Title
Bayesian Methods in the Quantitative Risk Assessment and Toxicity Profiling of Engineered Nanomaterials

Author
Patel, Trina Ramesh

Publication Date
2012

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Bayesian Methods in the Quantitative Risk Assessment and Toxicity Profiling of Engineered Nanomaterials

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Public Health

by

Trina Ramesh Patel

2012

Abstract of the Dissertation

Bayesian Methods in the Quantitative Risk Assessment and Toxicity Profiling of Engineered Nanomaterials

by

Trina Ramesh Patel

Doctor of Public Health

University of California, Los Angeles, 2012

Professor Donatello Telesca, Chair

Until recently, very little research has been conducted to assess the potential human health hazards associated with engineered nanomaterials (ENMs). *In-vitro* high-throughput screening (HTS) assays for the assessment of engineered nanomaterials provide new opportunities to learn how these particles interact at the cellular level, and may aid in reducing the demand for *in-vivo* testing. The large number of potential factors that could link nanomaterials to adverse human health impacts, create an imperative need to develop a stronger foundation for quantitate risk assessment in nanotoxicology.

In this dissertation we propose a probability model for the analysis of high-throughput cellular assays. In particular, we develop a method that builds a balance between model complexity and interpretability as a tool to be used by subject-matter specialists for assessing cytotoxicity. The resulting multivariate surface-response model allows for joint inference on dose and time kinetics, and associated classical risk assessment parameters of interest. We illustrate the proposed methodology by profiling a multivariate screening study of eight metal-oxide nanomaterials. Next, we present loss-function-based methods for the hazard ranking of engineered nanomaterials. Specifically, we provide a decision-making tool for prioritizing extensive *in-vivo* testing of emerging nanomaterials. The proposed framework allows for the aggregation of ranks across different sources of evidence while allowing for

differential weighting of this evidence based on its reliability and importance in risk ranking. We illustrate the methodology by ranking particles from a multivariate cytotoxicity screening study of eight metal oxides, conducted in two human cell-lines. Finally, we propose methodology for modeling the relationship between physicochemical properties of ENMs and their observed cytotoxicity, as an initial step in the development of a framework for predictive nanotoxicology. In particular, the proposed approach introduces a new measure of toxicity that is seamlessly integrated into a multi-dimensional model that accounts for dose and duration kinetics jointly using a flexible smooth surface fit. Moreover, the designed approach is appropriate for small sample size, and includes data integration and a framework for advanced dimension reduction through variable selection. The proposed method was applied to a library of 24 engineered nanomaterials.

The dissertation of Trina Ramesh Patel is approved.

Tom Belin

Catherine Sugar

André E. Nel

Donatello Telesca, Committee Chair

University of California, Los Angeles

2012

To my mom & dad,
who struggled through the
hardships of a foreign country
so that I could succeed;

and

to my niece Sia,
whose smile makes everyday
seem so much brighter.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my deep gratitude to Professor Donatello Telesca for his full support and guidance in the preparation of this dissertation. I am also very thankful to Professors Tom Belin, Catherine Sugar, André Nel, and Hilary Godwin for kindly serving on my doctoral committee and providing insights and suggestions during the development of this project.

Primary support was provided by the U.S. Public Health Service Grant U19 ES019528 (UCLA Center for Nanobiology and Predictive Toxicology). This work was also supported by the National Science Foundation and the Environmental Protection Agency under Cooperative Agreement Number DBI-0830117. Any opinions, findings, conclusions or recommendations expressed herein are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the Environmental Protection Agency. This work has not been subjected to an EPA peer and policy review.

Finally, I would like to thank my family and friends, especially my parents Ramesh and Panna, my sister Purvi, my brother Kartik, my brother-in-law Brijesh, and my best friends Minal and Jashmina, for their unfailing support during my graduate studies.

Vita

1997-2001	High School, Western H.S., Anaheim, California.
2005	B.S. (Bioengineering), UCLA, Los Angeles, California.
2004–2006	Staff Research Associate, UCSD Department of Bioengineering
2008	M.S. (Biostatistics), UCLA, Los Angeles, California.
2007–2008	Teaching Assistant, UCLA Biostatistics, Los Angeles, California
2007–2008	Data Manager and Clinical Coordinator, Infosphere Clinical Research, Los Angeles, California
2007–2008	Biostatistics Graduate Student Researcher, UCLA Health Services Research Center, Los Angeles, California
2009	Biostatistics Intern, Amgen Inc., Thousand Oaks, California
2010–2012	Biostatistics Graduate Student Researcher, UC Center for Environmental Implications of Nanotechnology, Los Angeles, California
2012–present	Senior Biostatistician, Edwards Lifesciences Inc., Irvine, California

Publications

Trina Patel, Donatello Telesca, Saji George, Andre E. Nel. (2012). Toxicity Profiling of Engineered Nanomaterials via Multivariate dose-response-surface Modeling. In Press, *Annals of Applied Statistics*.

Trina Patel, Donatello Telesca, Robert Rallo, Saji George, Tian Xia, Andre E. Nel. (2012). Hierarchical Rank Aggregation with Applications to Nanotoxicology. In Review, *Journal of Agricultural, Biological, and Environmental Statistics*.

CHAPTER 1

Introduction

Nanomaterials are a large class of substances engineered at the molecular level to achieve unique mechanical, optical, electrical, and magnetic properties. Nanotechnology is a rapidly growing field with over 800 consumer products on the market and possibly thousands of engineered nanomaterials (ENMs) under investigation (Tsuji et al. 2005, Society 2004). Examples of products manufactured using nanomaterials include sunscreens, cosmetics, electronics, and stain resistant clothing, among others (Nel et al. 2006). The materials that result from nanotechnology are projected to play a key role in meeting many fundamental needs currently facing society. Areas of impact include, but are not limited to, health care, food, safety, energy, environmental protection, and technology. Some of the current and potential research areas include new treatment methods for cancer, drug delivery, earth-quake safe buildings, food packaging to improve shelf-life, and more efficient energy conversion based on renewable energy sources (DECHEMA/VCI 2011)

The unusual properties of nanomaterials can be attributed to their small size (with one dimension less than 100 nm), chemical composition, surface structure, solubility, shape, and aggregation. The structural and chemical properties of ENMs allow for increased uptake and biological interaction with tissues, which would not otherwise be possible with the bulk form of the material. Many biological processes take place at the nanoscale level, and the introduction of ENMs into living organisms could lead to harmful interactions with molecular and cellular processes that are critical to life. In particular, many nanomaterials are able to move throughout the body, deposit in organs, and penetrate cell membranes, with the potential for cytotoxicity among other injury responses. For example, inhaled ultra-fine particles have been long understood to be capable of inducing oxidative stress and

pulmonary inflammation. Similarly, preliminary experiments in animal models indicate that exposure to titanium dioxide (TiO_2) and carbon black nanoparticles lead to oxidative stress and cytotoxic responses (Nel et al. 2006, Nel et al. 2009).

Nanotechnology does have the opportunity for great impact, especially in public health and medicine, but it confers enormous potential for human exposure and environmental contamination. In order to make this technology and the opportunities it presents sustainable, it is crucial to ensure the safety of nanomaterials (DECHEMA/VCI 2011). The potential for exposure as well as hazard, has spurred recent interest in early identification of potentially hazardous nanomaterials. Of utmost importance is the need to make decisions regarding the safety and potential toxicity of these particles to humans and the environment. Knowledge about the potential hazard of nanomaterials is still lacking and extensive study is required to understand how ENM properties such as size, shape, agglomeration state, solubility, and surface properties could lead to hazard generation at the nano-bio interface (Stern and McNeil 2008, Nel et al. 2006).

Current research in nanotoxicology includes new generation high-throughput screening (HTS) assays, which enable the simultaneous observation of multiple cellular injury pathways across an array of doses and times of exposure. These rapid screening approaches include the use of fluorescence-based cellular assays that assess key signals of nanoparticle toxicity in various cell types. An example of such an assay is one developed to measure the cytotoxicity of metal and metal oxide nanoparticles in relation to various cellular injury responses. This assay is particularly important, as oxidative stress constitutes one of the principal injury mechanisms though which ENMs can induce adverse health effects (Zhang et al. 2012). The assay is derived from the hierarchical oxidative stress model of Nel et al. (2006) and Xia et al. (2006), which postulates that low levels of oxidative stress (Tier 1) lead to a protective cellular response, while higher levels of oxidative stress will overwhelm antioxidant defense mechanisms and could trigger a Tier 2 response, characterized by pro-inflammatory effects. Finally toxic oxidative stress (Tier 3), the highest level, can trigger intracellular calcium flux, increased permeability of the mitochondria, and possibly cytotoxicity. The

hierarchical oxidative stress model can be useful for screening nanomaterials that are capable of generating oxygen radicals, such as metal and metal oxide nanomaterials, which can lead to cell damage via interference in important intracellular processes (George et al. 2009).

HTS assays provide an opportunity for the rapid testing of a large number of nanomaterials. Cellular assays help to define biological relationships and can be used to assess cytotoxicity by determining important measures of hazard. A large battery of HTS *in-vitro* studies can also be used as a screening tool for the hazard ranking and subsequent prioritization of a reduced set of *in-vivo* studies. While HTS assays cannot replace traditional animal studies, they are less costly and labor intensive, and they can be used to explore the large number of potential nanomaterial variables that can influence human health hazard (Meng et al. 2010, Stanley et al. 2008, Maynard et al. 2006). Finally, HTS assays aid in the development of a framework for predictive toxicology, as the ultimate goal is to limit the demand for *in-vivo* studies (Nel et al. 2006). Specifically, HTS assays provide an opportunity to discover how specific physicochemical characteristics of ENMs contribute to the definition of injury pathways, a required step in the predictive approach. The feasibility and utility of HTS assays have been illustrated in various fields such as functional genomics, with the use of microarray technology, as well as in pharmacology with the rapid screening of potential drug targets (Hoheisel 2006, White 2000).

HTS data are often characterized by high dimensionality, relatively small sample sizes, and high measurement error. From a statistical perspective, some of the inferential challenges include data quality and control, the formulation of robust probability models that are flexible yet interpretable, integration of heterogeneous sources of data, making non-standard inferences such as ranks, and the simultaneous testing of a large number of hypotheses. Bayesian Analysis has been shown to provide an advantageous inferential framework to deal with these challenges. For example, Telesca and Inoue (2009) introduced a Bayesian probability model for the analysis of high-throughput time course microarray data with a similar structure to the HTS data described above. Specifically, the model was used to select genes with differential activity as well as to provide a measure of similarity between genes.

Traditionally methodological research in risk assessment has involved both epidemiological and animal lab studies in order to establish safe levels for human exposure. Proper risk assessment involves the processes of hazard identification, hazard characterization, and exposure assessment (Edler et al. 2002). Numerous quantitative methods have been developed for the purposes of hazard identification and characterization using these sources of data.

Toxicity assessment is traditionally performed by calculating various measures of hazard, often using a dose escalation study. For example, in many instances it is believed that a exposure threshold exists, either biologically or statistically, below which no significant effect exists. In these cases, methods which estimate the threshold are used to derive exposures below which the toxin is considered to have no biologically significant adverse health effect. Some classical measures of the exposure threshold include the no-observed-adverse-effect level (NOAEL) and the lowest observable adverse effect level (LOAEL). The NOAEL is defined as the highest level of exposure at which no statistically significant effect occurs. Usually this is calculated by comparing treated animals to untreated animals. Most often, the entire dose-response curve is not utilized in the quantitative estimation of the NOAEL. In fact, only the threshold value is used in the calculation, and all other values only contribute in ensuring an appropriate number and spacing of data points. This method of estimation leads to imprecise estimates that are not very robust. The lowest-observed-adverse-effect level (LOAEL) is a similar measure and is used in situations where all tested exposure levels produce some statistically significant effect when compared to the control group (Edler et al. 2002). The benchmark dose (BMD) is an alternative method introduced by the EPA in which a dose-response model is used to estimate the lower statistical confidence limit for the dose at which adverse effects rise above a predetermined amount as compared to the background. In this way, the BMD attempts to use more of the data in obtaining an estimate of risk (Edler et al. 2002). Various methods for modeling dose-response data are discussed in Chapter 2.

The classical risk assessment parameters described above, summarize different aspects of the response trajectory, and are typically assessed disjointly for dose and time kinetics. In fact, there is still disagreement in the HTS setting on the best measures of risk (Stern

and McNeil 2008 and Maynard et al. 2006). Furthermore, some of these summaries can become even more problematic in the nanotoxicology setting due to issues such as dosimetry, where the administered doses are confounded by different particle bioavailability, effecting the concentration of material that is actually absorbed by the cell.

Methods for decision making and prioritization for further testing of materials often take the form of hazard ranking. The most common approaches include simple scoring methods or methods based on ordering estimates of risk assessment parameters.

The hazard ranking of chemicals has also been extensively studied, especially using partial order techniques and multi-criteria analysis (see Lerche et al. (2002) for a recent comparison of these methods). For example, Lerche and S∅rensen (2003) consider the ranking of objects using partial order theory and random linear extensions, and Lerche et al. (2004) apply these ideas to the ranking of chemicals. These methods are extensions of simple scoring methods, and again, are designed to work with summaries of previously analyzed data sets. Their direct applicability to the general HTS setting, described earlier, is therefore limited.

Finally, structure-activity relationships (SAR) are often used as a tool for risk assessment in situations where adequate toxicity data is unavailable. Quantitative structure-activity relationships (QSAR) are models that relate the biological activity of a material to its chemical and physicochemical properties. These models can yield valuable predictive power and can be used to understand the mechanism of toxicity of a material. The use of a combination of molecular properties to predict a compound's behavior with respect to biological end-points is a well accepted concept in the predictive toxicology of chemicals (Schultz et al. 2003). From a regulatory standpoint these methods are primarily used in circumstances where the risk of exposure is extremely low and require a strict validation process (Barlow et al. 2002, Edler et al. 2002). In contrast to QSARs for chemicals, the idea of QSARs for nano-sized materials is still in early development (Puzyn et al. 2009). The primary efforts to date have been on adapting the idea of classical chemical QSAR models to nanomaterials. The limited knowledge associated with nano-sized materials, especially in terms of physicochemical and toxicity data, the lack of a well defined toxicological endpoint, and the high variability in

molecular structures, make it difficult to implement the classical QSAR approach. Therefore, new methodology that is appropriate for small sample sizes and includes data integration as well as advanced dimension reduction, is an important area of research.

The objective of this dissertation is to use the information obtained from high-throughput screening studies to develop a quantitative foundation for risk assessment in nanotoxicology. In particular, we aim to develop a framework for toxicity assessment, prioritization of *in-vivo* testing of ENMs, and predictive nanotoxicology. To our knowledge the methodologies proposed in this dissertation are the first to provide formal statistical methodology for the analysis of the data structures introduced by the high-throughput screening assays described above. As the field of nanotoxicology, especially in regards to high-throughput screening assays, is relatively new, we are currently aware of only ad hock procedures, often limited to simple data pre-processing and visualization.

In Chapter 2 we propose a probability model for the analysis of high-throughput cellular assays. We make an effort to provide an analytic tool that builds a balance between model complexity and interpretability, as we hope that the proposed framework will be adopted by subject-matter scientists as a method for assessing cytotoxicity. From a methodological perspective, we propose a multivariate surface-response model that allows for joint inference on dose and time kinetics, along with associated classical risk assessment parameters. We illustrate the proposed methodology by profiling a multivariate screening study of metal oxides. In Chapter 3 we discuss loss-function based ranking methods and apply them to the analysis of HTS assays. We seek to provide an analytic tool that can be used as a decision making tool for prioritization of extensive *in-vivo* testing of emerging nanomaterials. Specifically, we propose a framework for the aggregation of ranks across different sources of evidence, while allowing for differential weighting of this evidence based on its reliability and importance in risk ranking. We applied the proposed methodology to the hazard ranking of nanomaterials using data from a multivariate cytotoxicity screening study of metal oxides, conducted in two different cell-lines. Finally, in Chapter 4 we propose methodology for modeling the relationship between physicochemical properties of ENM and their observed cytotoxicity. The

model described aims to develop a foundation for predictive nanotoxicology. In particular, the proposed methodology introduces a new measure of toxicity that is seamlessly integrated into a multi-dimensional model that accounts for dose and duration kinetics jointly using a flexible smooth surface fit. Moreover, the designed approach is appropriate for small sample sizes, and includes a framework for advanced dimension reduction through variable selection. We applied the proposed methodology to a material library of 24 metal oxide nanomaterials.

CHAPTER 2

Toxicity Profiling of ENMs

2.1 Overview

New generation *in-vitro* high-throughput screening (HTS) assays for the assessment of engineered nanomaterials provide an opportunity to learn how these particles interact at the cellular level, particularly in relation to injury pathways. These types of assays are often characterized by small sample sizes, high measurement error and high dimensionality, as multiple cytotoxicity outcomes are measured across an array of doses and durations of exposure. In this chapter we propose a probability model for the toxicity profiling of engineered nanomaterials. A hierarchical structure is used to account for the multivariate nature of the data by modeling dependence between outcomes and thereby combining information across cytotoxicity pathways. In this framework we are able to provide a flexible surface-response model that provides inference and generalizations of various classical risk assessment parameters. We discuss applications of this model to data on eight nanoparticles evaluated in relation to four cytotoxicity parameters.

2.2 Introduction

Current research in nano-toxicology includes new generation high-throughput screening (HTS) assays that enable the simultaneous observation of multiple cellular injury pathways across an array of doses and times of exposure. In this chapter, for example, we analyze data on eight metal and metal-oxide nanoparticles, monitored in relation to four cellular injury responses derived from the hierarchical oxidative stress model of Nel et al. (2006) and Xia

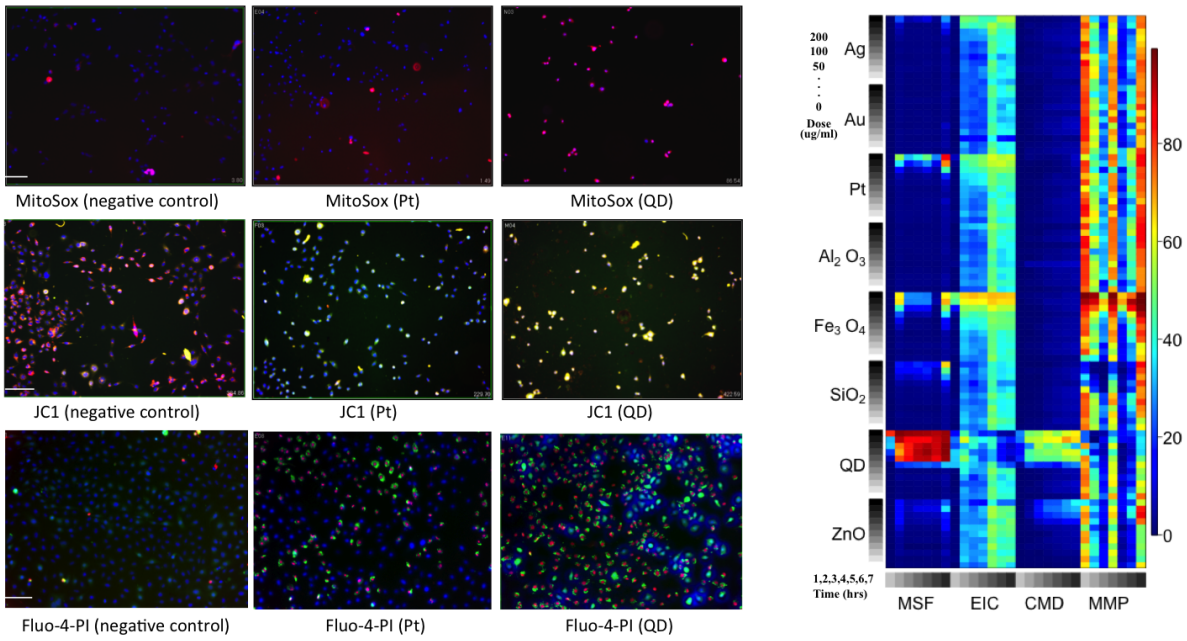

Figure 2.1: **Fluorescence images and heat map of raw data.** On the left are fluorescence images of RAW cells treated with various nanomaterials (quantum dot, platinum, and a negative control) and dyed with compatible dye combinations including MitoSox, JC1, PI, and Fluo-4. The subsequence fluorescence read-out, measured at varying wavelengths, provides a measure of the number of cells positive for the response. On the right is a heat map of the raw data for each particle and outcome. Colder colors indicate a smaller percentage of cells positive for the response and warmer colors indicate a larger percentage of cells positive for the response.

et al. (2006). All four outcomes are measured contemporaneously over a grid of ten doses and seven hours of exposure (see Figure 2.1). The four measured responses include mitochondrial superoxide formation, loss of mitochondrial membrane potential, elevated intracellular calcium, and membrane damage (George et al. 2009). For increasing dosage and duration of exposure, we observe typical dose-response kinetics with outcomes possibly depending on one another.

In toxicology, risk assessment involves the characterization of hazard as well as the potential for exposure. The HTS framework provides a wealth of information about cellular injury pathways, but proves a challenge for the classic risk assessment paradigm. In fact, there is still disagreement in the HTS setting on how to define, and the appropriateness of,

classical risk assessment parameters such as the no observable adverse effect level (NOAEL), the lowest observable adverse effect level (LOAEL), and the dose that produces 50% of the maximum response (EC50), among others.

Parametric functions such as families of sigmoidal curves are frequently used to fit dose-response data. Some commonly used sigmoidal models include log-logistic models, log-normal models, and Weibull models (see Ritz 2010 for a recent review of these models). The log-logistic functions are the most frequently used for modeling dose-response data in toxicology. The four parameter log-logistic model can be expressed as follows:

$$f(x;\ b, c, d, h) \quad = \quad c + \frac{d - c}{1 + \exp[b\{\log(x) - \log(h)\}]}. \tag{2.1}$$

Here h, the inflection point in the curve, provides a convenient risk assessment parameter since it can be interpreted as the 50% effective or inhibitory dose (EC50, IC50) (Emmens 1940). Other special cases of this model, frequently used in toxicology, include the 3 parameter log-logistic model which leads to the famous Hill equation (Hill 1910) and special cases of the Michaelis-Menton kinetics. Further extensions of these models include the five parameter log-logistic function, which provide a bit more flexibility by allowing the function to be asymmetric (Finney 1979), and the Brain-Cousens model, which includes an extra parameter to account for a possible favorable response to a toxin at low concentrations (Calabrese and Baldwin 2003). In general these models assume that the dose-response function is completely known apart from the few parameters to be estimated.

Several other methods have been proposed to model nonlinear dose-response relationships relaxing strictly parametric assumptions. Ramsay (1988) proposed the use of monotone regression splines to model a dose-response function. In this case, piecewise polynomials or splines can allow greater flexibility, while achieving monotonicity by imposing constraints on the estimated function. Li and Hunt (2004) proposed the use of linear B-splines with one random interior knot to model a nonlinear dose-response curve. In this context, the random interior knot provides inference on the dose at which the toxin begins to takes effect and thereby provides a useful parameter for risk assessment. Kong and Eubank (2006) suggested

the use of functions that combine smoothing spline techniques and the non-negativity properties of cubic B-splines to estimate the dose-response curve. The use of non-parametric techniques to estimate dose-response curves often provides a more realistic representation of the data generating process. At the same time, however, some of these techniques make it more difficult to interpret the model in terms of classical risk assessment.

Recent literature advocates the simultaneous use of multiple outcomes to assess risk. Regan and Catalano (1999) proposed a bivariate dose-response model that accounts for the dependence among outcomes of developmental toxicity using generalized estimating equations. Geys et al. (2001) proposed a similar model for risk assessment of developmental toxicity, but approached the problem using latent variables. Yu and Catalano (2005) suggested a model for quantitative risk assessment of bivariate continuous measures of neurotoxicity using percentile regression. These methods are often aimed at the analysis of one potentially toxic agent as it relates to adverse events or continuous outcomes observed in association with exposure over a range of doses. Their direct applicability to the general HTS setting described earlier is therefore limited.

From a statistical perspective, cellular interrogation data based on high-throughput platforms can be characterized as multivariate dependent observations. Each nanoparticle is indeed associated with a multiple set of cellular outcomes recorded both longitudinally, in relation to different exposure durations, and cross-sectionally, in relation to a dose escalation design. This particular design structure suggests that valid statistical inference must account for potentially complex patterns of dependence between different observations. A reasonable dependence scheme might, for example, assume data to be dependent within outcome and particle, as well as between outcomes for the same particle.

In conjunction with considerations related to the joint sampling distribution of these data structures, appropriate statistical treatment must account for non-linearities in the mean response associated with dose and duration dynamics. While, in principle, one can choose to define a random response surface in a completely non-parametric fashion, it is important to maintain a certain degree of interpretability, especially in relation to standard hazard

assessment quantities of interest to substantive scientists. In summary, perhaps reductively, the overall modeling challenge lies in the definition of a flexible and interpretable probabilistic representation for a family of dependent dose-response random surfaces.

In this chapter, we propose a hierarchical dose-response model for the analysis of HTS data from nanotoxicology. Our model builds on earlier work (Hastie and Tibshirani 1986, Li and Hunt 2004), expanding on them to account for the multivariate nature of the data and to address the estimation of a series of two dimensional dose-response surfaces. We provide a flexible framework for modeling dose and duration-response kinetics jointly, while providing inference on several risk assessment parameters of interest. We utilize a hierarchical structure to define dependence between outcomes and thereby borrow strength across injury pathways, providing the basis for a comprehensive risk assessment paradigm in HTS studies. We account for outlying observations via a t-distributed error model and describe how to carry out inference for the model parameters and their functions on the basis of simulated draws from their posterior distribution. To our knowledge we are the first to propose a principled statistical methodology for the joint analysis of this new generation of *in-vitro* data.

The remainder of the chapter is organized as follows. In Section 2.3 we introduce the proposed model. In Section 2.4 we discuss parameter estimation and associated inferential details. Section 2.5 employs the proposed model for the analysis of 8 metal-oxide nanomaterials and describes inference for various risk assessment parameters of interest. We conclude with a critical discussion of the limitations and possible extensions of our method in Section 2.6.

2.3 Model Formulation

2.3.1 Model Description

In this section we describe a dose-response model for a general HTS study, where we monitor a multivariate continuous outcome y, corresponding to J cytotoxicity parameters, in associa-

tion with the exposure of a number of cells to I different ENMs. More precisely, let $y_{ijk}(d,t)$ denote a multivariate response corresponding to ENM i ($i = 1, ..., I$), cytotoxicity parameter j ($j = 1, ..., J$), and replicate k ($k = 1, ..., K$) at dose $d \in [0, D]$ and time $t \in [0, T]$. In typical applications one observes y over a discrete set of doses $\tilde{d} = (d_1, ..., d_{m_1})'$ and exposure times $\tilde{t} = (t_1, ..., t_{m_2})'$. However, for clarity of exposition, we simplify our notation and without loss of generality refer to a general dose $d \in [0, D]$ and time $t \in [0, T]$. We introduce the following 4-stage hierarchical model.

Stage 1: Sampling Model

The observed response of particle i, cytotoxicity parameter j, and replicate k is modeled as:

$$y_{ijk}(d,t) = m_{ij}(d,t) + \epsilon_{ijk}(d,t), \tag{2.2}$$

where $\epsilon_{ijk}(d,t) \sim N(0, \sigma^2_{\epsilon_j}/\tau_i)$. Here $m_{ij}(d,t)$ denotes the response surface for particle i and outcome j. The proposed response surface describes dose and duration kinetics for all $d \in [0, D]$ and $t \in [0, T]$ and is expected to exhibit a non-linear dynamic over these domains. The distribution of y_{ijk} is modeled in terms of the error term ϵ_{ijk} as a scaled mixture of normal random variables to account for outlying observations. The error variance is defined in terms of the measurement error variance $\sigma^2_{\epsilon_j}$, specific to cytotoxicity parameter j, and on ENM-specific variance inflation parameter τ_i. If we define the joint distribution of $\epsilon_{ijk}(d,t)$ and τ_i as $P(\epsilon_{ijk}(d,t), \tau_i) = P(\epsilon_{ijk}(d,t) \mid \tau_i, \sigma_{\epsilon_j}) P(\tau_i \mid \nu)$, choosing $\epsilon_{ijk}(d,t) \mid \tau_i, \sigma_{\epsilon_j} \sim N(0, \sigma^2_{\epsilon_j}/\tau_i)$ and $\tau_i \mid \nu \sim Gamma(\nu/2, \nu/2)$, it can be shown that the marginal density of $\epsilon_{ijk}(d,t) \mid \sigma^2_{\epsilon_j}$ is distributed as a $T(\sigma^2_{\epsilon_j}, \nu)$ (West 1984). Under this framework, we can borrow strength across all ENMs by assuming the error variance is the same, but retain robustness in the model by allowing ENM-specific departures from normality. We allow the measurement error σ_{ϵ_j} to vary between cytotoxicity parameters due to heterogeneity in the cytotoxicity outcomes.

Stage 2: Response model at the ENM by cytotoxicity parameter level

The dose-response surface $m_{ij}(d,t)$ spans two dimensions (dose and time), and is modeled in an additive fashion as described by Hastie and Tibshirani (1986). If we let $(\alpha_{ij}, \boldsymbol{\beta}'_{ij}, \boldsymbol{\phi}'_{ij}, \boldsymbol{\gamma}'_{ij},$

$\boldsymbol{\psi}'_{ij}, \boldsymbol{\delta}'_{ij}, \boldsymbol{\chi}'_{ij})'$ be a parameter vector indexing the dose-response surface $m_{ij}(d,t)$, we can then define

$$m_{ij}(d,t) = \alpha_{ij} + f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) + g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) + h_{ij}(d,t;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}). \tag{2.3}$$

Here $f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij})$ is a function modeling the effect of dose d on response j for ENM i. Similarly, $g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij})$ is the function modeling the effect of time t and $h_{ij}(d,t;\, \boldsymbol{\chi}_{ij}, \delta_{ij})$ is the function modeling the interactive effect of dose and time. More specifically, we model the interaction of dose and time in a semi-parametric fashion as $h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij})$. This parameterization allows us to retain direct interpretation of the model parameters, while avoiding over-fitting of sparse data. To ensure likelihood identifiability, we require, without loss of generality, that $f_{ij}(d=0;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) = 0$, $g_{ij}(t=0;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) = 0$, and $h_{ij}(dt=0;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}) = 0$. The parameters α_{ij} can therefore be interpreted as the background response level for each particle and outcome.

We model dose-response curves $f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij})$, duration-response curves $g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij})$, and dose-time response curves $h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij})$ as linear combinations of basis functions. Specifically, we use linear B-splines with two random interior knots as points where the slope changes in a piecewise linear fashion. Let $\mathcal{B}(x, \boldsymbol{\eta})$ denote a 4-dimensional B-spline basis with interior knots $\boldsymbol{\eta} = (\eta_1, \eta_2)'$. Also, let $\boldsymbol{\beta}_{ij} = (\beta_{ij1}, .., \beta_{ij4})'$, $\boldsymbol{\gamma}_{ij} = (\gamma_{ij1}, ..., \gamma_{ij4})'$, and $\boldsymbol{\delta}_{ij} = (\delta_{ij1}, ..., \delta_{ij4})'$ be 4-dimensional vectors of spline coefficients. The functions $f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij})$, $g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij})$, and $h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij})$ can then be represented as follows:

$$\begin{aligned} f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) &= \mathcal{B}(d, \boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij}, \\ g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) &= \mathcal{B}(t, \boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij}, \\ h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}) &= \mathcal{B}(dt, \boldsymbol{\chi}_{ij})'\boldsymbol{\delta}_{ij}. \end{aligned} \tag{2.4}$$

Identifiability restrictions, $f_{ij}(d=0;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) = 0$, $g_{ij}(t=0;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) = 0$, and $h_{ij}(dt = 0;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}) = 0$ are implemented by fixing $\beta_{ij1} = 0$, $\gamma_{ij1} = 0$, and $\delta_{ij1} = 0$, for all particles and outcomes (see Figure 2.2 for an illustration).

Modeling dose and duration-response curves as piecewise linear functions allows for considerable flexibility while maintaining direct interpretability of the model parameters. Recall

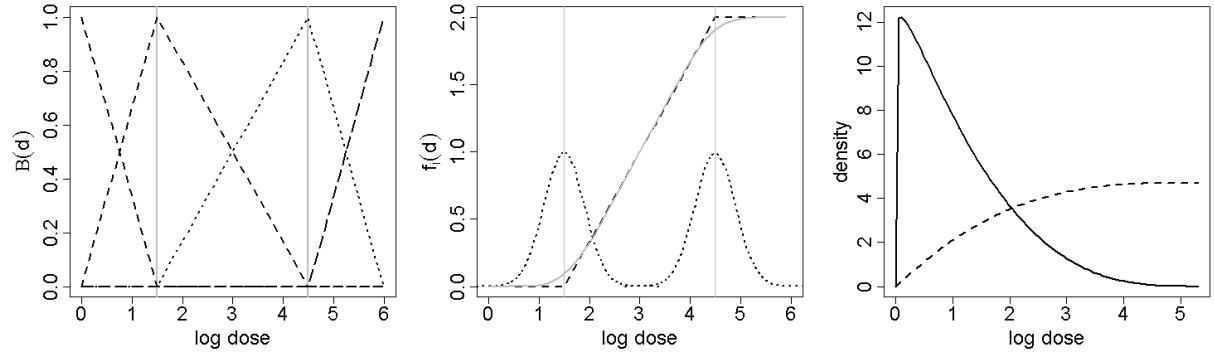

Figure 2.2: **dose-response as a change point model.** *(left)* B-spline basis function of degree 1, corresponding to change points (interior knots) at log doses of 1.5 and 4.5. *(Middle)* Example dose-response curve. The basis function on the left corresponds to a spline function with 2 change points. Each random change point has a corresponding distribution, resulting in a smooth dose-response curve. *(Right)* Example of a marginal prior distribution on the change points corresponding to the dose-response curve on the left. This formulation favors (a-priori) the choice of conservative values for the location of the first change-point *(solid line)*, and a relatively diffuse prior for our second change-point *(dotted line)*.

that in our formulation the interior knots are estimated as random quantities. This allows, marginally, for a smooth dose-response trajectory that is automatically adjusted to fit the data. The main advantage of the proposed functional representation is that, in the absence of a dose-time interaction, one can interpret the first interior knot ϕ_{ij1} as the dose at which ENM i becomes toxic in relation to cytotoxicity parameter j (Maximal Safe Dose - similar to the classical NOAEL concept). A similar interpretation can be given to ψ_{ij1}, in relation to duration-response. Note that the foregoing interpretation is contingent on fixing $\beta_{ij2} = 0$, $\gamma_{ij2} = 0$, and $\chi_{ij2} = 0$ when assuming no effect before the first change-point, and $\beta_{ij2} \leq 0$, $\gamma_{ij2} \leq 0$, and $\chi_{ij2} \leq 0$ when assuming a tonic effect before the first change-point. In the presence of a dose-time interaction, interpretation changes slightly and we instead consider the idea of safe exposure regions, which represent doses and time exposure combinations that do not induce cytotoxicity. Finally, in the absence of an interaction, the parameters ϕ_{ij2} and ψ_{ij2} are respectively interpreted as the dose and time at which the response stabilizes, or cells start a possible recovery process.

We can expand the model further to allow for the exclusion of interaction functions, where

not needed. To do that we include a latent indicator variable ρ_{ij}, so that for each particle i and outcome j

$$m_{ij}(d,t) \;\; = \begin{cases} \alpha_{ij} + f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) + g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) & \text{if } \rho_{ij} = 0 \\ \alpha_{ij} + f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) + g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) + h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}) & \text{if } \rho_{ij} = 1, \end{cases} \tag{2.5}$$

where $\rho_{ij} \sim Bern(\pi)$ and $\pi \sim U(0,1)$. We require that if $\rho_{ij} = 0$, $h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}) > 0$, to ensure identifiability. The indicator variable ρ_{ij}, can then be used to test explicitly for the dose-time interactions. The exchangeable Bernoulli trials prior on ρ_{ij} is designed to account for multiplicities (Scott and Berger 2006). This trans-dimensional parameterization is key to avoid overfitting, facilitates parameter interpretation, and allows for testing of specific scientific hypotheses related to the biological interference of nanomaterials.

For each ENM i and response j, we define the following prior distributions for α_{ij}, $\boldsymbol{\beta}_{ij}$, $\boldsymbol{\gamma}_{ij}$, and $\boldsymbol{\delta}_{ij}$:

$$\begin{aligned} \alpha_{ij} &\sim N(\alpha_{o_i}, \sigma^2_{\alpha_i}), \\ \boldsymbol{\beta_{ij}} &\sim N_4(\boldsymbol{\beta_{o_i}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta_i}}) I\{\beta_{ij1} = 0;\; \beta_{ij2} \le 0;\; (\beta_{ij3}, \beta_{ij4}) \ge 0\}, \\ \boldsymbol{\gamma_{ij}} &\sim N_4(\boldsymbol{\gamma_{o_i}}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma_i}}) I\{\gamma_{ij1} = 0;\; \gamma_{ij2} \le 0;\; (\gamma_{ij3}, \gamma_{ij4}) \ge 0\}, \\ \boldsymbol{\delta_{ij}} \mid \rho_{ij} = 1 &\sim N_4(\boldsymbol{m_{\delta_{ij}}}, \boldsymbol{v_{\delta_{ij}}}) I\{\delta_{ij1} = 0;\; \delta_{ij2} \le 0;\; (\delta_{ij3}, \delta_{ij4}) > 0\}. \end{aligned} \tag{2.6}$$

The truncated support for $\boldsymbol{\beta_{ij}}$, $\boldsymbol{\gamma_{ij}}$, and $\boldsymbol{\delta_{ij}}$ imposes functional constraints on $f(\cdot)$, $g(\cdot)$, and $h(\cdot)$, which are consistent with the expected behavior of canonical dose and duration kinetics. At the same time, however, it allows for the system to recover by permitting a decreasing slope after the second change-point. The covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta_i}}$ has diagonal elements $\boldsymbol{\sigma}_{\boldsymbol{\beta_{i\ell}}}$, $\ell = 1, .., 4$, and off diagonal elements equal to 0. Similarly for $\boldsymbol{\Sigma}_{\boldsymbol{\gamma_i}}$.

Prior distributions for $\boldsymbol{\phi}_{ij}$, $\boldsymbol{\psi}_{ij}$, and $\boldsymbol{\chi}_{ij}$ are defined to satisfy the following constraints: $(0 < \phi_{ij1} < \phi_{ij2} < D)$, $(0 < \psi_{ij1} < \psi_{ij2} < T)$, and $(0 < \chi_{ij1} < \chi_{ij2} < DT)$. More precisely, we assume that the joint distribution of the interior dose and duration knots follows a

generalized bivariate Beta density function, so that

$$\begin{aligned} \phi_{ij} &\sim B_2(a_{\phi_1}, b_{\phi_1}, a_{\phi_2}, b_{\phi_2}, D), \\ \psi_{ij} &\sim B_2(a_{\psi_1}, b_{\psi_1}, a_{\psi_2}, b_{\psi_2}, T), \\ \chi_{ij} &\sim B_2(a_{\chi_1}, b_{\chi_1}, a_{\chi_2}, b_{\chi_2}, DT). \end{aligned} \tag{2.7}$$

Here we assume that a random vector $\mathbf{x} = (x_1, x_2)'$ is distributed according to a generalized bivariate Beta distribution function ($\mathbf{x} \sim B_2(a_1, b_1, a_2, b_2, m)$), with support $\mathcal{S}(\mathbf{x}) = \{(x_1, x_2) : \ 0 < x_1 < x_2 < m\}$, if and only if:

$$\begin{aligned} p(\mathbf{x} \mid a_1, b_1, a_2, b_2, m) &= p(x_1 \mid a_1, b_1, m)\, p(x_2 \mid x_1, a_2, b_2, m) \\ &= \frac{1}{B(a_1, b_1)} \frac{x_1^{a_1-1}(m - x_1)^{b_1-1}}{m^{a_1+b_1-1}} \frac{1}{B(a_2, b_2)} \frac{(x_2 - x_1)^{a_2-1}(m - x_2)^{b_2-1}}{(m - x_1)^{a_2+b_2-1}}. \end{aligned} \tag{2.8}$$

The foregoing formulation, can be seen as a generalization of the Dirichlet distribution over a two-dimensional simplex. This general formulation can be simplified further, in order to achieve a right-skewed marginal distribution for x_1 and a uniform conditional distribution for x_2 given x_1. This is achieved by assuming $b_1 > a_1 > 1$ and $a_2 = b_2 = 1$.

Making use of this construction, we simplify the prior distribution in (2.7) as follows:

$$\begin{aligned} \phi_{ij} &\sim B_2(1, \lambda_{\phi_{i1}}, \lambda_{\phi_{i2}}, 1, 1, D) I\{\lambda_{\phi_{i2}} > \lambda_{\phi_{i1}} > 1\}, \\ \psi_{ij} &\sim B_2(1, \lambda_{\psi_{i1}}, \lambda_{\psi_{i2}}, 1, 1, T) I\{\lambda_{\psi_{i2}} > \lambda_{\psi_{i1}} > 1\}, \\ \chi_{ij} &\sim B_2(1, l_{\chi_{i1}}, l_{\chi_{i2}}, 1, 1, T) I\{l_{\chi_{i2}} > l_{\chi_{i1}} > 1\}. \end{aligned} \tag{2.9}$$

From a regulatory standpoint, this formulation favors (a-priori) the choice of conservative values for the location of the first change-point and a relatively diffuse prior distribution for our second change-point (see Figure 2.2).

Stage 3: Response model at the ENM level

For each ENM i, we exploit conditional conjugacy to define the following prior distributions

for population level parameters:

$$\alpha_{o_i} \sim N(m_{\alpha_i}, v_{\alpha_i}), \quad \boldsymbol{\beta_{o_i}} \sim N_4(\boldsymbol{m_{\beta_i}}, \boldsymbol{v_{\beta_i}}), \quad \boldsymbol{\gamma_{o_i}} \sim N_4(\boldsymbol{m_{\gamma_i}}, \boldsymbol{v_{\gamma_i}}). \tag{2.10}$$

In the absence of an interaction, the parameters $\boldsymbol{\beta_{o_i}}$ and $\boldsymbol{\gamma_{o_i}}$ represent summaries of the dose and duration-response trajectories across all outcomes and the α_{o_i} parameters represent a summary of the baseline response across all outcomes. In the presence of an interaction, we may construct these summaries conditionally on specific doses and durations of exposure.

Finally, considering the distribution introduced in (2.9), we define a prior model for population level parameters $\boldsymbol{\lambda_{\phi_i}} = (\lambda_{\phi_{i1}}, \lambda_{\phi_{i2}})$ and $\boldsymbol{\lambda_{\psi_i}} = (\lambda_{\psi_{i1}}, \lambda_{\psi_{i2}})$ as follows:

$$\lambda_{\phi_{i\ell}} \sim Gamma(a_{\lambda_{\phi i\ell}}, b_{\lambda_{\phi i\ell}}), \quad \lambda_{\psi_{i\ell}} \sim Gamma(a_{\lambda_{\psi i\ell}}, b_{\lambda_{\psi i\ell}}), \tag{2.11}$$

where $\ell = 1, 2$. The parameters $\boldsymbol{\lambda_{\phi_i}}$ and $\boldsymbol{\lambda_{\psi_i}}$ can be used to construct summaries of dose and duration-response change-points across all outcomes. Shape hyperparameters $(a_{\lambda_{\phi i\ell}}, b_{\lambda_{\phi i\ell}})$ and $(a_{\lambda_{\psi i\ell}}, b_{\lambda_{\psi i\ell}})$ can be tuned to favor more or less conservative values for the change-point locations at the particle level.

Stage 4: Hyperpriors

We complete the model by specifying prior distributions on our hyperparameters as follows:

$$\begin{array}{rclrcl} 1/\sigma^2_{\epsilon_j} & \sim & Gamma(a_{\epsilon_j}, b_{\epsilon_j}), & 1/\sigma^2_{\alpha_i} & \sim & Gamma(a_{\alpha_i}, b_{\alpha_i}), \\ 1/\sigma^2_{\beta_i} & \sim & Gamma(a_{\beta_i}, b_{\beta_i}), & 1/\sigma^2_{\gamma_i} & \sim & Gamma(a_{\gamma_i}, b_{\gamma_i}). \end{array} \tag{2.12}$$

We model our precision parameters as gamma distributions, exploiting conditional conjugacy. Again, prior parameters can be tuned to define more or less informative distributions consistent with the scale of the outcomes (Gelman 2006). Note that in our formulation, $x \sim Gamma(a, b)$ denotes a Gamma distributed random quantity with shape a and rate b, such that $E(x) = a/b$.

