4=3195.6;
finish minimumvariances;

run model_definition;
run read_data;
run original_covmat;
run minimumvariances;

*%lambdaloop(m=8,lambda_low=1,lambda_high=31,lambda_by=10,sortcol=34,ncol=37);
*lambdaloop(m=8,lambda_low=1,lambda_high=3,lambda_by=1,integer=0,divisor=4,sortcol=34,ncol=37);
quit;

*For an 8 point design;
data temp;
    set penalized_ds8;
    rename      bestdesign1 =count      bestdesign2 =lambda
                bestdesign3 =index      bestdesign4 =t1
                bestdesign5 =q1         bestdesign6 =n1
                bestdesign7 =t2         bestdesign8 =q2
                bestdesign9 =n2         bestdesign10=t3
                bestdesign11=q3         bestdesign12=n3
                bestdesign13=t4         bestdesign14=q4
                bestdesign15=n4         bestdesign16=t5
                bestdesign17=q5         bestdesign18=n5
                bestdesign19=t6         bestdesign20=q6
                bestdesign21=n6         bestdesign22=t7
                bestdesign23=q7         bestdesign24=n7
                bestdesign25=t8         bestdesign26=q8
                bestdesign27=n8         bestdesign28=det
                bestdesign29=actual_var bestdesign30=practical_var
                bestdesign31=actual_d   bestdesign32=practical_d
                bestdesign33=fn_value   bestdesign34=practical_fn_value
                bestdesign35=flag1      bestdesign36=flag2
                bestdesign37=flag3;

```

```
data temp2;  
  set pnlzd.penalized_ds8 temp;run;  
proc print data=temp2;run;  
  
data pnlzd.penalized_ds8;  
  set temp2;run;  
proc print data=pnlzd.penalized_ds8; run;
```

Appendix D

Appendix D

SAS Optimal Design Code

```
*Create data set of starting values for design search;
data doses4;
  do x1=0;
    do x2=0.9, 1.1, 2 ;
      do x3=4, 6, 8 ;
        do x4=12, 16, 20;
          output;
        end;
      end;
    end;
  end;
run;*27 sets of starting values;

*Create data set to contain results of design search;
data lrci4pt;
  input index count2 _dchk_2 _dchk_22 x1 q1 n1 x2 q2 n2 x3 q3 n3 x4 q4 n4 initial final lbest
  xflag qflag ;
datalines;

;
run;
```

```

*Macro to fit nonlinear exponential model using SIMPLEX;
%macro parameter_estimates(fixdelta=,delta=);    *fixdelta=1 ==> delta fixed;
                                                *fixdelta=0 ==> estimate delta;

    run initial;
    delta=&delta;
    fixdelta=&fixdelta;
    run optima;

%mend parameter_estimates;

*Macro to find confidence interval;
%macro findci;

    run target;
    run find_lb;
    run find_ub;

%mend findci;

*Macro to find lower confidence bound;
%macro findlb;

    run target;
    run find_lb;

%mend findlb;

*Macro to create candidate design vector;
%macro create_designx;

    *Create x vector associated with candidate design;
    designx=.;
    %do i=1 %to &m;
        if n&i>0 then designx=designx // j(n&i,1,x&i);
    %end;
    designx=designx[2:nrow(designx)];

%mend create_designx;

```

```

*Macro to predict responses for a candidate design;
%macro create_designdata;

    *Randomly select residuals from the original experiment;
    residualnumber=54#ranuni(j(nrow(designx),1,70100))+1;
    designres=stdres_sorted[residualnumber];

    *Predict responses for the candidate design based only randomly selected residuals;
    muhat=alphafixed+gammafixed#exp(betafixed#(designx-deltafixed)#(designx>=deltafixed));
    predy=muhat+designres#(sqrt(taufixed#muhat));

    *Combine data from the original experiment with design data;
    x=orig_x // designx;
    y=orig_y // predy;

%mend create_designdata;

*Macro to find the bootstrap estimate of the median lower bound;
%macro bootstraplb;

    %create_designx;

    *Estimate lower confidence bound on the threshold using a bootstrap-type procedure;
    lbs=j(100,1,0);
    do i=1 to 100;
        %create_designdata;
        %findlb;
        lbs[i]=lb;
    end;

    *Sort lower bounds;
    temp=lbs;
    lbs[rank(lbs),]=temp;
    unsortedlbs=temp; sortedlbs=lbs;
    lbest=sortedlbs[51];

%mend bootstraplb;

```

```

*Macro to list the design parameters;
%macro listdesignparms;

    %do i=2 %to &m;
        x&i //
    %end;
    q2
    %do i=3 %to &m;
        // q&i
    %end;