2.4 Estimation and Inference

2.4.1 Posterior Simulation via MCMC

Using the B-spline representation introduced in Section 2.3.1, we can write the expected j-th response level associated with ENM i, at dose d, and exposure time t as

$$m_{ij}(d,t;\ \alpha_{ij},\boldsymbol{\beta}_{ij},\cdots) = \begin{cases} \alpha_{ij} + \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} + \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} & \text{if } \rho_{ij}=0 \\ \alpha_{ij} + \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} + \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} + \mathcal{B}(dt,\boldsymbol{\chi}_{ij})'\boldsymbol{\delta}_{ij} & \text{if } \rho_{ij}=1 \end{cases}$$

Let $\boldsymbol{\beta} = \{\boldsymbol{\beta}_{ij} :\ i=1,\ldots,I,\ j=1,\ldots,J\}$ and define $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ in a similar fashion. These parameters denote the full set of spline coefficients. Furthermore, consider knot parameters $\boldsymbol{\phi} = \{\boldsymbol{\phi}_{ij} :\ i=1,\ldots,I, j=1,\ldots,J\}$, with $\boldsymbol{\psi}$ and $\boldsymbol{\chi}$ similarly defined, and background response parameters $\boldsymbol{\alpha} = \{\alpha_{ij} :\ i=1,\ldots,I, j=1,\ldots,J\}$. Finally, let $\boldsymbol{\sigma}^2_\epsilon = (\sigma^2_{\epsilon_1},\ldots,\sigma^2_{\epsilon_J})'$ and $\boldsymbol{\tau} = (\tau_1,\ldots,\tau_I)'$. If we denote with $\mathbf{Y}$ the complete set of response values for all particles and cytotoxicity outcomes, the likelihood function can be written as follows:

$$L(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\chi},\boldsymbol{\alpha},\boldsymbol{\sigma}^2_\epsilon,\boldsymbol{\tau},\boldsymbol{\rho}\mid\mathbf{Y}) \propto \prod_{i,j,k,d,t}\left[\left(\frac{\sigma^2_{\epsilon_j}}{\tau_i}\right)^{-\frac{1}{2}}\exp\left\{-\frac{(\,y_{ijk}(d,t) - m_{ij}(d,t;\ldots)\,)^2}{2\sigma^2_{\epsilon_j}/\tau_i}\right\}\right], \tag{2.13}$$

where the product is taken over all replicates k, particles i, outcomes j, doses d and times t. We are interested in the posterior distribution

$$\begin{aligned} P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\chi},\boldsymbol{\alpha},\boldsymbol{\sigma}^2_\epsilon,\boldsymbol{\tau},\boldsymbol{\rho}\mid\mathbf{Y}) \propto\ & L(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\chi},\boldsymbol{\alpha},\boldsymbol{\sigma}^2_\epsilon,\boldsymbol{\tau},\boldsymbol{\rho}\mid\mathbf{Y}) \\ & \times P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\chi},\boldsymbol{\alpha},\boldsymbol{\sigma}^2_\epsilon,\boldsymbol{\tau},\boldsymbol{\rho}), \end{aligned} \tag{2.14}$$

where the prior model $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\chi},\boldsymbol{\alpha},\boldsymbol{\sigma}^2_\epsilon,\boldsymbol{\tau},\boldsymbol{\rho})$, is fully described in Section 2.3.1. This quantity is, however, unavailable in closed analytic form, therefore we base our inference on Markov Chain Monte Carlo ($MCMC$) simulations.

The proposed posterior simulation algorithm combines Gibbs steps within Metropolis-Hastings steps in a hybrid sampler, where we update parameters component-wise (Tierney 1994). We directly sample components when closed-form full conditional distributions are

available using a Gibbs sampling algorithm (Geman and Geman 1984, Gelfand and Smith 1990); otherwise, we use the Metropolis-Hastings (MH) approach (Metropolis et al., 1953). Available full conditional distributions are given in Appendix A.1. As we are considering selection of interaction functions in a trans-dimensional setting, we implement a reversible jumps algorithm to move between models with and without the dose-time interaction function $h_{ij}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij})$ (Green 1995). The model indicator ρ_{ij} and corresponding model parameters $\boldsymbol{\delta}_{ij}$ and $\boldsymbol{\chi}_{ij}$ are updated jointly using reversible jump MCMC steps. After the model structure has been specified, the model parameters are updated from their corresponding conditional posterior distributions. The proposed sampling scheme can be summarized as follows.

1. Fixed dimensional updates. Given the current state of the latent interaction indicators ρ_{ij}, response surfaces are uniquely defined as in (2.5). Posterior sampling is standard here and proceeds by updating spline coefficients $\boldsymbol{\beta}, \boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ from their conditional posterior via direct simulation (Gibbs step - Appendix A.1). Knot parameters $\boldsymbol{\phi}, \boldsymbol{\psi}$ and $\boldsymbol{\chi}$ are updated via a MH step. For example, when sampling the interior knot parameters $\boldsymbol{\phi}$ we use an appropriate proposal kernel $q(\phi^0_{ij\ell}, \phi^1_{ij\ell})$ to efficiently construct Markov chains with the desired stationary distribution. While accounting for the fact that $\phi_{ij1} < \phi_{ij2}$, we consider uniform proposal densities of the form

$$q(\phi^1_{ij\ell} \mid \phi^0_{ij\ell}) = U(\phi^0_{ij\ell} - w_{\phi ij\ell}, \phi^0_{ij\ell} + w_{\phi ij\ell}) I(S_\phi), \tag{2.15}$$

where $\ell = 1, 2$. Here S_ϕ denotes the appropriate support and must satisfy the constraints $0 < \phi_{ij1} < \phi_{ij2} < D$. Proposed values of $\phi_{ij\ell}$ are accepted with the following probabilities:

$$\min\left\{1;\ \frac{p(\phi^1_{ij\ell} \mid y_{ijk}, \boldsymbol{\theta}_{\setminus\phi})}{p(\phi^0_{ij\ell} \mid y_{ijk}, \boldsymbol{\theta}_{\setminus\phi})} \frac{q(\phi^0_{ij\ell} \mid \phi^1_{ij\ell})}{q(\phi^1_{ij\ell} \mid \phi^0_{ij\ell})}\right\}, \quad \ell = 1, 2. \tag{2.16}$$

To tune proposal kernels, each $\phi_{ij\ell}$ was sampled using an initial value of w that was recalibrated throughout the burn-in period to achieve an acceptance rate between 30% and

70% (Roberts and Rosenthal 2001). Specifically, the acceptance rate of $\phi_{ij\ell}$ was monitored every 200 iterations throughout the burn-in period with $w_{\phi ij\ell}$ adjusted appropriately if the acceptance rate did not fall within the desired range. A similar Metropolis-Hastings scheme was adapted for sampling the duration-response parameters $\boldsymbol{\psi}$, dose-time interaction parameters $\boldsymbol{\chi}_{ij} \mid \rho_{ij} = 1$, as well as for population level knot parameters.

2. Trans-dimensional updates. We sample the model space by randomly proposing the birth or death of dose-time interaction functions $h_{ij}(\cdot)$. This is accomplished by selecting a particle i and outcome j at random and by jointly updating ρ_{ij}, $\boldsymbol{\delta}_{ij}$ and $\boldsymbol{\chi}_{ij}$. In detail:

1. For uniformly random $i \in (1, ...I)$ and $j \in (1, ...J)$, propose a systematic change $\rho_{ij}^0 \rightarrow \rho_{ij}^1 = 1 - \rho_{ij}^0$. We assume for the moment that we propose moving from $\rho_{ij}^0 = 0$ to $\rho_{ij}^1 = 1$, implying the birth of a new interaction function $h_{ij}(\cdot)$.

2. Propose new knots and spline coefficients $\boldsymbol{\delta}_{ij}^1 \sim q(\boldsymbol{\delta}_{ij}^1)$ and $\boldsymbol{\chi}_{ij}^1 \sim q(\boldsymbol{\chi}_{ij}^1)$.

3. Accept the proposed move with probability $\tau_b = min(1, R_b)$, where

$$R_b = \frac{p(y_{ijk} \mid \boldsymbol{\delta}_{ij}^1, \boldsymbol{\chi}_{ij}^1, \rho_{ij}^1, \boldsymbol{\theta}_{\setminus \boldsymbol{\delta_{ij}}, \boldsymbol{\chi_{ij}}, \rho_{ij}})}{p(y_{ijk} \mid \rho_{ij}^0, \boldsymbol{\theta}_{\setminus \boldsymbol{\delta_{ij}}, \boldsymbol{\chi_{ij}}, \rho_{ij}})} \frac{p(\boldsymbol{\delta}_{ij}^1 \mid \rho_{ij}^1) p(\boldsymbol{\chi}_{ij}^1 \mid \rho_{ij}^1)}{q(\boldsymbol{\delta}_{ij}^1) q(\boldsymbol{\chi}_{ij}^1)} \frac{p(\rho_{ij}^1)}{p(\rho_{ij}^0)}, \tag{2.17}$$

   where we use $\boldsymbol{\theta}_{\setminus \omega}$ to denote all model parameters, with the exception of ω.

In the case where the proposed move would imply a death of an interaction function ($\rho_{ij}^0 = 1 \rightarrow \rho_{ij}^1 = 0$), the acceptance probabilty would simply be $\tau_d = 1/\tau_b$.

While the proposal densities $q(\boldsymbol{\delta}_{ij})$ $q(\boldsymbol{\chi}_{ij})$ in (2.17) can in theory be defined almost arbitrarily, to guarantee efficient exploration of the model space we consider truncated multivariate normal proposals for $\boldsymbol{\delta}_{ij}$ and $\boldsymbol{\chi}_{ij}$ centered around regions of high posterior probability. Efficient optimization within the MCMC iterations is achieved using standard profile likelihood ideas (Severini and Stainwalis 1994).

2.4.2 Posterior Inference

In this section we discuss inference on ENM-specific risk assessment parameters, based on draws from the posterior distribution described in Section 2.4.1. Table 2.1 summarizes several quantities of interest including the maximal safe dose, maximal safe exposure time, and the maximal response. This list is not exhaustive. However, other risk assessment parameters of interest, such as benchmark doses (BMD) or effective concentrations (ECα) are easily obtained from our model output in a numerical fashion. In the case of a dose-time interaction, these quantities are defined conditionally on specific doses and durations of exposure.

Let $\boldsymbol{\phi_{ij}^{(n)}}$, $\boldsymbol{\psi_{ij}^{(n)}}$, $\boldsymbol{\chi_{ij}^{(n)}}$, $\boldsymbol{\beta_{ij}^{(n)}}$, $\boldsymbol{\gamma_{ij}^{(n)}}$, $\boldsymbol{\delta_{ij}^{(n)}}$, $\alpha_{ij}^{(n)}$ and $\rho_{ij}^{(n)}$, $n = 1, ..., N$, denote N MCMC draws

Parameter	Model function	Parameter interpretation
β_{3ij}^{*}	dose-response slope between ϕ_{ij1} and ϕ_{ij2}	Overall dose effect
γ_{3ij}^{*}	duration-response slope between ϕ_{ij1} and ϕ_{ij2}	Overall exposure time effect
ϕ_{1ij}	dose-response change point 1	Maximal safe dose
ψ_{1ij}	duration-response change point 1	Maximal safe exposure time
m_{ij}^{*}	Evaluated numerically	Maximal response

Table 2.1: **Risk assessment parameters.** ENM level risk assessment parameters associated with the hierarchical model introduced in 2.3.1. For each parameter we summarize its function in the model and the related interpretation as a cytotoxicity risk factor.

from the posterior distribution of $\boldsymbol{\phi_{ij}}$, $\boldsymbol{\psi_{ij}}$, $\boldsymbol{\chi_{ij}}$, $\boldsymbol{\beta_{ij}}$, $\boldsymbol{\gamma_{ij}}$, α_{ij} and ρ_{ij}. In the absence of an interaction term, posterior samples $\phi_{ij1}^{(n)}$ and $\psi_{ij1}^{(n)}$ directly provide us with an approximation of the posterior distribution for the maximal safe dose and maximal safe exposure time. We can also obtain the posterior samples for the overall dose effect, $\beta_{ij3}^{*(n)} = \beta_{ij3}^{(n)}/(\phi_{ij2}^{(n)} - \phi_{ij1}^{(n)})$, which is the slope of the dose-response curve between ϕ_{ij1} and ϕ_{ij2}. Similarly, we can obtain the posterior distribution for the overall time effect, using posterior samples $\gamma_{ij3}^{*(n)} = \gamma_{ij3}^{(n)}/(\psi_{ij2}^{(n)} - \psi_{ij1}^{(n)})$. In the presence of a dose-time interaction we can define any of the summaries described above conditionally on a given dose and time. For example, the maximal safe dose conditional on exposure time can be defined as $\min\{\phi_{ij1}, \chi_{ij1}/t\}$, and posterior samples can be obtained from $\min\{\phi_{ij1}^{(n)}, \chi_{ij1}^{(n)}/t\}$. Given posterior draws, one can proceed with

the straightforward construction of standard posterior summaries, such as means, maxima a posteriori, modes, quantiles, and credible regions. We may also be interested in testing for a dose-time interaction. The expected inclusion probability of the dose-time interaction function can be estimated using posterior draws $\rho_{ij}^{(n)}$ as $\hat{p}_{ij} = \sum_n \rho_{ij}^{(n)}/N$. Given the prior distribution described in (2.5), this posterior probability is known to adjust for multiplicities and can be used to test for a dose-time interaction. Scott and Berger (2006), for example, recommend selecting the median model, that is including all interactions for which $\hat{p}_{ij} > 0.5$. Also of interest is an estimate of the dose-response surface, $m_{ij}(d,t)$. for particle i and outcome j. This surface is, of course, defined in an infinite-dimensional space. However, given the basis-function representation introduced in Section 2.3.1, we only need finite draws from the parameter set of interest. More precisely, draws from the marginal posterior distribution of the dose-response surface for any dose $d \in [0, D]$ and time $t \in [0, T]$ are given by

$$m_{ij}^{(n)}(d,t) = \begin{cases} \alpha_{ij}^{(n)} + \mathcal{B}(d, \boldsymbol{\phi}_{ij}^{(n)})'\boldsymbol{\beta}_{ij}^{(n)} + \mathcal{B}(t, \boldsymbol{\psi}_{ij}^{(n)})'\boldsymbol{\gamma}_{ij}^{(n)} & \text{if } \rho_{ij}^{(n)} = 0 \\ \alpha_{ij}^{(n)} + \mathcal{B}(d, \boldsymbol{\phi}_{ij}^{(n)})'\boldsymbol{\beta}_{ij}^{(n)} + \mathcal{B}(t, \boldsymbol{\psi}_{ij}^{(n)})'\boldsymbol{\gamma}_{ij}^{(n)} + \mathcal{B}(dt, \boldsymbol{\chi}_{ij}^{(n)})'\boldsymbol{\delta}_{ij}^{(n)} & \text{if } \rho_{ij}^{(n)} = 1. \end{cases} \tag{2.18}$$

For each $\boldsymbol{\phi}_{ij}^{(n)}$, $\boldsymbol{\psi}_{ij}^{(n)}$, $\boldsymbol{\beta}_{ij}^{(n)}$, $\boldsymbol{\phi}_{ij}^{(n)}$, and $\alpha_{ij}^{(n)}$, $n = 1, ..., N$, we evaluate the dose-response function given in (4.19) over a grid of values $\tilde{D} = (d_1, ..., d_n)'$ and $\tilde{T} = (t_1, ..., t_n)'$. The posterior mean of the samples $m_{ij}^{(n)}$, $n = 1, ..., N$, at each value of $\tilde{D}$ and $\tilde{T}$ can be used to summarize the fit of the dose-response surface, as shown in Figures 2.3 to 2.6. Other quantities of interest include the posterior distribution of the dose-response function $f_{ij}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij})$, duration-response function $g_{ij}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij})$, and dose-time interaction function $h_{ij}(dt;\, \chi_{ij}, \boldsymbol{\delta}_{ij})$. Draws from the marginal posterior distribution of these functions for any dose $d \in [0, D]$ and time $t \in [0, T]$ are given by

$$\begin{aligned} f_{ij}^{(n)}(d;\, \boldsymbol{\phi}_{ij}, \boldsymbol{\beta}_{ij}) &= \mathcal{B}(d, \phi_{ij}^{(n)})'\boldsymbol{\beta}_{ij}^{(n)}, \\ g_{ij}^{(n)}(t;\, \boldsymbol{\psi}_{ij}, \boldsymbol{\gamma}_{ij}) &= \mathcal{B}(t, \psi_{ij}^{(n)})'\boldsymbol{\gamma}_{ij}^{(n)}, \\ h_{ij}^{(n)}(dt;\, \boldsymbol{\chi}_{ij}, \boldsymbol{\delta}_{ij}) &= \mathcal{B}(dt, \chi_{ij}^{(n)})'\boldsymbol{\delta}_{ij}^{(n)}. \end{aligned} \tag{2.19}$$

For each draw, we evaluate the dose-response functions over a grid of values $d \in \tilde{D}$, and the duration-response functions over a grid of values $t \in \tilde{T}$. As described before, standard point-wise posterior summaries can be obtained in a straightforward fashion. Simultaneous

confidence bands for the functional effect of interest can be constructed following the Monte Carlo approximation suggested by Baladandayuthapani et al. (2005).

Additional summaries of interest can be obtained in a numerical fashion. For example, the posterior distribution for the maximal response value $m^*_{ij} = \max\{m_{ij}(d,t);\ d \in [0,D], t \in [0,T]\}$, may be obtained evaluating $m^{(n)}_{ij}(d,t)$ over a fine grid of doses $\tilde{D}$ and times $\tilde{T}$. An approximate posterior draw from m^*_{ij} can be defined as $m^{*(n)}_{ij} = \max\{m^{(n)}_{ij}(d,t);\ d \in \tilde{D}, t \in \tilde{T}\}$. Given smoothness constraints on $m_{ij}(d,t)$, defined in Section 2.3.1, the foregoing procedure is likely to provide a good approximation to the posterior distribution of the maximal response value, provided $\tilde{D}$ and $\tilde{T}$ define a sufficiently detailed evaluation grid. Similar procedures may be adopted to obtain inference on other risk assessment parameters like ECαs or BMDs.

2.5 Applications

2.5.1 Synthetic Data

To assess estimation of the model presented in Section 2.3, we present a simulation study in Appendix A.2. The dose and time kinetics were simulated in an additive fashion, from various parametric functions, including both canonical and non-canonical profiles that are still reasonably interpretable under a toxicity framework. We also placed increasingly conservative priors on the population level parameters $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ in order to assess the sensitivity of the model results to our choice of prior parameters. In Appendix A.3, we provide an additional sensitivity analysis assessing model results to our choice of prior model for the change-point parameters. We compare our prior model results to both a truncated normal prior and a parameterization of the bivariate beta prior that results in a uniform prior on the simplex.

Simulation results indicate that our model is robust to model mis-specification and is not very sensitive to our choice of prior. We do, however, maintain that using the bivariate beta prior defined in (8) is likely to be more appropriate in data analytic frameworks, as the

implied stochastic behavior of the response surface, *a priori*, reflects more closely the usual biological mechanisms of toxicity. More specifically, it assigns zero probability of toxicity to zero dose and time, where toxicity is indeed not expected to occur. Furthermore, this prior accounts for issues such as dosimetry, in which the administered doses are confounded by different particle bioavailability. Therefore, in some particles toxicity is not expected to occur for doses and times greater than zero.

2.5.2 Case Study Background

We illustrate the proposed methodology by analyzing data on macrophage cells (RAW cells) exposed to eight different metal and metal-oxide nanoparticles, monitored in relation to four cytotoxicity parameters. All four outcomes are measured over a grid of ten doses and seven times (hours) of exposure (see Figures 2.3 to 2.6). Cytotoxicity screening is based on the hierarchical oxidative stress model (George et al. 2009). More specifically, a multi-parametric assay that utilizes four compatible dye combinations and subsequent change in fluorescence read-out was used to measure four responses relating to the highest tier of oxidative stress (toxic oxidative stress). The four measured responses include mitochondrial superoxide formation (MSF), loss of mitochondrial membrane potential (MMP), elevated intracellular calcium (EIC), and cellular membrane damage (CMD). Figure 2.1 provides fluorescence images of cells exposed to various nanomaterials (50 μg/mL and 3 hours), including quantum dot, platinum, and a negative control consisting of no nanomaterials. *Row 1* includes images of cells treated with a dye combination including MitoSox, which permeates the mitochondria and fluoresces red when oxidized by superoxide. Red fluorescence measured in cells treated with MitoSox is therefore a measure of mitochondrial superoxide formation. Similarly, in *Row 2* cells are treated with a dye combination including JC1, which stains the cytoplasm red in healthy cells, but forms a monomer in cells with decreased membrane potential and consequently stains the cytoplasm green. Finally, in *Row 3* cells are strained with a dye combination including Fluo-4 and Propidium Iodide (PI). In cells with damaged membranes, PI is able to permeate the cell and bind to DNA where it causes the nucleus to

emit a red florescence. Fluo-4 is a dye that emits a green fluorescence in the cytoplasm in cells with elevated intracellular calcium. Each sample was also stained with a Hoechst dye which causes all cell nuclei to emit a blue florescence, allowing for a count of the total number of cells. An analysis of the fluorescence readout, monitored at varying wavelengths, results in a measure of the percentage of cells positive for each response. Figure 2.1 also provides a heat map of the raw responses for each particle and outcome, where colder colors (blues and greens) indicate a smaller percentages of cells positive for the response and warmer colors (oranges and reds) indicate a higher percentage of cells positive for the response. The final data was normalized using a logit transformation to unconstrain the support so that it can take on values between -∞ and ∞. Our inferences are based on 20,000 MCMC samples from the posterior distribution in (4.18), after discarding a conservative 60,000 iterations for burn-in. MCMC sampling was performed in R version 2.10.0, and convergence diagnostics were performed using the package CODA (Convergence Diagnostics and Output Analysis), (Plummerm et al. 2006).

2.5.3 Case Study Analysis and Results

We fit the model described in Section 2.3.1 to the metal-oxide data-set described in the previous section. The prior on the interior knot parameters was modeled using the simplified density described in (2.9). A set of relatively non-informative $Gamma(2,1)$ and $Gamma(3,1)$ priors were considered for the components of both $\boldsymbol{\lambda_{\phi_i}}$ and $\boldsymbol{\lambda_{\psi_i}}$, along with a vague $B_2(2,3,1,1,DT)$ prior for our dose-time interaction change-point parameter $\boldsymbol{\chi}_{ij}$. We also fixed $\beta_{ij2} = 0$ and $\gamma_{ij2} = 0$, assuming no effect before ϕ_{ij1} and ψ_{ij1}, thereby allowing, in the absence of a dose-time interaction, the interpretation of ϕ_{ij1} as the maximal safe dose and ψ_{ij1} as the maximal safe exposure time. Similarly, when $\rho_{ij} = 1$, we fixed $\delta_{ij2} = 0$. We placed $Gamma(.01,.01)$ priors on the $1/\sigma_{\epsilon_j}$ parameters, $Gamma(1,.1)$ priors on all remaining precision parameters, and $N(0,10)$ priors on the α_{oi} parameters. The parameters $\boldsymbol{\beta_{oi}}$ and $\boldsymbol{\gamma_{oi}}$ are modeled as truncated multivariate normals with mean $\mathbf{1}$ and a covariance matrix with diagonal elements 10 and off diagonal elements 0. Finally, we placed a prior distribution

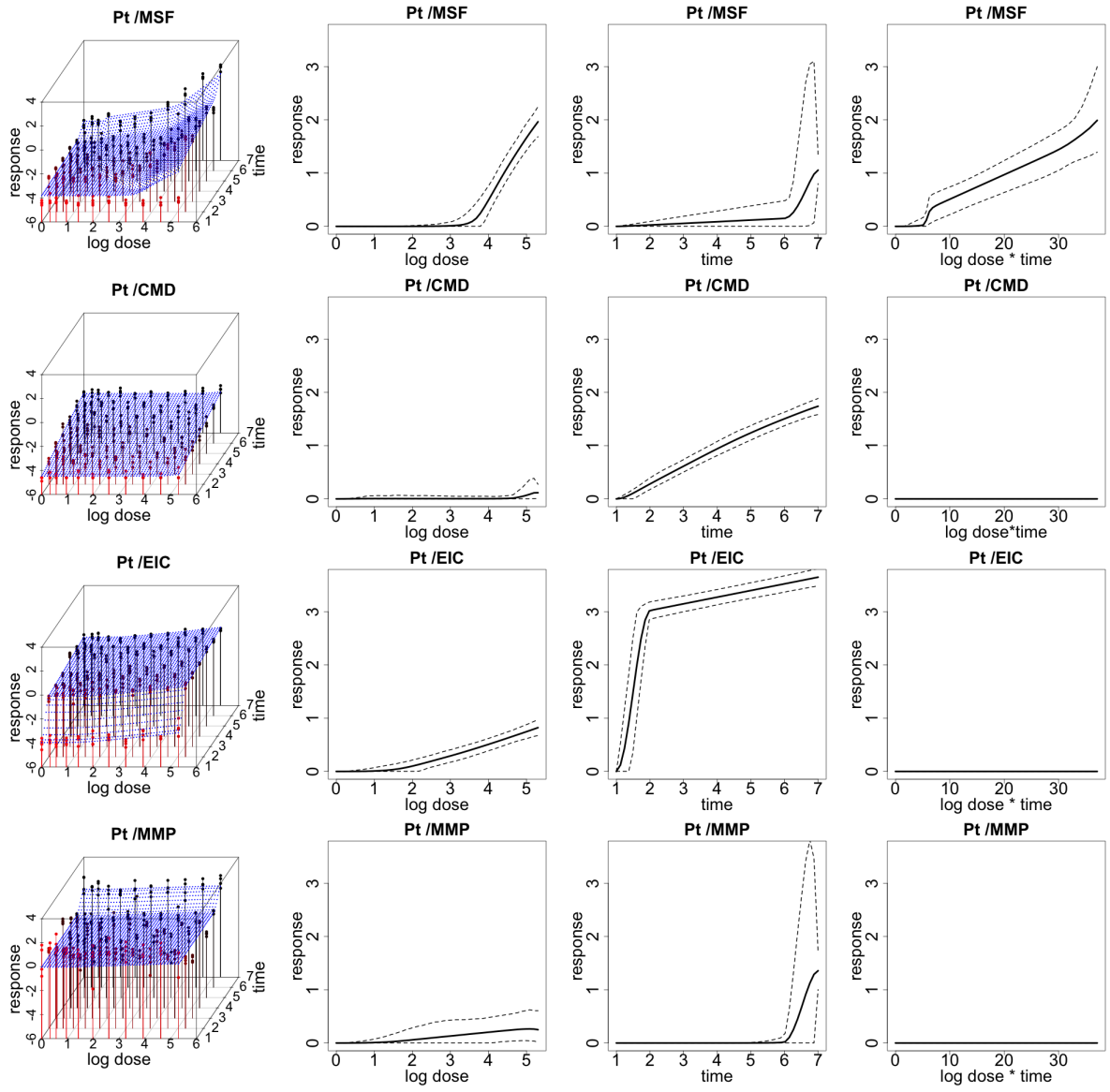

Figure 2.3: **Fitted response curves for the platinum (Pt) ENM.** Fitted response surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

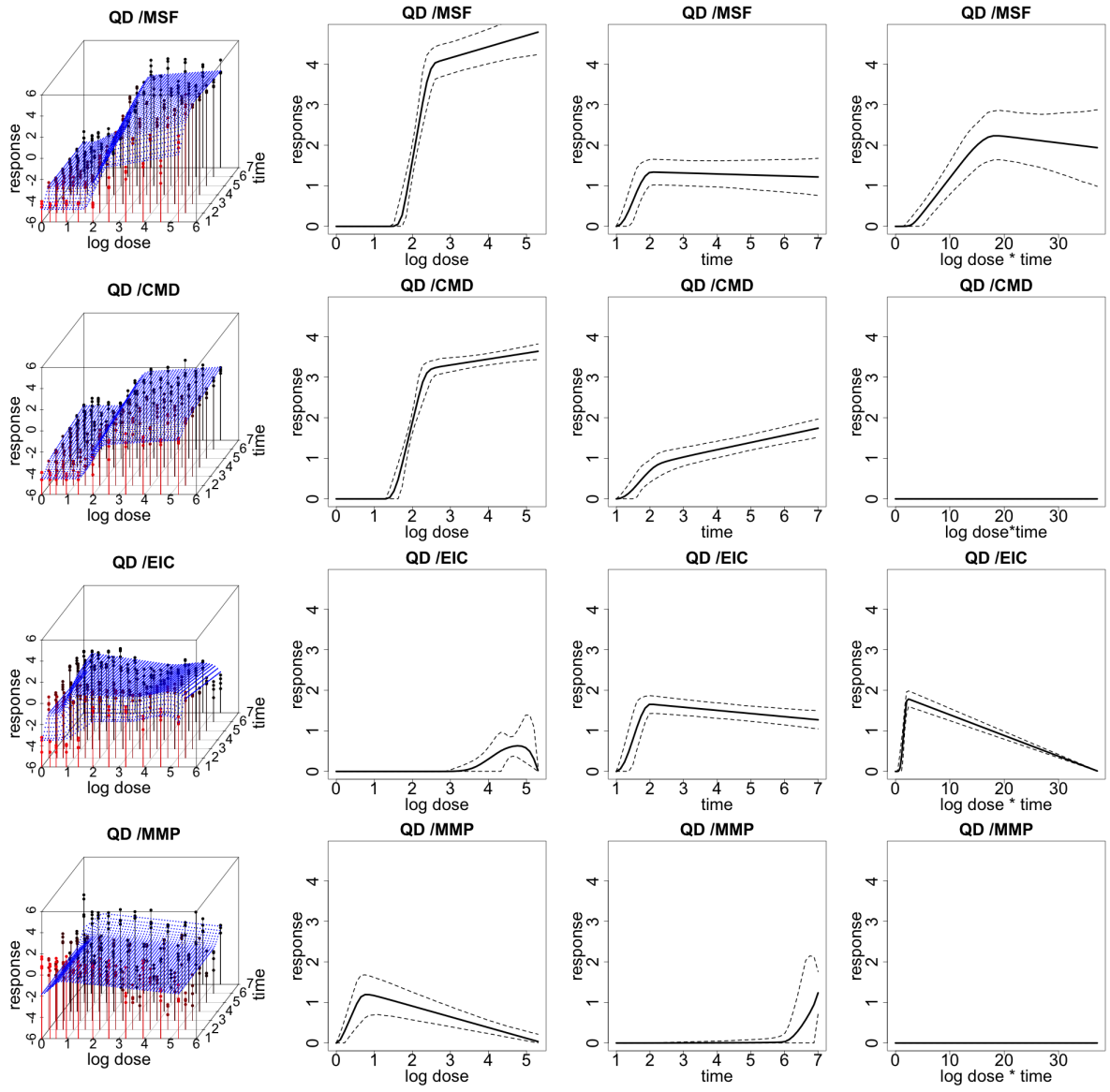

Figure 2.4: **Fitted response curves for the quantum dot (QD) ENM.** *(left)* Fitted response surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

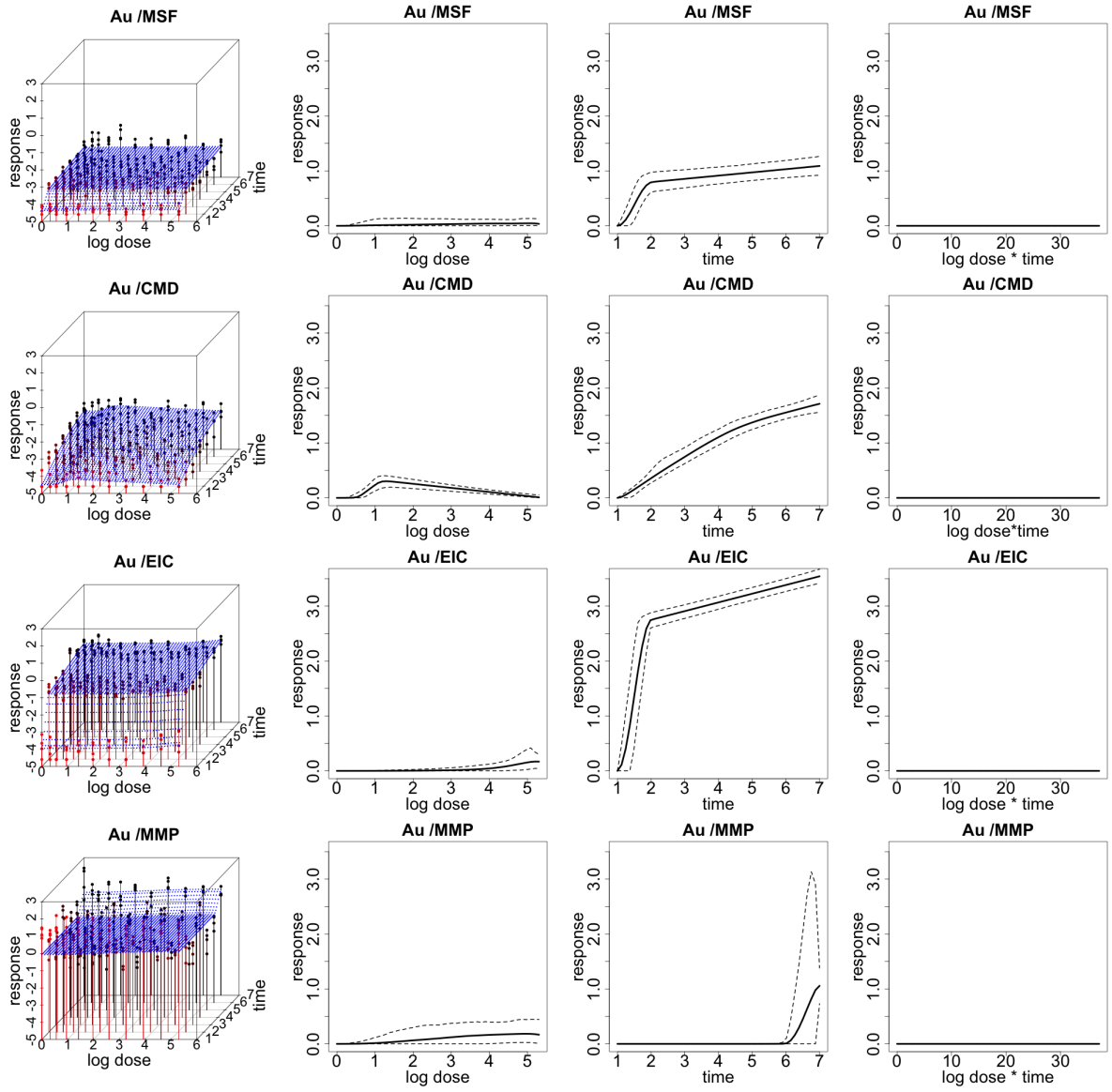

Figure 2.5: **Fitted response curves for the gold (Au) ENM.** Fitted response surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

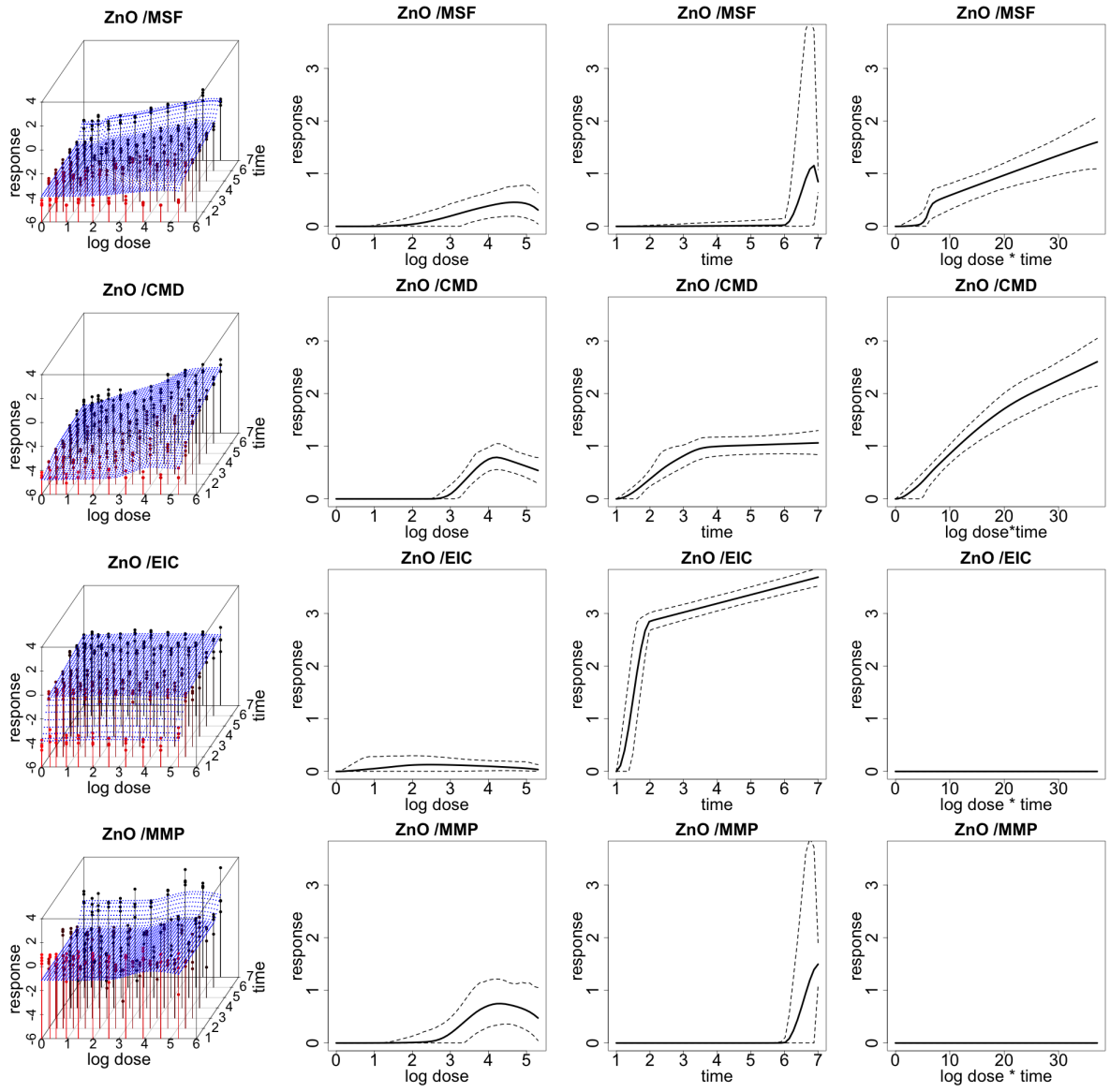

Figure 2.6: **Fitted response curves for the quantum dot (ZnO) ENM.** Fitted response surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

on the degrees of freedom parameter ν, for the t-distributed error described in Section 2.3.1. We specified the prior to be uniform on 1, 2, 4, 8, 16, and 32 degrees of freedom (Besag and Higdon 1999). In concordance with our synthetic data experiments, a sensitivity analysis on the case study data-set proved robust to reasonable variations in the prior specification.

We provide graphical summaries of goodness of fit and posterior predictive performance in Figure 2.7. The top panel shows the mean and 95% posterior intervals of the posterior predictive mean response across all doses and times of exposure (black), along with the empirical mean response (red), for each particle and outcome. In all cases the empirical mean response is contained within the 95% posterior intervals of the posterior predictive mean distribution, indicating good average posterior coverage across doses and times of exposure. The bottom panel provides a plot of the probability integral transform histogram for the entire model (Gneiting et al. 2007). Visual assessment of the plot indicates that it is close to uniformity, suggesting relatively good posterior predictive calibration. Additional summaries and diagnostic tools are detailed in Appendix A.5.

Figures 2.3 to 2.6 illustrate data and results associated with four of the particles examined in this HTS study. Particularly, we report inference for platinum, quantum dot, zinc oxide, and gold nanomaterials for each of the four cytotoxicity outcomes. Inferences for the remaining four particles are reported in Appendix A.4. Specifically, *column 1* shows expected posterior response surfaces across dose and time for all outcomes. As the posterior expectation marginalizes over the interior knots, smooth surfaces reflect the uncertainty about the location of the change-points, and provide an illustration of how the proposed technique will adjust for smoothness in an unsupervised fashion. Also included are functional posterior expectations associated with dose-response curves $f_{ij}(d)$ *(column 2)*, representing the effect due to dose; duration-response curves $g_{ij}(t)$ *(column 3)*, representing the effect due to exposure time; and the expected dose-time interaction function $h_{ij}(t)$ *(column 4)*.

Figure 2.8 provides a plot of the estimated median response, relative to the background, for different doses and times of exposure. Blue colors indicate safety regions or areas of reduced risk to the cells, while red regions indicate increased risk of cytotoxicity. Finally, Figure 2.9 provides posterior summary estimates, including mean and 95% posterior inter-

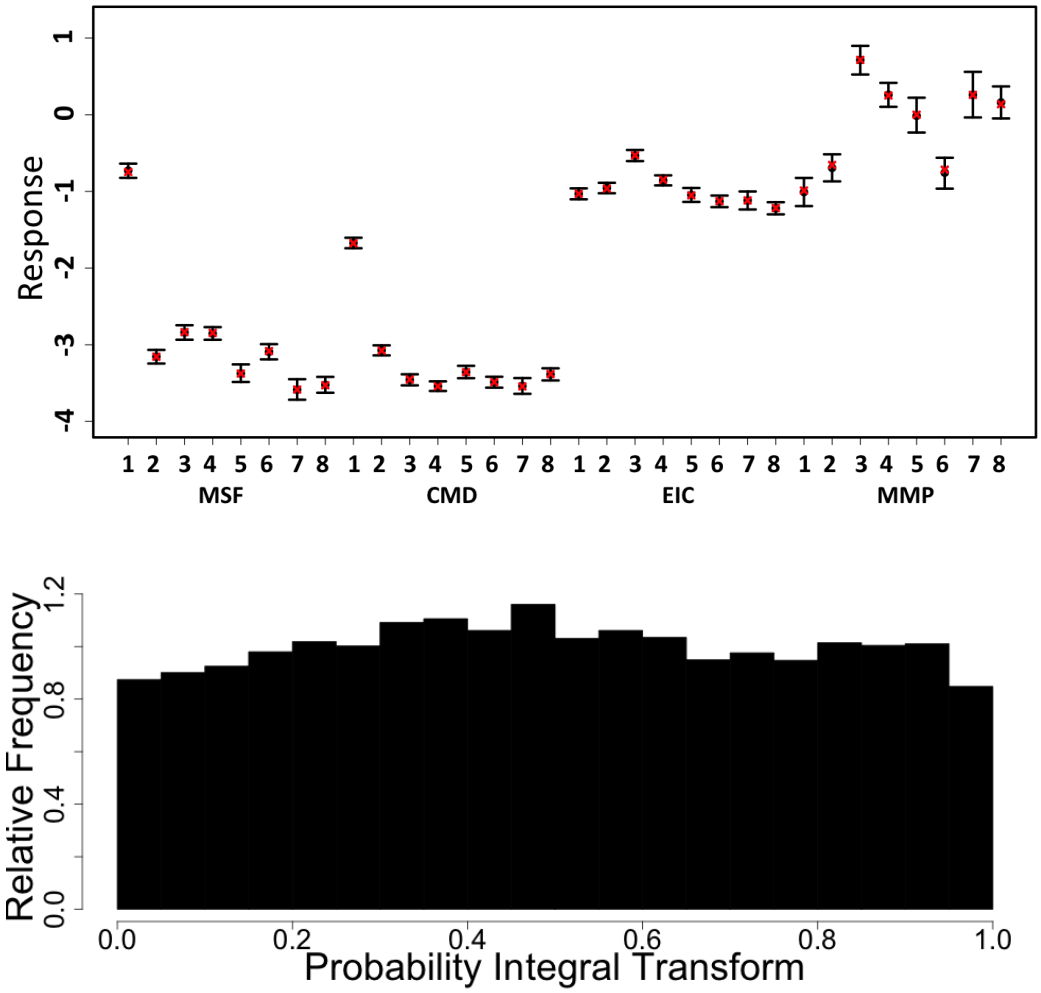

Figure 2.7: **Graphical model diagnostics.** (Bottom) Probability Integral Transform assessing empirical calibration of the posterior predictive distribution. (Top) Mean and 95% posterior intervals of the posterior predictive mean response across all doses and times of exposure, for all outcomes and particles 1 through 8 (QD, ZnO, Fe_3O_4, Pt, Ag, SiO_2, Al_2O_3, Au). Also included are the empirical mean responses across all doses and times of exposure (red).

vals, for the maximal safe dose conditional on the duration of exposure. Note that in the absence of a dose-time interaction, the maximal safe dose is the same across all exposure times.

Quantum dot (QD) shows a relatively high toxic response for plasma membrane damage and mitochondrial superoxide formation. In particular, we see a more pronounced dose effect for membrane damage and both a time, dose and significant dose-time interaction ($\hat{\rho} = .99$) effect for mitochondrial superoxide formation. This supports previous findings with conventional assays that QD nanoparticles stabilized by toluene are capable of inducing tier 2 and 3 oxidative stress responses induced by the toluene (George et al. 2011). Platinum (Pt) shows a pronounced dose, time, and dose-time interaction effect ($\hat{\rho} = .99$) for mitochondrial superoxide formation. Platinum also shows a notable time effect for elevated calcium but not for mitochondrial depolarization or membrane damage, indicating that the particle induced sub-lethal effects to the cell without cytotoxicity. The Zinc oxide nanoparticle (ZnO) shows a relatively high toxic response for plasma membrane damage, elevated calcium and mitochondrial depolarization. In particular, we see a more pronounced time effect for the elevated calcium and both a time and dose effect for membrane damage and mitochondrial depolarization. This again verifies what has previously been demonstrated in conventional assays, in which ZnO nanoparticles have been found to be capable of inducing tier 2 and 3 oxidative stress responses through Zn_2^+ release (George et al. 2009). In contrast, the gold nanoparticle (Al) shows very little response for all outcomes, indicating that, compared to the other particles, it has small risk of inducing a sublethal or lethal cytotoxic response.

2.6 Discussion

In this chapter, we propose a statistical framework for modeling dependent dose-response surfaces over multivariate outcomes. The proposed methodology accounts for dose and duration kinetics jointly using a flexible model that does not compromise interpretability. We account for the multivariate nature of the data using the hierarchical framework and

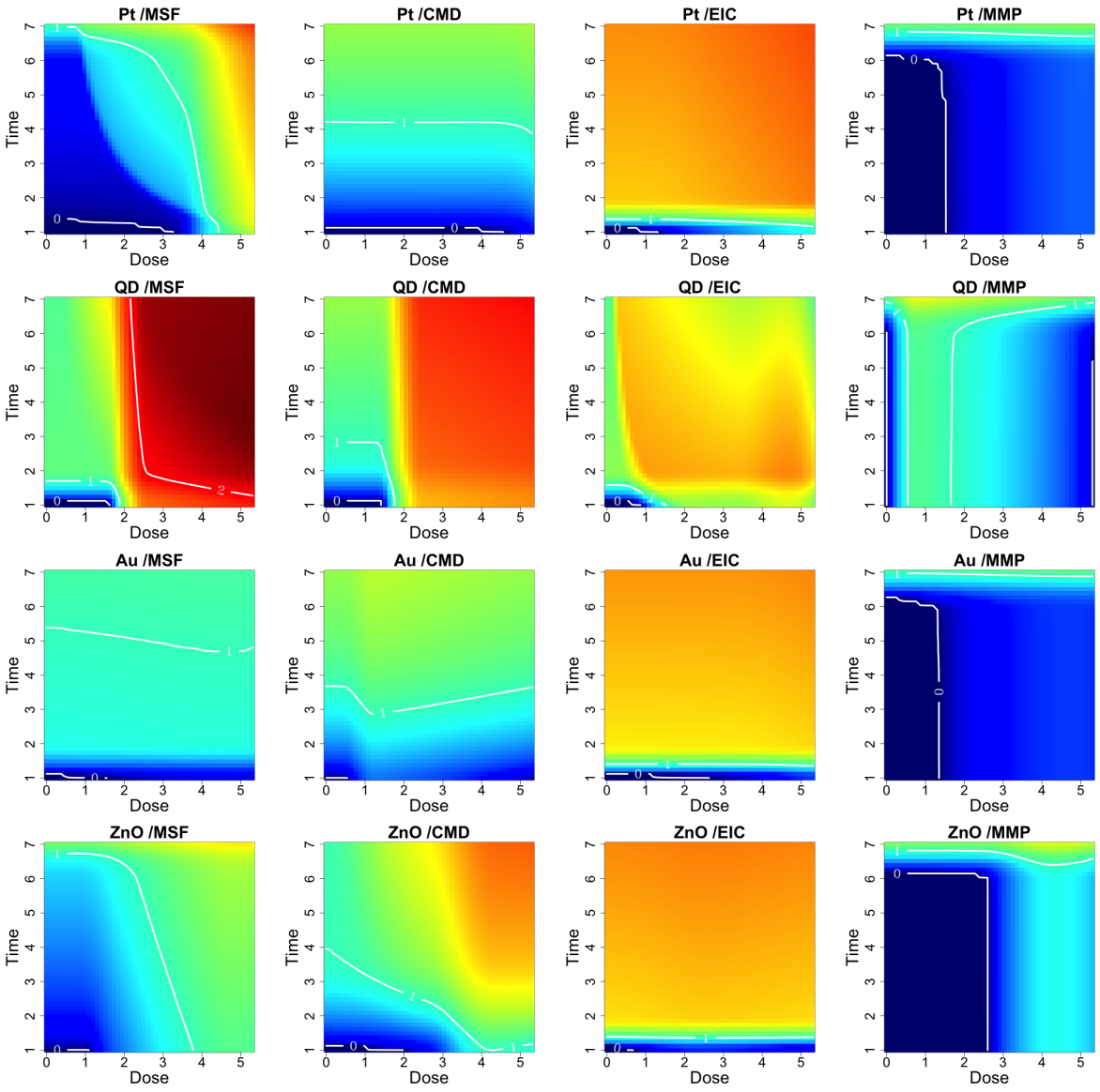

Figure 2.8: **Safe exposure regions for the quantum dot (QD), platinum (Pt), gold (Au), and zinc oxide (ZnO) nanomaterials.** For each particle and outcome we can define dose and time exposure regions which do not induce cytotoxicity. Red colored regions indicate greater cytotoxicity to the cells, whereas blue colored regions indicate reduced risk. Contour lines quantitate the median estimated response, relative to the background, where zero response areas can be interpreted as safe exposure regions.

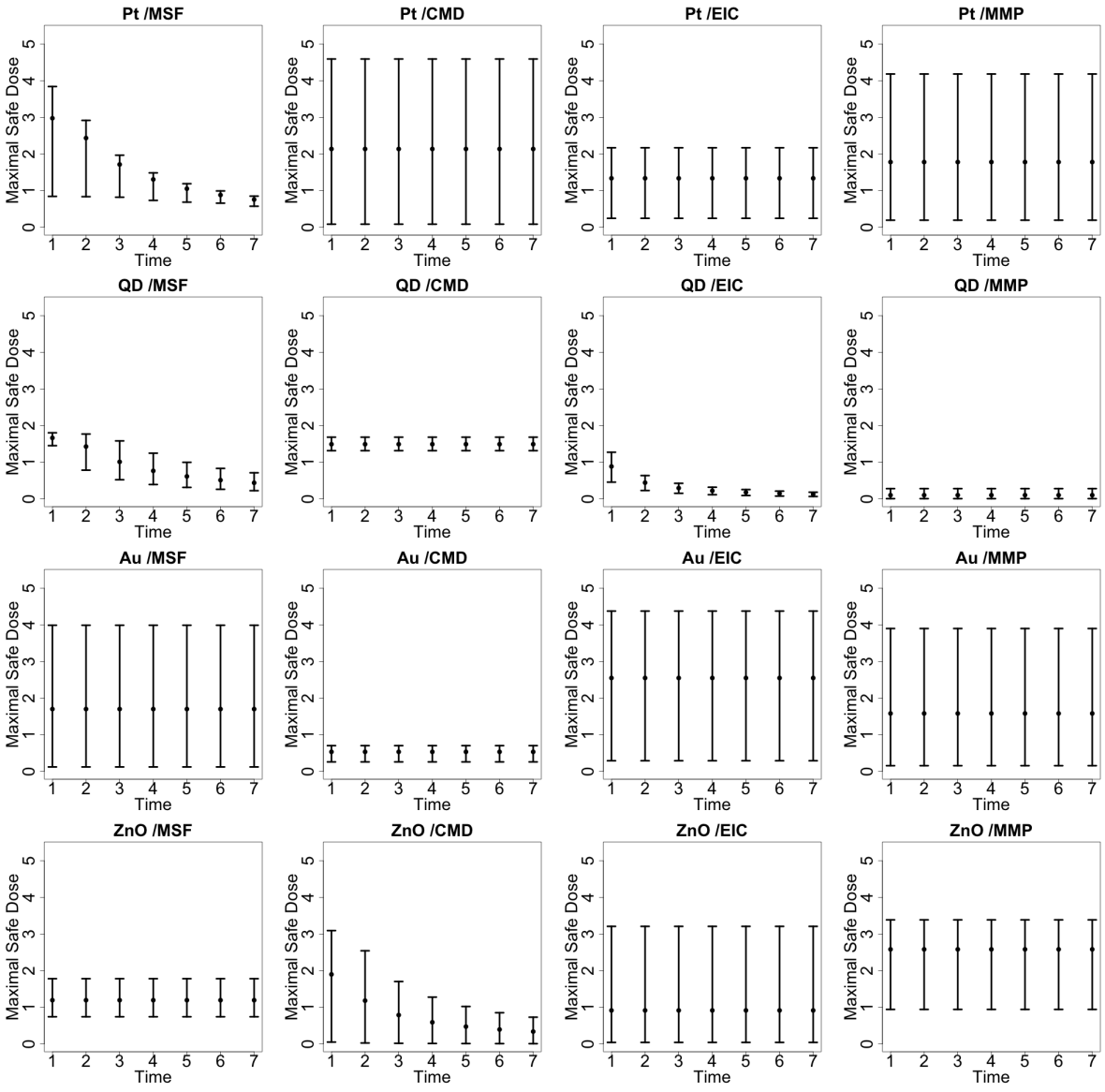

Figure 2.9: **Maximal Safe Dose for the quantum dot (QD), platinum (Pt), gold (Au), and Zinc Oxide (ZnO) nanomaterials.** Posterior summary estimates of the maximal safe dose, conditional on exposure time, including the posterior mean and associated 95% posterior intervals. In the case of no interaction, the maximal safe dose is the same across all times.

thereby efficiently combine information and borrow strength across cellular injury patterns. We account for the non-robust nature of the data by allowing for particle specific variance inflation, resulting in a t-distributed model for the error structure.

The main challenge associated with the class of models proposed in this chapter is finding the right balance between model complexity and model interpretability. An alternative formulation of the dose-response surface would seek inference for a general smooth surface $m_{ij}(d,t)$. However, our simplified approach, based on the assumptions of additivity and linearity, maintains a very appealing level of interpretability, allowing for the definition of specific risk assessment parameters while maintaining an adequate level of flexibility. A related generalization of the proposed additive framework would include a more general class of functional interactions to account for a possible synergistic effect between dose and duration of exposure. This would come at the cost of reduced interpretability, but, at the same time, could be of clear scientific interest in some contexts. In this initial modeling effort, we chose to work with a t-distributed error structure and therefore normalized our response to unconstrain the support so that it could take on values between -∞ and ∞. An alternative formulation could retain the original scale of the data, but rather define a generalized multivariate model such that the outcome distribution can be described using binomial or beta random quantities. This extension would require a substantial increase in computational complexity, with the possible need to consider numerical or analytical approximations, but it is clearly worthy of further methodological exploration.

The hierarchical formulation introduced in this chapter is easily adapted to the case where multiple cell lines are used to test for cytotoxicity. A natural integration strategy would perhaps find motivation in the meta analytic framework, with information shared between experiments via the inclusion of one extra level in the hierarchy.

Finally, the proposed model can also be expanded by the inclusion of covariates. This is naturally defined as an extension to stage 3 of the model introduced in Section 2.3. The addition of covariates is especially important for relating specific ENM properties to toxicity, and is therefore an important area for future work.

CHAPTER 3

Hierarchical Rank Aggregation

3.1 Overview

The development of high-throughput screening (HTS) assays in the field of nanotoxicology provide new opportunities for the hazard assessment and ranking of engineered nanomaterials (ENMs). It is often necessary to rank lists of materials based on multiple risk assessment parameters, often aggregated across several measures of toxicity and possibly spanning an array of experimental platforms. Bayesian models coupled with the optimization of loss functions have been shown to provide an effective framework for conducting inference on ranks. In this article we present various loss function based ranking approaches and apply them to the ranking of particles for combinations of experiments and toxicity outcomes. Additionally, we propose a framework for the aggregation of ranks across different sources of evidence, which also allows for the differential weighting of this evidence based on its reliability and importance in risk ranking. We apply these methods to high-throughput toxicity data on two human cell lines exposed to eight different nanomaterials and measured in relation to four cytotoxicity outcomes.

3.2 Introduction

This chapter considers hazard ranking of engineered nanomaterials (ENMs) from high-throughput screening (HTS) studies. Decision support, specifically involving hazard ranking, can be used to develop a framework for the prioritization of extensive *in-vivo* testing of emerging nanomaterials. Given ethical and economic considerations associated with an-

imal experiments, initial prioritization schemes must indeed rely on high content *in-vitro* screening of a large number of particles (Lilienblum et al. 2008).

The statistical challenge associated with hazard ranking from HTS data lies in its richness and heterogeneity, as multi-dimensional measurements are often taken over a small number of replicates with relatively low signal. Inferential goals include toxicity ranking of ENMs, as well as associated measures of uncertainty, both within and aggregated across the many sources of evidence. A heat-map visualization of a sample HTS data set is provided in Figure 3.1, where the toxic response of two human cell lines, exposed to eight different nanoparticles, is measured in relation to four cellular response outcomes monitored across an array of doses and durations of exposure.

We propose an approach to ranking particles aimed at achieving three goals: to use a hierarchical dose-response model to rank particles within outcomes and experiments; to derive an aggregate or consensus ranking that summarizes information across outcomes and experiments; and finally to account for the varying levels of reliability and importance of the outcomes and experiments. An aggregate ranking across different outcomes and experiments can aid in decision-making for future testing. Although the rankings within outcomes or experiments are expected to be positively correlated this does not guarantee that the ranked lists of ENMs will be in complete agreement. As an example, Figure 3.2 shows dose-response-surfaces fit to HTS toxicity data for quantum dots, nano-zinc oxide and nano-platinum. Nano-platinum shows higher responses for mitochondrial superoxide formation and almost no response for membrane damage. Conversely, zinc oxide has a very strong response for membrane damage. Furthermore, different sources of information come with varying levels of reliability and importance in terms of assessing overall toxicity. For example, while membrane damage is viewed as a lethal response to the cell, mitochondrial superoxide formation is only a sublethal indicator of cellular oxidative stress. Therefore, the information derived from these outcomes might need to be weighted differently when ranking a material's hazard potential.

We illustrate the proposed framework for analyzing and ranking data on two cell lines

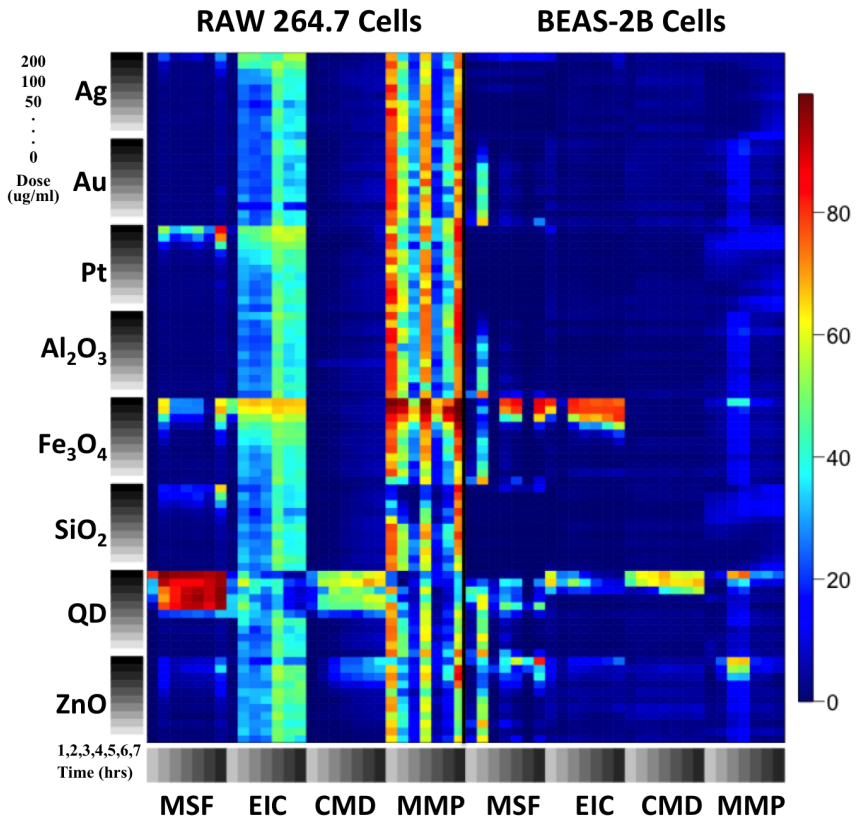

Figure 3.1: **Heatmap of HTS data assessing the cytotoxicity of eight different ENMs.** The rows and columns correspond to the doses and times of exposure, respectively, for each cell type and cytotoxicity parameter measured. The four cytotoxicity parameters measured include mitochondrial superoxide formation (MSF), loss of mitochondrial membrane potential (MMP), elevated intracellular calcium (EIC), and cellular membrane damage (CMD). Blue colors indicate the least harmful activity, while yellow indicates a high cellular response.

exposed to eight metal and metal-oxide nanoparticles, monitored for the presence of four cytotoxicity outcomes, and measured across an array of doses and times of exposure.

The multi-level structure of the data coupled with the non-standard inferential goals can be naturally accounted for under the Bayesian hierarchical framework. Shen and Louis (1998) and Lockwood et al. (2002) describe rank estimation under a Bayesian setting and show that rather than ranking posterior means, estimation based on minimizing a loss function specific to ranks is more appropriate. Lin et al. (2006) explore ranking based on optimizing various loss functions using a two-stage hierarchical model. More specifically, they review

ranking based on squared-error loss and extend this to other loss functions that are tuned to application specific goals. For example, when the goal is to identify relatively high or low rankers, a loss function that penalizes classification errors produces estimates which minimize error. Lin et al. (2009) apply various loss functions to ranking health service providers based on standardized mortality ratios estimated from a two-stage hierarchical model. Noma et al. (2010) extend some of these ideas to the analysis of microarray data and develop three empirical Bayes methods for ranking genes based on a hierarchical mixture model for differential expression.

The problems of comparing lists and rank aggregation have also been considered in the computer science and bioinformatics literature in relation to meta-search (Dwork et al. 2001 and Fagin et al. 2003). Fagin et al. (2003) defined a set of distance measures that could be used to quantify dissimilarities between lists. In the context of rank aggregation, they are interested in finding the collection of measures that has the minimum total distance with respect to the given lists. Dwork et al. (2001) considered the problem of aggregating across lists using a Markov process approach. First, pairwise majority preferences are summarized across lists and then the matrix of pairwise preferences are used to produce a Markov Chain (MC) transition matrix. The aggregate ranking can then be defined according to the stationary distribution of this MC. DeSemet et al. (2002) used a similar approach to model the aggregate behavior of a large number of decision makers. These techniques, however, are designed to work with summaries of previously analyzed data sets, and to our knowledge are not directly applicable to a comprehensive data-fusion exercise.

In this chapter we present various loss-function-based ranking methods, including those suggested by Shen and Louis (1998) and Lin et al. (2006), and apply them to the ranking of HTS data within outcomes and experiments. Additionally, we build on this approach, describing loss-functions that can be used to provide an overall ranking of particles, and thereby propose a framework for the prioritization of further testing of high risk materials.

The remainder of this chapter is organized as follows. In Section 3.3 we introduce a general statistical model for toxicity profiling and estimation of various risk assessment parameters.

In Section 3.4 we present various loss function based methods for ranking and discuss the applicability of these methods to nanoparticle toxicology. In Section 3.5 we apply these methods to the analysis and ranking of eight metal-oxide nanomaterials. Finally, we conclude with a critical discussion of the limitations and possible extensions of these methods in Section 3.6.

3.3 Statistical Models of Toxicity

In this section we describe a basic framework for a dose-response model for a general HTS study, in which we monitor a multivariate continuous outcome y, corresponding to J cytotoxicity parameters, in association with the exposure of a number of cells to I different ENMs. We are also interested in combining results across K different experiments, often consisting of multi-outcome HTS assays conducted over K cell lines. More precisely, let $y_{ijk\ell}(d,t)$ denote the response corresponding to ENM i ($i = 1, ..., I$), cytotoxicity parameter j ($j = 1, ..., J$), experiment k ($k = 1, ..., K$) and replicate ℓ ($\ell = 1, ..., L$), for a cell population exposed to dose $d \in [0, D]$ for a duration $t \in [0, T]$. In practice, observations are obtained over a discrete set of doses and durations of exposure. However, for simplicity of notation and without loss of generality, we will assume that doses and times are defined over continuous intervals. Hazard ranking will focus on the expected response $E\{y_{ijk\ell}(d,t)\} = m_{ijk}(d,t)$, where $m_{ijk}(\cdot)$ defines the exposure-duration dynamic for the process under study and may be specified in a parametric or non-parametric fashion. The proposed hazard ranking framework will not depend on specific representations of this quantity. We therefore maintain the discussion on fairly general grounds, leaving specific modeling considerations to the case study in Section 3.5. A detailed discussion of these issues is also reported in Chapter 2.

For notational convenience we will simply identify the full exposure-duration surface with $m_{ijk} = \{m_{ijk}(d,t) : \ d \in [0,T], \ t \in [0,T]\}$. We assume that the joint distribution of an integrated HTS experiment can be represented by the following multi-stage hierarchical

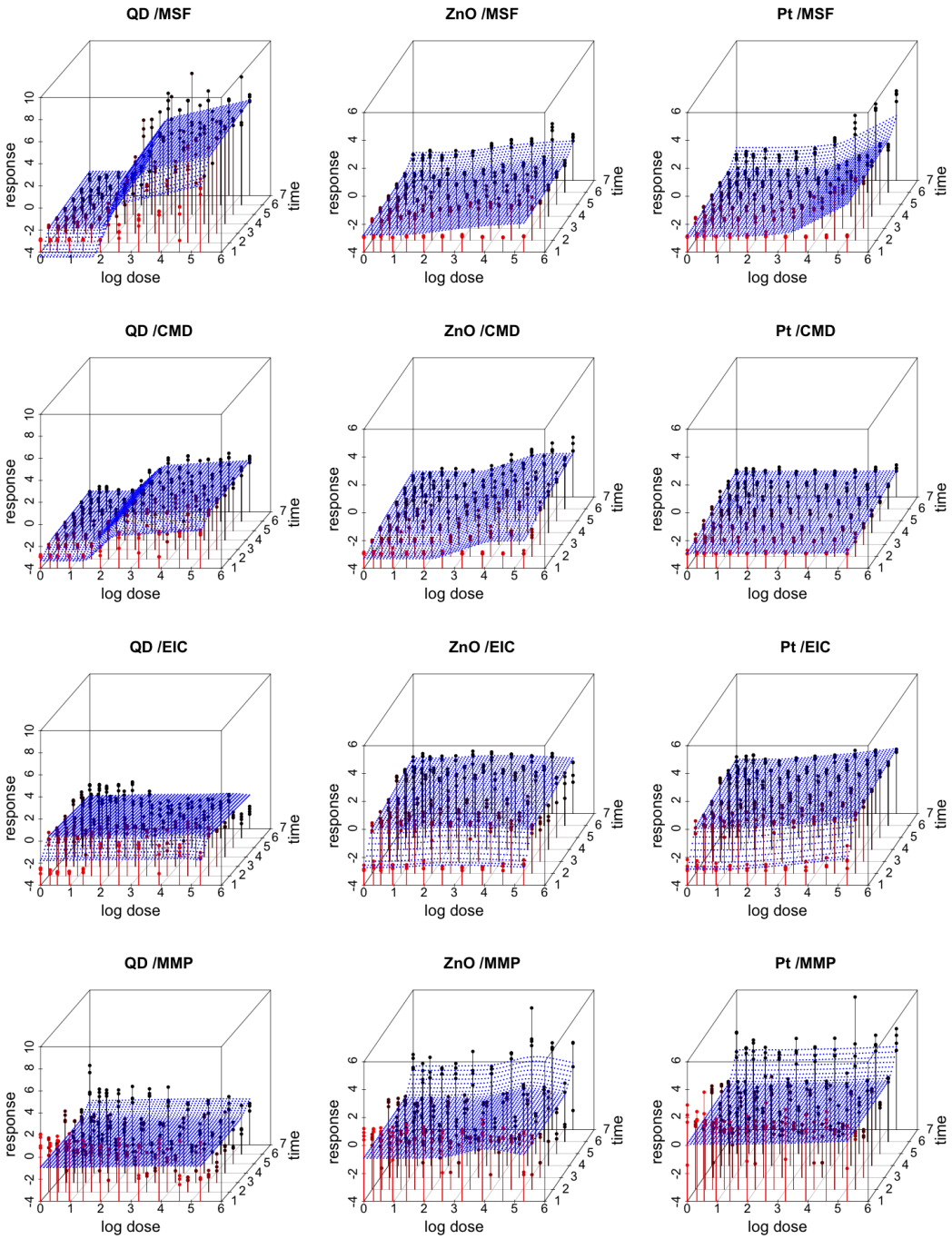

Figure 3.2: **Fitted response-surfaces for the quantum dot (QD), zinc oxide (ZnO) and platinum (Pt) nanomaterials exposed to the macrophage (RAW** 264.7**) cell line.** Each of the four responses measured include: mitochondrial superoxide formation (MSF), loss of mitochondrial membrane potential (MMP), elevated intracellular calcium (EIC), and cellular membrane damage (CMD).

model

$$\begin{aligned} y_{ijk\ell}(d,t) \mid m_{ijk}, \sigma^2_{ijk} &\sim p(y_{ijk\ell}(d,t) \mid m_{ijk}, \sigma^2_{ijk}) \\ m_{ijk} \mid m_{ij}, \sigma^2_{ij} &\sim p(m_{ijk} \mid m_{ij}, \sigma^2_{ij}) \\ m_{ij} \mid m_i, \sigma^2_i &\sim p(m_{ij} \mid m_i, \sigma^2_i) \\ m_i &\sim p(m_i), \end{aligned} \tag{3.1}$$

where m_{ij} and m_i denote exposure-duration surfaces at the integrated particle-by-outcome (ij) and integrated particle (i) levels. The model is completed with prior distributions on variance components and possible nuisance parameters.

Let, $\boldsymbol{m} = \{m_{ijk} : \ i = 1, \ldots, I; \ j = 1, \ldots, J; \ k = 1, \ldots, K\}$ denote the set of all smooth surfaces and let $\boldsymbol{\sigma^2} = \{\sigma^2_{ijk} : \ i = 1, \ldots, I; \ j = 1, \ldots, J; \ k = 1, \ldots, K\}$ be the set of all variance parameters. Furthermore, if we denote by $\boldsymbol{Y}$ the complete set of response values for all particles, cytotoxicity parameters, and experiments, all evidence available over exposure-duration dynamics and associated variability at the different levels of the hierarchy is contained in the posterior distribution

$$p(\boldsymbol{m}, \boldsymbol{\sigma} \mid \boldsymbol{Y}) \propto p(\boldsymbol{Y} \mid \boldsymbol{m}, \boldsymbol{\sigma^2}) p(\boldsymbol{m}, \boldsymbol{\sigma^2}). \tag{3.2}$$

This quantity, or most often Monte Carlo samples from this joint distribution, can be used to calculate the joint posterior distribution of various functionals $g(m_{ijk})$. These functions provide easily interpretable summaries of full exposure-duration dynamics, which may be used as overall measures of hazard. Some examples of risk assessment summaries frequently used in toxicology include exposure levels or time thresholds below which no significant effect exists, doses or times at which adverse effects rise above some predetermined amount as compared to the background (benchmark doses), and summaries of the slope of the dose-response trajectory including the dose or time that produces an $\alpha\%$ inhibitory response (ECα) (Edler et al. 2002). These measures summarize different aspects of a response-surface, and in most cases are computed separately for dose and time kinetics. Although no sufficient univariate measure exists that can synthesize every aspect of the response-surface, in this paper we will focus on the area under the response-surface as a possible summary measure

(Section 3.5). However, we maintain that the proposed method is directly applicable to any functional of the response-surface set $\boldsymbol{m}$.

3.4 Decision Theoretic Approaches to Hazard Ranking

3.4.1 Estimating Ranks

Most common ranking methods are based on ordering estimates of target parameters, such as MLEs and posterior means, or on the ordering of statistics testing some null hypothesis of interest. As described, among others, by Louis and Shen (1999) and Lin et al. (2006), these methods perform poorly when the posterior distributions of the target parameters are not stochastically ordered and are not invariant under monotone transformation of the target parameters. A more appropriate method of ranking is based on calculating the joint posterior distribution of the ranks, followed by inference guided by loss functions appropriate for analytic goals.

The rank of some target parameter $g(m_{ijk})$, for each particle i, outcome j, and experiment k can be defined as the sum represented by $\sum_{h=1}^{I} \mathbf{1}_{\{g(m_{ijk}) \geq g(m_{hjk})\}}$. A direct extension of this definition to the aggregate ranking of particles, for example, across K experiments, might be defined by the ranks of population level parameters $g(m_{ij})$. A caveat of this procedure is that, due to technical and/or biological sources of variability, data on heterogeneous measures of toxicity may support high posterior variances for population level parameters, which are not necessarily reflected in small rank correlations amongst lower level quantities. This procedure may in fact result in the artificial inflation of reported aggregate rank variability.

In order to avoid this paradox, we define aggregate ranks across experiments but within outcomes, and across outcomes and experiments, as a weighted average of individual ranks within both outcomes and experiments. More formally, the rank of our target parameter,

$g(m_{ijk})$, at each level of hierarchy, can be described as follows:

$$\begin{aligned} R_{ijk} &= rank(g(m_{ijk})) = \sum_{h=1}^{I} \mathbf{1}_{\{g(m_{ijk}) \geq g(m_{hjk})\}} \\ R_{ij} &= \sum_{k=1}^{K} w_{e_k} R_{ijk} \qquad \sum_{k=1}^{K} w_{e_k} = 1 \\ R_{i} &= \sum_{j=1}^{J} w_{e_k} \sum_{k=1}^{K} w_{o_j} R_{ijk} \qquad \sum_{j=1}^{J} w_{o_j} = 1 \end{aligned} \tag{3.3}$$

where, a rank of 1 corresponds to the highest rank, or most hazardous particle. Here, the R_{ijk} allow us to determine the rank of each particle within each cytotoxicity outcome and experimental group. R_{ij} is the rank of each particle within each cytotoxicity parameter, but aggregated across all experiments. R_i is an overall summary of the rank of each particle across all cytotoxicity parameters and experimental groups. Finally, w_{o_j} and w_{e_k} are weight functions that allow us to assign a measure of importance to each cytotoxicity parameter j and experiment k.

Provided that draws from $p(m_{ijk} \mid \boldsymbol{Y})$ are available via Markov Chain Monte Carlo $(MCMC)$ or otherwise, all knowledge about R_{ijk} is easily summarized in the posterior distribution $p(R_{ijk} \mid \boldsymbol{Y})$, and similarly for R_{ij} and R_i. In the following sections we describe in detail optimal point estimators of ranks at the different levels of the hierarchy from a decision theoretic perspective, as well as potential substantive strategies to select weights w_{o_j} and w_{e_k}.

3.4.2 Rank Estimates Based on Squared-Error Loss Functions

Following Shen and Louis (1998) and Lockwood et al. (2002), a natural choice for loss functions may be based on the squared-error associated with posterior ranks. Let R_{ijk}^{est} be a point estimator for the posterior rank of particle i, outcome j, experiment k, and similarly for R_{ij}^{est} and R_{i}^{est}. For each particle at each level of hierarchy squared-error loss functions

can be described as follows:

$$\begin{aligned} L_{jk}^{(SEL)} &= L_{jk}^{(SEL)}\left(R_{ijk}^{est}, R_{ijk}\right) &&= \tfrac{1}{I}\sum_i \left(R_{ijk}^{est} - R_{ijk}\right)^2, \\ L_{j}^{(SEL)} &= L_{j}^{(SEL)}\left(R_{ij}^{est}, R_{ij}\right) &&= \tfrac{1}{I}\sum_i \left(R_{ij}^{est} - \sum_{k=1}^{K} w_{ek} R_{ijk}\right)^2, \\ L^{(SEL)} &= L^{(SEL)}\left(R_{i}^{est}, R_{i}\right) &&= \tfrac{1}{I}\sum_i \left(R_{i}^{est} - \sum_{j=1}^{J} w_{oj} \sum_{k=1}^{K} w_{oj} R_{ijk}\right)^2. \end{aligned} \tag{3.4}$$

These loss functions are minimized by the following posterior means,

$$\begin{aligned} \bar{R}_{ijk}(\boldsymbol{Y}) &= E_{g(m_{ijk})|\boldsymbol{Y}}[R_{ijk} \mid \boldsymbol{Y}] &&= \sum_{h=1}^{I} P(g(m_{ijk}) \geq g(m_{hjk}) \mid \boldsymbol{Y}), \\ \bar{R}_{ij}(\boldsymbol{Y}) &= E_{g(m_{ijk})|\boldsymbol{Y}}[R_{ij} \mid \boldsymbol{Y}] &&= \sum_{k=1}^{K} w_{ek} \sum_{h=1}^{I} P(g(m_{ijk}) \geq g(m_{hjk}) \mid \boldsymbol{Y}), \\ \bar{R}_{i}(\boldsymbol{Y}) &= E_{g(m_{ijk})|\boldsymbol{Y}}[R_{i} \mid \boldsymbol{Y}] &&= \sum_{j=1}^{J} w_{oj} \sum_{k=1}^{K} w_{ek} \sum_{h=1}^{I} P(g(m_{ijk}) \geq g(m_{hjk}) \mid \boldsymbol{Y}). \end{aligned} \tag{3.5}$$

The $\bar{R}_i$ are generally not integer values. For optimal integer ranks we use $\hat{R}_i = rank(\bar{R}_i)$, and similarly, for $\bar{R}_{ij}$ and $\bar{R}_{ijk}$. In some cases it is more convenient to define percentiles instead of ranks (Lockwood et al. 2002). Therefore, we let $Q_{ijk} = R_{ijk}/(I+1)$ denote the percentile rank for particle i, outcome j, and experiment k. Again, we can write similar expressions for Q_i and Q_{ij}.

3.4.3 Rank Estimates Based on Upper $100(1-\gamma)\%$ or Lower $100(\gamma)\%$ Loss Functions

In many situations, interest may focus on identifying some fraction of particles with the highest (or lowest) likelihood of conferring adverse effects. Lin et al. (2006) proposed a loss function which addresses this goal by penalizing misclassification by an amount that depends on the distance of the estimated percentile from some cut-point γ. Let γ $(0 < \gamma < 1)$ denote the most toxic fraction of total ENMs that we would like to identify. Also let Q_{ijk}^{est} be a point estimator for the posterior percentile rank of particle i, outcome j, and experiment k, and similarly for Q_{ij}^{est} and Q_{i}^{est}. In order to rank ENMs at each level of hierarchy, loss functions

based on the upper $100(1-\gamma)\%$ classification error can be described as follows:

$$\begin{aligned} L_{jk}^{(PLF)} &= \tfrac{1}{I}\sum_i \left(\gamma - Q_{ijk}^{est}\right)^2 \left(\mathbf{1}_{\{Q_{ijk}>\gamma, Q_{ijk}^{est}<\gamma\}} + I_{\{Q_{ijk}<\gamma, Q_{ijk}^{est}>\gamma\}}\right), \\ L_{j}^{(PLF)} &= \tfrac{1}{I}\sum_i \left(\gamma - Q_{ij}^{est}\right)^2 \left(\mathbf{1}_{\{Q_{ij}>\gamma, Q_{ij}^{est}<\gamma\}} + I_{\{Q_{ij}<\gamma, Q_{ij}^{est}>\gamma\}}\right), \\ L^{(PLF)} &= \tfrac{1}{I}\sum_i \left(\gamma - Q_{i}^{est}\right)^2 \left(\mathbf{1}_{\{Q_{i}>\gamma, Q_{i}^{est}<\gamma\}} + I_{\{Q_{i}<\gamma, Q_{i}^{est}>\gamma\}}\right). \end{aligned} \tag{3.6}$$

For notational convenience we will assume γI is an integer and let

$$\begin{aligned} \pi_{ijk}(\gamma) &= P(R_{ijk} > \gamma(I+1) \mid \boldsymbol{Y}) &&= \sum_n^I P(R_{ijk} = n \mid \boldsymbol{Y}), \\ \pi_{ij}(\gamma) &= P(\sum_{k=1}^K w_{o_j} R_{ijk} > \gamma(I+1) \mid \boldsymbol{Y}) &&= \sum_n^I P(\sum_{k=1}^K w_{o_j} R_{ijk} = n \mid \boldsymbol{Y}), \\ \pi_{i}(\gamma) &= P(\sum_{j=1}^J w_{e_k} \sum_{k=1}^K w_{o_j} R_{ijk} > \gamma(I+1) \mid \boldsymbol{Y}) &&= \sum_n^I P(\sum_{j=1}^J w_{e_k} \sum_{k=1}^K w_{o_j} R_{ijk} = n \mid \boldsymbol{Y}). \end{aligned} \tag{3.7}$$

where $n = \gamma I + 1$. The quantities that minimize the posterior risk induced by the loss function described in (3.6) can be described by: $\tilde{R}_{ijk}(\gamma) = rank(\pi_{ijk}(\gamma))$, for particle i, cytotoxicity parameter j, and experiment k. We can write similar expressions for $\tilde{R}_{ij}(\gamma)$ and $\tilde{R}_{i}(\gamma)$. Here we have optimized the classification errors for ranking the most hazardous particles. Ranking based on particles least likely to be hazardous will follow a similar procedure.

3.4.4 Assigning Weights

Different sources of information come with varying levels of reliability and importance in terms of hazard ranking. Our definitions of aggregate ranks include weights w_{o_j}, averaging across J possible outcomes, and w_{e_k}, averaging across K experiments. These quantities may be used directly in the definition of aggregate ranks in a way that reflects possibly differing measures of importance to be assigned to each outcome and experiment.

The definition of differentiated aggregate hazard measures is especially important in HTS cytotoxicity studies. A typical HTS assay will, in fact, include outcomes measuring only sublethal effects, often used as hypothesis generation and confirmatory tools by biologists, as well as outcomes measuring lethal effects, which are more directly related to cytotoxicity.

Intuitively, we might want to give a higher weight to particles that induce lethal effects relative to particles that only induce sublethal effects without cytotoxicity. For example, in cytotoxicity studies of metal-oxide nanomaterials, PI uptake, a measure of cellular membrane damage, is the outcome most often chosen to carry out risk ranking (George et al. 2011). Similar considerations are valid for the differential weighting of experiments. For example George et al. (2009) report an experiment carried out on two cell lines: bronchial epithelial cells and a macrophage cells. In this case, one cell line may prove more relevant for the prioritization of animal inhalation toxicity experiments as compared to the other. When no particular experiments or outcomes are considered to be more or less relevant as a measure of hazard, uniform weights can be applied.

While differential weighting is, in principle, completely general and can be determined using expert's criteria, we found that it is often easier to obtain expert opinion on the ordering of different outcomes and experiments. If we let $(\pi_1, ...\pi_k...\pi_K)$ be the ordering of the K experiments based on relative importance, then one possibility for assigning weights is to find $w_{e_1}, ...w_{e_{\pi_K}}$, such that

$$w_{e_{\pi_{k-1}}} = \delta w_{e_{\pi_k}}, \qquad \sum_{k=1}^{K} \delta^{k-1} w_{e_{\pi_1}} = 1, \quad \text{for } k = 1, \ldots, K; \;\; \delta \geq 1; \tag{3.8}$$

leading to $w_{ek} = \delta^{k-1} / \sum_{h=1}^{K} \delta^{h-1}$. Aggregate weight determination across J outcomes can be obtained in the same fashion. In this formulation, differential weighting depends solely on a one dimensional, easily interpreted parameter δ. A value of δ equal to 1 assigns equal importance to each experiment, while a value of δ equal to .5 assigns twice the importance to experiment π_{k-1} as experiment π_k. Alternatively, weights w_{oj} and w_{ek} can be given arbitrary values, for example based on expert elicitation, as long as they satisfy the constraints $\sum_{j=1}^{J} w_{oj} = 1$ and $\sum_{k=1}^{K} w_{ek} = 1$.

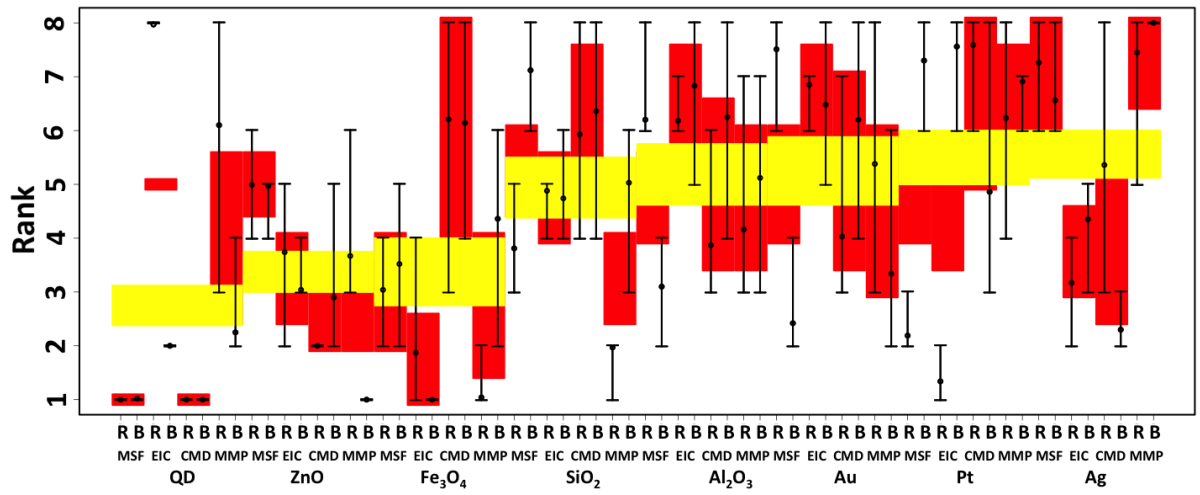

Figure 3.3: **Rankings based on squared-error loss using equal weights for all outcomes.** Posterior expected ranks and 95% posterior intervals computed by minimizing squared-error loss. Each cell-line and outcome is given equal weight. Ranks are displayed individually within outcomes and experiments (black), aggregated across experiments, but within outcomes (red), and aggregated across all outcomes and experiments (yellow).

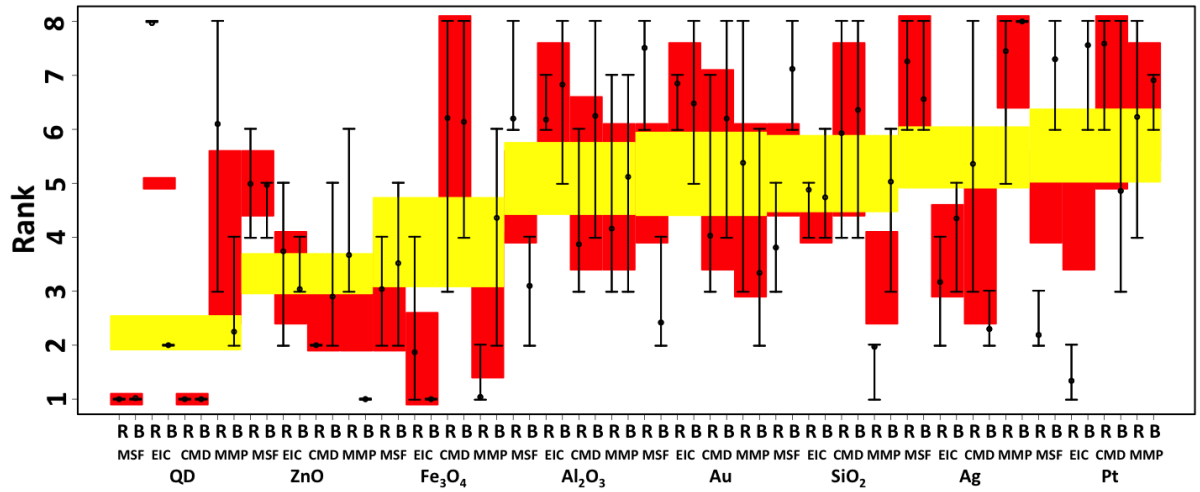

Figure 3.4: **Rankings based on squared-error loss using moderately aggressive weights favoring important outcomes.** Posterior expected ranks and 95% posterior intervals computed by minimizing squared-error loss. Each cell-line is given equal weight, and each outcome is given weights (.37, .27, .21, .15), in the order (CMD, MSF, MMP, EIC). Ranks are displayed individually within outcomes and experiments (black), aggregated across experiments, but within outcomes (red), and aggregated across all outcomes and experiments (yellow).