%mend listdesignparms;

*Optimal design macro;
%macro design(m=, startvals=, design_n=);

    *Break up vector of starting values;
    m=&m; design_n=&design_n;
    %do i=1 %to &m;
        x&i=&startvals[&i];
        q&i=1/m;
        n&i=design_n*q&i;
    %end;

    *Define objective function for use within simplex2;
    start function2;
    sumq=0; sumn=0; xcheck=0; qcheck=0;
    %do i=2 %to &m;
        x&i=parms2[&i-1];
        q&i=parms2[&m + (&i-2)];
        n&i=round(design_n#q&i);
        sumq=sumq+q&i;    sumn=sumn+n&i;
        xcheck=xcheck + (x&i<0);    qcheck=qcheck + (q&i<0) + (q&i>1);
    %end;
    q1=1-sumq; n1=design_n-sumn;
    qcheck=qcheck + (q1<0) + (q1>1);

```

```

if ((xcheck+qcheck)>0)then do;
    fn_value2=10**30;
    if xcheck>0 then xflag=1;
    if qcheck>0 then qflag=1;
end;
else do;
    %bootstraplb;
    fn_value2=-lbest;
end;
finish function2;

*Define procedure for finding the lower bound associated with a candidate design;
%include simplex2;
start optima2;
    xflag=0; qflag=0;
    in_parms2=%listdesignparms;
    in_steps2=in_parms2#0.5;
    %bootstraplb; initial=lbest;
    print 'The initial design is given by ';
    print 'doses ' %do i=1 %to &m;
        x&i
        %end;;
    print 'proportions ' %do i=1 %to &m;
        q&i
        %end;;
    print 'The expected lower bound on the threshold associated with the initial design is '
        lbest;
    run simplex2; final=-fn_value2;
    run function2;
    print 'After ' count2 ' iterations,';
    print 'the resulting design is given by ';
    print 'doses ' %do i=1 %to &m;
        x&i
        %end;;
    print 'proportions ' %do i=1 %to &m;
        q&i
        %end;;
    print 'The associated lower bound on the threshold is ' final lbest;

```

```

finish optima2;

run optima2;

%mend design;

*Grid search;
%macro designloop;
  %do index=1 %to 27 ;*nrow(start_doses);
    proc iml;

      *Read in deltamethrin data;
      use deltamethrin;
      read all var {dose} into orig_x;
      read all var {percent} into orig_y;

      *Assign starting value for alpha,beta;
      start initial;
        alpha=25; gamma=100-alpha;
        beta=-0.2;
        p=3;
        n=nrow(y);
      finish initial;

      *Define ql according to  $v(\mu)=t*\mu$ ;
      start ql;
        gamma=100-alpha;
        mu=alpha+gamma#exp(beta#(x-delta)#(x>=delta));
        ql=sum(y#log(mu)-mu);
        tau=sum(((y-mu)##2)/mu)/(n-p);
      finish ql;

      *Define objective function for use within simplex;
      %include simplex;
      start function;
        if fixdelta=0 then do;
          alpha=parms[1]; beta=parms[2]; delta=parms[3];
        end;

```

```

else do;
    alpha=params[1]; beta=params[2];
end;
if alpha<0 then fn_value=10**30;
else do;
    run ql;
    fn_value=-ql;
end;
finish function;

*Define procedure for finding mgle;
start optima;
if fixdelta=0 then do;
    in_params=alpha // beta // delta;
end;
else do;
    in_params=alpha // beta;
end;
in_steps=in_params#0.01;
run ql;
run simplex; run function;
/*
print 'After ' count ' iterations, the achieved ql is ' ql;
print 'Parameter Estimates -----';
if fixdelta=0 then do;
    print 'alpha ' alpha ' beta ' beta ' delta ' delta;
end;
else do;
    print 'alpha ' alpha 'beta ' beta ' fixed delta ' delta;
end;
*/
finish optima;

```

```

*Define procedure for calculating target ql (for constructing confidence interval);
start target;
  %parameter_estimates(fixdelta=0,delta=1);
  qfull=ql;deltahat=delta; taufull=tau;
  f=finv(0.95,1,n-p);
  compare=qfull-0.5*taufull*f;
  *print 'Compare restricted qls to ' compare;
finish target;