3.5 A Case Study in Nanotoxicology

3.5.1 Background

We analyze and rank data on cells exposed to eight different metal and metal-oxide nanoparticles and monitored in relation to four cytotoxicity parameters. All four outcomes are measured over a grid of ten doses $(0, .375, .750, 1.6, 3.12, 6.25, 12.5, 25, 50, 100, 200\ \mu g/ml)$, seven times (hours) of exposure, and four replicates at every dose-time combination (see Figure 3.2). Cytotoxicity screening is based on the hierarchical oxidative stress model described in Xia et al. (2006). More specifically, a multi-parametric, epiflourescence assay was used to measure four responses relating to toxic oxidative stress. Three of the responses measured sublethal effects to the cell, including mitochondrial superoxide formation (MSF), loss of mitochondrial membrane potential (MMP), and elevated intracellular calcium (EIC), and the last outcome measured cellular membrane damage (CMD), a lethal effect to the cell. The nanoparticles measured included: silver (Ag), gold (Au), platinum (Pt), iron oxide (Fe_3O_4), aluminum oxide (Al_2O_2), silicon dioxide (SiO_2), zinc oxide (ZnO), and quantum dot (QD). Experiments were conducted in two different cell lines related to inhalation toxicity: bronchial epithelial cell lines (BEAS-2B) and macrophage (RAW 264.7) cell lines.

3.5.2 Analysis and Results

We fit response-surfaces for the metal-oxide data set using the model described in Section 3.1. Specifically, we model the data in a generalized additive fashion and parameterized using linear basis spline functions. Details for a similar reduced experiment were described in Chapter 2. To summarize briefly, the response-surface $m_{ijk}(d,t)$, corresponding to ENM i $(i = 1, ..., I)$, cytotoxicity parameter j $(j = 1, ..., J)$, and replicate k $(k = 1, ..., K)$ at dose $d \in [0, D]$ and time $t \in [0, T]$, was modeled as follows:

$$m_{ijk}(d,t) = \alpha_{ijk} + \mathcal{B}(d, \boldsymbol{\phi}_{ijk})'\boldsymbol{\beta}_{ijk} + \mathcal{B}(t, \boldsymbol{\psi}_{ijk})'\boldsymbol{\gamma}_{ijk}. \tag{3.9}$$

Equal Weights	
ENM	Rank
QD	2.79 (2.38,3.12)
ZnO	3.29 (3,3.75)
Fe_3O_4	3.4 (2.75,4)
SiO_2	4.98 (4.38,5.5)
Al_2O_3	5.21 (4.62,5.75)
Au	5.28 (4.62,5.88)
Pt	5.5 (5,6)
Ag	5.56 (5.12,6)

$wo_j = (.37, .27, .21, .15)$	
ENM	Rank
QD	2.27 (1.92,2.54)
ZnO	3.25 (2.96,3.69)
Fe_3O_4	3.95 (3.09,4.73)
Al_2O_3	5.08 (4.43,5.75)
Au	5.15 (4.41,5.94)
SiO_2	5.2 (4.48,5.88)
Ag	5.47 (4.92,6.04)
Pt	5.63 (5.03,6.37)

$wo_j = (.54, .27, .12, .07)$	
ENM	Rank
QD	1.66 (1.46,1.82)
ZnO	3.18 (2.88,3.72)
Fe_3O_4	4.65 (3.51,5.67)
Al_2O_3	5 (4.2,5.89)
Au	5.09 (4.15,6.13)
Ag	5.12 (4.41,5.87)
SiO_2	5.55 (4.58,6.44)
Pt	5.74 (4.94,6.77)

Table 3.1: **Rankings based on squared-error loss using aggressive weights favoring important outcomes.** Aggregated ranks across each outcome and cell-line. Posterior expected ranks and 95% posterior intervals computed by minimizing squared-error loss. Each cell-line is given equal weight and each outcome (CMD,MSF,MMP,EIC) is given varying weights w_{oj}.

where, $\mathcal{B}(d, \boldsymbol{\phi}_{ijk})$ and $\mathcal{B}(d, \boldsymbol{\psi}_{ijk})$ denote two 4-dimensional B-spline basis with interior knots $\boldsymbol{\phi}_{ijk} = (\phi_{ijk1}, \phi_{ijk2})'$ and $\boldsymbol{\psi}_{ijk} = (\psi_{ijk1}, \psi_{ijk2})'$. Also, $\boldsymbol{\beta}_{ij} = (\beta_{ij1}, .., \beta_{ij4})'$ and $\boldsymbol{\gamma}_{ij} = (\gamma_{ij1}, ..., \gamma_{ij4})'$ are two 4-dimensional vectors of spline coefficients. Identifiability restrictions are implemented by fixing $\beta_{ijk1} = 0$ and $\gamma_{ijk1} = 0$, allowing us to interpret α_{ij} as the background response level for each particle, outcome, and experiment. We also fix $\beta_{ij2} = 0$ and $\gamma_{ij2} = 0$, assuming no effect before ϕ_{ij1} and ψ_{ij1}. See Figure 3.2 for an example of surfaces fit to the data on RAW 264.7 cells exposed to the zinc oxide (ZnO), quantum dot (QD), and platinum (Pt) nanomaterials, and measured across the four cytotoxicity outcomes. Our inferences are based on 20,000 MCMC samples from the posterior distribution in (3.2), after discarding a conservative 60,000 iterations for burn-in. MCMC sampling was performed in R version 2.10.0, and convergence diagnostics were performed using the package CODA (Convergence Diagnostics and Output Analysis), (Plummerm et al. 2006).

We define our target parameter, $g(m_{ijk})$, as the area under the response-surface, excluding any background response, for each particle, outcome, and experiment. Any number of classical summaries can be derived from this model and used for risk assessment such as benchmark doses (BMD), effective concentrations (ECα), no observable adverse effect level (NOAEL), among others. These measures summarize different aspects of the response-surface, and typically disjointly for dose and time kinetics. In fact, there is still disagreement in the HTS setting on the best measures of risk (Stern and McNeil 2008 and Maynard et al. 2006). Furthermore, some of these summaries can become even more problematic in the

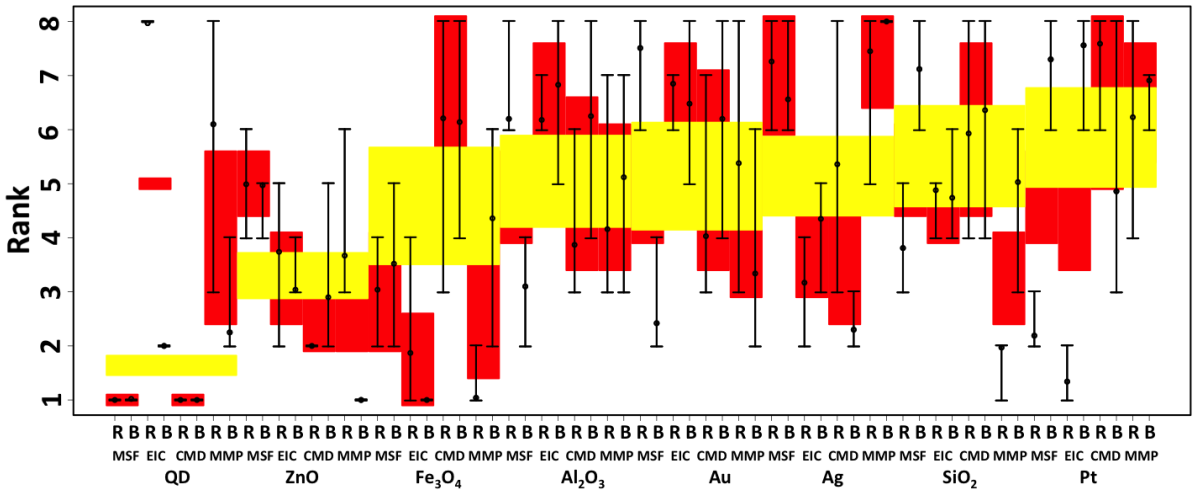

Figure 3.5: **Rankings based on squared-error loss using aggressive weights favoring important outcomes.** Posterior expected ranks and 95% posterior intervals computed by minimizing squared-error loss. Each cell-line is given equal weight, and each outcome is given weights (.54, .27, .12, .07), in the order (CMD, MSF, MMP, EIC). Ranks are displayed individually within outcomes and experiments (black), aggregated across experiments, but within outcomes (red), and aggregated across all outcomes and experiments (yellow).

MSF		EIC		CMD		MMP	
ENM	Rank	ENM	Rank	ENM	Rank	ENM	Rank
QD	1.01 (1,1)	Fe3O4	1.43 (1,2.5)	QD	1 (1,1)	ZnO	2.34 (2,3.5)
Fe_3O_4	3.28 (2,4)	ZnO	3.39 (2.5,4)	ZnO	2.45 (2,3.5)	Fe3O4	2.7 (1.5,4)
Al_2O_3	4.65 (4,5.5)	Ag	3.76 (3,4.5)	Ag	3.83 (2.5,5)	SiO2	3.5 (2.5,4)
Pt	4.75 (4,5.5)	Pt	4.45 (3.5,5)	Al2O3	5.06 (3.5,6.5)	QD	4.17 (2.5,5.5)
Au	4.96 (4,6)	SiO2	4.81 (4,5.5)	Au	5.12 (3.5,7)	Au	4.36 (3,6)
ZnO	4.98 (4.5,5.5)	QD	4.99 (5,5)	SiO2	6.14 (4.5,7.5)	Al2O3	4.64 (3.5,6)
SiO_2	5.47 (4.5,6)	Al2O3	6.5 (5.5,7.5)	Fe3O4	6.18 (4,8)	Pt	6.57 (5.5,7.5)
Ag	6.91 (6,8)	Au	6.66 (6,7.5)	Pt	6.23 (5,8)	Ag	7.72 (6.5,8)

Table 3.2: **Rankings within outcomes based on squared-error loss.** Aggregated ranks across each each outcome and aggregated across cell-lines. Posterior expected ranks and 95% posterior intervals computed by minimizing squared-error loss. Each cell-line is given equal weight.

nanotoxicology setting due to issues such as dosimetry, where the administered doses may in fact differ from the dose absorbed by the cell. The area under the response-surface is an overall summary of the entire dose and duration-response dynamic. Although no sufficient one-dimensional summary of the dose and duration-response profile exists, the area under the surface may be a more comprehensive summary of risk. Using the model described in (3.9), we define area under the surface as follows:

$$\begin{aligned} AUS &= \int_0^T \int_0^D m_{ijk}(d,t) - \alpha_{ijk}\,\mathrm{d}d\,\mathrm{d}t \\ &= \int_0^T \int_0^D \mathcal{B}(d,\boldsymbol{\phi}_{ijk})'\boldsymbol{\beta}_{ijk} + \mathcal{B}(t,\boldsymbol{\psi}_{ijk})'\boldsymbol{\gamma}_{ijk}\,\mathrm{d}d\,\mathrm{d}t \\ &= T[\tfrac{1}{2}(D-\phi_{ijk1})\beta_{ijk3} + \tfrac{1}{2}(D-\phi_{ijk2})\beta_{ijk4}] + \\ &\quad D[\tfrac{1}{2}(T-\psi_{ijk1})\gamma_{ijk3} + \tfrac{1}{2}(T-\psi_{ijk2})\gamma_{ijk4}] \end{aligned} \tag{3.10}$$

Using the ranking methods described in Section 3.4, we rank the eight nanoparticles within cell lines and outcomes, within outcomes but aggregated across cell-lines, and aggregated across cell-lines and outcomes. Table 3.1 (*left panel*) provides posterior expected ranks, and associated 95% posterior intervals ranks, for each particle, aggregated across all cytotoxicity outcomes and cell-lines, with each outcome and cell-line weighted equally. Based on knowledge about oxidative stress pathways and the assays used to measure cytotoxicity outcomes, it is believed that the outcomes measured can be ranked in order of importance as follows: (CMD, MSF, MMP, EIC). Using the weight function described in (3.8) and a value of $\delta = .75$, yields weights $(.37, .27, .21, .15)$, for the four outcomes. Similarly, using a slightly more aggressive $\delta = .5$ gives weights $(.54, .27, .12, .07)$. Table 3.1 (*middle panel* and *right panel*) provides overall summaries, aggregated across outcomes and cell-lines, for the posterior ranks using the weight functions described. Additionally, Table 3.2 and 3.3 provide rankings for each particle, within outcomes but across cell lines, and within outcomes and experiments, respectively. In each of the cases we compute expected ranks and 95% posterior intervals by minimizing squared-error loss. Figures 3.3 - 3.5 provide graphical summaries of the expected posterior ranks and associated 95% posterior intervals at each level of hierarchy. The ranks are computed by minimizing squared-error loss and using equal weights for each cell line, but using different weights for each outcome as described above.

	MSF		EIC		CMD		MMP	
Cell-line	ENM	Rank	ENM	Rank	ENM	Rank	ENM	Rank
RAW 264.7	QD	1 (1,1)	Pt	1.34 (1,2)	QD	1 (1,1)	Fe3O4	1.04 (1,2)
	Pt	2.19 (2,3)	Fe3O4	1.87 (1,4)	ZnO	2 (2,2)	SiO2	1.97 (1,2)
	Fe_3O_4	3.04 (2,4)	Ag	3.17 (2,4)	Al2O3	3.87 (3,6)	ZnO	3.67 (3,6)
	SiO_2	3.81 (3,5)	ZnO	3.74 (2,5)	Au	4.03 (3,7)	Al2O3	4.16 (3,7)
	ZnO	4.99 (4,6)	SiO2	4.88 (4,5)	Ag	5.36 (3,8)	Au	5.38 (3,8)
	Al_2O_3	6.2 (6,8)	Al2O3	6.18 (6,7)	SiO2	5.93 (4,8)	QD	6.1 (3,8)
	Ag	7.26 (6,8)	Au	6.85 (6,7)	Fe3O4	6.21 (3,8)	Pt	6.23 (4,8)
	Au	7.51 (6,8)	QD	7.97 (8,8)	Pt	7.59 (6,8)	Ag	7.45 (5,8)
BEAS-2B	QD	1.02 (1,1)	Fe3O4	1 (1,1)	QD	1 (1,1)	ZnO	1 (1,1)
	Au	2.42 (2,4)	QD	2 (2,2)	Ag	2.3 (2,3)	QD	2.25 (2,4)
	Al_2O_3	3.1 (2,4)	ZnO	3.04 (3,4)	ZnO	2.9 (2,5)	Au	3.34 (2,6)
	Fe_3O_4	3.52 (2,5)	Ag	4.35 (3,5)	Pt	4.86 (3,8)	Fe3O4	4.36 (2,6)
	ZnO	4.97 (4,5)	SiO2	4.74 (4,6)	Fe3O4	6.14 (4,8)	SiO2	5.03 (3,6)
	Ag	6.56 (6,8)	Au	6.48 (5,8)	Au	6.2 (4,8)	Al2O3	5.12 (3,7)
	SiO_2	7.12 (6,8)	Al2O3	6.83 (5,8)	Al2O3	6.25 (4,8)	Pt	6.91 (6,7)
	Pt	7.3 (6,8)	Pt	7.56 (6,8)	SiO2	6.36 (4,8)	Ag	8 (8,8)

Table 3.3: **Individual rankings based on squared-error loss.** Posterior expected ranks and 95% posterior intervals computed by minimizing squared-error loss. Each cell-line is given weight and each outcome (CMD,MSF,MMP,EIC) is given weights: $(.37, .27, .21, .15)$.

In Figure 3.5, the posterior intervals reflect the uncertainty about the hazard rankings, and provide an illustration of how the proposed technique combines information and provides more precise estimates of ranks when aggregated across outcomes and cell lines. Although it is difficult to distinguish between more or less toxic particles when looking at individual ranks, we can discern that aggregated across outcomes and cell lines, quantum dot nanoparticles show significantly higher responses than all other particles. Zinc oxide nanoparticles also show a significantly higher toxic response than nano platinum, silver, aluminum oxide, silicon dioxide, and gold particles. Table 3.2 (*column 3*) provides rankings for cellular membrane damage, aggregated across the two cell lines. In terms of cellular membrane damage, quantum dot and zinc oxide show significantly higher cytotoxic responses then the remaining six nanoparticles. Previous studies have shown that while platinum, silver, aluminum oxide, silicon dioxide, and gold particles nanoparticles have been shown to trigger sublethal responses, they do not increase PI uptake, an indicator of cellular membrane damage, leading to cell death (George et al. 2011). In contrast, it has been shown using conventional assays that QD nanoparticles stabilized by toluene are capable of inducing tier 2 and 3 oxidative stress responses induced by the toluene (George et al. 2011), and are therefore significantly more toxic. Similarly, it has been demonstrated that ZnO nanoparticles are capable of inducing

tier 2 and 3 oxidative stress responses through Zn_2^+ release (George et al. 2009).

3.6 Discussion

In this article we present various loss function based ranking approaches and apply them to the hazard ranking of nanomaterials, using multivariate toxicity data obtained from HTS assays. Furthermore, we extend these methods to the aggregation of ranks across different sources of evidence. We account for the multivariate nature of the data using a Bayesian hierarchical framework, and coupled with a loss function, are thereby able to derive a rank estimate and its associated uncertainty. The proposed methodology accounts for the variability in the scale of the response across cytotoxicity measures and experimental platforms and allows for the differential weighting of these measures in the estimation of an aggregate rank distribution.

As described by Louis and Shen (1999), Lin et al. (2006), when the posterior distributions of the target parameters are stochastically ordered and are invariant under monotone transformation, the choice of ranking method does not matter. However, in many cases there is a clear benefit involved with using an optimal procedure which is clearly defined by inferential goals. In this chapter we present the most commonly used squared-error loss function, which optimizes the overall ranking of all particles. We also present the upper $100(1-\gamma)\%$ (or lower $100(\gamma)\%$) loss function, which is useful when the goal is identifying the most (or least) toxic fraction of particles. Many other loss functions can be considered. For example, Lin et al. (2006) suggest the use of a weighted combination of several loss functions in order to broaden the class of all loss functions.

One advantage of a ranked list of ENMs is that it allows for the prioritization of further *in-vivo* testing of the materials, especially when resources are limited. Although, loss function based ranking methods are optimal, in many cases the results may not be conclusive as in our example indicated by the large, often overlapping, confidence intervals around the rank estimates. Aggregated ranks across different sources of information can be used to

combine information and reduce the uncertainty of the results. Although the ranked lists are expected to be correlated, this does not guarantee that the ranked lists of ENMs will be in complete agreement. In some cases this disagreement may in fact increase the uncertainty of the results.

Another important consideration when ranking particles is the use of a proper estimate of toxicity. In this chapter, we focused on one summary, the area under the response-surface, that we believe is an adequate measure of risk. An alternative formulation would be to construct multiple ranked lists of particles using more then one summary of risk, followed by the aggregation of these lists. Furthermore, in this chapter we have focused on ENM hazard as the sole factor in determining risk. In toxicology, risk assessment involves the characterization of hazard as well as the potential for exposure. Currently, there is not enough exposure information available to perform a traditional risk assessment, but an important area of future research involves the aggregation of hazard and exposure rankings (Maynard et al. 2006).

CHAPTER 4

Relating ENM Properties to Toxicity

4.1 Introduction

Knowledge about the potential hazard of nanomaterials is still lacking and extensive study is required to understand how ENM properties such as size, shape, agglomeration state, solubility, and surface properties, could lead to hazard generation at the nano-bio interface (Stern and McNeil 2008, Nel et al. 2006). High-throughput screening (HTS) assays provide an opportunity to study biological relationships, and may suggest which nanoparticles are likely to have an *in-vivo* effect. While HTS assays cannot replace traditional animal studies, they can be used to explore the large number of potential nanomaterial variables that can influence human health hazard (Meng et al. 2010, Stanley et al. 2008, Maynard et al. 2006).

In this chapter, we analyze data on 24 metal-oxide nanoparticles, monitored in relation to a cellular injury response measuring cellular membrane damage. The outcome was measured contemporaneously over a grid of ten doses and seven hours of exposure. All particles were characterized in terms of their size, dissolution, crystal structure, conduction energies, as well as many other particle descriptors (Zhang et al. 2012).

Quantitative structure-activity relationships (QSAR) models provide a way to understand how combinations of physical and chemical characteristics predict the toxicity of nanomaterials with respect to its biological activity. In contrast to QSARs for chemicals, very little work has been done to develop nano-QSAR models. The primary efforts to date have been on adapting the idea of classical chemical QSAR models to nanomaterials, which are restricted due to limited information for ENMs in terms of physicochemical and toxicity data, and the

lack of a well defined toxicological endpoint.

The use of a combination of molecular properties to predict a compound's behavior, with respect to biological end-points, is a well accepted concept in the predictive toxicology of chemicals (Schultz et al. 2003). Stanley et al. (2008) present a nano-QSAR model based on an advanced data mining algorithms. The authors apply a hierarchical and consensus clustering methods to classify nanoparticles into groups based on similarity in their toxic activity profiles. Subsequent clustering patterns can then be compared in relation to ENM properties that tend to be assigned to the same class. Unsupervised clustering methods, although useful, lack specificity and do not take full account of dose and time response profiles.

Puzyn et al. (2011) present a nano-QSAR model for metal-oxide nanoparticles based on methods commonly used in the predictive toxicology of chemicals. The authors modeled the relationship between ENM toxicity and various structural descriptors, using a modified multiple linear regression algorithm that accounts for non-linear effects. ENM toxicity was characterized by a summary that estimated the dose concentration that brought about 50% reduction in bacterial cell viability (EC50). Similarly, Liu et al. (2011) present a nano-QSAR model that uses logistic regression to to model the probability that a particle is toxic, given structural covariates. A particle was defined as toxic, at a given concentration, based on a statistically significant difference in mean response, as compared to the background response in unexposed cells. These techniques, however, are designed to work with previously defined summaries of the dose-response trajectory, and to our knowledge, fail to take into account the uncertainty inherent in the estimation of these summaries when relating them to structural covariates.

In this chapter, we introduce a new new class of models that relate ENM physicochemical properties to cytotoxicity profiles, in order to initiate a framework for predictive toxicology. The proposed methodology introduces a new measure of toxicity that is seamlessly integrated into a multi-dimensional model that accounts for dose and duration kinetics jointly using a flexible smooth surface fit. This toxicological endpoint measures the probability of toxicity,

an intuitively more appealing summary of hazard, as it is comparable across experiments, and can be used to seamlessly integrate multiple dose and time escalation studies. Furthermore, the probability of toxicity is used to link nanomaterial physicochemical properties to non-linear and multi-dimensional cytotoxicity profiles, while accounting for the variability in the estimation of this summary. Finally, this methodology is appropriate for limited data sets, and includes data integration and a framework for advanced dimension reduction through variable selection.

The remainder of this chapter is organized as follows. In Section 4.2 we introduce the proposed model. In Section 4.3 we discuss parameter estimation and associated inferential details. Section 4.4 employs the proposed model for the analysis of 24 metal-oxide nanomaterials, and describes inference for cytotoxicity profiles and their relationship to ENM-specific physicochemical properties. We conclude with a critical discussion of the limitations and possible extensions of the proposed method in Section 4.5.

4.2 Model Formulation

4.2.1 Toxicity Model

In this section we describe a smooth response model, where we monitor a continuous outcome y corresponding to the exposure of a number of cells to I different ENMs. Let $y_{ij}(d,t)$ denote a response corresponding to ENM i $(i = 1, ..., I)$ and replicate j $(j = 1, ..., J)$, at dose $d \in [0, D]$ and time $t \in [0, T]$. In practice, observations are obtained over a discrete set of doses and durations of exposure. However, for ease of presentation and without loss of generality, we will assume that doses and times are defined over continuous intervals. We introduce the following hierarchical model.

Stage 1: Sampling Model

The observed response of particle i and replicate j is modeled as:

$$y_{ij}(d,t) = m_i(d,t) + \epsilon_{ij}(d,t), \tag{4.1}$$

where $\epsilon_{ij}(d,t) \sim N(0, \sigma^2_\epsilon/\tau_i)$ and $\tau_i \mid \nu \sim Gamma(\nu/2, \nu/2)$. Here $m_i(d,t)$ denotes the response surface for each particle i. This quantity describes dose and duration kinetics for all $d \in [0, D]$ and $t \in [0, T]$, and is expected to exhibit a non-linear dynamic over these domains. The distribution of y_{ij} is modeled in terms of the error term ϵ_{ij} as a scaled mixture of normal random variables to account for outlying observations. The error variance is defined in terms of the measurement error σ^2_ϵ, and on ENM-specific variance inflation parameter τ_i. This formulation results in a t-distributed error (see West 1984), where we borrow strength across all ENM by assuming the same error variance, but retain robustness in the model by allowing ENM-specific departures from normality.

Stage 2: Surface Response model at the ENM level

The dose-response surface $m_i(d,t)$ spans two dimensions (dose and time), and is modeled as a non-parametric smooth fit. Let $(\alpha_i, \boldsymbol{\beta}_i')'$ be a parameter vector indexing the response surface $m_i(d,t)$. We define

$$m_i(d,t) = \alpha_i + f_i(d,t;\boldsymbol{\beta}_i), \tag{4.2}$$

where $f_i(d,t;\boldsymbol{\beta}_i)$ is a smooth surface modeling the effect of dose d and time t, for each ENM i. To ensure likelihood identifiability we require, without loss of generality, that $f_{ij}(d = 0, t = 0; \boldsymbol{\beta}_i) = 0$.

The surface response function $f_i(d,t;\boldsymbol{\beta}_i)$ is modeled using two-dimensional P-splines, as described by Lang and Brezger (2006). We assume that the surface $f_i(d,t;\boldsymbol{\beta}_i)$ can be approximated by the tensor product of two 1-dimensional B-splines. Specifically, we use polynomial splines of degree r and equally spaced knots over the domain of d and t. let $\mathcal{B}_m(x)$ denote the mth basis of a M-dimensional B-spline basis with interior knots $x_{min} = \eta_0 < \eta_1, ... < \eta_{s-1} < \eta_s = x_{max}$. Also, let $\boldsymbol{\beta}_i = (\beta_{i1}, .., \beta_{i,m_d m_t}, ...\beta_{i,M_d M_t})'$ be a $(M_d M_t)$-dimensional vector of spline coefficients. The function $f_i(d,t;\boldsymbol{\beta}_i)$ can then be represented as follows:

$$f_i(d,t;\boldsymbol{\beta}_i) \;=\; \sum_{m_d}^{M_d} \sum_{m_t}^{M_t} \beta_{i,m_d m_t} \mathcal{B}_{m_t}(t) \mathcal{B}_{m_d}(d) = \boldsymbol{H}_i \boldsymbol{\beta}_i, \tag{4.3}$$

where M_d is the number of basis functions over the domain of d, M_t is the number of basis functions the domain of t, and $\boldsymbol{H}$ is $(DTK \times M_d M_t)$ dimensional design matrix, which can be defined as $(\mathcal{B}_1(d)\mathcal{B}_1(t), \ldots, \mathcal{B}_{m_d}(d)\mathcal{B}_{m_t}(t), \ldots \mathcal{B}_{M_d}(d)\mathcal{B}_{M_t}(t))$. Identifiability restriction $f_{ij}(d = 0, t = 0; \boldsymbol{\beta}_i) = 0$, is implemented by fixing $\beta_{i1} = 0$, for all particles.

We can expand the model further, to include a summary measure of overall toxicity. To do that we include a latent indicator variable γ_i, so that for each particle i

$$m_i(d,t) \quad = \begin{cases} \alpha_i & \text{if } \gamma_i = 0 \\ \alpha_i + f_i(d,t;\boldsymbol{\beta}_i) & \text{if } \gamma_i = 1, \end{cases} \tag{4.4}$$

where $\gamma_i \mid \boldsymbol{\theta} \sim Bern(p_i)$. The indicator variable γ_i, can then be used as a measure of toxicity for particle i. This trans-dimensional parameterization is key to linking the toxicity of ENM i to it's physicochemical properties, where $\boldsymbol{\theta}$ is the parameter vector relating the toxicity of particle i to their corresponding physicochemical properties. These quantities will be described further in Section 4.2.2.

For each ENM i, we define the following prior distribution for α_i:

$$\alpha_i \quad \sim \quad N(m_{\alpha_i}, v_{\alpha_i}). \tag{4.5}$$

The priors for $\boldsymbol{\beta}_i$ are based on spacial smoothness priors (Besag and Kooperberg 1995, Lang and Brezger 2004). Specifically, we use a prior based on the Kronecker product of the penalty matrices of the dose and time effects, and is defined as follows:

$$\boldsymbol{\beta}_i \quad \sim \quad N_{M_d M_t}(0, \sigma^2_{\beta i}(K_d \otimes K_t)) \tag{4.6}$$

The penalty matrices K_d and K_t, can be constructed based on a first or second order random walk process, as described in detail by Lang and Brezger (2004). To summarize, penalty matrices K_d and K_t are used to define prior distributions on the $\boldsymbol{\beta}$ parameters of our non-linear response function, and are the Bayesian analog of difference penalties used in penalized likelihood estimation. First order difference penalties correspond to a first-order random walk

process and similarly, second order differences corresponds to a second-order random walk process. For example, for the dose component we obtain

$$\begin{aligned} \beta_{im_d} &= \beta_{im_{d-1}} + \zeta_{im_d} && (first-order) \\ \beta_{im_d} &= 2\beta_{im_{d-1}} - \beta_{im_{d-2}} + \zeta_{im_d} && (second-order) \end{aligned} \tag{4.7}$$

where, $\zeta_{im_d} \sim N(0, \sigma^2_{\beta i})$. The smoothness of the response surface is controlled by the variance parameter $\sigma^2_{\beta i}$. For example, for a second-order random walk process, the penalty matrix has the following banded structure

$$K_d = \begin{bmatrix} 1 & -2 & 1 & & & & & & & \\ -2 & 5 & -4 & 1 & & & & & & \\ 1 & -4 & 6 & -4 & 1 & & & & & \\ & & & \ddots & \ddots & \ddots & \ddots & & & \\ & & & & & 1 & -4 & 6 & -4 & 1 \\ & & & & & & 1 & -4 & 5 & -2 \\ & & & & & & & 1 & -2 & 1 \end{bmatrix} \tag{4.8}$$

The foregoing penalty matrix leads to a rank deficient covariance matrix. However, adding a small positive constant to components (1,1) and (D,D) reinstates propriety without fundamentally changing the smoothing properties of the proposed stochastic prior. Another choice of prior for $\boldsymbol{\beta}_{ij}$ can be based on the four nearest neighbors (see Lang and Brezger 2004). We prefer the prior described in 4.6 because it provides a smoother fit to the data (Clayton 1996, Lang and Brezger 2004), a desirable property in this context due to the high measurement error associated with HTS data.

Stage 3: Hyperpriors

We complete the model by specifying prior distributions on the hyperparameters as follows:

$$1/\sigma^2_\epsilon \sim Gamma(a_\epsilon, b_\epsilon), \quad 1/\sigma^2_{\alpha_i} \sim Gamma(a_{\alpha_i}, b_{\alpha_i}), \quad 1/\sigma^2_{\beta_i} \sim Gamma(a_{\beta_i}, b_{\beta_i}). \tag{4.9}$$

We model precision parameters as gamma distributions, again exploiting conditional conjugacy. Note that in our formulation, $x \sim Gamma(a, b)$ denotes a Gamma distributed random quantity with shape a and rate b, such that $E(x) = a/b$.

4.2.2 Covariate Model

For each ENM (i=1,...I), we would like to model a set of P physicochemical properties as a vector of covariates $\boldsymbol{x_i} = (x_{i0}, ...x_{ip}, ..., x_{iP})'$, where $p = 0$ corresponds to the model intercept. We relate cytotoxicity vector γ_i to covariate vector $\boldsymbol{x_i}$ through a regression at the ENM level, such that

$$g(\gamma_i \mid \boldsymbol{x_i}) \quad = \quad \boldsymbol{x_i'\theta}, \tag{4.10}$$

where $\boldsymbol{\theta} = (\theta_0, ...\theta_p, ...\theta_P)$ is a vector of regression coefficients and g is a standard link function. The regression coefficients $\boldsymbol{\theta}$, summarize how variability in ENM physicochemical properties explain the variability in the probability of toxicity, as determined by the ENM toxicity profiles.

Following Albert and Chib (1993), we introduce a latent-variable structure into the model. We define a latent continuous measure of toxicity Z_i, which is positive when a particle is toxic, and negative when a particle is non-toxic. More precisely, for each ENM ($i = 1, ...I$), we denote a continuous variable Z_i, where

$$\gamma_i(d, t) \quad = \begin{cases} 0 & \text{if } Z_i \leq 0 \\ 1 & \text{if } Z_i > 0. \end{cases} \tag{4.11}$$

Furthermore, the toxicity of each ENM i is related to the P physicochemical properties by the normal regression model

$$Z_i \quad = \quad \boldsymbol{x'\theta} + \zeta_i, \tag{4.12}$$

where all ζ_i ($i = 1, \ldots, I$) are assumed to be independent and identically distributed accord-

ing to g. We can then define the probability of toxicity, given covariate vector $\boldsymbol{x_i}$, as

$$p_i = P(\gamma_i = 1) = P(Z_i > 0) = g(\boldsymbol{x_i'\theta}). \tag{4.13}$$

Specifically, the last equality holds true when $g(\cdot)$ is a symmetric distribution such as the probit model. Introducing a latent structure when modeling binary data, along with a probit link function, where $g(\cdot)$ is the normal cumulative distribution function, allows for an automatic Gibbs sampling algorithm to be constructed (Albert and Chib 1993 and Johnson and Albert 1999).

In some instances we may like to define only a subset of the P regressors, which are specifically related to the probability of toxicity. We can introduce variable selection into the model above by defining a binary indicator vector

$$\rho_p \quad = \begin{cases} 0 & \text{if } \theta_p = 0 \text{ (physicochemical property } p \text{ selected)} \\ 1 & \text{if } \theta_p \neq 0 \text{ (physicochemical property } p \text{ not selected),} \end{cases} \tag{4.14}$$

where $p = 1, ...P$. Given $\boldsymbol{\rho}$, we can define $\boldsymbol{\theta}_\rho$ as a vector that includes all nonzero elements of $\boldsymbol{\theta}$, and $\boldsymbol{X_\rho}$ as the design matrix, which only includes the columns of X corresponding to the non-zero elements of $\boldsymbol{\rho}$. In this context, we assume that $\rho_0 = 1$, therefore θ_0, model intercept, is included in every model.

Finally, we complete the model by defining a prior for $\boldsymbol{\theta}_\rho$ and $\boldsymbol{\rho}$ as

$$\begin{aligned} \boldsymbol{\theta}_\rho \quad &\sim \quad N_q(0, c(\boldsymbol{X_\rho'X_\rho})^{-1}) \\ P(\rho_p = 1) \quad &= \quad \pi_p, 0 \leq \pi_p \geq 1, p = 1, \ldots, P, \end{aligned} \tag{4.15}$$

where c is a constant positive scaling factor and $q = \sum_{p=0}^{P} \rho_p$. Large values of c lead to a non-informative prior distribution on $\boldsymbol{\theta}$. As described by Lee et al. (2003), choosing small values for π_p, leads to parsimonious models by restricting the number of covariates included. The values of π_p can also be tuned to include prior knowledge about the importance of certain

physicochemical properties. Most commonly, a prior is placed on π_p, such that

$$\pi_p \sim Beta(a_\pi, b_\pi), \tag{4.16}$$

as described by Scott and Berger (2006).

4.3 Estimation and Inference

4.3.1 Posterior Simulation via MCMC

Using the P-spline representation introduced in Section 4.2.1, we can write the response level associated with ENM i, at dose d, and exposure time t as

$$m_i(d,t;\ \alpha_i, \boldsymbol{\beta}_i, \cdots) = \begin{cases} \alpha_i + \sum_{m_d}^{M_d} \sum_{m_t}^{M_t} \beta_{i,m_d m_t} \mathcal{B}_{m_t}(t) \mathcal{B}_{m_d}(d) & \text{if } \gamma_i = 1 \\ \alpha_i & \text{if } \gamma_i = 0. \end{cases}$$

Furthermore, we can write the probability that ENM i is toxic, given the vector of physicochemical properties contained in $\boldsymbol{x_i}$, as

$$P(\gamma_i = 1 \mid \boldsymbol{\rho}) = P(Z_i > 0) = \boldsymbol{x'_{i\rho}} \boldsymbol{\theta}_\rho.$$

Let $\boldsymbol{\beta} = \{\boldsymbol{\beta}_i :\ i = 1, \ldots, I\}$ denote the full set of spline coefficients, $\boldsymbol{\alpha} = \{\alpha_i :\ i = 1, \ldots, I\}$ the background response parameters, and $\boldsymbol{\gamma} = \{\gamma_i :\ \ i = 1, \ldots, I\}$ the latent indicator of toxicity. Furthermore, consider regression coefficients $\boldsymbol{\theta} = \{\boldsymbol{\theta}_i :\ \ i = 1, \ldots, I\}$, latent continuous measure of toxicity $\boldsymbol{Z} = \{Z_i :\ \ i = 1, \ldots, I\}$, and latent indicator of covariate selection $\boldsymbol{\rho} = \{\rho_p :\ \ p = 0, \ldots, P\}$. Finally, let $\boldsymbol{\tau} = (\tau_i, \ldots, \tau_I)'$. If we denote with $\mathbf{Y}$ the complete set of response values for all particles, we can write the likelihood function as follows:

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \sigma^2_\epsilon, \boldsymbol{\tau} \mid \mathbf{Y}) \propto \prod_{i,j,k,d,t} \left[(\sigma^2_\epsilon / \tau_i)^{-\frac{1}{2}} \exp\left\{ -\frac{(\, y_{ij}(d,t) - m_i(d,t;\ldots)\,)^2}{2\sigma^2_\epsilon / \tau_i} \right\} \right], \tag{4.17}$$

where the product is taken over all replicates k, particles i, doses d and times t. We seek inference on the full set of parameters, described above, through the posterior probability

$$\begin{aligned} P(\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \sigma_\epsilon^2, \boldsymbol{\tau} \mid \mathbf{Y}) \propto \; & L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \boldsymbol{\sigma}_\epsilon^2, \boldsymbol{\tau} \mid \mathbf{Y}) \\ & \times P(\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \sigma_\epsilon^2, \boldsymbol{\tau}) \end{aligned} \tag{4.18}$$

where the prior model, $P(\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \sigma_\epsilon^2, \boldsymbol{\tau})$, is a product of the conditionally independent prior distributions described in Section 4.2.

The posterior distribution is unavailable in closed analytic form, therefore we base our inference on Markov Chain Monte Carlo ($MCMC$) simulations. Closed-form full conditional distributions are available for all parameters, therefore the proposed posterior simulation algorithm utilizes a Gibbs sampling algorithm to directly sample and update parameters component-wise (Geman and Geman 1984, Gelfand and Smith 1990); Full conditional distributions are given in Appendix C.1.

As we are considering selection of response functions in a trans-dimensional setting, we will first sample from the marginal distribution of the model indicator γ, integrating out $\boldsymbol{\beta}$, allowing us to jump between models with and without a dose-time response function $f_i(d, t; \boldsymbol{\beta}_i)$. Similarly, for the covariate model described in Section 4.2.2, we begin by sampling from the marginal distribution of $\boldsymbol{\rho}$, integrating out $\boldsymbol{\theta}$, allowing us to jump between models with a subset of the P covariates. The model indicators $\boldsymbol{\gamma}$ and $\boldsymbol{\rho}$ are updated using Gibbs sampling steps. After the model structure has been specified, the model parameters are updated from their corresponding conditional posterior distributions. The proposed sampling scheme can be summarized as follows.

1. Trans-dimensional updates

We begin by drawing from $\boldsymbol{\gamma} \mid \boldsymbol{Y}, \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \sigma_\epsilon^2, \boldsymbol{\tau}$; the marginalized conditional distribution of $\boldsymbol{\gamma}$, obtained by integrating $\boldsymbol{\beta}$ out of the full conditional distribution of $\boldsymbol{\gamma}$. More precisely,

$$
\begin{aligned}
P(\boldsymbol{Y} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\sigma}_\epsilon^2, \boldsymbol{\tau}) &\propto \int_\beta P(\boldsymbol{Y} \mid \boldsymbol{\beta}_\gamma \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\sigma}_\epsilon^2, \boldsymbol{\tau}) P(\boldsymbol{\beta}_\gamma \mid \boldsymbol{\gamma}) \, \mathrm{d}\boldsymbol{\beta}_\gamma \\
&\propto \prod_{i=1}^{I} \mid \boldsymbol{\Sigma_{\beta i}} \mid^{-\frac{1}{2}} \mid (\tfrac{\tau_i}{\sigma_{\epsilon i}^2} \boldsymbol{H'_{\gamma i} H_{\gamma i}} + \boldsymbol{\Sigma_{\beta i}^{-1}}) \mid^{-\frac{1}{2}} \\
&\quad exp\left\{ \tfrac{1}{2} \left((\tfrac{\tau_i}{\sigma_{\epsilon i}^2} \boldsymbol{H'_{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})' (\tfrac{\tau_i}{\sigma_{\epsilon i}^2} \boldsymbol{H'_{\gamma i} H_{\gamma i}} + \boldsymbol{\Sigma_{\beta i}^{-1}})^{-1} (\tfrac{\tau_i}{\sigma_{\epsilon i}^2} \boldsymbol{H'_{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}}) \right) \right\},
\end{aligned}
$$

where $\tilde{\boldsymbol{Y}}_{\boldsymbol{i}} = \boldsymbol{Y_i} - \boldsymbol{1_{DTK}} \boldsymbol{\alpha}_{\boldsymbol{i}}'$. The derivation of this integral is provided in Appendix C.1. The conditional distribution $P(\gamma \mid \boldsymbol{Y}, \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\sigma}_\epsilon^2, \boldsymbol{\tau})$ is then given by

$$
\begin{aligned}
P(\boldsymbol{\gamma} \mid \boldsymbol{Y}, \boldsymbol{\alpha}, \boldsymbol{\theta}, \sigma_\epsilon^2, \boldsymbol{\tau}) &\propto P(\boldsymbol{Y} \mid \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\sigma}_\epsilon^2, \boldsymbol{\tau}) P(\boldsymbol{\gamma} \mid \boldsymbol{\theta}) \\
&\propto \prod_{i=1}^{I} \mid \boldsymbol{\Sigma_{\beta i}} \mid^{-\frac{1}{2}} \mid (\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{H'_{\gamma i} H_{\gamma i}} + \boldsymbol{\Sigma_{\beta i}^{-1}}) \mid^{-\frac{1}{2}} \\
&\quad exp\left\{ \tfrac{1}{2} \left((\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{H'_{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})' (\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{H'_{\gamma i} H_{\gamma i}} + \boldsymbol{\Sigma_{\beta i}^{-1}})^{-1} (\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{H'_{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}}) \right) \right\} \\
&\quad \Phi(-\boldsymbol{x_i \theta}_i)^{\gamma_i} \Phi(\boldsymbol{x_i \theta}_i)^{1-\gamma_i},
\end{aligned}
$$

where $\Phi(\cdot)$ is the cdf of a standard normal distribution. When the number of particles get large, it is convenient to draw each γ_i component-wise from $P(\gamma_i \mid \boldsymbol{Y_i}, \gamma_{k \neq i}, \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\rho}, \sigma_\epsilon^2, \boldsymbol{\tau})$.

A similar scheme was adapted for sampling from $\boldsymbol{\rho} \mid \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\gamma}$; the marginalized conditional distribution of $\boldsymbol{\rho}$, obtained by integrating $\boldsymbol{\theta}$ out of the full conditional distribution of $\boldsymbol{\rho}$. Following Lee et al. (2003), the conditional distribution of $P(\rho_p \mid \rho_{k \neq p}, \boldsymbol{Z})$ is given by

$$
\begin{aligned}
P(\rho_p \mid \rho_{k \neq p}, \boldsymbol{Z}) &\propto P(\boldsymbol{Z} \mid \boldsymbol{Z}) P(\boldsymbol{\rho}_p) \\
&\propto exp\left[\tfrac{1}{2} (\boldsymbol{Z'Z} - \tfrac{c}{1+c} \boldsymbol{Z' X}_\rho (\boldsymbol{X}_\rho' \boldsymbol{X}_\rho)^{-1} \boldsymbol{X}_\rho' \boldsymbol{Z}) \right] \pi_p^{\rho_p} (1 - \pi_p)^{1-\rho_p}.
\end{aligned}
$$

2. Fixed dimensional updates.

Given the current state of the latent indicators $\boldsymbol{\gamma}$, response surfaces are uniquely defined as in (5.3). Posterior sampling here is standard, and proceeds by updating spline coefficients $\boldsymbol{\beta}_\gamma$, background response parameters $\boldsymbol{\alpha}$, and variance parameters $\boldsymbol{\tau}$ and σ_ϵ^2, from their full conditional distributions via direct simulation (Gibbs step - Appendix C.1).

Similarly given the latent indicators $\boldsymbol{\rho}$, the covariate model is uniquely defined as in (4.14). Again, posterior sampling is standard and proceeds by updating regression coefficients $\boldsymbol{\theta}_\rho$ from $\boldsymbol{\theta}_\rho \mid \boldsymbol{\rho}$, and, $\boldsymbol{Z}$ from $\boldsymbol{Z} \mid \boldsymbol{\theta}, \boldsymbol{\rho}, \boldsymbol{\gamma}$, via Gibbs sampling. Again, full conditional distributions

are given in Appendix C.1.

4.3.2 Posterior Inference

In this section we discuss inference for the model described in Section 4.2, based on draws from the posterior distribution described in (4.18). Let $\boldsymbol{\beta_i^{(n)}}$, $\alpha_i^{(n)}$, $\gamma_i^{(n)}$, $\boldsymbol{\theta_i^{(n)}}$ and $\rho_p^{(n)}$, $n = 1, ..., N$, denote N MCMC draws from the posterior distribution of $\boldsymbol{\beta_i}$, α_i, γ_i, $\boldsymbol{\theta_i}$ and ρ_p.

One quantity of interest is an estimate of the dose-response surface $m_i(d, t)$ for each ENM i. This surface is defined in an infinite-dimensional space; however, given the P-spline representation introduced in Section 4.2.1, we only need finite draws from the parameter set of interest. More precisely, draws from the marginal posterior distribution of the dose-response surface for any dose $d \in [0, D]$ and time $t \in [0, T]$ are given by

$$m_i^{(n)}(d,t) \;\; = \begin{cases} \alpha_i^{(n)} + \sum_{m_d}^{M_d} \sum_{m_t}^{M_t} \beta_{i,m_d m_t}^{(n)} \mathcal{B}_{m_t}(t) \mathcal{B}_{m_d}(d) & \text{if } \gamma_i^{(n)} = 1 \\ \alpha_i^{(n)} & \text{if } \gamma_i^{(n)} = 0. \end{cases} \tag{4.19}$$

For each $\boldsymbol{\beta_i^{(n)}}$, $\alpha_i^{(n)}$, and $\gamma_i^{(n)}$, $n = 1, ..., N$, we evaluate the dose-response function given in (4.19) over a grid of values $\tilde{D} = (d_1, ..., d_n)'$ and $\tilde{T} = (t_1, ..., t_n)'$. The posterior mean of the samples $m_{ij}^{(n)}$, $n = 1, ..., N$, at each value of $\tilde{D}$ and $\tilde{T}$, can be used to summarize the fit of the dose-response surface, as shown in Figures 4.1 and 4.2.

The expected inclusion probability of the dose/duration-response surface $\boldsymbol{\beta}_i$ can be estimated using posterior draws $\gamma_i^{(n)}$, as $\hat{p}_{\gamma i} = \sum_n \gamma_i^{(n)}/N$. Similarly, the expected inclusion probability of each of the covariates can be estimated using posterior draws $\rho_p^{(n)}$, as $\hat{p}_{\rho p} = \sum_n \rho_p^{(n)}/N$. This posterior probability can be used for model selection, Scott and Berger (2006) for example, recommend selecting the median model, that is including all interactions for which $\hat{p_{\rho p}} > 0.5$.

Posterior samples $\theta_p^{(n)}$, $p = 0, ..., P$, can provide us with approximations of the posterior distribution for the probability of toxicity, for an ENM with a given a set of physicochemical properties. More precisely, draws from the marginal posterior distribution of the probability

of toxicity for any particle k, with a corresponding set of physicochemical properties $\boldsymbol{x_k} = (x_{0k}, ...x_{pk}, ..., x_{k(P)})'$, are given by

$$p_k^{(n)} = \Phi(\boldsymbol{x_k'}\boldsymbol{\theta}^{(n)}) \tag{4.20}$$

Given posterior draws, one can proceed with the straightforward construction of standard posterior summaries, like means, maxima a posteriori, modes, quantiles and credible regions. Another quantity of interest may be an estimate of the probability of toxicity, as a function of one or more physicochemical properties. Given posterior samples $\boldsymbol{\theta}_p^{(n)}$, $n = 1, ..., N$, we evaluate the probability given in (4.20) over a plausible grid of values $\tilde{\boldsymbol{x}}$, for a particular set of physicochemical properties. The posterior mean of the samples $p^{(n)}$, $n = 1, ..., N$, at each value of $\tilde{\boldsymbol{x}}$, can be used to summarize the probability of toxicity as a function of a set of covariates, as shown in Figure 4.5.

4.4 Applications

4.4.1 Case Study Background

We illustrate the proposed methodology by analyzing data on bronchial epithelial cells (BEAS-2B), exposed to a library of 24 nanoparticles. This data is aimed at screening for toxicity of metal-oxide nanoparticles, including ZnO, CuO, CoO, Fe_2O_3, Fe_3O_4, WO_3, Cr_2O_3, Mn_2O_3, Ni_2O_3, SnO_2, CeO_2, Al_2O_3, among others. The amount of oxidative stress and inflammation induced by metal-oxide nanoparticles, the principal injury mechanism through which ENMs can induce adverse health effects, was measured using a multi-parametric assay that utilizes four compatible dye combinations and subsequent change in fluorescence read-out. In this chapter, we focus on one specific cytotoxic response measuring the amount of membrane damage to the cell. Specifically, this outcome is measured by treating BEAS-2B cells with Propidium Iodide (PI) and measuring the resulting florescence. In cells with damaged membranes PI is able to permeate the cell and bind to DNA, where it causes the

nucleus to emit a red florescence. Each sample was also stained with a Hoechst dye, which causes all cell nuclei to emit a blue florescence, and allows for a count of the total number of cells. An analysis of the fluorescence readout, measured at varying wavelengths, results in a measure of the percentage of cells positive for each response. The outcome is measured over a grid of ten doses and seven times (hours) of exposure (see Figures 4.1 and 4.2), and the final response was normalized using a logit transformation to unconstrain the support so that it can take on values between -∞ and ∞.

Furthermore, all particles were characterized in terms of their size, dissolution, crystal structure, conduction energies, as well as many other particle descriptors (Zhang et al. 2012). Burello and Worth (2011) hypothesized that comparing the conduction and valance band energies to the redox potential of the reactions occurring within a cell could predict nanoparticle toxicity of oxides. Specifically, the mechanistic interpretation is based on the idea that oxidizing and reducing substances can create an imbalance in the normal intracellular state of the cell, either by the production of oxygen radicals or by reducing antioxidant levels. In particular, the potential for oxidative stress might be predicted by comparing the energy structure of oxides, as measured by their conduction and valance band energy levels, to the redox potential of the cell. When these two energy levels are comparable, it can possibly allow for the transfer of electrons, and subsequent imbalance in the normal intracellular state. The normal cellular redox potential is in the range $(-4.12 to - 4.84eV)$. The conduction band energy of a particle is a measure of the energy sufficient to free an electron from an atom, and allow it to move freely within the material. Therefore, one measure of interest is the relationship between particles with conduction band energies within and outside the range of the cellular redox potential, to their cytotoxicity profiles.

Another measure of interest is the metal dissolution rate of a particle. Particles that are highly soluble have the ability to shed metal ions, which can lead to nanoparticle toxicity (Zhang et al. 2012). Other measures that might be of interest include the primary size of the particle, a measure of the crystal structure (b(Å)), lattice energy ($\Delta H_{lattice}$), which measures the strength of the bonds in the particles, and the enthalpy of formation ($\Delta H_{Me^{n+}}$), which

is a combined measure of the energy required to convert a solid to a gas and the energy required to remove n elections from that gas.

4.4.2 Case Study Analysis and Results

The model described in Section 4.2 was fit to the metal-oxide data-set described in the previous section. We placed relatively diffuse $Gamma(.01, .01)$ priors on the $1/\sigma_\epsilon$ parameter, $Gamma(1, .1)$ priors on all remaining precision parameters, and $N(0, 100)$ priors on the α_i parameters. We also placed a prior distribution on the degrees of freedom parameter ν, for the t-distributed error described in Section 4.2.1. Specifically, the prior was modeled using a discrete uniform distribution on 1,2,4,8,16, and 32 degrees of freedom (Besag and Higdon 1999). Finally, for the covariate model we placed a relatively diffuse prior on the regression coefficient $\boldsymbol{\theta}$, by fixing the scaling factor c to a value of 100, and set a $Beta(1, 8)$ prior on π_p, favoring relatively parsimonious models.

Furthermore, the model was initially fit to the ENM library and corresponding six chemical and structural properties described above. In particular, we model log(dissolution) in a non-linear fashion as a spline function with a change point at log(dissolution) of 2.3, conduction band energy as a binary covariate, and all other ENM characteristics described above as continuous linear predictors. In modeling log-dissolution, we really have two separate classes of particles, those with low dissolution and those with high dissolution, therefore it is natural to allow for a discontinuity at the boundary between the two classes. An estimate of the inclusion probability, $\hat{\boldsymbol{\rho}}$ was used to select a single model, specifically we followed Scott and Berger (2006) by selecting the median model, that is including all covariates for which $\hat{\rho}_p > 0.5$. The final model included log(metal dissolution) and conduction band energy.

Our inferences are based on 20,000 MCMC samples from the posterior distribution in (4.18), after discarding a conservative 60,000 iterations for burn-in. MCMC sampling was performed in R version 2.10.0, and convergence diagnostics were performed using the package CODA (Convergence Diagnostics and Output Analysis), (Plummerm et al. 2006).

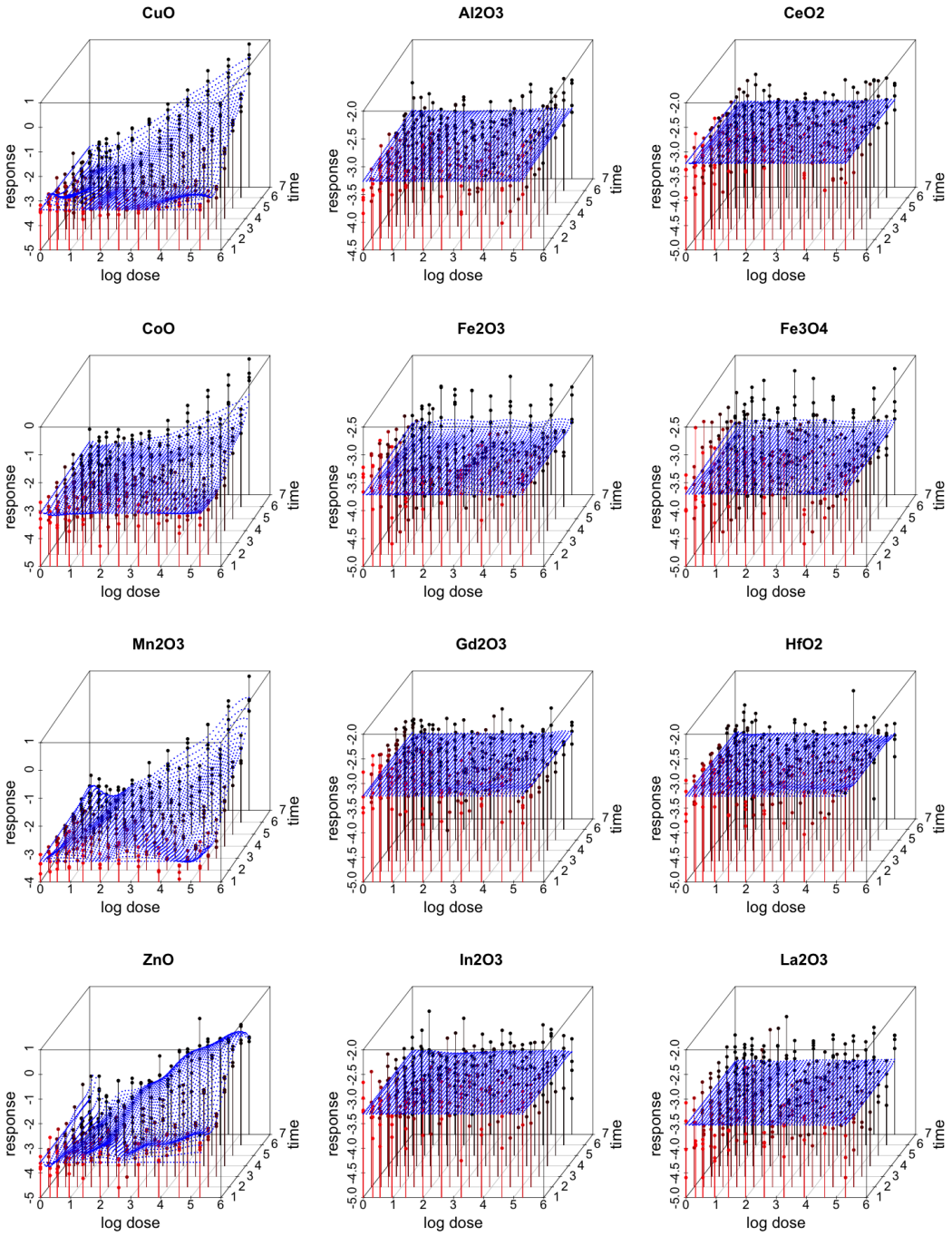

Figure 4.1: **Fitted response surfaces for CuO, Al_2O_3, CeO_2, CoO, Fe_2O_3, Fe_3O_4, Mn_2O_3, Gd_2O_3, HfO_2, ZnO, In_2O_3, and La_2O_3 ENMs.** Estimated smooth response surfaces, where the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

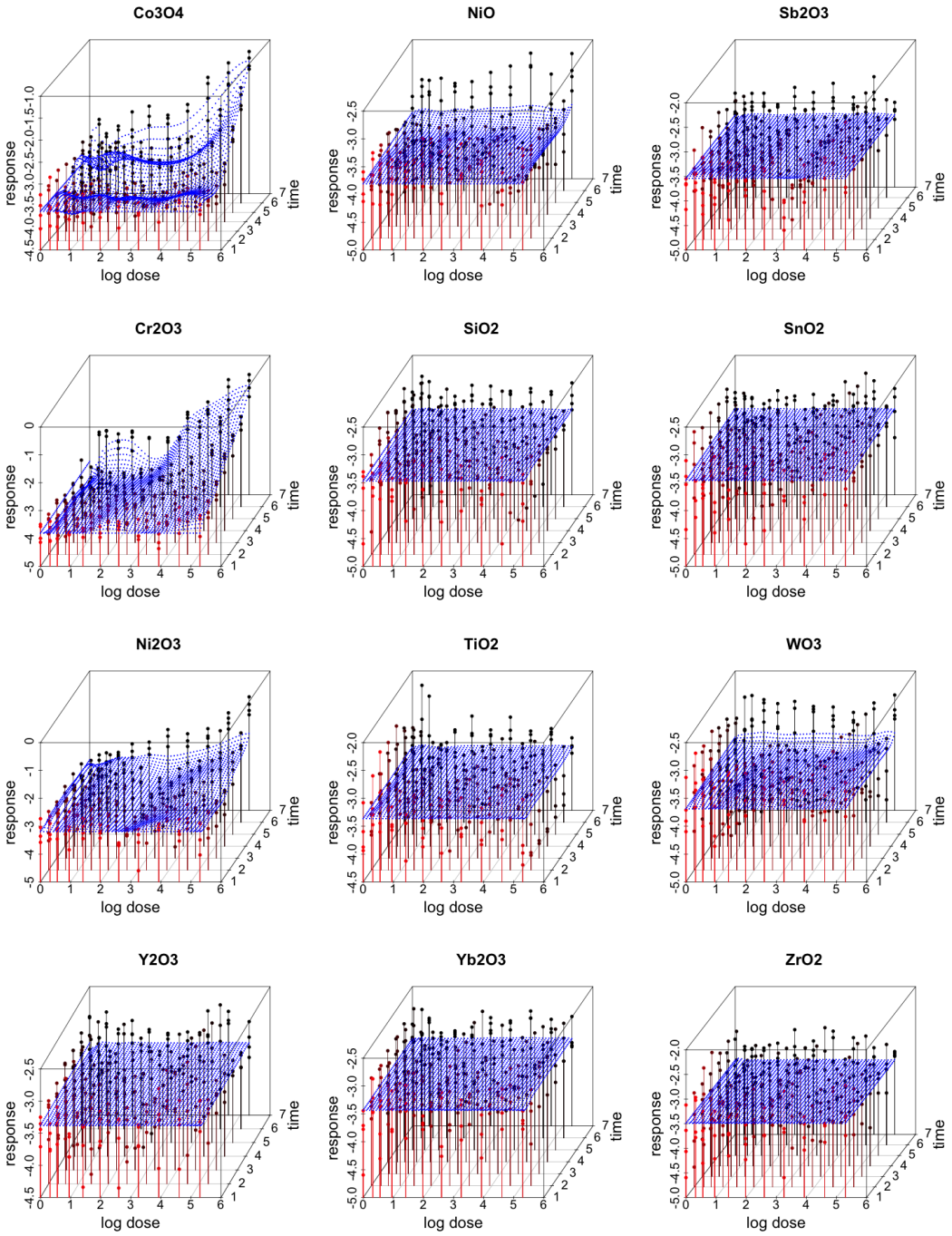

Figure 4.2: **Fitted response surfaces for Co_3O_4, NiO, Sb_2O_3, Cr_2O_3, SiO_2, SnO_2, Ni_2O_3, TiO_2, WO_3, Y_2O_3, Yb_2O_3, and ZrO_2 ENMs.** Estimated smooth response surfaces, where the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

We formally present model assessment and posterior predictive performance in Appendix C.2. To summarize, we assess goodness of fit using cross-validation and posterior predictive checks. The conditional predictive ordinate (cpo), as defined by Geisser (1980), is the predictive density of observation ℓ, given all other observations, and can be used as diagnostic tool for detecting observations with poor model fit. Large values of -log(CPO_k) indicate observations that are not being fitted well. The *top* panel of Figure C.1, provides a plot of -log($\hat{CPO}_i(d,t)$) for our final model. In general, low values of -log($CPO_i(d,t)$) indicate good model fit. The *middle* panel indicates that the largest values of -log($\hat{CPO}_i(d,t)$) tend to be observations with large exposure times, This is expected, as cell death is followed after sometime by the dissolution of cell nuclei, hindering the measurement of cellular responses. Next, we plotted the probability integral transform histogram for the entire model, as described by Gneiting et al. (2007). The plot is provided in the *bottom* panel of Figure C.1, and visual assessment indicates that it is close to uniformity, suggesting relatively good posterior predictive calibration. Additional summaries and diagnostic tools are detailed in Appendix C.2.

Figures 4.1 and 4.2 illustrate data and the posterior expected dose-time response for the 24 particles examined in this HTS study. Figure 4.4 provides a plot of the estimated probability of toxicity for each of the 24 ENMs.

Table 4.1 provides estimates of the posterior mean inclusion probabilities and posterior summaries for the averaged model, which includes all seven physicochemical properties described in Section 4.4.1. Table 4.2 provides posterior summaries for the final model, which includes conduction band energy and log(metal dissolution rate). Figure 4.5 also includes summary plots of the posterior probability of toxicity as a function of conduction band energy and metal dissolution. In particular, the *top* panel displays the posterior mean probability of toxicity, and associated 95% posterior intervals, for different values of the log(metal dissolution rate), given the conduction band energy is outside the range of the cellular redox potential. The *middle* panel displays the posterior mean probability of toxicity, and 95% posterior intervals, for conduction band energies inside and outside the range of the cellular

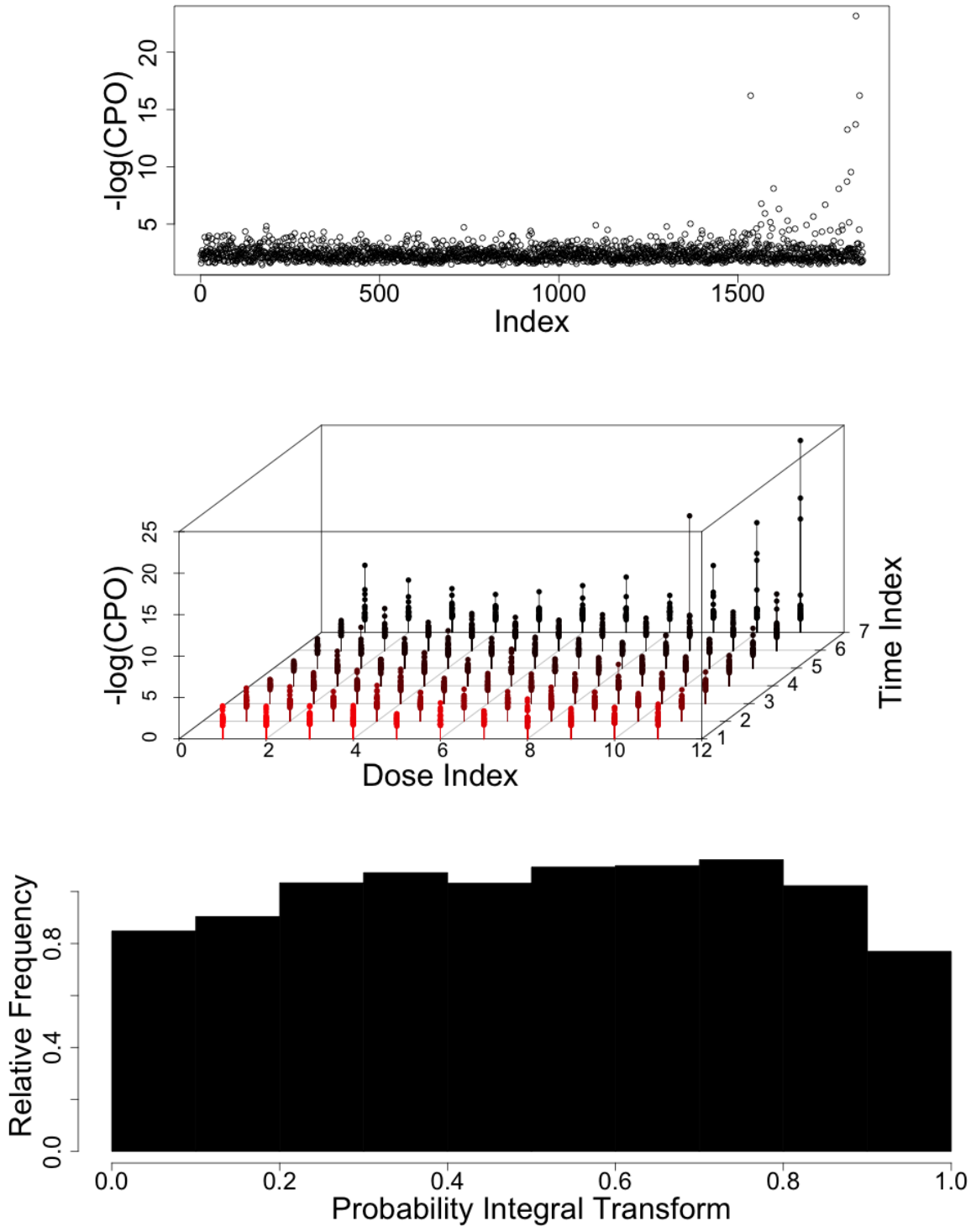

Figure 4.3: **Graphical model diagnostics.** *(*Top) Estimate of $-\log(cpo_i(d,t))$ for detecting observations with poor model fit. *(*Middle) Plot of -$log(cpo_i(d,t))$ as a function of dose and time, indicating any relationship between outlying observations and the administered dose or duration of exposure. *(*Bottom) Probability Integral Transform assessing empirical calibration of the posterior predictive distribution.

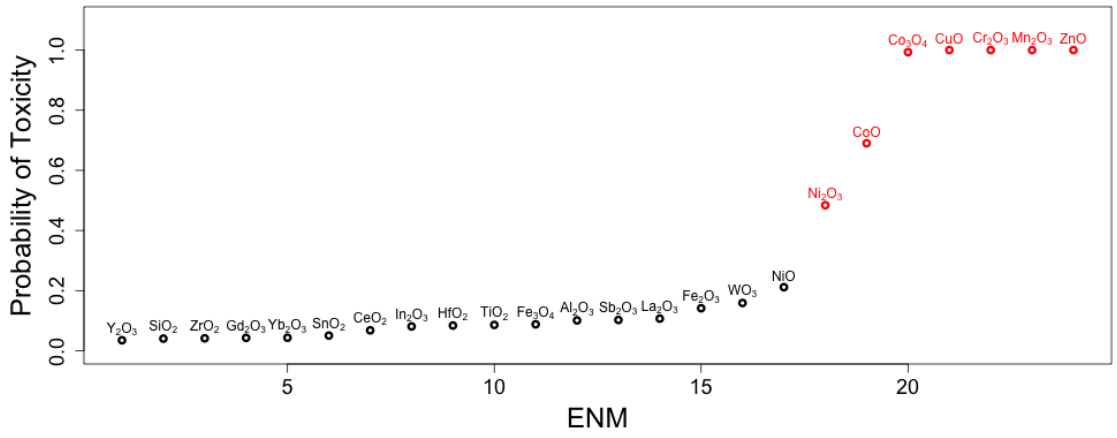

Figure 4.4: **Posterior mean probability of toxicity for each ENM.** Particles with a probability of toxicity greater then .5 are indicated in red.

redox potential, given no metal dissolution. The *bottom* panel of Figure 4.5 provides the posterior mean probability of toxicity, as a function of both conduction band energy and log(metal dissolution). Red colored regions indicate greater probability of toxicity, whereas blue colored regions indicate low probability of toxicity.

From Figures 4.1, 4.2, and 4.4, we see that CuO, Mn_2O_3, ZnO, Cr_2O_3, CoO, CO_3O_4, and Ni_2O_3 nanomaterials have a pronounced dose/time effect, as compared to the other 17 materials. These seven particles all have posterior mean probabilities of toxicity above .5, suggesting that they are capable of inducing cytotoxicity. In particular, conduction band energies inside the redox potential of the cell, predict high probability of cytotoxicity. Similarly, particles with a high metal dissolution rate, seem to to predict high probability of cytotoxicity. Our model confirms, from a statistical perspective, the conduction band energy and ENM metal dissolution properties, in the definition of cytotoxicity to BEAS-2B cells (Burello and Worth 2011, Zhang et al. 2012).

Parameter	Inclusion Probability	Posterior Summaries
Intercept	1.00	-0.50(-16.17,18.31)
Conduction Band Energy	0.77	1.62(0.00, 3.81)
log(Metal Dissolution Rate)<2.3	0.69	0.07(-0.19, 0.39)
log(Metal Dissolution Rate)>2.3	0.69	2.81(-0.19,12.90)
log(Primary Size)	0.31	0.00(-0.72, 0.71)
log(Crystal Structure (b(Å)))	0.35	-0.03(-2.11, 1.97)
log(Enthalpy of Formation ($\Delta H_{Me^{n+}}$))	0.38	-0.20(-2.38, 1.37)
log(Lattice Energy ($\Delta H_{lattice}$))	0.36	-0.18(-1.89, 2.10)

Table 4.1: **Posterior summaries for regression coefficients corresponding to the model which includes predictors for seven different ENM physicochemical properties.** Posterior mean inclusion probabilities as well as posterior mean and associated 95% posterior intervals are provided for the model regression coefficients.

Parameter	Posterior Summaries
Intercept	-1.22(-2.45,-0.20)
Conduction Band Energy	2.02(0.49, 3.75)
log(Metal Dissolution Rate)<2.3	0.11(-0.19, 0.41)
log(Metal Dissolution Rate)>2.3	3.52(-0.18, 8.49)

Table 4.2: **Posterior summaries for regression coefficients corresponding to the final model.** The final model includes model includes predictors for to conduction band energy and log(metal dissolution rate). Posterior mean estimates and associated 95% posterior intervals are provided for the model regression coefficients.

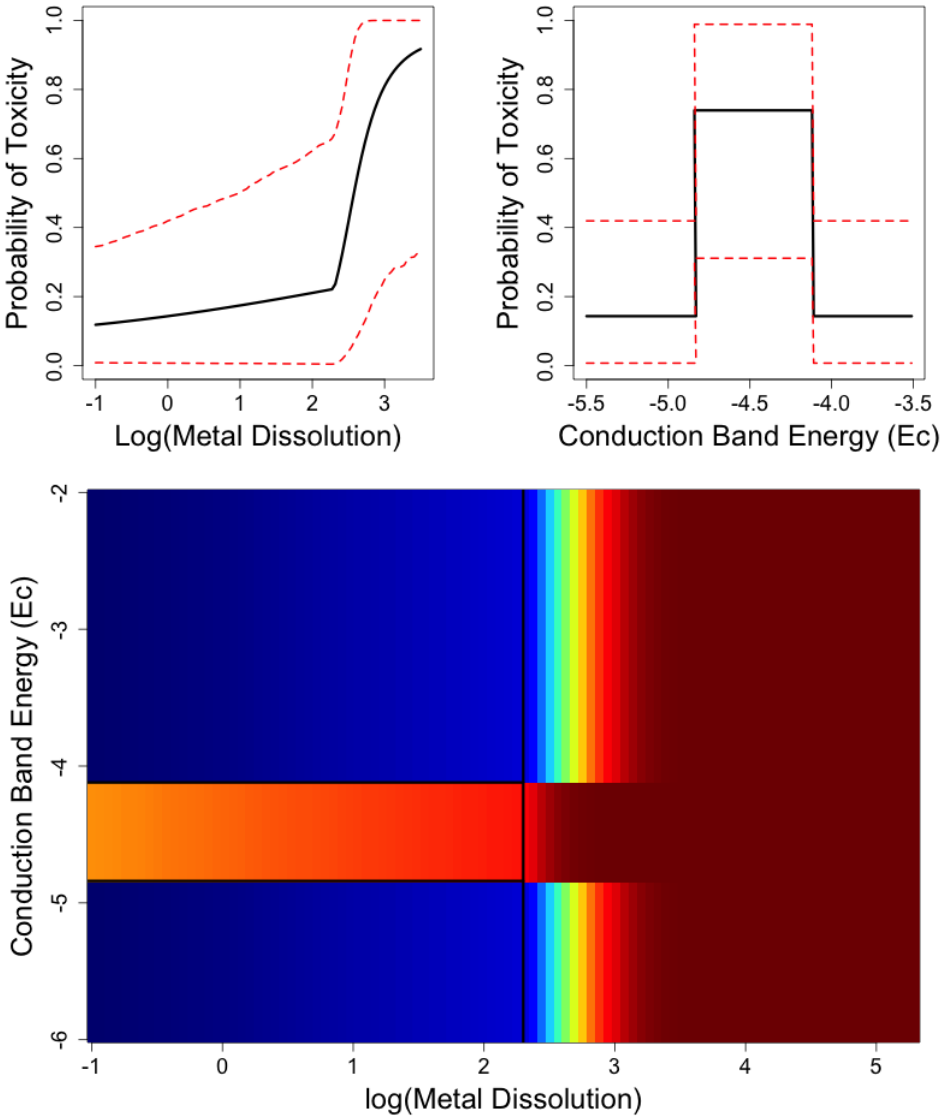

Figure 4.5: **Posterior summaries of the probability of toxicity as a function of conduction band energy and metal dissolution.** *(Top)* Posterior mean (black) and 95% posterior intervals (red) for the probability of toxicity as a function of log(metal dissolution rate), given conduction band energy outside the range of the cell redox potential. *(Middle)* Posterior mean (black) and 95% posterior intervals (red) for the probability of toxicity as a function of conduction band energy, given no metal dissolution. *(Bottom)* Posterior mean probability as a function of conduction band energy and log(metal dissolution rate). Red colored regions indicate greater probability of toxicity, whereas blue colored regions indicate low probability of toxicity.

4.5 Discussion

In this chapter, we propose a model that assesses how nanomaterial physicochemical properties explain toxicity *in-vitro*. The proposed methodology introduces a new measure of toxicity that is seamlessly integrated into any model structure, including a multi-dimensional model that accounts for dose and duration kinetics jointly using a flexible smooth surface fit. This measure is intuitively more appealing, as it is a measures the probability of toxicity, rather then single univariate summaries of the response surface trajectory. Furthermore, the indicator of toxicity is used to link ENM physicochemical properties to non-linear and multidimensional cytotoxicity profiles, while accounting for the uncertainty in the estimation of this summary. Measures such as the (ECα), benchmark doses, and the area under the curve, vary between experiments, and therefore do not easily allow for a comprehensive data-fusion exercise.

The methodology is also appropriate for limited data sets, as it includes data integration and a framework for advanced dimension reduction through variable selection. We account for the non-robust nature of the data by allowing for particle specific variance inflation, resulting in a t-distributed model for the error structure. Finally, the hierarchical framework is easily adapted to include a multivariate response, associated with a multiple set of cellular responses, which can thereby efficiently combine information and borrow strength across cellular injury patterns. One limitation of using variable selection in a settling where the sample size is extremely small, includes increased sensitivity of the results to the selection prior placed on the inclusion parameters.

An important extension to the work described in this chapter includes the inclusion of multiple cytotoxicity parameters, in the form of multivariate dependent observations. A reasonable dependence scheme can, for example, assume data to be dependent within outcome and particle, as well as between outcomes for the same particle. A further extension would involve a modeling framework that combines multiple data sources in a meta-analytic fashion. This technique is also easily adapted into the hierarchical modeling framework. In nanotoxicology, these extensions could provide an opportunity for addressing questions that

cannot be currently answered, due to limited sample sizes, and are therefore an important area for future work. The proposed model is only an initial step in an effort to develop a framework for predictive nanotoxicology. A further extension, toward the ultimate goal of developing a QSAR model, might include the use of functional regression to relate physico-chemical properties to cytotoxicity profiles using more sophisticated functional forms, as the relationship to cytotoxicity is often highly non-linear.

CHAPTER 5

Discussion

In this dissertation, we propose methodologies that provide a formal statistical framework for the analysis of the high dimensional data structures that are generated by new-generation high-throughput screening assays. In particular, we develop a foundation for toxicity assessment, prioritization of *in-vivo* testing of ENM, and predictive nanotoxicology.

In Chapter 2 we propose a probability model for the analysis of high-throughput cellular assays. We provide an analytic tool, that builds a balance between model complexity and model interpretability, as a method for assessing cytotoxicity. In particular, we developed a statistical framework for modeling dependent dose-response-surfaces over multivariate outcomes. We account for dose and duration kinetics jointly using a flexible model which does not compromise interpretability. We also account for the multivariate nature of the data using a hierarchical structure, and thereby efficiently combine information and borrow strength across cellular injury patterns. Finally, we account for the non-robust nature of the data by allowing for particle specific variance inflation, resulting in a t-distributed model for the error structure.

In Chapter 3 we discuss loss-function-based ranking methods and apply them to the analysis of HTS assays. We seek to provide a decision-making tool that we hope will be used for the prioritization of *in-vivo* animal testing of emerging nanomaterials. Specifically we present various loss function based ranking approaches, and apply them to the hazard ranking of nanomaterials. Moreover, we extend these methods to the aggregation of ranks across different sources of evidence, while allowing for the differential weighting of this evidence based on its reliability and importance in risk ranking. We account for the multivariate

nature of the data using a Bayesian hierarchical model, and coupled with a loss function, are thereby able to derive a rank estimate and a measure of its uncertainty.

Finally in Chapter 4, we propose a framework for modeling the relationship between ENM physicochemical properties and their observed cytotoxicity, as an initial step in the development of a foundation for predictive nanotoxicology. We introduce a new measure of toxicity that is seamlessly integrated into any model structure, including a multi-dimensional model that accounts for dose and duration kinetics jointly using a flexible smooth surface fit. Furthermore, the probability of toxicity is used to link ENM physicochemical properties to non-linear and multidimensional cytotoxicity profiles, while accounting for the uncertainty in the estimation of this summary. The methodology is also appropriate for limited data sets, as it includes data integration, and a formulation for dimension reduction through variable selection.

Important areas of future work include the development of statistical methods for the quantitative assessment of data reliability, more refined sampling models that accounts for the digital processing of measurements, and the further development of models for predictive nanotoxicology.

High-throughput cellular assays allow the measurement of possible toxic effects of nanomaterials at the cellular level. These measurements often translate into a multivariate response that summarizes different levels of cellular injury associated with the nanomaterials. The nanoparticle environment, particle aggregation, as well as other properties of the particles under study, are sometimes associated with aberrations in the florescent readings, leading to high measurement errors. In the context of high-throughput cellular assays it becomes important to provide quantitative tools for rigorous data quality assessment. Potential strategies could range from the detection of outliers in an unsupervised manner, to the integration of prior knowledge about known biological interactions with information coming from high-throughput scans, in the definition of quantitative models for the validation of measurements.

In this initial modeling effort, we choose to work with digital fluorescence measurements

that were used to define the percent of cells positive for the measured response. In order to use a t-distributed error structure, we normalize our response to unconstrain the support so that it can take on values between -∞ and ∞. In high-throughput toxicity studies, prolonged exposure to toxic nanomaterials leads to cellular death, and often, subsequent dissolution of the cell nuclei, leading to reduced reliability in cellular measurement. A more refined formulation of the model could retain the total cell count, along with a count of the total number of cells positive for the measured response. This would naturally lead to the definition of a sampling model that assigns less certainty to observations with reduced reliability. Retaining the original structure of the data would involve the definition of a generalized multivariate model, such that the outcome distribution can be described using binomial random quantities. This extension would require a substantial increase in computational complexity, with the possible need to consider numerical or analytical approximations, but it is clearly worthy of further consideration.

An important extension to the model described in Chapter 4 is the inclusion of multiple cytotoxicity parameters, in the form of multivariate dependent observations. A reasonable dependence scheme can, for example, assume data to be dependent within outcome and particle, as well as between outcomes for the same particle. We can also expand the model further by defining a more refined sampling model that allows for different measurement errors across exposure times, as by nature these measurements are taken at different times and consequently on different arrays. For example, if we let $y_{ijk}(d,t)$ denote a multivariate response corresponding to ENM i ($i = 1, ..., I$), cytotoxicity parameter j ($j = 1, ..., J$), and replicate k ($k = 1, ..., K$), at administered dose d in $\boldsymbol{d} = (d_1, ...d_D)'$ and time t in $\boldsymbol{t} = (t_1, ...t_T)'$. The observed response for particle i, outcome j, and replicate k can be modeled as:

$$y_{ijk}(d,t) = m_{ij}(d,t) + \epsilon_{ijk}(d,t). \tag{5.1}$$

Let a $J \times T$ matrix $\boldsymbol{E}_i(d = d_\ell) = \{\epsilon_i(d = d_\ell) : \ i = 1, \ldots, I, \ell = 1, ...L\}$, and similarly let $\boldsymbol{Y}_i$ denote the complete set of response values for particle i, then the distribution of Y_i can be modeled in terms of the error term $\boldsymbol{E_i}(d = d_\ell)$, as a matrix normal random variable as

follows:

$$\boldsymbol{E_i}(d = d_\ell) \sim MN\left(0, \underset{J\times J}{\boldsymbol{\Sigma}_1}, \frac{1}{\tau_i}\underset{T\times T}{\boldsymbol{\Sigma}_2}\right). \tag{5.2}$$

Under this framework, we can borrow strength across all ENM by assuming the error variance is the same, but retain robustness in the model by allowing ENM-specific departures from normality. Furthermore, the measurement error $\boldsymbol{\Sigma}_1$ accounts for heterogeneity in the cytotoxicity outcomes, while allowing for outcomes to be correlated. Similarly, the measurement error $\boldsymbol{\Sigma}_2$ accounts for possible variability in the measurement error across exposure times, while allowing for correlation between these measurements.

The dose-response-surface $m_{ij}(d,t)$ can be modeled as a non-parametric smooth fit, while allowing for a quantitative measure of toxicity γ_{ij} . Let $(\alpha_{ij}, \boldsymbol{\beta}'_{ij}, \gamma_{ij})'$ be a parameter vector indexing the response-surface $m_{ij}(d,t)$. We define

$$m_{ij}(d,t) \quad = \begin{cases} \alpha_{ij} & \text{if } \gamma_{ij} = 0 \\ \alpha_{ij} + f_{ij}(d,t;\boldsymbol{\beta}_{ij}) & \text{if } \gamma_{ij} = 1, \end{cases} \tag{5.3}$$

where $f_{ij}(d,t;\boldsymbol{\beta}_{ij})$ is a smooth surface modeling the effect of dose d and time t, for each ENM i and cytotoxicity parameter j, and γ_{ij} is a latent indicator variable that can be used as a measure of toxicity for each particle i and cytotoxicity parameter j. To ensure likelihood identifiability we must require, without loss of generality, that $f_{ij}(d = 0, t = 0; \boldsymbol{\beta}_{ij}) = 0$. Similarly to the model described in Chapter 4, the response-surface $f_{ij}(d,t;\boldsymbol{\beta}_{ij})$ can be modeled using two-dimensional P-splines, as described by Lang and Brezger (2006). We can complete the model by defining appropriate priors for $\boldsymbol{\beta}_{ij}$, α_{ij}, γ_{ij}, and all remaining precision parameters.

Again, we can directly extend the model described in Chapter 4, and define a model linking cytotoxicity profiles to measured ENM-specific physicochemical properties as follows. For each ENM (i=1,...I), we would like to model a set of P physicochemical properties as a vector of covariates $\boldsymbol{x_i} = (x_{i0}, ...x_{ip}, ..., x_{i(P)})'$. We relate our cytotoxicity vector γ_{ij} to our covariate

vector $\boldsymbol{x_i}$ through a regression at the ENM by cytotoxicity parameter level, such that

$$g(\gamma_{ij} \mid \boldsymbol{x_i}) \quad = \quad \boldsymbol{x_i'\theta} \tag{5.4}$$

where $\boldsymbol{\theta} = (\theta_0, ...\theta_p, ...\theta_P)$ is our vector of regression coefficients and g is a standard link function. We may proceed as in the univariate case, by introducing a continuous latent variable Z_i and modeling g as probit model, allowing the construction of a Gibbs sampling algorithm (Albert and Chib 1993 and Johnson and Albert 1999). Again, in the case where P is large, we may proceed as in Chapter 4 by allowing for dimension reduction in the form of variable selection. One limitation of using variable selection in a settling where the sample size is extremely small, is the increased sensitivity of the results to the chosen selection prior.

A further extension of the model described above would involve a modeling framework that combines multiple experiments or data sources in a meta-analytic fashion. This technique is easily adapted into the hierarchical modeling framework, with information shared between studies via the structuring of one extra level in the hierarchy. In nanotoxicology, these extensions could provide an opportunity for addressing questions that cannot be currently answered due to limited sample sizes, and are therefore an important area of future work.

Finally, the proposed model is only an initial step in an effort to develop a framework for predictive nanotoxicology. A further extension, toward the ultimate goal of developing a QSAR model, includes the further development of the regression model relating the probability of toxicity to ENM-specific physicochemical properties. Due to the limited samples size in the case-study described in Chapter 4, the modeling of functional covariates was kept relatively simple. As the relationship between cytotoxicity and ENM structural and chemical properties is often non-linear, one approach might include the use of functional regression to relate physicochemical properties to cytotoxicity profiles, using more sophisticated functional forms. For example, nanoparticle size measurements are most often measured as a distribution of sizes rather then a single value. In this case it might be important to relate the observed cytotoxicity of a material to its measured size distribution. This is naturally defined as an extension of the covariate model described in Section 4.2.2 of Chapter 4, as

follows:

$$g(\gamma_i \mid \boldsymbol{x_i}, c(s)) \quad = \quad \boldsymbol{x_i'\theta} + \int c_i(s)\zeta(s)\ \mathrm{d}s, \tag{5.5}$$

where $\boldsymbol{\theta} = (\theta_0, ...\theta_p, ...\theta_P)$ is our usual vector of regression coefficients, s is a continuous index for size, $c(s)$ is the covariate function, and $\zeta(s)$ is a functional parameter, which models how variability in the size distribution explains variability in the probability of toxicity. Further details on this type of model can be found in Ramsey and Silverman (2005). One way to model the functional component of the expression above might be through the use of generalized additive regression based on Bayesian P-splines, as detailed in Lang and Brezger (2004). Modeling the relationship between cytotoxicity profiles and ENM-specific physicochemical properties using functional covariates would require substantial increase in computational complexity, but as the library of available data of ENMs rapidly expands, it is clearly worthy of further exploration.

APPENDIX A

Appendix A: Toxicity Profiling of ENMs

A.1 Full Conditional Distributions

In this appendix we describe some of the full conditional distributions for the model described in Chapter 2. Let $y_{ijk}(d,t)$ denote a multivariate response corresponding to ENM i ($i = 1,...,I$), cytotoxicity parameter j ($j = 1,...,J$), and replicate k ($k = 1,...,K$) at some dose d in $d = (d_1,...d_D)'$ and some time t in $\boldsymbol{t} = (t_1,...t_T)'$.