*Define procedure for finding the lower bound on the dose threshold;
start find_lb;
  delta_low=0; delta_high=deltahat;
  %parameter_estimates(fixdelta=1,delta=delta_low);
  qred=ql;
  criterion=abs(qred-compare);
  if criterion<0.5 then lb=delta_low;
  else do;
    if qred>compare then do until (qred<compare | criterion<0.5);
      temp_low=delta_low-1;
      %parameter_estimates(fixdelta=1,delta=temp_low);
      qred=ql; criterion=abs(qred-compare);
      delta_low=temp_low;
    end;
    if qred<compare then do until (criterion<0.5);
      delta_mid=(delta_low+delta_high)/2;
      %parameter_estimates(fixdelta=1,delta=delta_mid);
      qred=ql; criterion=abs(qred-compare);
      if qred>compare then delta_high=delta_mid;
      else delta_low=delta_mid;
    end;
    if criterion<0.5 then lb=delta_mid;
  end;
  *print 'The lower confidence bound on the threshold is ' lb '.';
finish find_lb;

```



```

*Define procedure for finding the upper bound on the dose threshold;
start find_ub;
  delta_low=deltahat; delta_high=3;
  %parameter_estimates(fixdelta=1,delta=delta_high);
  qred=ql;
  criterion=abs(qred-compare);
  if criterion<0.5 then ub=delta_high;
  else do;
    if qred>compare then do until (qred<compare | criterion<0.5);
      temp_high=delta_high+1;
      %parameter_estimates(fixdelta=1,delta=temp_high);
      qred=ql; criterion=abs(qred-compare);
      delta_high=temp_high;
    end;
    if qred<compare then do until (criterion<0.5);
      delta_mid=(delta_low+delta_high)/2;
      %parameter_estimates(fixdelta=1,delta=delta_mid);
      qred=ql; criterion=abs(qred-compare);
      if qred>compare then delta_low=delta_mid;
      else delta_high=delta_mid;
    end;
    if criterion<0.5 then ub=delta_mid;
  end;
*print 'The upper confidence bound on the threshold is ' ub '.';
finish find_ub;

*Find mqls for original experiment;
x=orig_x; y=orig_y;
%parameter_estimates(fixdelta=0,delta=1);
taufixed=round(tau,0.01); alphafixed=round(alpha,1); gammafired=100-alphafixed;
betafixed=round(beta,0.01); deltafixed=round(delta,0.01);
print 'The Original Experiment -----';
print 'There were ' n ' observations contained in the original experiment.';
print 'After ' count ' iterations, the achieved quasi-likelihood is ' ql '.';
print 'The resulting parameter estimates are alpha ' alphafixed ', beta '
      betafixed;
print 'and delta ' deltafixed '.';

```

```

print 'The associated estimate of tau is ' taufixed '.';
*Calculate and sort residuals from original experiment;
stdres=(y-mu)#(1/(sqrt(tau#mu)));
temp=stdres;
stdres[rank(stdres),]=temp;
stdres_unsorted=temp; stdres_sorted=stdres;
*print stdres_unsorted stdres_sorted;

*Find qlr-ci for original experiment;
run target;
%findlb;
print 'The lower confidence bound on the dose threshold is ' lb n;

use doses4;
read all var {x1 x2 x3 x4} into start_doses;

*tempd={0 0.7 7.35 14};
index=&index;
tempd=start_doses[index,];
%design(m=4,startvals=tempd,design_n=100);
edit lrci4pt var {index count2 _dchk_2 _dchk_22
                x1 q1 n1
                x2 q2 n2
                x3 q3 n3
                x4 q4 n4
                initial final lbest
                xflag qflag};
append var {index count2 _dchk_2 _dchk_22
            x1 q1 n1
            x2 q2 n2
            x3 q3 n3
            x4 q4 n4
            initial final lbest
            xflag qflag};
%put 'Design for starting value set ' &index ' completed.';

quit;

%end;

```

```
%end designloop;
```

```
%designloop;
```

Vita

Sharon Dziuba Yeatts was born on August 10, 1978, in Perth Amboy, New Jersey. After a number of moves, she graduated from Midlothian High School in Midlothian, Virginia in May of 1996. She graduated Magna Cum Laude from the University of South Carolina in Columbia, South Carolina, with a Bachelor's degree in Statistics in May of 2000. During the summer of 2000, Sharon completed a statistical internship with ProMetrics Consulting, Inc. She returned to the University of South Carolina for graduate school and received her Master of Science in Statistics in May, 2002. In August of 2002, Sharon began a doctoral program in Biostatistics at Virginia Commonwealth University in Richmond, Virginia. She spent the summer of 2003 working as a statistical intern under contract with the Environmental Protection Agency. She has been funded for two years (2004-2006) by a National Institute of Environmental Health Sciences training grant entitled "Integration of Chemical Mixtures Toxicology and Statistics," #T32 ES007334. Sharon has accepted a position as Research Assistant Professor with the Department of Biostatistics, Bioinformatics and Epidemiology at the Medical University of South Carolina, in Charleston, South Carolina, upon completion of her doctoral program.