Also, let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\alpha_o}, \boldsymbol{\beta_o}, \boldsymbol{\gamma_o}, \boldsymbol{\tau}, \boldsymbol{\sigma_\epsilon}, \boldsymbol{\sigma_\alpha}, \boldsymbol{\sigma_\beta}, \boldsymbol{\sigma_\gamma})$ denote the full parameter vector and, let $\boldsymbol{\theta_{\backslash\delta}}$ denote the vector containing all components of $\boldsymbol{\theta}$ except for some parameter δ in $\boldsymbol{\theta}$. Moreover, we denote $\boldsymbol{Y_{ij}}$ as the complete set of response values for particle i and outcome j. Using the notation above we define the full conditional distributions for all available parameters as follows.

A.1.1: Full conditional distributions for the random effect parameters α_{ij}, $\boldsymbol{\beta_{ij}}$, $\boldsymbol{\gamma_{ij}}$, and $\boldsymbol{\delta_{ij}}$.

$$\alpha_{ij} \mid \boldsymbol{Y_{ij}}, \boldsymbol{\theta}_{\backslash\alpha_{ij}} \sim N\left(\frac{\sum_{d,t,k} \tilde{y}_{\alpha_{ijk}}(d,t)}{\frac{\sigma^2_{\epsilon_j}/\tau_i}{\sigma^2_{\epsilon_j}/\tau_i} + \frac{\alpha_{o_i}}{\sigma_{\alpha_i}}}, \frac{1}{\frac{D\times T\times K}{\sigma^2_{\epsilon_j}/\tau_i} + \frac{1}{\sigma_{\alpha_i}}} \right),$$

where

$$\tilde{y}_{\alpha_{ijk}}(d,t) = \begin{cases} y_{ijk}(d,t) - \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} - \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} & \text{if } \rho_{ij} = 0 \\ y_{ijk}(d,t) - \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} - \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} - \mathcal{B}(d*t,\boldsymbol{\chi}_{ij})'\boldsymbol{\delta}_{ij} & \text{if } \rho_{ij} = 1. \end{cases}$$

$$\boldsymbol{\beta_{ij}} \mid \boldsymbol{Y_{ij}}, \boldsymbol{\theta}_{\backslash \boldsymbol{\beta_{ij}}} \sim N_4(M_\beta, V_\beta),$$

where

$$M_\beta = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_i}} + \sum_{d,t,k} \frac{\mathcal{B}(d,\boldsymbol{\phi}_{ij})\mathcal{B}(d,\boldsymbol{\phi}_{ij})'}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_i}} \boldsymbol{\beta_{o_i}} + \sum_{d,t,k} \frac{\mathcal{B}(d,\boldsymbol{\phi}_{ij})\tilde{y}_{\beta_{ijk}}(d,t)}{\sigma^2_{\epsilon_j}/\tau_i} \right),$$

$$V_\beta = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_i}} + \sum_{d,t,k} \frac{\mathcal{B}(d,\boldsymbol{\phi}_{ij})\mathcal{B}(d,\boldsymbol{\phi}_{ij})'}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1}$$

$$\tilde{y}_{\beta_{ijk}}(d,t) = \begin{cases} y_{ijk}(d,t) - \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} - \alpha_{ij} & \text{if } \rho_{ij} = 0 \\ y_{ijk}(d,t) - \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} - \mathcal{B}(d*t,\boldsymbol{\chi}_{ij})'\boldsymbol{\delta}_{ij} - \alpha_{ij} & \text{if } \rho_{ij} = 1. \end{cases}$$

$$\boldsymbol{\gamma_{ij}} \mid \boldsymbol{Y_{ij}}, \boldsymbol{\theta}_{\backslash \boldsymbol{\gamma_{ij}}} \sim N_4(M_\gamma, V_\gamma),$$

where

$$M_\gamma = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} + \sum_{d,t,k} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ij})\mathcal{B}(t,\boldsymbol{\psi}_{ij})'}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} \boldsymbol{\gamma_{o_i}} + \sum_{d,t,k} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ij})\tilde{y}_{\gamma_{ijk}}(d,t)}{\sigma^2_{\epsilon_j}/\tau_i} \right),$$

$$V_\gamma = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} + \sum_{d,t,k} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ij})\mathcal{B}(t,\boldsymbol{\psi}_{ij})'}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1},$$

$$\tilde{y}_{\gamma_{ijk}}(d,t) = \begin{cases} y_{ijk}(d,t) - \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} - \alpha_{ij} & \text{if } \rho_{ij} = 0 \\ y_{ijk}(d,t) - \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} - \mathcal{B}(d*t,\boldsymbol{\chi}_{ij})'\boldsymbol{\delta}_{ij} - \alpha_{ij} & \text{if } \rho_{ij} = 1. \end{cases}$$

$$\boldsymbol{\delta_{ij}} \mid \boldsymbol{Y_{ij}}, \boldsymbol{\theta}_{\backslash \boldsymbol{\delta_{ij}}} \sim N_4(M_\delta, V_\delta),$$

where

$$M_\delta = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\delta_i}} + \sum_{d,t,k} \frac{\mathcal{B}(t,\boldsymbol{\chi}_{ij})\mathcal{B}(d*t,\boldsymbol{\chi}_{ij})'}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\delta_i}} \boldsymbol{m_{\delta_{ij}}} + \sum_{d,t,k} \frac{\mathcal{B}(d*t,\boldsymbol{\chi}_{ij})\tilde{y}_{\delta_{ijk}}(d,t)}{\sigma^2_{\epsilon_j}/\tau_i} \right),$$

$$V_\delta = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} + \sum_{d,t,k} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ij})\mathcal{B}(t,\boldsymbol{\psi}_{ij})'}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1},$$

$$\tilde{y}_{\delta_{ijk}}(d,t) = y_{ijk}(d,t) - \mathcal{B}(d,\boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ij} - \mathcal{B}(t,\boldsymbol{\psi}_{ij})'\boldsymbol{\gamma}_{ij} - \alpha_{ij}.$$

A.1.2: Full conditional distributions for ENM level parameters α_{o_i}, β_{o_i}, and γ_{o_i}.

$$\alpha_{o_i} \mid \boldsymbol{\theta}_{\backslash \alpha_{ij}} \sim N\left(\frac{\frac{\sum_{j=1}^{J} \alpha_{ij}}{\sigma^2_{\alpha_i}} + \frac{m_{\alpha_i}}{s^2_{\alpha_i}}}{\frac{J}{\sigma^2_{\alpha_i}} + \frac{1}{s^2_{\alpha_i}}}, \frac{1}{\frac{J}{\sigma^2_{\alpha_i}} + \frac{1}{s^2_{\alpha_i}}}\right),$$

$$\boldsymbol{\beta_{o_i}} \mid \boldsymbol{\theta}_{\backslash \boldsymbol{\beta_{ij}}} \sim N_4\left(\left(J\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_i}} + \boldsymbol{v}^{-1}_{\boldsymbol{\beta_i}}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_i}} \textstyle\sum_{j=1}^{J} \boldsymbol{\beta_{ij}} + \boldsymbol{v}^{-1}_{\boldsymbol{\beta_i}} \boldsymbol{m}_{\boldsymbol{\beta_i}}\right), \left(J\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_i}} + \boldsymbol{v}^{-1}_{\boldsymbol{\beta_i}}\right)^{-1}\right),$$

$$\boldsymbol{\gamma_{o_i}} \mid \boldsymbol{\theta}_{\backslash \boldsymbol{\gamma_{ij}}} \sim N_4\left(\left(J\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} + \boldsymbol{v}^{-1}_{\boldsymbol{\gamma_i}}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} \textstyle\sum_{j=1}^{J} \boldsymbol{\gamma_{ij}} + \boldsymbol{v}^{-1}_{\boldsymbol{\gamma_i}} \boldsymbol{m}_{\boldsymbol{\gamma_i}}\right), \left(J\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_i}} + \boldsymbol{v}^{-1}_{\boldsymbol{\gamma_i}}\right)^{-1}\right).$$

A.1.3: Full conditional distributions for the error variance parameter $\sigma^2_{\epsilon_j}$ and variance inflation parameter τ_i.

$$1/\sigma^2_{\epsilon_j} \mid \boldsymbol{Y_{ij}}, \boldsymbol{\theta}_{\backslash \sigma_{\epsilon_j}} \sim Gamma\left(a_{\epsilon_j} + \tfrac{I \times D \times T \times K}{2}, \tfrac{1}{2} \sum_{d,t,k,i} (y_{ijk}(d,t) - m_{ij}(d,t))^2 \tau_i + b_{\epsilon_j}\right),$$

where

$$m_{ij}(d,t) = \begin{cases} \mathcal{B}(d, \boldsymbol{\phi}_{ij})' \boldsymbol{\beta}_{ij} + \mathcal{B}(t, \boldsymbol{\psi}_{ij})' \boldsymbol{\gamma}_{ij} + \alpha_{ij} & \text{if } \rho_{ij} = 0 \\ \mathcal{B}(d, \boldsymbol{\phi}_{ij})' \boldsymbol{\beta}_{ij} + \mathcal{B}(t, \boldsymbol{\psi}_{ij})' \boldsymbol{\gamma}_{ij} + \mathcal{B}(d * t, \boldsymbol{\chi}_{ij})' \boldsymbol{\gamma}_{ij} + \alpha_{ij} & \text{if } \rho_{ij} = 1. \end{cases}$$

$$\tau_i \mid \boldsymbol{Y_{ij}}, \boldsymbol{\theta}_{\backslash \tau_i} \sim Gamma\left(\tfrac{\nu}{2} + \tfrac{J \times K \times D \times T}{2}, \tfrac{1}{2} \sum_{d,t,k,j} \tfrac{(y_{ijk}(d,t) - m_{ij}(d,t))^2}{\sigma^2_{\epsilon_j}} + \tfrac{\nu}{2}\right),$$

where

$$m_{ij}(d,t) = \begin{cases} \mathcal{B}(d, \boldsymbol{\phi}_{ij})' \boldsymbol{\beta}_{ij} + \mathcal{B}(t, \boldsymbol{\psi}_{ij})' \boldsymbol{\gamma}_{ij} + \alpha_{ij} & \text{if } \rho_{ij} = 0 \\ \mathcal{B}(d, \boldsymbol{\phi}_{ij})' \boldsymbol{\beta}_{ij} + \mathcal{B}(t, \boldsymbol{\psi}_{ij})' \boldsymbol{\gamma}_{ij} + \mathcal{B}(d * t, \boldsymbol{\chi}_{ij})' \boldsymbol{\gamma}_{ij} + \alpha_{ij} & \text{if } \rho_{ij} = 1. \end{cases}$$

A.1.4: Full conditional distributions for other variance parameters

$$1/\sigma^2_{\alpha_i} \mid \boldsymbol{\theta}_{\backslash \sigma_{\alpha_i}} \sim Gamma\left(a_{\alpha i} + \tfrac{J}{2}, b_{\alpha_i} + \tfrac{1}{2} \textstyle\sum_{j=1}^{J} (\alpha_{ij} - \alpha_{o_i})^2\right),$$

$$1/\sigma^2_{\beta_{i\ell}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\beta_{i\ell}}} \sim Gamma\left(a_{\beta_i} + \tfrac{J}{2}, b_{\beta_i} + \tfrac{1}{2} \textstyle\sum_{j=1}^{J} (\beta_{ij\ell} - \beta_{o_i \ell})^2\right),$$

$$1/\sigma^2_{\gamma_{i\ell}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\gamma_{i\ell}}} \sim Gamma\left(a_{\gamma_i} + \tfrac{J}{2}, b_{\gamma_i} + \tfrac{1}{2} \textstyle\sum_{j=1}^{J} (\gamma_{ij\ell} - \gamma_{o_i \ell})^2\right),$$

where $\ell = 1, .., 4$.

A.2 Simulation Study: Assessing Model Fit

To assess estimation of the model presented in Chapter 2, we simulated four sets, ($i = 1, ..., 4$), of four independent surfaces, ($j = 1, ..., 4$), each evaluated at ten doses ($d \in [0, D]$) and seven times of exposure ($t \in [0, T]$). The dose and time kinetics were simulated in an additive fashion from various parametric functions. In addition for each set of response-surfaces, we simulated a dose*time interaction function for two of the four surfaces. If we let $y_{ijk}(d, t)$ denote a multivariate response corresponding to a set of surfaces i ($i = 1, ..., 4$), outcome j ($j = 1, ..., 4$), and replicate k ($k = 1, ..., 4$), at dose $d \in [0, D]$ and time $t \in [0, T]$, then each simulated surface can be described by

$$y_{ijk}(d, t) = s_{ij}(d, t) + \xi_{ijk}(d, t). \tag{A.1}$$

Here $s_{ij}(d, t) = \alpha_{ij} + r_{ij}(d) + u_{ij}(t) + w_{ij}(dt)I(r_{ij} = 1)$ represents a smooth dose-response-surface and r_{ij} an indicator that determines whether or not a dose-time interaction was simulated. The functions $r_{ij}(d)$, $u_{ij}(t)$, and $w_{ij}(dt)$ were simulated from various parametric functions such as log-logistic, weibull, polynomial, and cubic spline models. These parametric functions were used to construct both canonical and non-canonical profiles that are still reasonably interpretable in a toxicity framework. Note that the simulated response curves do not follow our model which uses a B-spline representation with two random interior knots. Dose-response surfaces were also simulated with varying levels of noise, across each outcome j. This is as expected in high-throughput screening studies, where different assay systems are able to capture measurements with varying levels of precision. More specifically, $\xi_{ijk}(d, t) \sim N(0, \sigma_{\xi_j})$, where $\sigma_{\xi_j} = .6, .8, 1, 1.2$ for $j = 1, ..., 4$ respectively.

The model described in the Chapter 2 was fit to the simulated surfaces. We placed relatively diffuse priors on all parameters. More precisely, we placed $Gamma(.01, .01)$ priors on the $1/\sigma_{\epsilon_j}$ parameters, $Gamma(1, .1)$ priors on all remaining precision parameters, and a discrete uniform prior on 1,2,4,8,16, and 32, for the degrees of freedom parameter ν. The population level parameters α_{o_i} are modeled as $N(0, 10)$ and the $\boldsymbol{\beta_{o_i}}$ and $\boldsymbol{\gamma_{o_i}}$ parameters as truncated $N_4(\mathbf{0}, 10\boldsymbol{I})$. We fixed $\beta_{ij2} = 0$, $\gamma_{ij2} = 0$, and $\delta_{ij2} = 0$, assuming no effect before

the first change-point. Note that this constraint can be changed to $\beta_{ij2} \leq 0$, $\gamma_{ij2} \leq 0$ and $\delta_{ij2} \leq 0$, when assuming a tonic effect before the first change-point. We also placed a diffuse $N_4(\mathbf{0}, 100\boldsymbol{I})$ prior on our dose-time interaction amplitude parameters $\boldsymbol{\delta}_{ij}$. Finally, in order to assess the sensitivity of the model results to our choice of prior parameters, we specified increasingly informative priors on our population level change-point parameters $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$, and on our dose-time interaction change-point parameters $\boldsymbol{\chi}_{ij}$. Our inferences are based on 20,000 MCMC samples, after discarding a conservative 60,000 iterations for burn-in.

Figure A.1 gives examples of increasingly informative prior distributions on our dose-response change-point parameters ϕ_{ij1} and ϕ_{ij2}. We impose a right-skewed prior distribution on ϕ_{ij1}, which favors (a-priori) conservative values for the location of the first change-point parameter, and allow a relatively diffuse prior distribution for our second change-point parameter ϕ_{ij2}. Similar examples can be constructed for our duration-response change-points ψ_{ij1} and ψ_{ij2}, and our dose-time interaction change-points χ_{ij1} and χ_{ij2}.

Figures A.2 through A.5 provide results from our simulation study using increasingly informative priors on our population level change-point parameters $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\psi_i}}$, and our dose-time interaction change-point parameters $\boldsymbol{\chi}_{ij}$. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves are superimposed with expected dose and duration-response curves, and point-wise 95% posterior intervals estimated from our model. Prior 1 (blue) is the least informative, and corresponds to a set of relatively diffuse $Gamma(2, 1)$ and $Gamma(3, 1)$ priors on $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$, and a diffuse $B_2(2, 3, 1, 1, DT)$ prior on $\boldsymbol{\chi}_{ij}$. Prior 2 is a moderately informative $Gamma(2.5, 1)$ and $Gamma(7, 1.5)$ prior on $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$, and a $B_2(2.5, 4.7, 1, 1, DT)$ prior on $\boldsymbol{\chi}_{ij}$. Finally prior 3 is the most informative, and corresponds to a $Gamma(2, 1)$ and $Gamma(7, 1.5)$ prior on $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$ and $\boldsymbol{\lambda}_{\boldsymbol{\phi_i}}$, and a $B_2(2, 4.7, 1, 1, DT)$ prior on $\boldsymbol{\chi}_{ij}$. The *left panel* shows fitted dose-response-surfaces, spanning dose and time, for our least informative prior. Table A.1 also provides posterior mean estimates of ρ_{ij}, which are the expected inclusion probabilities of the dose-time interaction function, after adjusting for multiplicity, and can be used to test for a dose-time interaction.

The simulation results show that the posterior mean response trajectories are able to adequately estimate the true dose and duration-response curves, while capturing the most

important features. In particular, our model is able to appropriately model both canonical and non-canonical dose-response trajectories, while providing interpretable risk assessment parameters. It is also evident that our model is relatively insensitive to the choice of prior parameters for our change-points. Priors 1-3 are, in most cases, indistinguishable, especially in those cases where the data sufficiently provide information on the location of the change-points. Finally, although we do not include a formal simulation study to determine the false positive and negative rate for correctly identifying an interaction function using our model, we do notice that in all the simulated cases the model is able to correctly distinguish between data simulated with and without a dose*time interaction function.

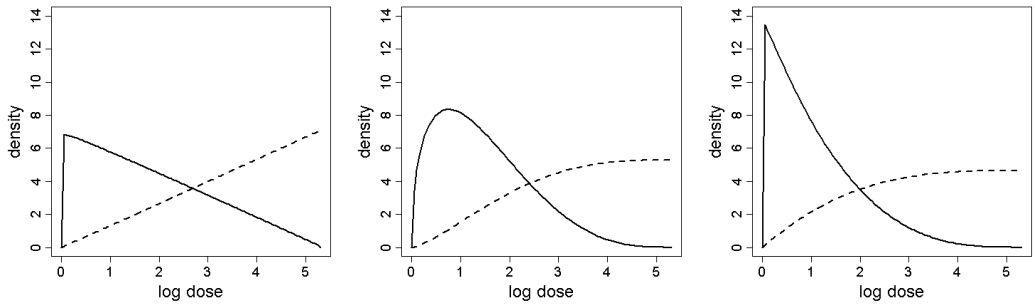

Figure A.1: **Marginal prior distributions on the change-point parameters.** Examples of increasingly informative prior distributions on the change-point parameters ϕ_{ij1}, ϕ_{ij2}. We favor (a-priori) the choice of conservative values for the location of the first change-point *(solid line)* and a relatively diffuse prior for our second change-point *(dotted line)*.

Prior	ENM	Outcome 1	Outcome 2	Outcome 3	Outcome 4
Prior1	1	0.01	0.99	0.99	0.00
	2	0.00	0.00	0.99	0.99
	3	0.99	0.00	0.99	0.00
	4	0.99	0.99	0.01	0.00
Prior2	1	0.00	0.98	0.99	0.00
	2	0.00	0.00	0.99	0.99
	3	0.99	0.00	0.99	0.00
	4	0.99	0.99	0.00	0.00
Prior3	1	0.00	0.99	0.99	0.00
	2	0.00	0.01	0.99	0.99
	3	0.99	0.00	0.99	0.11
	4	0.99	0.99	0.00	0.00

Table A.1: **Expected inclusion probabilities of the dose-time interaction function using priors 1-3.** Posterior mean estimates of ρ_{ij} describe the expected inclusion probabilities of the dose-time interaction function, and can be used to test for an interaction.

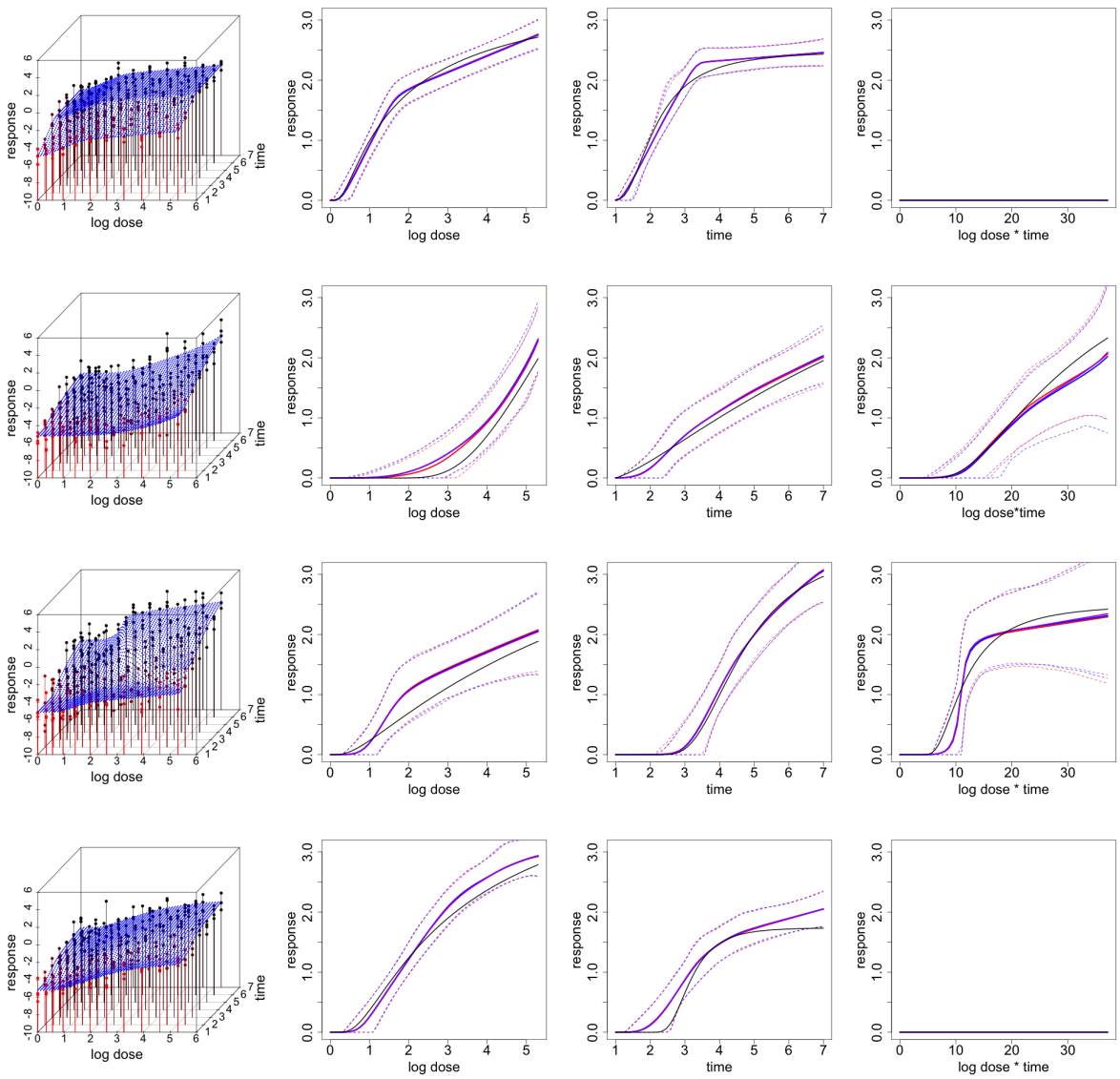

Figure A.2: **Simulation study to assess model fit and prior sensitivity (weibull).** (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the least informative, prior 2 (red) is moderately informative and priors 3 (purple) is the most informative. (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data.

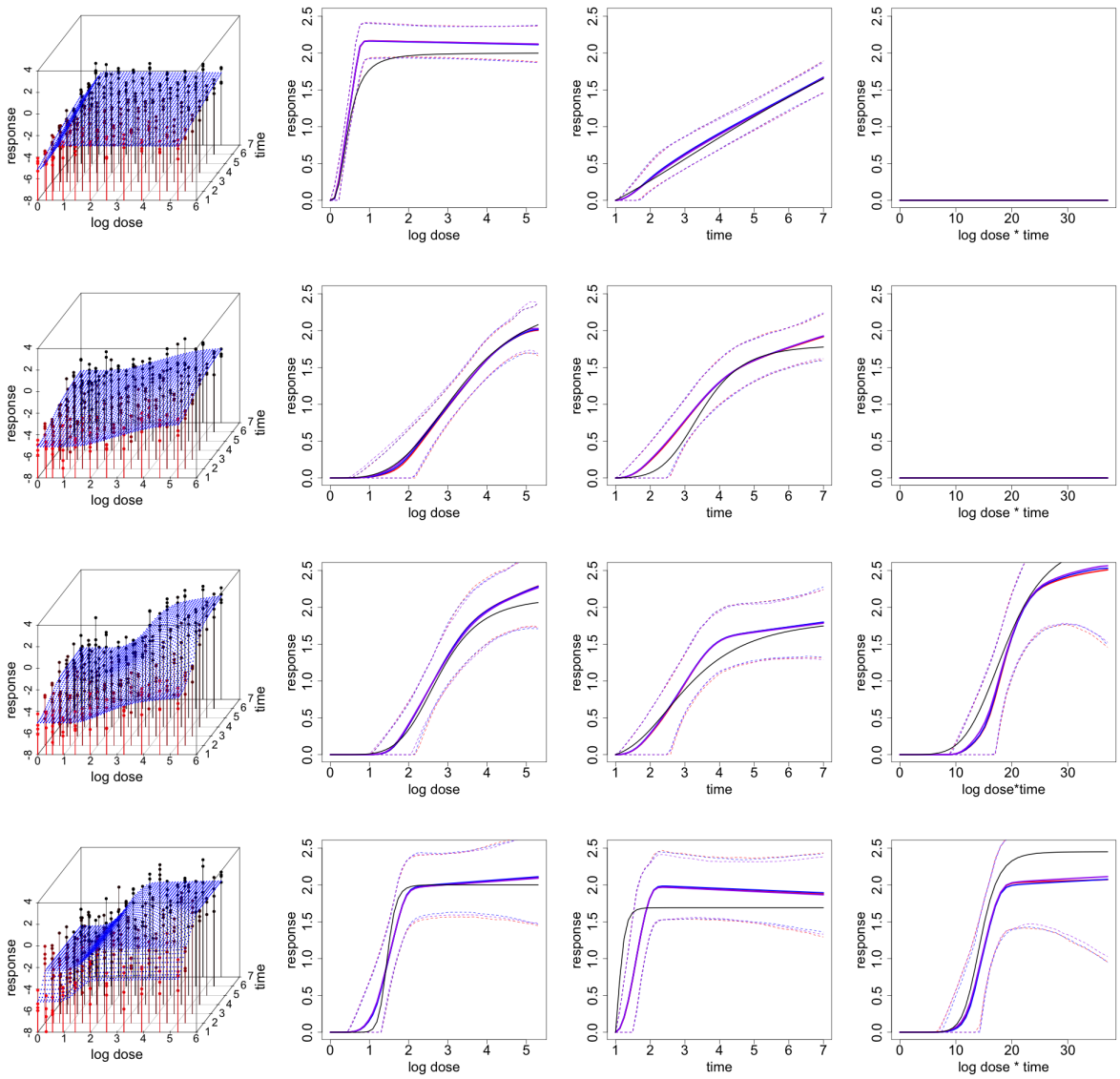

Figure A.3: **Simulation study to assess model fit and prior sensitivity (log-logistic).** (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the least informative, prior 2 (red) is moderately informative and priors 3 (purple) is the most informative. (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data.

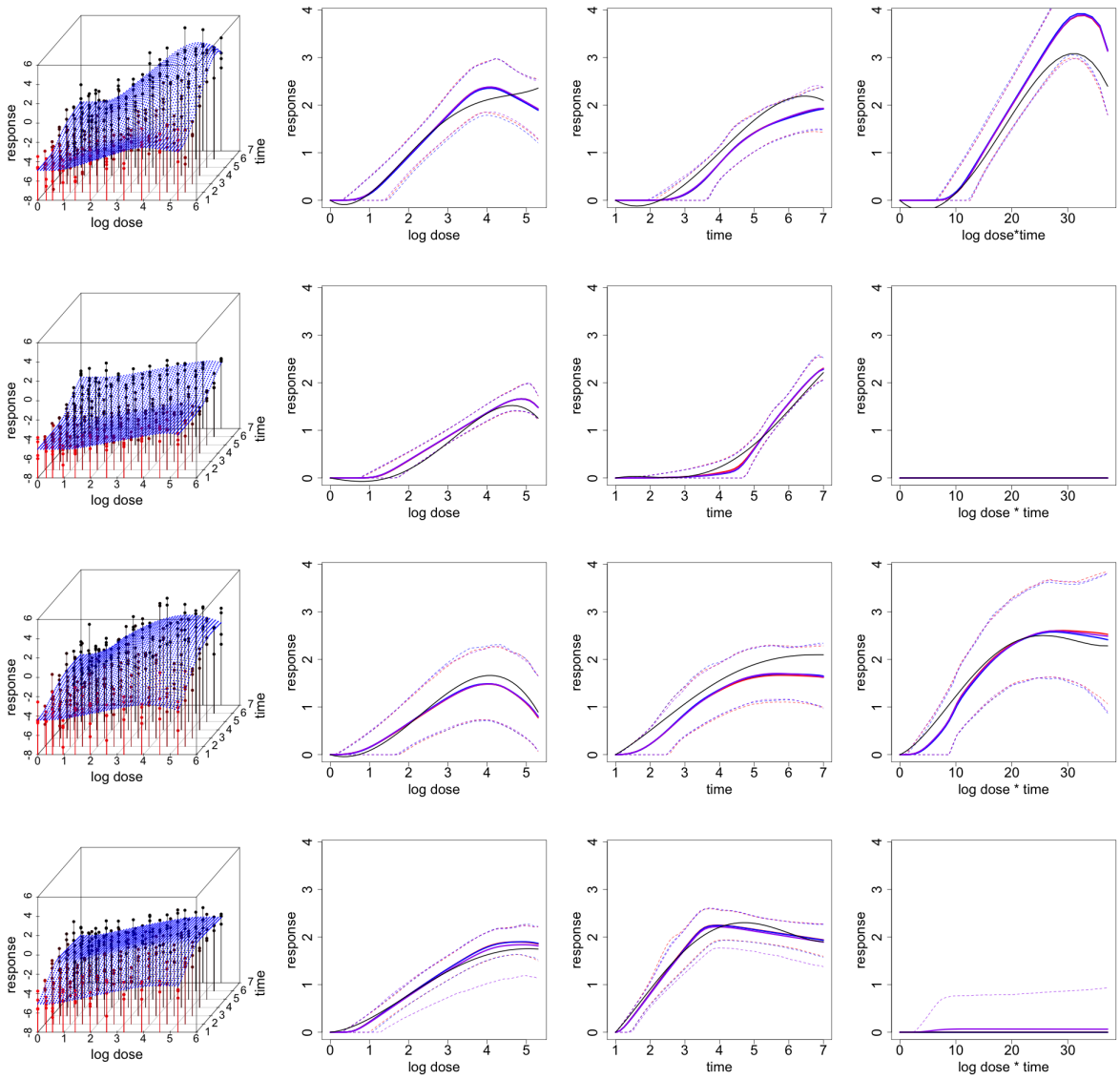

Figure A.4: **Simulation study to assess model fit and prior sensitivity (polynomial).** (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the least informative, prior 2 (red) is moderately informative and priors 3 (purple) is the most informative. (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data.

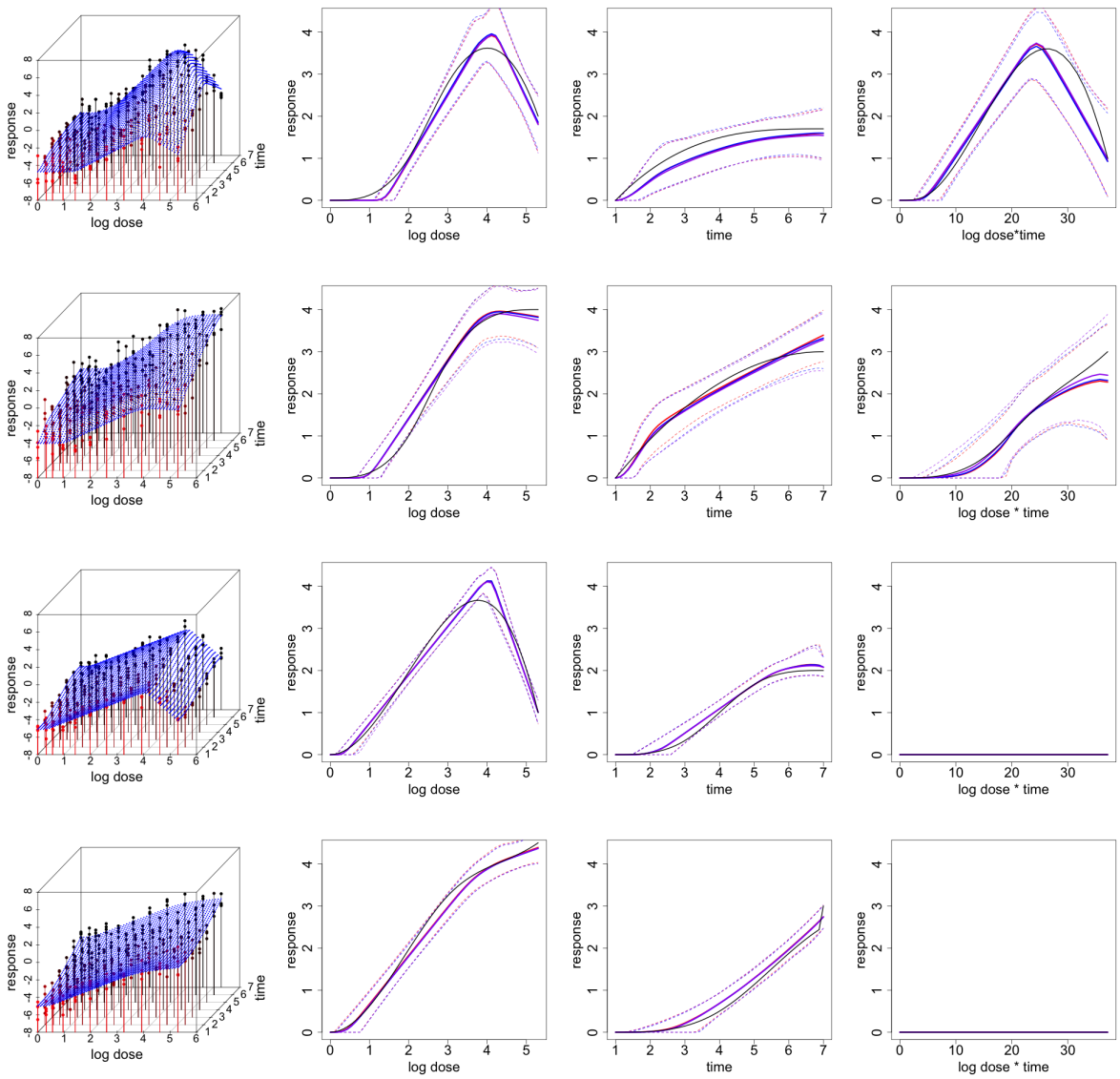

Figure A.5: **Simulation study to assess model fit and prior sensitivity (cubic splines).** (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the least informative, prior 2 (red) is moderately informative and priors 3 (purple) is the most informative. (*column 1*) Fitted response-surfaces estimated using prior 1, superimposed with simulated data.

A.3 Simulation Study: Assessing Prior Model Sensitivity

To assess sensitivity of the model to our choice of a prior model on the change-point parameters, we conducted a sensitivity analysis using the simulated data described above. We fit the simulated data to models with three different priors on the change-point parameters. The first prior was the bivariate beta prior parameterization described above, where we impose a right-skewed prior distribution on ϕ_{ij1}, ψ_{ij1}, and χ_{ij1}, and allow a conditionally uniform prior distribution for our second change-point parameters ϕ_{ij2}, ψ_{ij2}, and χ_{ij2}. Here we specified a set of relatively diffuse $Gamma(2,1)$ and $Gamma(3,1)$ priors for our population level change-point parameters $\boldsymbol{\lambda_{\phi_i}}$ and $\boldsymbol{\lambda_{\phi_i}}$, and a diffuse $B_2(2,3,1,1,DT)$ prior for our dose-time interaction change-point parameter $\boldsymbol{\chi}_{ij}$. The second prior is also a bivariate beta prior, with extremely informative, perhaps unreasonable, $Gamma(60,3)$ and $Gamma(1000,10)$ priors for our population level change-point parameters $\boldsymbol{\lambda_{\phi_i}}$ and $\boldsymbol{\lambda_{\phi_i}}$, and a $B_2(20,100,1,1,DT)$ prior for our dose-time interaction change-point parameter $\boldsymbol{\chi}_{ij}$.

The third prior is another parameterization of the bivariate beta distribution which allows for a uniform prior on the simlex, and can be described as follows:

$$\begin{aligned} \boldsymbol{\phi}_{ij} &\sim B_2(1,1,1,1,D), \\ \boldsymbol{\psi}_{ij} &\sim B_2(1,1,1,1,T), \\ \boldsymbol{\chi}_{ij} &\sim B_2(1,1,1,1,T), \end{aligned} \tag{A.2}$$

with support $\mathcal{S}(\boldsymbol{\phi_{ij}}) = \{(\phi_{ij1}, phi_{ij2}) : \ 0 < \phi_{ij1} < \phi_{ij2} < D\}$ and $\mathcal{S}(\boldsymbol{\psi_{ij}}) = \{(\psi_{ij1}, psi_{ij2}) : 0 < \psi_{ij1} < \psi_{ij2} < T\}$. The marginal prior distribution of ϕ_{ij1} is uniform and the conditional prior distribution of ϕ_{ij2} given ϕ_{ij2} is also uniform. Similarly for $\boldsymbol{\psi_{ij}}$ and $\boldsymbol{\chi_{ij}}$.

Finally, the fourth prior consists of two truncated normal priors, where for each ENM i

and response j, we define the following prior distributions for $\boldsymbol{\phi}_{ij}$, $\boldsymbol{\psi}_{ij}$, and $\boldsymbol{\chi}_{ij}$.

$$\begin{aligned}
\phi_{ij1} &\sim N(\phi_{o_{i1}}, \sigma^2_{\phi_{i1}})I(0 < \phi_{ij1} < D), & \phi_{ij2} &\sim N(\phi_{o_{i2}}, \sigma^2_{\phi_{i2}})I(\phi_{ij2} < \phi_{ij2} < D),\\
\psi_{ij1} &\sim N(\psi_{o_{i1}}, \sigma^2_{\psi_{i1}})I(0 < \psi_{ij1} < T), & \psi_{ij2} &\sim N(\psi_{o_{i2}}, \sigma^2_{\psi_{i2}})I(\psi_{ij2} < \psi_{ij2} < T),\\
\chi_{ij1} &\sim N(\tfrac{1}{3}DT, 100)I(0 < \chi_{ij1} < DT), & \chi_{ij2} &\sim N(\tfrac{2}{3}DT, 100)I(\chi_{ij2} < \chi_{ij2} < DT).
\end{aligned} \tag{A.3}$$

For each ENM i, we also define the following prior distributions for our population level parameters

$$\begin{aligned}
\phi_{o_{i1}} &\sim N(\tfrac{1}{3}D, 10), & \phi_{o_{i1}} &\sim N(\tfrac{2}{3}D, 10),\\
\psi_{o_{i1}} &\sim N(\tfrac{1}{3}T, 10), & \psi_{o_{i1}} &\sim N(\tfrac{2}{3}T, 10),
\end{aligned} \tag{A.4}$$

and similarly the following priors on our hyperparameters

$$\begin{aligned}
1/\sigma^2_{\phi_{i1}} &\sim Gamma(a_{\phi_{i1}}, b_{\phi_{i1}}), & 1/\sigma^2_{\phi_{i2}} &\sim Gamma(a_{\phi_{i2}}, b_{\phi_{i2}}),\\
1/\sigma^2_{\psi_{i1}} &\sim Gamma(a_{\psi_{i1}}, b_{\psi_{i1}}), & 1/\sigma^2_{\psi_{i2}} &\sim Gamma(a_{\psi_{i2}}, b_{\psi_{i2}}).
\end{aligned} \tag{A.5}$$

For all other parameters we used relatively diffuse priors, as described in Appendix A.2. Our inferences are based on 20,000 MCMC samples, after discarding a conservative 60,000 iterations for burn-in.

Figures A.6 through A.9 provide results from our simulation study, using the four different prior models for our change-point parameters described above. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves are superimposed with expected dose and duration-response curves, and point-wise 95% posterior intervals estimated from our model. Finally, *column 1* shows fitted dose-response-surfaces, spanning dose and time, for prior 2.

In summary, our analysis shows that inferential results tend to be robust to the choice of prior model. Priors 1, 3, and 4 are, in most cases, indistinguishable. Prior 2, is an extremely informative prior, and therefore results do somewhat differ from the other 3 priors, although not unreasonably. The benefit of using the bivariate beta prior described in the Chapter 2 is that it reflects the biological mechanism of toxicity. More specifically, it assigns zero probability to zero dose and time where toxicity is not expected to occur. The bivariate

beta prior also takes a conservative standpoint by assigning high probability to low doses and times of exposure, above a certain threshold.

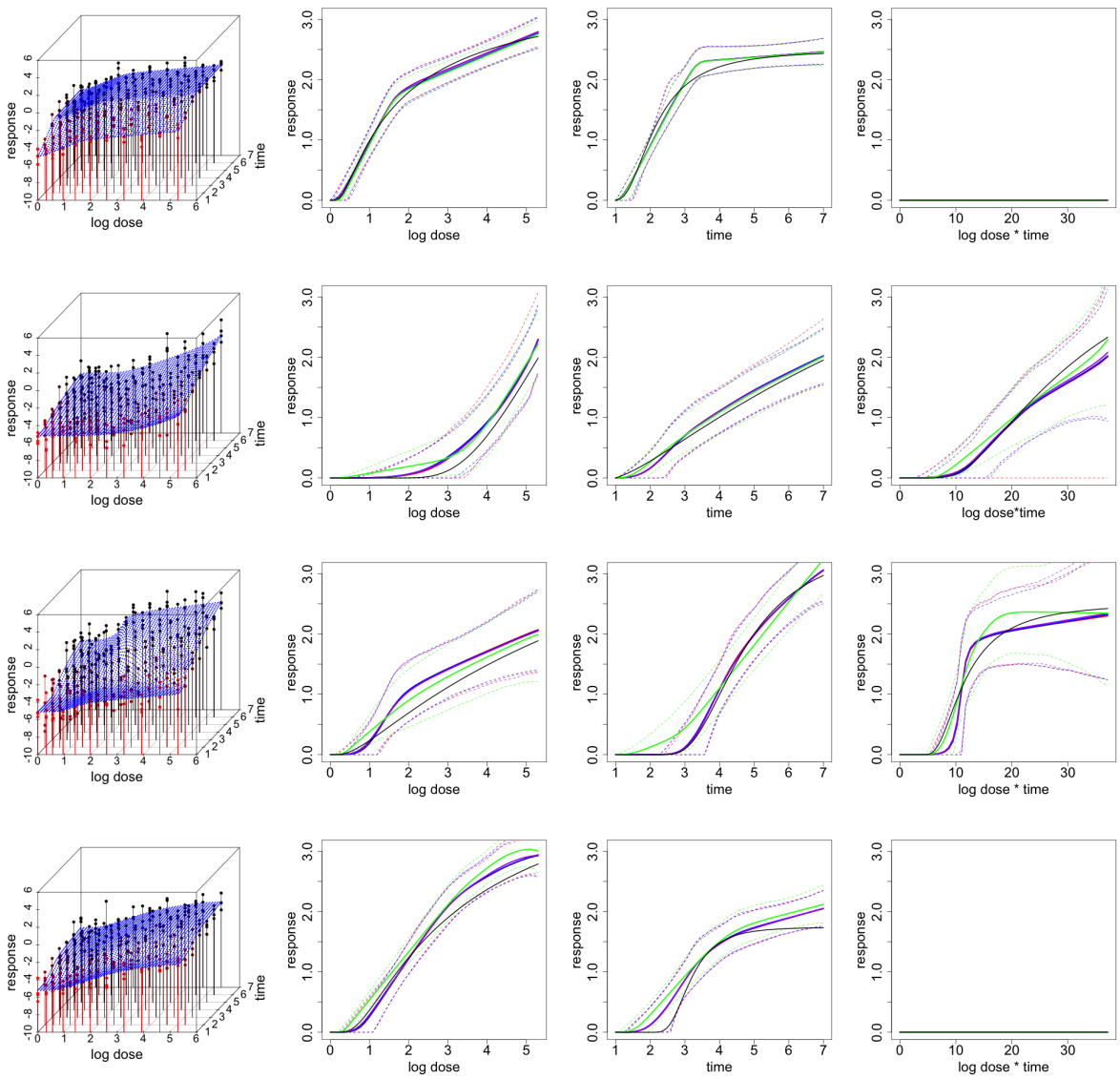

Figure A.6: **Simulation study to assess prior sensitivity to change-point model (weibull).** (*left*) Fitted response-surfaces estimated using prior 2, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the bivariate beta prior presented in Chapter 2, prior 2(green) is an extremely informative bivariate beta prior, prior 3 (red) is a uniform prior over the simplex and prior 4 (purple) is the truncated normal prior described above.

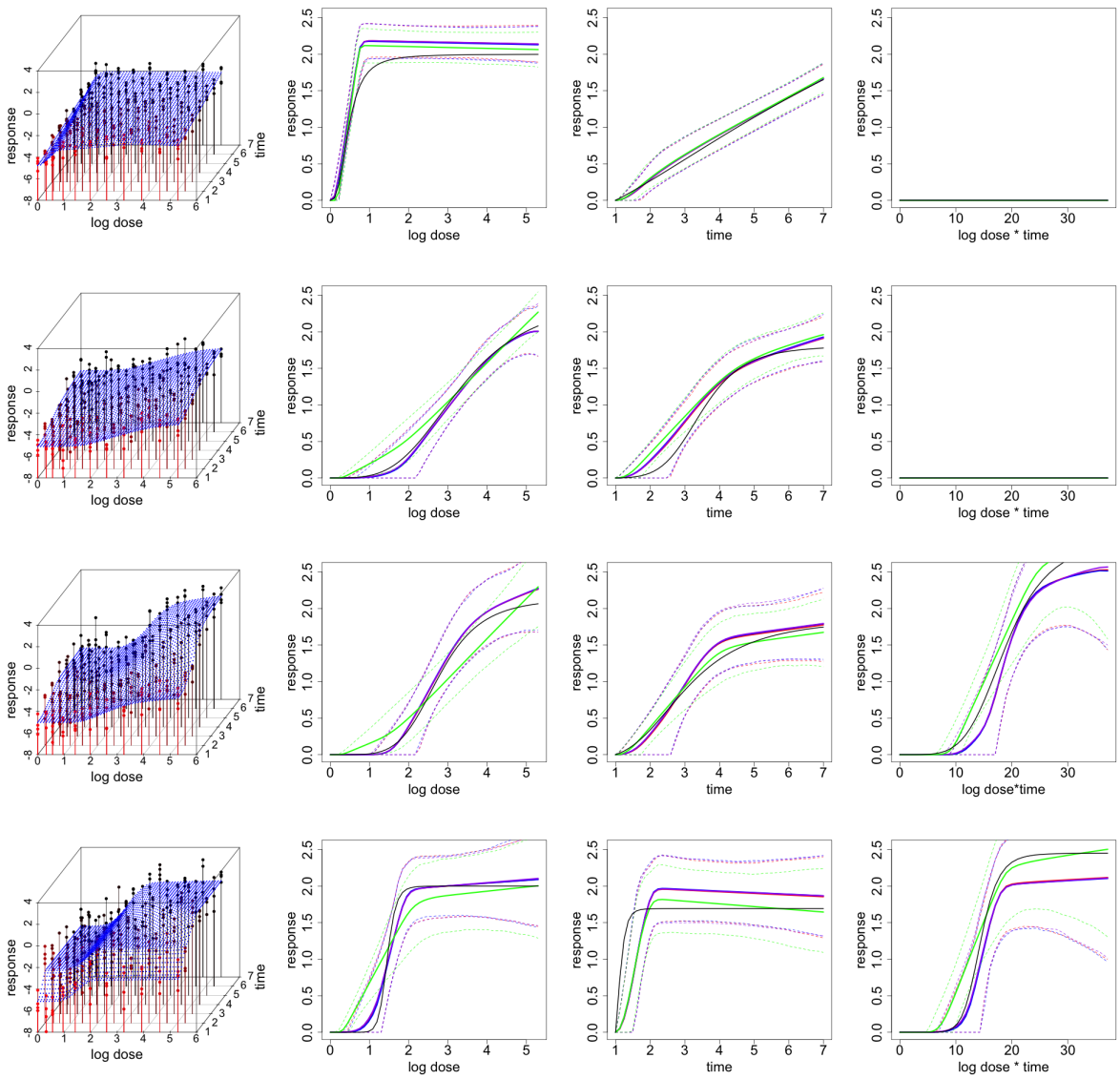

Figure A.7: **Simulation study to assess prior sensitivity to change-point model (log-logistic).** (*left*) Fitted response-surfaces estimated using prior 2, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the bivariate beta prior presented in Chapter 2, prior 2(green) is an extremely informative bivariate beta prior, prior 3 (red) is a uniform prior over the simplex and prior 4 (purple) is the truncated normal prior described above.

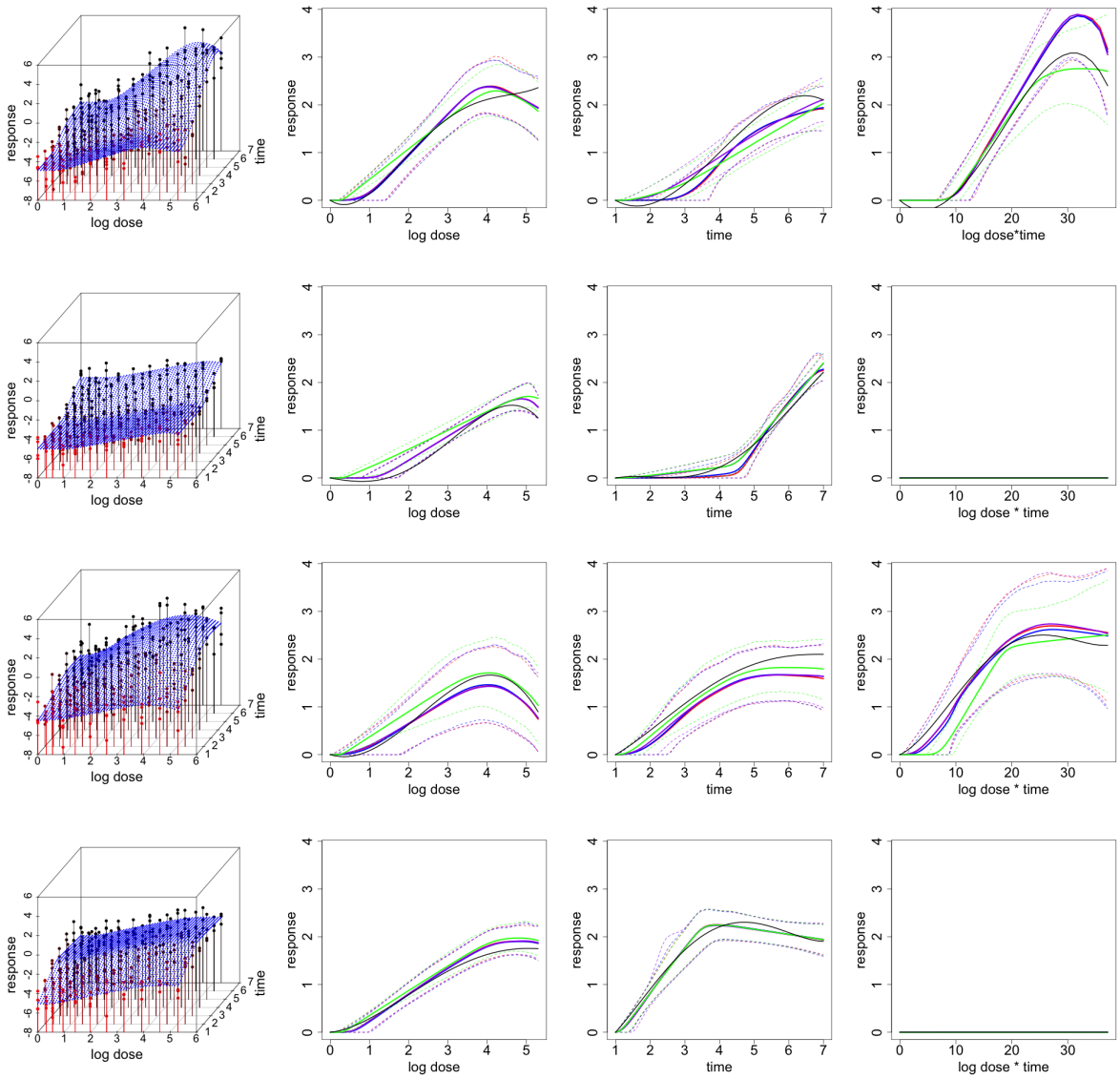

Figure A.8: **Simulation study to assess sensitivity to prior change-point model (polynomials).** (*left*) Fitted response-surfaces estimated using prior 2, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the bivariate beta prior presented in Chapter 2, prior 2(green) is an extremely informative bivariate beta prior, prior 3 (red) is a uniform prior over the simplex and prior 4 (purple) is the truncated normal prior described above.

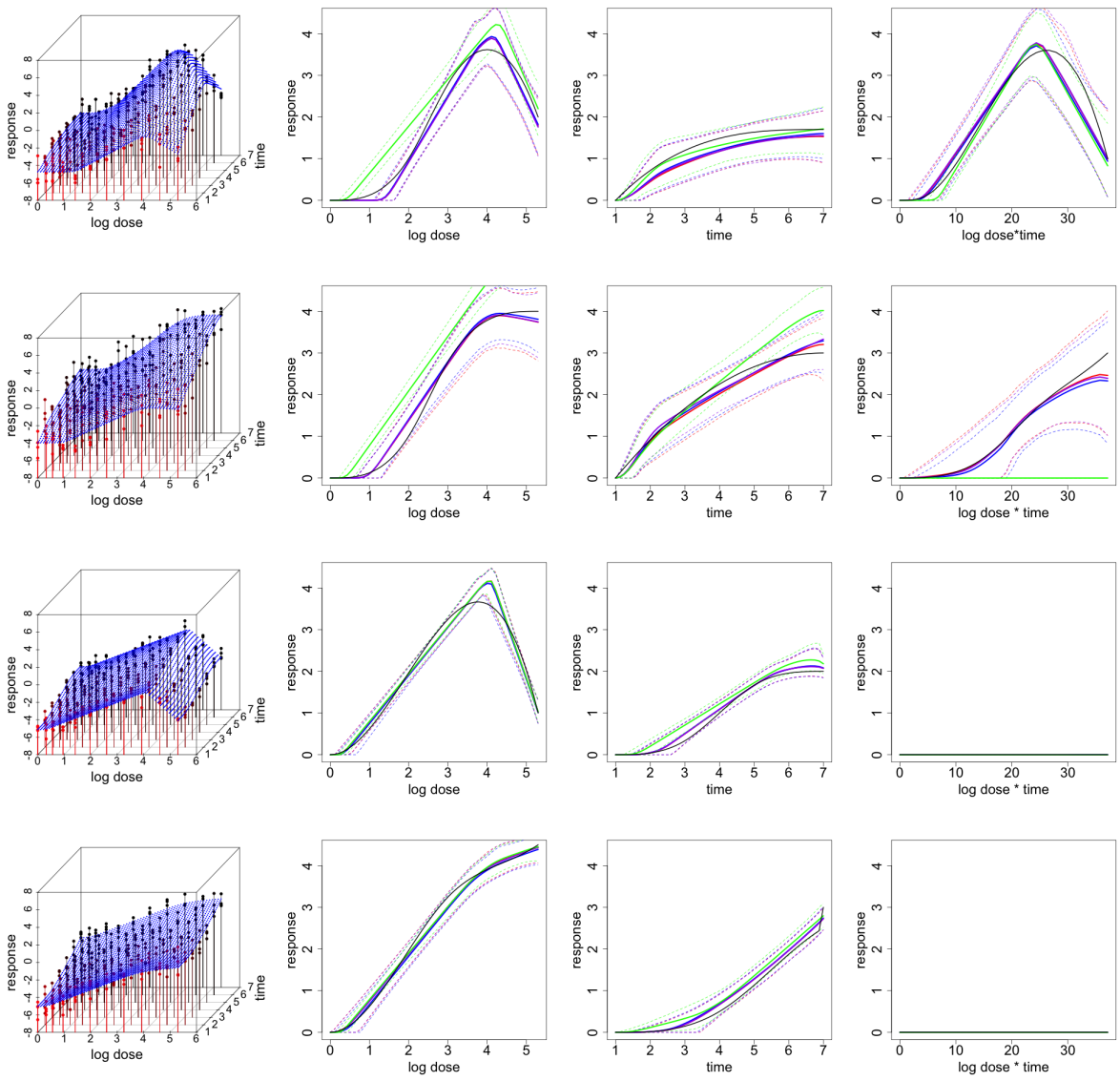

Figure A.9: **Simulation study to assess prior change-point model sensitivity (cubic splines.** (*left*) Fitted response-surfaces estimated using prior 2, superimposed with simulated data. Simulated dose (*column 2*), duration (*column 3*), and dose-time interaction (*column 4*) response curves (black) superimposed with expected curves estimated from our model (solid colored lines) and point-wise 95% posterior intervals (dashed lines). Prior 1 (blue) is the bivariate beta prior presented in Chapter 2, prior 2(green) is an extremely informative bivariate beta prior, prior 3 (red) is a uniform prior over the simplex and prior 4 (purple) is the truncated normal prior described above.

A.4 Additional Figures and Tables

The model presented in Chapter 2 was fit to a metal-oxide library consisting of eight nanoparticles, monitored for four cytotoxicity responses, which were measured over a grid of ten doses and seven times (hours) of exposure. The four responses include mitochondrial superoxide formation (MSF), loss of mitochondrial membrane potential (MMP), elevated intracellular calcium (EIC), and cellular membrane damage (CMD). Each outcome was measured as the percentage of cells positive for the response, and was normalized using a logit-like transformation.

In this appendix we present figures and tables for the remaining four particles not presented in Chapter 2, including silver (Ag), aluminum oxide (Al_2O_3), iron oxide (Fe_3O_4), and silicon dioxide (SiO_2) (see Figures A.10 to A.13). Specifically, *column 1* shows fitted dose-response-surfaces spanning dose and time. Also included are fitted curves for the expected dose-response function $f_{ij}(d)$ *(column 2)*, which represent the effect due to dose, the expected duration-response function $g_{ij}(t)$ *(column 3)*, which represent the effect due to time, and the expected dose-time interaction function $h_{ij}(t)$ *(column 4)*.

Figure A.14 provides a plot of the estimated median response, relative to the background, for a range of doses and times of exposure. Blue colors indicate safety regions, or areas of reduced risk to the cells, while red colored regions indicate increased risk of cytotoxicity. Table A.2 provides posterior estimates of the expected inclusion probability of the dose-time interaction function, for each ENM i and outcome j. The estimated inclusion probability can be used to test for a dose-time interaction, adjusted for multiplicity. Finally, Figure A.15 provides posterior summary estimates, including mean and 95% posterior intervals, for the maximal safe dose, conditional on the duration of exposure.

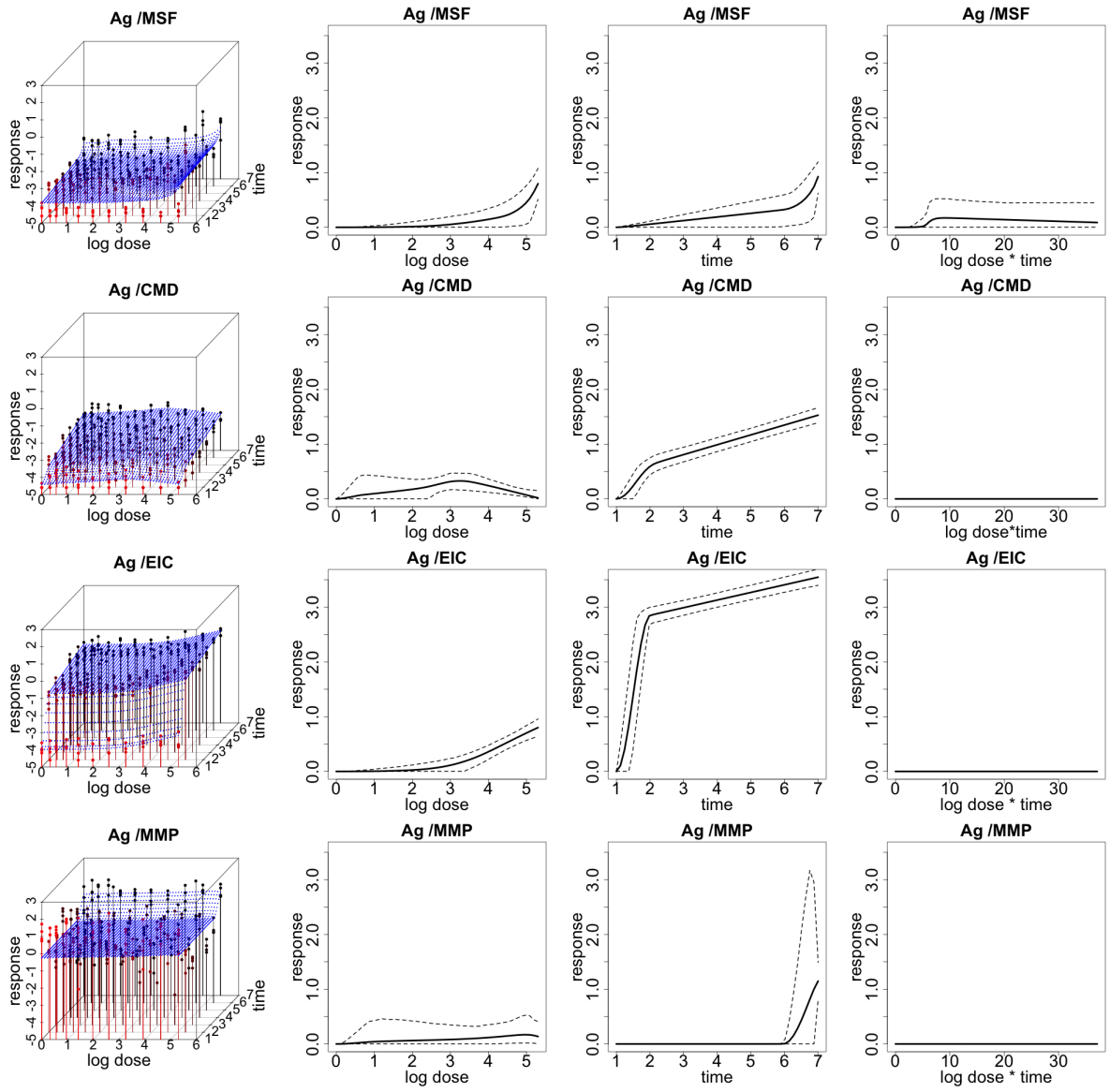

Figure A.10: **Fitted response curves for the silver (Ag) ENM.** Fitted response-surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

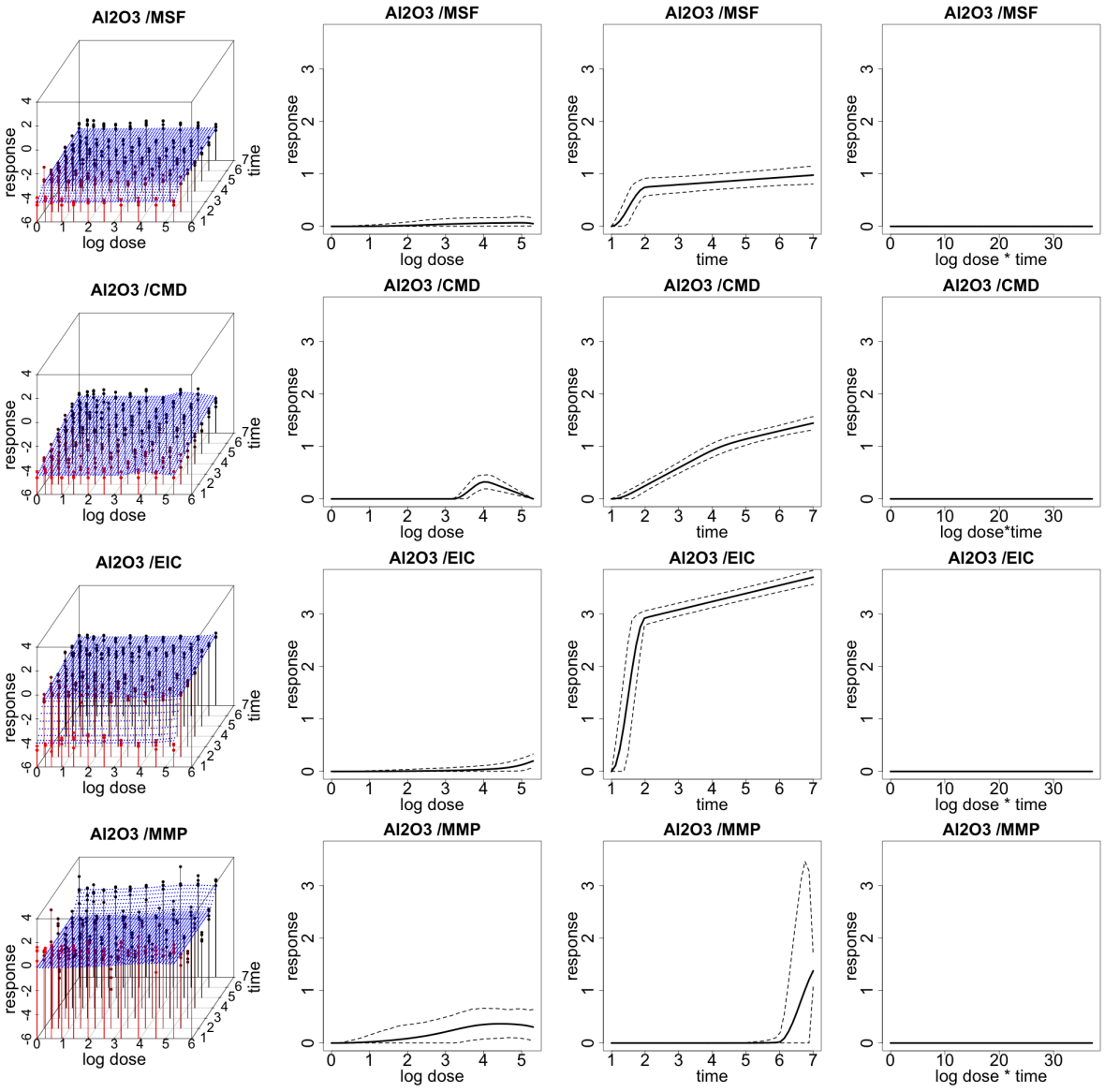

Figure A.11: **Fitted response curves for the aluminum oxide (Al_2O_3) ENM.** Fitted response-surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

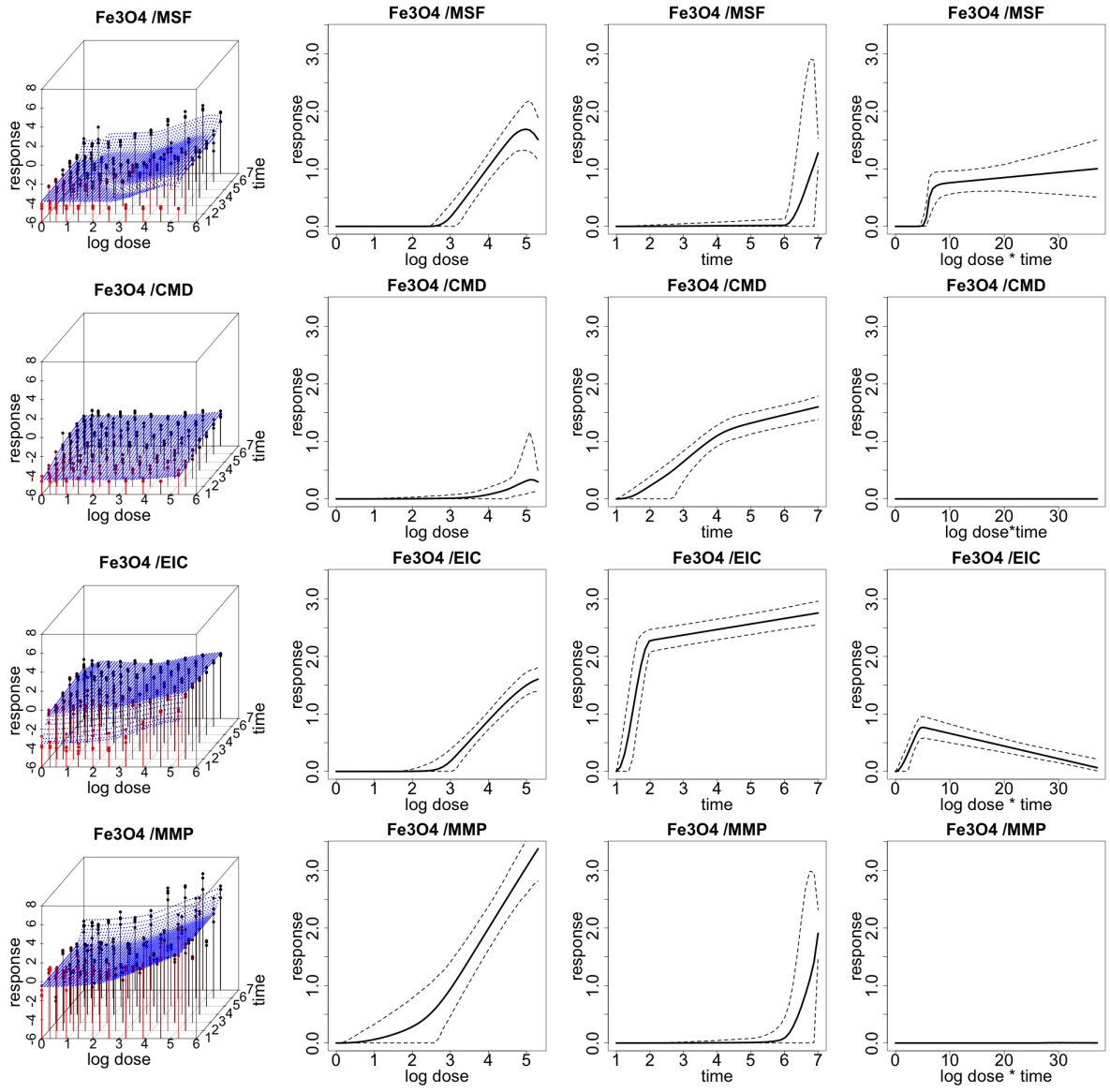

Figure A.12: **Fitted response curves for the iron oxide (Fe_3O_4) ENM.** Fitted response-surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

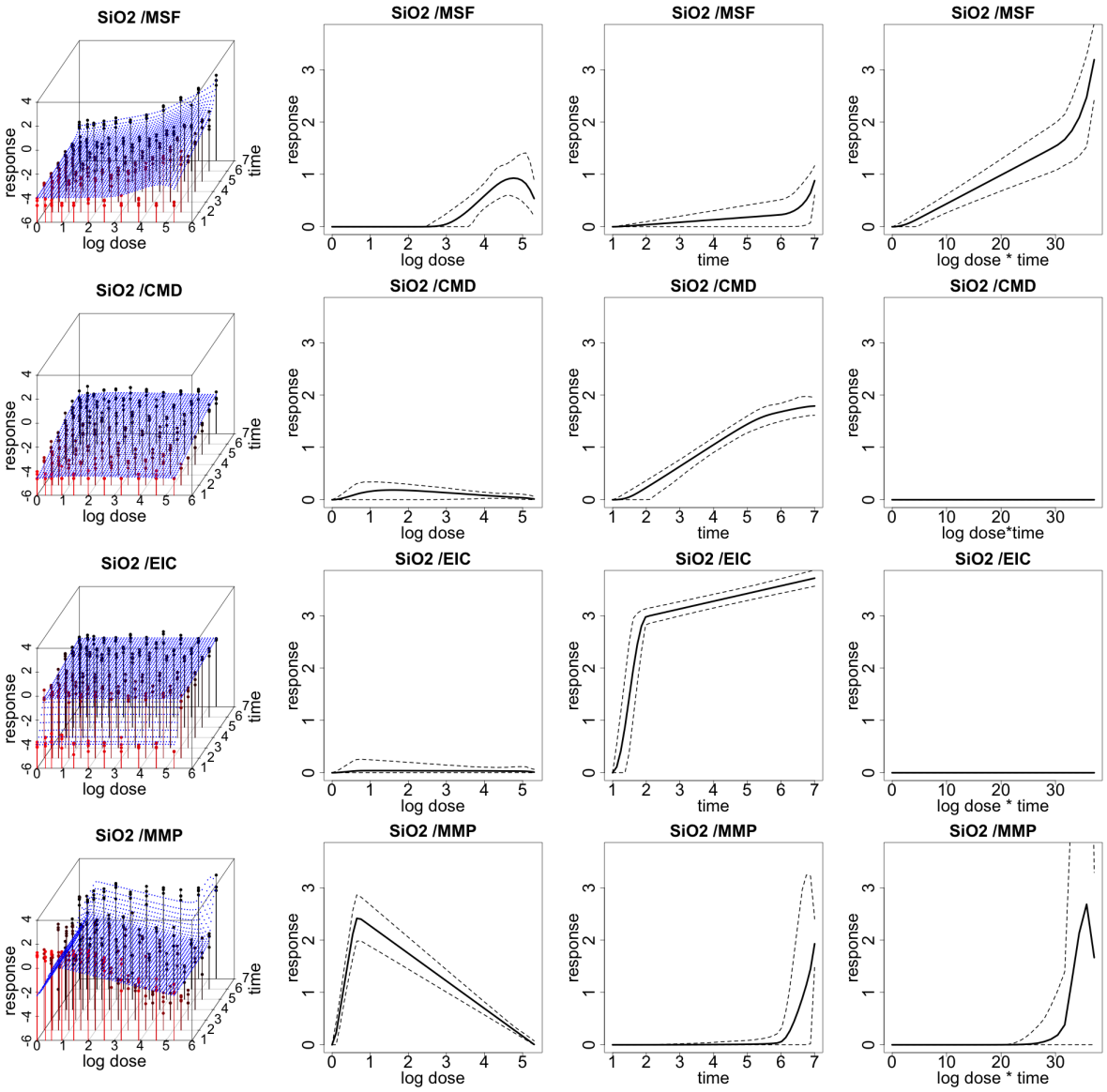

Figure A.13: **Fitted response curves for the silicon dioxide (SiO_2) ENM.** Fitted response-surfaces *(column 1)*, dose-response function, $f_{ij}(d)$ *(column 2)*, duration-response function, $g_{ij}(t)$ *(column 3)*, dose/duration interaction function, $h_{ij}(dt)$ *(column 4)* and associated 95% posterior intervals. In *(column 1)*, the color red represents response values corresponding to lower time points and the color black represents response values corresponding to higher time points.

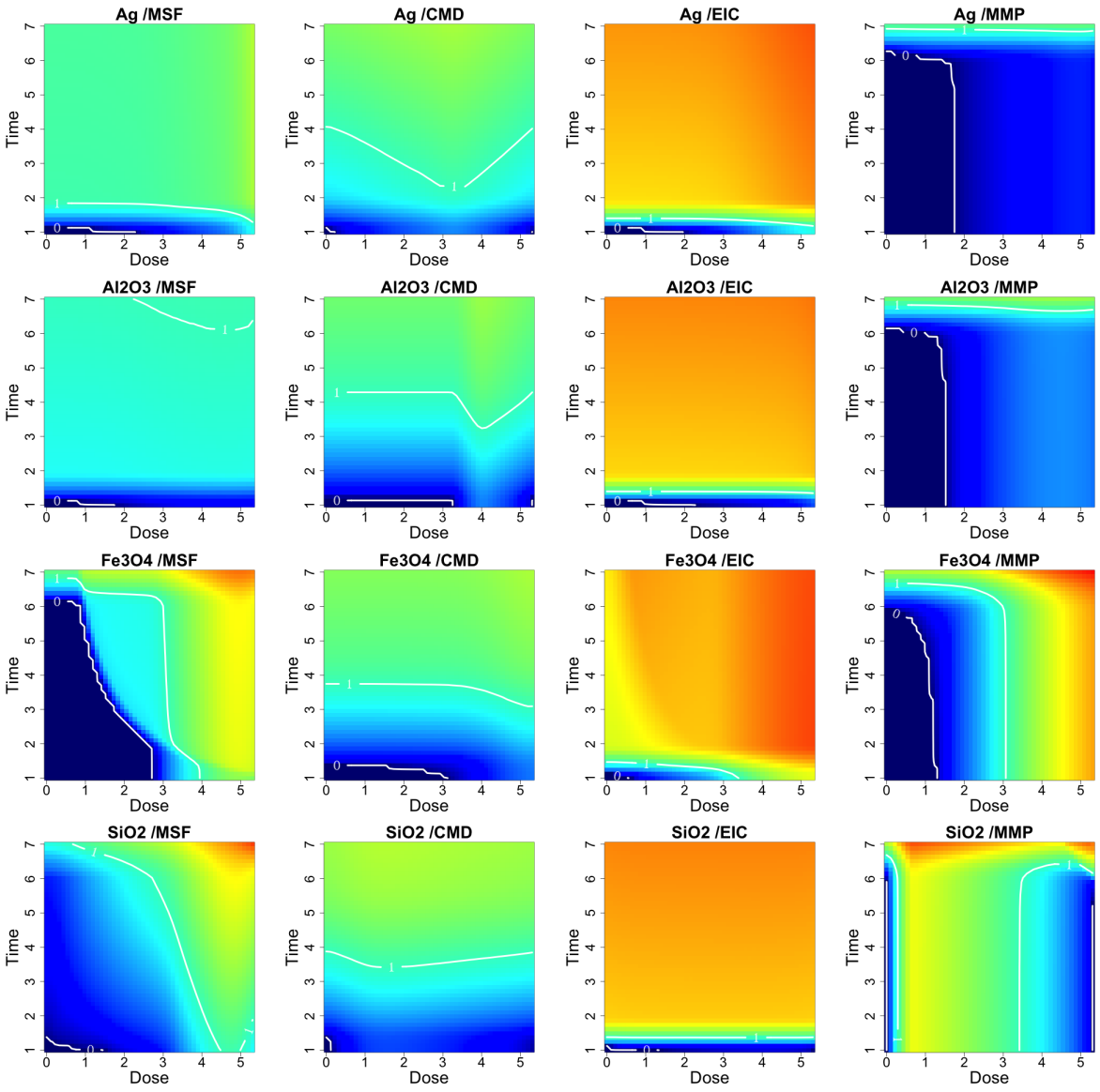

Figure A.14: **Safe exposure regions for silver (Ag), aluminum dioxide (Al_2O_3), iron oxide (Fe_3O_4), and silicon dioxide (SiO_2) ENMs..** For each particle and outcome we can define dose and time exposure regions which do not induce cytotoxicity. Lighter regions indicate greater cytotoxicity to the cells, whereas darker regions indicate reduced risk. Contour lines quantitate the median estimated response, relative to the background, where zero response areas can be interpreted as safe exposure regions.

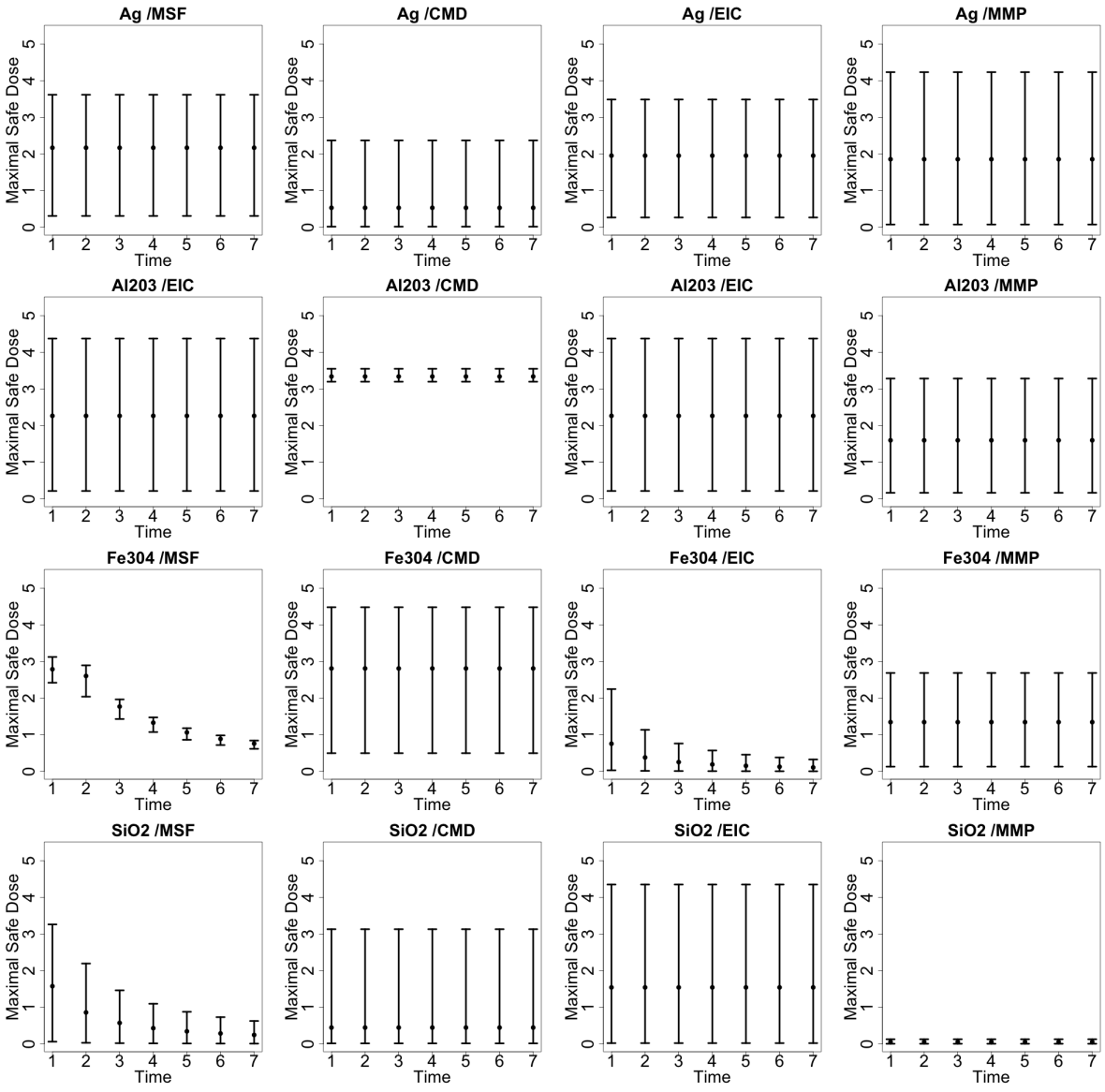

Figure A.15: **Maximal Safe Dose for silver (Ag), aluminum dioxide (Al_2O_3), iron oxide (Fe_3O_4), and silicon dioxide (SiO_2) ENMs.** Posterior summary estimates of the maximal safe dose, conditional on exposure time, including the posterior mean and associated 95% posterior intervals. In the case of no interaction, the maximal safe dose is the same across all times.

ENM	MSF	EIC	CMD	MMP
Ag	0.59	0.00	0.00	0.99
Au	0.00	0.00	0.00	0.00
Pt	0.99	0.00	0.00	0.00
Al_2O_3	0.00	0.00	0.00	0.00
Fe_3O_4	0.99	0.00	0.00	0.00
SiO_2	0.99	0.00	0.00	0.46
QD	0.99	0.00	0.00	0.00
ZnO	0.99	0.00	0.99	0.00

Table A.2: **Expected inclusion probabilities of the dose-time interaction function.** Posterior mean estimates of ρ_{ij} for each particle and outcome. The expected inclusion probability can be used to test for a dose-time interaction.

A.5 Model Assessment

In this Appendix we discuss model assessment. First we include a plot of the probability integral transform (PIT) histogram as measure of predictive performance, as described by Gneiting et al. (2007). Next, we check the average coverage of the posterior predictive distribution across all doses and times, by plotting the distribution of the posterior predictive mean response averaged across all doses and times of exposure, along with the empirical mean response across all doses and times of exposure.

Figure A.16 provides a plot of the PIT histogram for the entire model, including all doses, times, outcomes, and particles. Visual assessment indicates that the plot does tend toward uniformity, indicating good overall predictive performance. Figure A.17 also includes plots of the PIT histogram, separately for each ENM. The PIT histograms look approximately uniform for most ENM.

Figures C.2 and C.3 provide plots of the distribution of the posterior predictive mean response averaged across all doses and times of exposure (black), for each particle and outcome. The mean and associated 95% posterior intervals for this posterior predictive distribution are marked using vertical lines (black). Also included is the empirical mean response across all doses and times of exposure (red). Figure C.4 summarizes these results by plotting the mean and 95% posterior intervals of the posterior predictive mean response (black), along

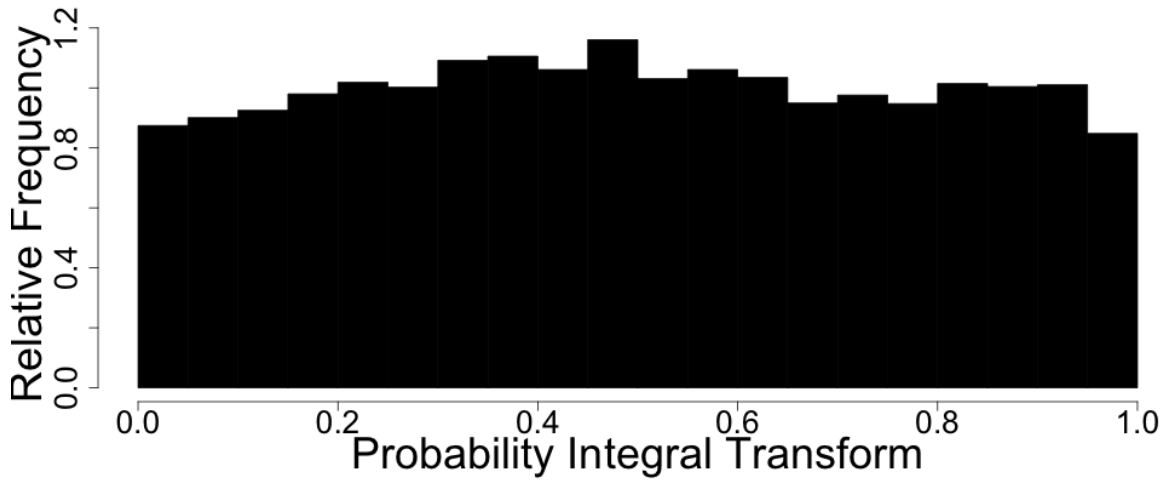

Figure A.16: **PIT histogram for the entire model.**

with the the empirical mean response across all doses and times (red), for each particle and outcome. In all cases the empirical mean response is contained within the 95% posterior intervals of the posterior predictive mean distribution, indicating relatively good posterior coverage across all doses and times of exposure. Further details on the implementation of model diagnostics described above can be found in Appendix C.2.

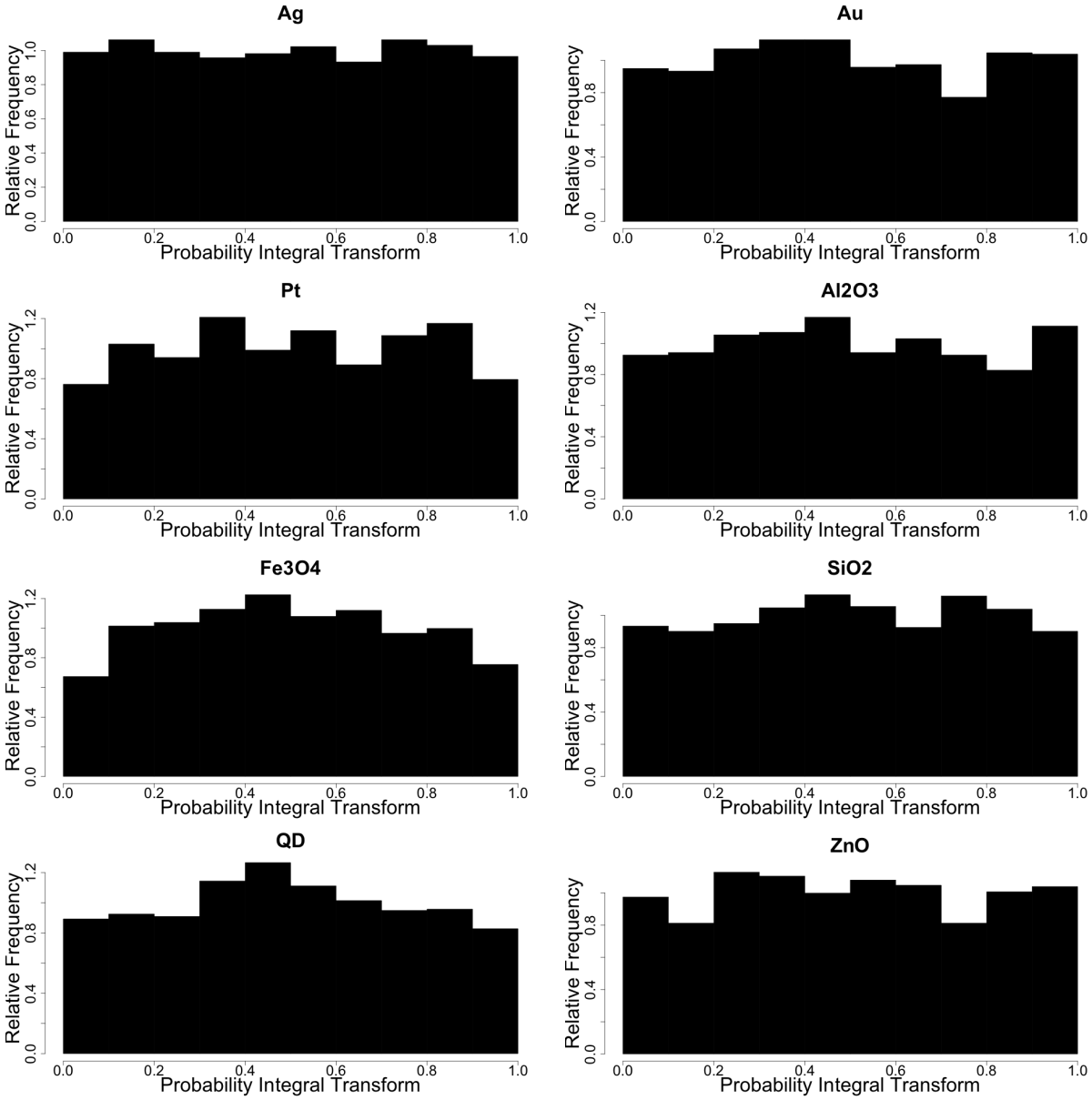

Figure A.17: **PIT histograms separately for each ENM.**

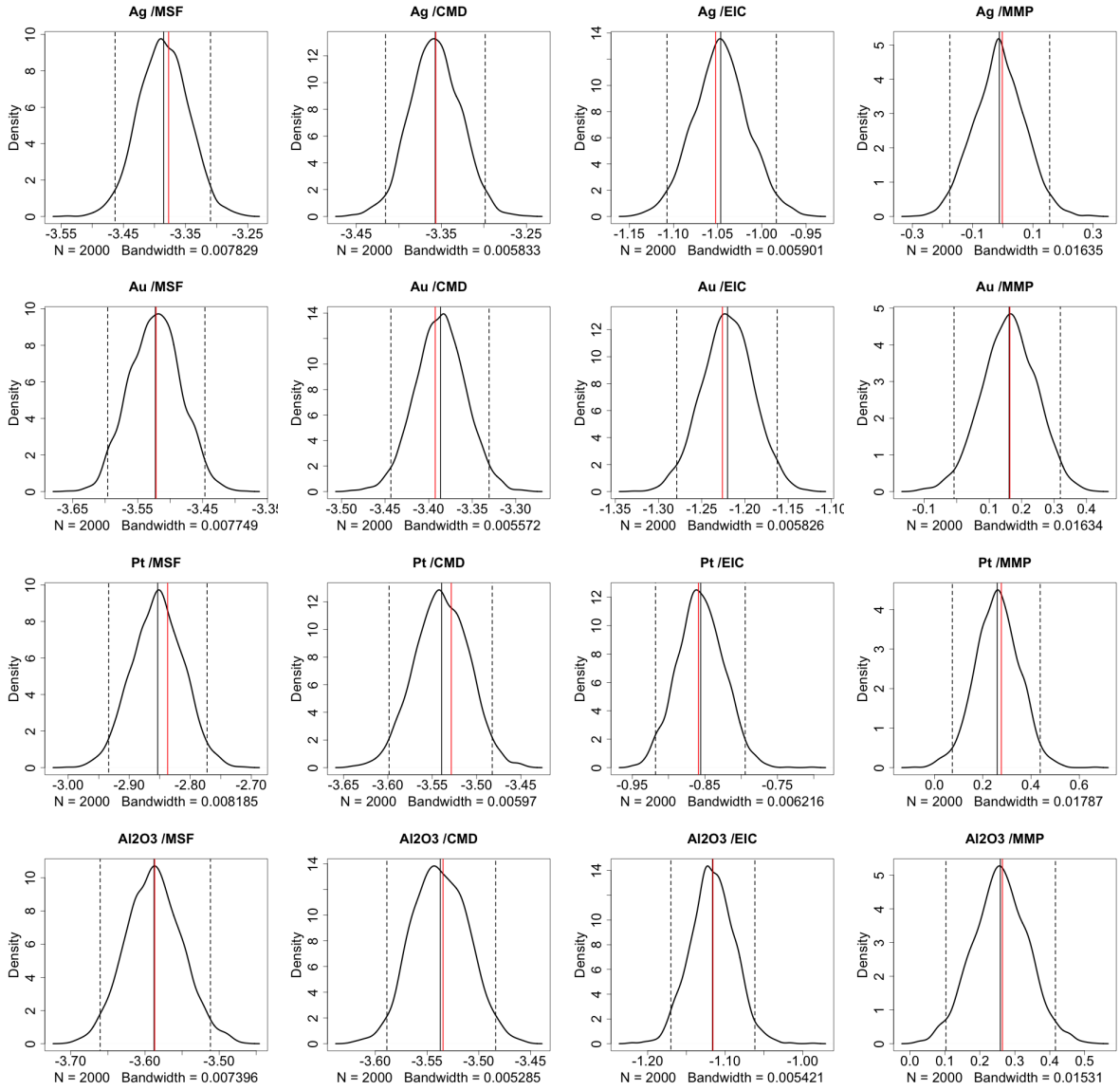

Figure A.18: **Posterior predictive mean distributions for Ag, Au, Pt and Al_2O_3 ENMs.** For each particle and outcome we plot the distribution of the posterior predictive mean response across all doses and times of exposure (black), along with the mean (solid black line) and associated 95% posterior intervals (dotted black lines) for this distribution. Also included is the empirical mean response across all doses and times of exposure (red).

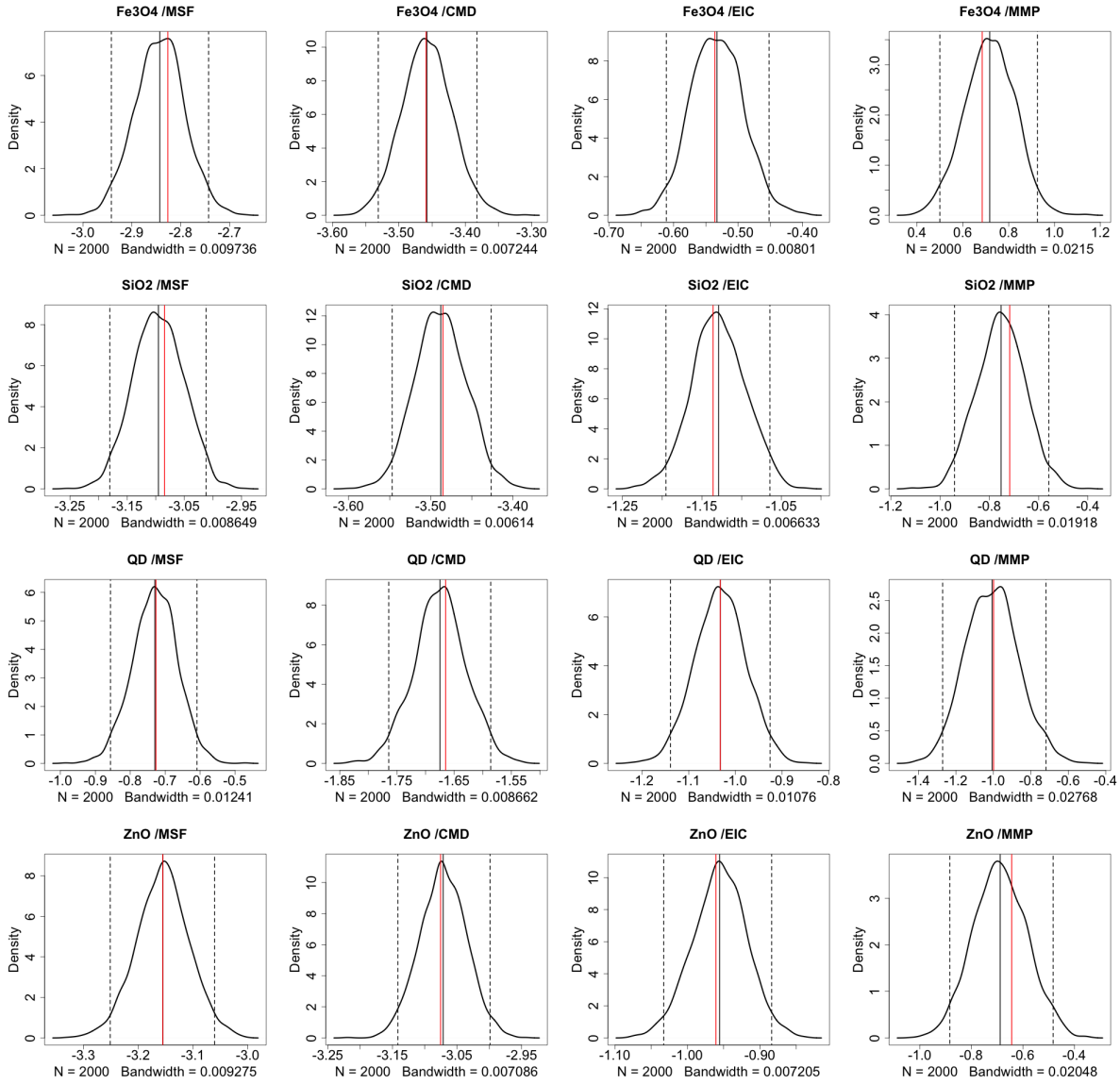

Figure A.19: **Posterior predictive mean distributions for Fe_3O_4, SiO_2, QD and ZnO ENMs.** For each particle and outcome we plot the distribution of the posterior predictive mean response across all doses and times of exposure (black), along with the mean (solid black line) and associated 95% posterior intervals (dotted black lines) for this distribution. Also included is the empirical mean response across all doses and times of exposure (red).

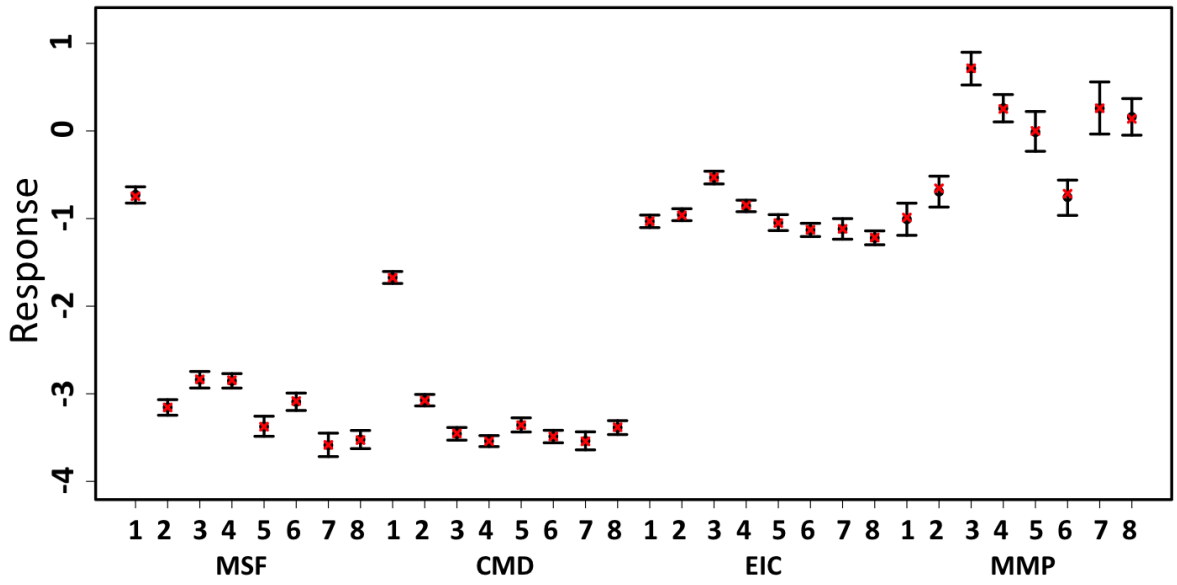

Figure A.20: **Summary of posterior predictive mean coverage.** Mean and 95% posterior intervals of the posterior predictive mean response across all doses and times of exposure, for all outcomes and particles 1 through 8 (QD, ZnO, Fe_3O_4, Pt, Ag, SiO_2, Al_2O_3, Au). Also included are the empirical mean responses across all doses and times of exposure (red).

APPENDIX B

Hierarchical Rank Aggregation

B.1 Toxicity Model

In this appendix we describe in detail the toxicity model used in the Chapter 3 case study. Let y denote a multivariate continuous outcome, corresponding to J cytotoxicity parameters, in association with the exposure of a number of cells to I different ENMs, and measured across K different experiments. More precisely, let $y_{ijk}(d,t)$ denote a multivariate response corresponding to ENM i ($i = 1, ..., I$), cytotoxicity parameter j ($j = 1, ..., J$), experiment k ($k = 1, ..., K$), and replicate ℓ ($\ell = 1, ..., L$), at dose $d \in [0, D]$ and time $t \in [0, T]$. Typically y is observed over a discrete set of doses $\tilde{d} = (d_1, ..., d_{m1})'$ and exposure times $\tilde{t} = (t_1, ..., t_{m2})'$. However, for clarity of exposition, we simplify our notation without loss of generality and refer to a general dose $d \in [0, D]$ and time $t \in [0, T]$. We introduce the following 5-stage hierarchical model.

Stage 1: Sampling Model

The observed response of particle i, cytotoxicity parameter j, experiment k, and replicate ℓ is modeled as:

$$y_{ijk\ell}(d,t) = m_{ijk}(d,t) + \epsilon_{ijk\ell}(d,t), \tag{B.1}$$

where $\epsilon_{ijk}(d,t) \sim N(0, \sigma^2_{\epsilon_{jk}}/\tau_i)$ and $\tau_i \sim Gamma(\nu/2, \nu/2)$. Here $m_{ij}(d,t)$ denotes the response-surface for particle i and outcome j. The error term ϵ_{ijk} is defined in terms of the measurement error variance $\sigma^2_{\epsilon_{jk}}$, specific to cytotoxicity parameter j and experiment k, and on ENM-specific variance inflation parameter τ_i. In this way we allow departures from normality and allow the measurement error σ_{ϵ_j} to vary between cytotoxicity parameters and

experiments, due to heterogeneity in the cytotoxicity outcomes and experimental platforms.

Stage 2: Response model at the ENM by cytotoxicity outcome by experimental level

The dose-response surface $m_{ijk}(d,t)$ spans two dimensions (dose and time), and is modeled in an additive fashion as described by Hastie and Tibshirani (1986). Let $(\alpha'_{ijk}, \boldsymbol{\beta}'_{ijk}, \boldsymbol{\phi}'_{ijk}, \boldsymbol{\gamma}'_{ijk}, \boldsymbol{\psi}'_{ijk})'$ be the parameter vector indexing the dose-response-surface $m_{ijk}(d,t)$. We define

$$m_{ijk}(d,t) = \alpha_{ijk} + f_{ijk}(d;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk}) + g_{ijk}(t;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk}). \tag{B.2}$$

Here $f_{ijk}(d;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk})$ is the function modeling the effect of dose d on response j, for ENM i and experiment k. Similarly, $g_{ijk}(t;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk})$ is the function modeling the effect of time t. To ensure likelihood identifiability we require, without loss of generality, that $f_{ijk}(d = 0;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk}) = 0$ and $g_{ijk}(t = 0;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk}) = 0$. The parameters α_{ijk} can therefore be interpreted as the background response level for each particle and outcome.

The dose-response curves $f_{ijk}(d;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk})$ and the duration-response curves $g_{ijk}(t;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk})$, are modeled as a linear combination of basis functions. Specifically, we use linear B-splines with two random interior knots. Let $\mathcal{B}(x, \boldsymbol{\eta})$ denote a 4-dimensional B-spline basis with interior knots $\boldsymbol{\eta} = (\eta_1, \eta_2)'$. Also, let $\boldsymbol{\beta}_{ijk} = (\beta_{ijk1}, .., \beta_{ijk4})'$ and $\boldsymbol{\gamma}_{ijk} = (\gamma_{ijk1}, ..., \gamma_{ijk4})'$ be two 4-dimensional vectors of spline coefficients. The functions $f_{ijk}(d;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk})$ and $g_{ijk}(t;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk})$ can then be represented as follows:

$$\begin{aligned} f_{ijk}(d;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk}) &= \mathcal{B}(d, \boldsymbol{\phi}_{ij})'\boldsymbol{\beta}_{ijk}, \\ g_{ijk}(t;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk}) &= \mathcal{B}(t, \boldsymbol{\psi}_{ijk})'\boldsymbol{\gamma}_{ijk}. \end{aligned} \tag{B.3}$$

Identifiability restrictions, $f_{ijk}(d = 0;\, \boldsymbol{\phi}_{ijk}, \boldsymbol{\beta}_{ijk}) = 0$ and $g_{ijk}(t = 0;\, \boldsymbol{\psi}_{ijk}, \boldsymbol{\gamma}_{ijk}) = 0$, are implemented by fixing $\beta_{ijk1} = 0$ and $\gamma_{ijk1} = 0$, for all particles and outcomes.

Similarly to the model described in Chapter 2, the first interior knot ϕ_{ijk1} can be interpreted as the dose at which ENM i becomes toxic in relation to cytotoxicity parameter j and experiment k (Maximal Safe Dose - similar to the classical NOAEL concept). A similar

interpretation can be given to ψ_{ijk1}, in relation to duration-response. As before, the foregoing interpretation is contingent on fixing $\beta_{ijk2} = 0$ and $\gamma_{ijk2} = 0$, when assuming no effect before ϕ_{ijk1} and ψ_{ijk1}, and $\beta_{ijk2} \leq 0$ and $\gamma_{ijk2} \leq 0$, when assuming a favorable effect before ϕ_{ijk1} and ψ_{ijk1}. The parameters ϕ_{ijk2} and ψ_{ijk2}, are respectively interpreted as the dose and time at which the response stabilizes or the cellular response starts a possible recovery process.

For each ENM i, response j, and experiment k, we define the following prior distributions for α_{ijk}, $\boldsymbol{\beta}_{ijk}$, and $\boldsymbol{\gamma}_{ijk}$

$$\begin{aligned}
\alpha_{ijk} &\sim N(\alpha_{o_{ij}}, \sigma^2_{\alpha_{ij}}), \\
\boldsymbol{\beta_{ijk}} &\sim N_4(\boldsymbol{\beta_{o_ij}}, \boldsymbol{\Sigma_{\beta_{ij}}}) I\{\beta_{ijj1} = 0;\ \beta_{ijj2} \leq 0;\ (\beta_{ijk3}, \beta_{ijk4}) \geq 0\}, \\
\boldsymbol{\gamma_{ijk}} &\sim N_4(\boldsymbol{\gamma_{o_ij}}, \boldsymbol{\Sigma_{\gamma_{ij}}}) I\{\gamma_{ijk1} = 0;\ \gamma_{ijk2} \leq 0;\ (\gamma_{ijk3}, \gamma_{ijk4}) \geq 0\}.
\end{aligned} \tag{B.4}$$

The covariance matrix $\boldsymbol{\Sigma_{\beta_{ij}}}$ has diagonal elements $\boldsymbol{\sigma_{\beta_{ijp}}}$, $p = 1, .., 4$, and off diagonal elements equal to 0; similarly for $\boldsymbol{\Sigma_{\gamma_{ij}}}$.

Prior distributions for $\boldsymbol{\phi}_{ijk}$ and $\boldsymbol{\psi}_{ijk}$ are defined to satisfy the following constraints: $(0 < \phi_{ijk1} < \phi_{ijk2} < D)$ and $(0 < \psi_{ijk1} < \psi_{ijk2} < T)$. More precisely, we model the joint distribution of the interior dose and duration knots using a generalized bivariate, as described in detail Chapter 2.

$$\begin{aligned}
\phi_{ij} &\sim B_2(1, \lambda_{\phi_{ij1}}, \lambda_{\phi_{ij2}}, 1, 1, D) I\{\lambda_{\phi_{ij2}} > \lambda_{\phi_{ij1}} > 1\}, \\
\psi_{ij} &\sim B_2(1, \lambda_{\psi_{ij1}}, \lambda_{\psi_{ij2}}, 1, 1, T) I\{\lambda_{\psi_{ij2}} > \lambda_{\psi_{ij1}} > 1\}.
\end{aligned} \tag{B.5}$$

This results in a right-skewed marginal distribution for x_1, and a uniform conditional distribution for x_2 given x_1.

Stage 3: Response model at the ENM by cytotoxicity outcome level

For each ENM i and outcome j, we define the following prior distributions for the population level parameters across experiments

$$\alpha_{o_{ij}} \sim N(\alpha_{oo_i}, \sigma^2_{\alpha_{o_i}}), \quad \boldsymbol{\beta_{o_{ij}}} \sim N_4(\boldsymbol{\beta_{oo_i}}, \boldsymbol{\Sigma}_{\beta_{o_i}}), \quad \boldsymbol{\gamma_{o_{ij}}} \sim N_4(\boldsymbol{\gamma_{oo_i}}, \boldsymbol{\Sigma}_{\gamma_{o_i}}). \tag{B.6}$$

The parameters $\boldsymbol{\beta_{o_{ij}}}$ and $\boldsymbol{\gamma_{o_{ij}}}$ are summaries of the dose and duration-response trajectories across all experiments and $\alpha_{o_{ij}}$ is a summary of the baseline response across all experiments.

We also define a prior model for population level parameters $\boldsymbol{\lambda_{\phi_{ij}}} = (\lambda_{\phi_{ij1}}, \lambda_{\phi_{ij2}})$ and $\boldsymbol{\lambda_{\psi_{ij}}} = (\lambda_{\psi_{ij1}}, \lambda_{\psi_{ij2}})$ as follows:

$$\lambda_{\phi_{ijr}} \sim Exponential(\lambda_{\phi oir}), \quad \lambda_{\psi_{ijr}} \sim Exponential(\lambda_{\psi oir}), \tag{B.7}$$

where $r = 1, 2$. The parameters $\boldsymbol{\lambda_{\phi_{ij}}}$ and $\boldsymbol{\lambda_{\psi_{ij}}}$ can be used to construct summaries of dose and duration-response change-points across all experiments.

Stage 4: Response model at the ENM level

For each ENM i, we define the following prior distributions for the population level parameters across experiments and outcomes

$$\alpha_{oo_i} \sim N(m_{\alpha_i}, v_{\alpha_i}), \quad \boldsymbol{\beta_{oo_i}} \sim N_4(\boldsymbol{m_{\beta_i}}, \boldsymbol{v_{\beta_i}}), \quad \boldsymbol{\gamma_{oo_i}} \sim N_4(\boldsymbol{m_{\gamma_i}}, \boldsymbol{v_{\gamma_i}}). \tag{B.8}$$

The parameters $\boldsymbol{\beta_{oo_i}}$ and $\boldsymbol{\gamma_{oo_i}}$ are summaries of the dose and duration-response trajectories across all outcomes and experiments. The α_{oo_i} parameters summarize the baseline response across all outcomes and experiments.

The population level parameters $\boldsymbol{\lambda_{\phi_i}} = (\lambda_{\phi_{i1}}, \lambda_{\phi_{i2}})$ and $\boldsymbol{\lambda_{\psi_i}} = (\lambda_{\psi_{i1}}, \lambda_{\psi_{i2}})$ can be defined as follows:

$$\lambda_{\phi_{o_{ir}}} \sim Gamma(a_{\lambda_{\phi ir}}, b_{\lambda_{\phi ir}}), \quad \lambda_{\psi_{o_{ir}}} \sim Gamma(a_{\lambda_{\psi ir}}, b_{\lambda_{\psi ir}}), \tag{B.9}$$

where $\ell = 1, 2$. The parameters $\boldsymbol{\lambda_{\phi_{o_i}}}$ and $\boldsymbol{\lambda_{\psi_{o_i}}}$ can be used to construct summaries of dose and duration-response change-points across all outcomes and experiments.

Stage 5: Hyperpriors

We complete the model by specifying prior distributions on our hyperparameters as follows:

$$\begin{array}{llllll} 1/\sigma^2_{\epsilon_{jk}} & \sim & Gamma(a_{\epsilon_{jk}}, b_{\epsilon_{jk}}), & \nu & \sim & Uniform(1,2,4,8,16,32), \\ 1/\sigma^2_{\alpha_{ij}} & \sim & Gamma(a_{\alpha_{ij}}, b_{\alpha_{ij}}), & 1/\sigma^2_{\beta_{ij}} & \sim & Gamma(a_{\beta_{ij}}, b_{\beta_{ij}}), \\ 1/\sigma^2_{\gamma_{ij}} & \sim & Gamma(a_{\gamma_{ij}}, b_{\gamma_{ij}}), & 1/\sigma^2_{\alpha_{oi}} & \sim & Gamma(a_{\alpha_{oi}}, b_{\alpha_{oi}}), \\ 1/\sigma^2_{\beta_{oi}} & \sim & Gamma(a_{\beta_{oi}}, b_{\beta_{oi}}), & 1/\sigma^2_{\gamma_{oi}} & \sim & Gamma(a_{\gamma_{oi}}, b_{\gamma_{oi}}). \end{array} \tag{B.10}$$

Note that in our formulation, $x \sim Gamma(a, b)$ denotes a Gamma distributed random quantity with shape a and rate b, such that $E(x) = a/b$.

B.2 Full Conditional Distributions

In this appendix we describe some of the full conditional distributions for the model described above. Again let $y_{ijk\ell}(d,t)$ denote a multivariate response corresponding to ENM i ($i = 1, ..., I$), cytotoxicity parameter j ($j = 1, ..., J$), experiment k ($k = 1, ..., K$), and replicate ℓ ($\ell = 1, ..., L$), at some dose d in $d = (d_1, ...d_D)'$ and some time t in $\boldsymbol{t} = (t_1, ...t_T)'$. Also, let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\alpha_o}, \boldsymbol{\beta_o}, \boldsymbol{\gamma_o}, \boldsymbol{\lambda_\phi}, \boldsymbol{\lambda_\psi}, \boldsymbol{\alpha_{oo}}, \boldsymbol{\beta_{oo}}, \boldsymbol{\gamma_{oo}}, \boldsymbol{\lambda_{\phi_o}}, \boldsymbol{\lambda_{\psi_o}}, \boldsymbol{\tau}, \nu, \boldsymbol{\sigma_\epsilon}, \boldsymbol{\sigma_\alpha}, \boldsymbol{\sigma_\beta}, \boldsymbol{\sigma_\gamma}, \boldsymbol{\sigma_{\alpha_o}}, \boldsymbol{\sigma_{\beta_o}}, \boldsymbol{\sigma_{\gamma_o}})$ denote the full parameter vector, and let $\boldsymbol{\theta}_{\backslash \boldsymbol{\delta}}$ denote the vector containing all components of $\boldsymbol{\theta}$ except for some parameter δ in $\boldsymbol{\theta}$. Moreover, we denote $\boldsymbol{Y_{ijk}}$ as the complete set of response values for particle i, outcome j and, experiment k. Using the notation above we define the full conditional distributions for all available parameters as follows.

B.2.1: Full conditional distributions for the random effect parameters α_{ijk}, $\boldsymbol{\beta_{ijk}}$, and $\boldsymbol{\gamma_{ijk}}$.

$$\alpha_{ijk} \mid \boldsymbol{Y_{ijk}}, \boldsymbol{\theta}_{\backslash \alpha_{ijk}} \sim N\left(\frac{\frac{\sum_{d,t,\ell} \tilde{y}_{\alpha_{ijk\ell}}(d,t)}{\sigma^2_{\epsilon_{jk}}/\tau_i} + \frac{\alpha_{o_{ij}}}{\sigma_{\alpha_{ij}}}}{\frac{D\times T\times \ell}{\sigma^2_{\epsilon_{jk}}/\tau_i} + \frac{1}{\sigma_{\alpha_{ij}}}}, \frac{1}{\frac{D\times T\times \ell}{\sigma^2_{\epsilon_{jk}}/\tau_i} + \frac{1}{\sigma_{\alpha_{ij}}}} \right),$$

where $\tilde{y}_{\alpha_{ijk\ell}}(d,t) = y_{ijk\ell}(d,t) - \mathcal{B}(d, \boldsymbol{\phi}_{ijk})'\boldsymbol{\beta}_{ijk} - \mathcal{B}(t, \boldsymbol{\psi}_{ijk})'\boldsymbol{\gamma}_{ijk}$.

$\boldsymbol{\beta_{ijk}} \mid \boldsymbol{Y_{ijk}}, \boldsymbol{\theta}_{\backslash \boldsymbol{\beta_{ijk}}} \sim N_4(M_\beta, V_\beta)$,

where

$$M_\beta = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{ij}}} + \sum_{d,t,\ell} \frac{\mathcal{B}(d,\boldsymbol{\phi}_{ijk})\mathcal{B}(d,\boldsymbol{\phi}_{ijk})'}{\sigma^2_{\epsilon_{jk}}/\tau_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{ij}}} \boldsymbol{\beta_{o_{ij}}} + \sum_{d,t,\ell} \frac{\mathcal{B}(d,\boldsymbol{\phi}_{ijk}) \tilde{y}_{\beta_{ijk\ell}}(d,t)}{\sigma^2_{\epsilon_{jk}}/\tau_i} \right),$$

$$V_\beta = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{ij}}} + \sum_{d,t,\ell} \frac{\mathcal{B}(d,\boldsymbol{\phi}_{ijk})\mathcal{B}(d,\boldsymbol{\phi}_{ijk})'}{\sigma^2_{\epsilon_{jk}}/\tau_i} \right)^{-1},$$

$$\tilde{y}_{\beta_{ijk\ell}}(d,t) = y_{ijk\ell}(d,t) - \mathcal{B}(t,\boldsymbol{\psi}_{ijk})'\boldsymbol{\gamma}_{ijk} - \alpha_{ijk}.$$

$\boldsymbol{\gamma_{ijk}} \mid \boldsymbol{Y_{ijk}}, \boldsymbol{\theta}_{\backslash \boldsymbol{\gamma_{ijk}}} \sim N_4(M_\gamma, V_\gamma)$,

where

$$M_\gamma = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_{ij}}} + \sum_{d,t,\ell} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ijk})\mathcal{B}(t,\boldsymbol{\psi}_{ijk})'}{\sigma^2_{\epsilon_{jk}}/\tau_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_{ij}}} \boldsymbol{\gamma_{o_{ij}}} + \sum_{d,t,\ell} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ijk}) \tilde{y}_{\gamma_{ijk\ell}}(d,t)}{\sigma^2_{\epsilon_{jk}}/\tau_i} \right),$$

$$V_\gamma = \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_{ij}}} + \sum_{d,t,\ell} \frac{\mathcal{B}(t,\boldsymbol{\psi}_{ijk})\mathcal{B}(t,\boldsymbol{\psi}_{ijk})'}{\sigma^2_{\epsilon_{jk}}/\tau_i} \right)^{-1},$$

$$\tilde{y}_{\gamma_{ijk\ell}}(d,t) = y_{ijk\ell}(d,t) - \mathcal{B}(d,\boldsymbol{\phi}_{ijk})'\boldsymbol{\beta}_{ijk} - \alpha_{ijk}.$$

B.2.2: Full conditional distributions for ENM by outcome level parameters $\alpha_{o_{ij}}$, $\beta_{o_{ij}}$, and $\gamma_{o_{ij}}$.

$$\alpha_{o_i} \mid \boldsymbol{\theta}_{\backslash \alpha_{ijk}} \sim N \left(\frac{\frac{\sum_{k=1}^K \alpha_{ijk}}{\sigma^2_{\alpha_{ij}}} + \frac{\alpha_{oo_i}}{\sigma_{\alpha o_i})^2}}{\frac{K}{\sigma^2_{\alpha_{ij}}} + \frac{1}{\sigma_{\alpha o_i})^2}}, \frac{1}{\frac{K}{\sigma^2_{\alpha_{ij}}} + \frac{1}{\sigma_{\alpha o_i})^2}} \right),$$

$$\boldsymbol{\beta_{o_i}} \mid \boldsymbol{\theta}_{\backslash \boldsymbol{\beta_{ijk}}} \sim N_4 \left(\left(K\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{ij}}} + \boldsymbol{\Sigma}_{\boldsymbol{\beta_{o_i}}} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{ij}}} \sum_{k=1}^K \boldsymbol{\beta_{ij}} + \boldsymbol{\Sigma}_{\boldsymbol{\beta_{o_i}}} \right)^{-1} \boldsymbol{\beta_{oo_i}} \right), \left(K\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{ij}}} + \boldsymbol{\Sigma}_{\boldsymbol{\beta_{o_i}}} \right)^{-1} \right),$$

$$\boldsymbol{\gamma_{o_i}} \mid \boldsymbol{\theta}_{\backslash \boldsymbol{\gamma_{ijk}}} \sim N_4 \left(\left(K\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_{ij}}} + \boldsymbol{\Sigma}_{\boldsymbol{\gamma_{o_i}}} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_{ij}}} \sum_{k=1}^K \boldsymbol{\gamma_{ij}} + \boldsymbol{\Sigma}_{\boldsymbol{\gamma_{o_i}}} \right)^{-1} \boldsymbol{\gamma_{oo_i}} \right), \left(K\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\gamma_{ij}}} + \boldsymbol{\Sigma}_{\boldsymbol{\gamma_{o_i}}} \right)^{-1} \right).$$

B.2.3: Full conditional distributions for ENM level parameters α_{oo_i}, β_{oo_i}, and γ_{oo_i}.

$$\alpha_{oo_i} \mid \boldsymbol{\theta}_{\backslash \alpha_{o_{ij}}} \sim N \left(\frac{\sum_{j=1}^J \sum_{k=1}^K \alpha_{o_{ij}}}{\frac{\sigma^2_{\alpha_{oi}}}{\sigma^2_{\alpha_{oi}}} + \frac{m_{\alpha_i}}{s^2_{\alpha_i}}}, \frac{1}{\frac{JK}{\sigma^2_{\alpha_{oi}}} + \frac{1}{s^2_{\alpha_i}}} \right),$$

$$\boldsymbol{\beta_{oo_i}} \mid \boldsymbol{\theta}_{\backslash \boldsymbol{\beta_{o_{ij}}}} \sim N_4 \left(\left(JK\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{oi}}} + \boldsymbol{v}^{-1}_{\boldsymbol{\beta}_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{oi}}} \sum_{j=1}^J \sum_{k=1}^K \boldsymbol{\beta_{o_{ij}}} + \boldsymbol{v}^{-1}_{\boldsymbol{\beta}_i} \boldsymbol{m}_{\boldsymbol{\beta}_i} \right), \left(JK\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\beta_{oi}}} + \boldsymbol{v}^{-1}_{\boldsymbol{\beta}_i} \right)^{-1} \right),$$

$$\boldsymbol{\gamma_{oo_i}} \mid \boldsymbol{\theta}_{\backslash \gamma_{o_{ij}}} \sim N_4 \left(\left(JK\boldsymbol{\Sigma}^{-1}_{\gamma_{oi}} + \boldsymbol{v}^{-1}_{\gamma_i} \right)^{-1} \left(\boldsymbol{\Sigma}^{-1}_{\gamma_{oi}} \sum_{j=1}^{J} \sum_{k=1}^{K} \boldsymbol{\gamma_{o_{ij}}} + \boldsymbol{v}^{-1}_{\gamma_i} \boldsymbol{m}_{\gamma_i} \right), \left(JK\boldsymbol{\Sigma}^{-1}_{\gamma_{oi}} + \boldsymbol{v}^{-1}_{\gamma_i} \right)^{-1} \right).$$

B.2.4: Full conditional distributions for the error variance parameter $\sigma^2_{\epsilon_{jk}}$ and variance inflation parameter τ_i.

$$1/\sigma^2_{\epsilon_j k} \mid \boldsymbol{Y_{ijk}}, \boldsymbol{\theta}_{\backslash \sigma_{\epsilon_j k}} \sim Gamma \left(a_{\epsilon_{jk}} + \tfrac{I \times D \times T \times L}{2}, \tfrac{1}{2} \sum_{d,t,\ell,i} (y_{ijk\ell}(d,t) - m_{ijk}(d,t))^2 \tau_i + b_{\epsilon_{jk}} \right),$$

where $m_{ijk}(d,t) = \mathcal{B}(d, \boldsymbol{\phi}_{ijk})' \boldsymbol{\beta}_{ijk} + \mathcal{B}(t, \boldsymbol{\psi}_{ijk})' \boldsymbol{\gamma}_{ijk} + \alpha_{ijk}$.

$$\tau_i \mid \boldsymbol{Y_{ijk}}, \boldsymbol{\theta}_{\backslash \tau_i} \sim Gamma \left(\tfrac{\nu}{2} + \tfrac{J \times K \times L \times D \times T}{2}, \tfrac{1}{2} \sum_{d,t,\ell,k,j} \tfrac{(y_{ijk\ell}(d,t) - m_{ijk}(d,t))^2}{\sigma^2_{\epsilon_j k}} + \tfrac{\nu}{2} \right),$$

where $m_{ijk}(d,t) = \mathcal{B}(d, \boldsymbol{\phi}_{ijk})' \boldsymbol{\beta}_{ijk} + \mathcal{B}(t, \boldsymbol{\psi}_{ijk})' \boldsymbol{\gamma}_{ijk} + \alpha_{ijk}$.

A.4: Full conditional distributions for other variance parameters.

$$1/\sigma^2_{\alpha_{ij}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\alpha_{ij}}} \sim Gamma \left(a_{\alpha_{ij}} + \tfrac{K}{2}, b_{\alpha_{ij}} + \tfrac{1}{2} \textstyle\sum_{k=1}^{K} (\alpha_{ijk} - \alpha_{o_{ij}})^2 \right),$$

$$1/\sigma^2_{\beta_{ijp}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\beta_{ijp}}} \sim Gamma \left(a_{\beta_{ij}} + \tfrac{K}{2}, b_{\beta_{ij}} + \tfrac{1}{2} \textstyle\sum_{k=1}^{K} (\beta_{ijkp} - \beta_{o_{ijp}})^2 \right),$$

$$1/\sigma^2_{\gamma_{ijp}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\gamma_{ijp}}} \sim Gamma \left(a_{\gamma_{ij}} + \tfrac{K}{2}, b_{\gamma_{ij}} + \tfrac{1}{2} \textstyle\sum_{k=1}^{K} (\gamma_{ijkp} - \gamma_{o_{ijp}})^2 \right),$$

$$1/\sigma^2_{\alpha_{oi}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\alpha_{oi}}} \sim Gamma \left(a_{\alpha_i} + \tfrac{J}{2}, b_{\alpha_i} + \tfrac{1}{2} \textstyle\sum_{j=1}^{J} (\alpha_{o_{ij}} - \alpha_{oo_i})^2 \right),$$

$$1/\sigma^2_{\beta_{oip}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\beta_{oip}}} \sim Gamma \left(a_{\beta_i} + \tfrac{J}{2}, b_{\beta_i} + \tfrac{1}{2} \textstyle\sum_{j=1}^{J} (\beta_{o_{ijp}} - \beta_{oo_{ip}})^2 \right),$$

$$1/\sigma^2_{\gamma_{oip}} \mid \boldsymbol{\theta}_{\backslash \sigma_{\gamma_{oip}}} \sim Gamma \left(a_{\gamma_i} + \tfrac{J}{2}, b_{\gamma_i} + \tfrac{1}{2} \textstyle\sum_{j=1}^{J} (\gamma_{o_{ijp}} - \gamma_{oo_{ip}})^2 \right),$$

where $p = 1, .., 4$.

B.3 Additional Ranking Results

The model presented in Appendix B.1, and the ranking methods presented in Chapter 3, were used to rank the eight nanoparticles within cell lines and outcomes, within outcomes but aggregated across cell-lines, and aggregated across cell-lines and outcomes. The four responses include mitochondrial superoxide formation (MSF), loss of mitochondrial membrane potential (MMP), elevated intracellular calcium (EIC), and cellular membrane damage (CMD). The nanomaterials include: silver (Ag), gold (Au), platinum (Pt), iron oxide

(Fe_3O_4), aluminum oxide (Al_2O_2), silicon dioxide (SiO_2), zinc oxide (ZnO), and quantum dot (QD). The two screened cell lines, both related to inhalation toxicity, include bronchial epithelial cell lines (BEAS-2B) and macrophage (RAW 264.7) cell lines. Ranks were initially constructed assuming equal weights for outcomes and cell-lines. Based on knowledge about oxidative stress pathways and the assays used to measure these outcomes, it is believed that the outcomes measured can be ranked in order of importance as follows: (CMD, MSF, MMP, EIC). Using the weight function described in (3.8) and a value of $\delta = .75$, we derived weights $(.37, .27, .21, .15)$, for the four outcomes. Similarly, using a slightly more aggressive $\delta = .5$, we derived weights $(.54, .27, .12, .07)$. In Chapter 3, we presented tables and figures of these ranks by minimizing squared-error loss. In this appendix, we present results for ranks constructed by minimizing the percentile loss functions presented in Chapter 3. Specifically, we would like to optimize ranks to identify the 25% most toxic fraction of total ENM (top 2 ENM).

Table B.1 (*left panel*) provides posterior expected ranks, and associated 95% posterior intervals ranks, for each particle, aggregated across all cytotoxicity outcomes and cell-lines, with each outcome and cell-line weighted equally. Table B.1 (*middle panel* and *right panel*) provide overall summaries, aggregated across outcomes and cell-lines, for the posterior ranks using the weight functions described above. Additionally, Table B.2 and B.3 provide rankings for each particle, within outcomes but across cell lines, and within outcomes and experiments, respectively. In each of these cases, we compute expected ranks and 95% posterior intervals by minimizing the upper 25% loss functions.

Similar to the results found in Chapter 3, we see that in terms of the overall ranks, when we use weights favoring the most important outcomes, QD is significantly more toxic then the remaining seven nanoparticles.

Equal Weights	
ENM	Rank
Ag	4.5 (4.5,4.5)
Au	4.5 (4.5,4.5)
Pt	4.5 (4.5,4.5)
Al_2O_3	4.5 (4.5,4.5)
Fe_3O_4	4.5 (4.5,4.5)
SiO_2	4.5 (4.5,4.5)
QD	4.5 (4.5,4.5)
ZnO	4.5 (4.5,4.5)

$w_j = (.37, .27, .21, .15)$	
ENM	Rank
QD	1 (1,4.5)
Ag	5 (5,5)
Au	5 (5,5)
Pt	5 (5,5)
Al_2O_3	5 (5,5)
Fe_3O_4	5 (5,5)
SiO_2	5 (5,5)
ZnO	5 (5,5)

$w_j = (.54, .27, .12, .07)$	
ENM	Rank
QD	1 (1,1)
Ag	5 (5,5)
Au	5 (5,5)
Pt	5 (5,5)
Al_2O_3	5 (5,5)
Fe_3O_4	5 (5,5)
SiO_2	5 (5,5)
ZnO	5 (5,5)

Table B.1: **Overall rankings based on upper** 25% **loss functions.** Aggregated ranks across each outcome and cell-line. Posterior expected ranks and 95% posterior intervals computed by minimizing squared error loss. Each cell-line is given equal weight and each outcome (CMD,MSF,MMP,EIC) is given varying weights w_j.

MSF		EIC		CMD		MMP	
ENM	Rank	ENM	Rank	ENM	Rank	ENM	Rank
QD	1 (1,1)	Fe3O4	1 (1,4.5)	QD	1 (1,1)	ZnO	1 (1,4.5)
Fe_3O_4	5 (1.5,5)	Ag	5 (5,5)	ZnO	5 (1.5,5)	Fe3O4	5 (1.5,5)
Ag	5 (5,5)	Au	5 (5,5)	Ag	5 (5,5)	Ag	5 (5,5)
Au	5 (5,5)	Pt	5 (5,5)	Au	5 (5,5)	Au	5 (5,5)
Pt	5 (5,5)	Al2O3	5 (5,5)	Pt	5 (5,5)	Pt	5 (5,5)
Al_2O_3	5 (5,5)	SiO2	5 (5,5)	Al2O3	5 (5,5)	Al2O3	5 (5,5)
SiO_2	5 (5,5)	QD	5 (5,5)	Fe3O4	5 (5,5)	SiO2	5 (5,5)
ZnO	5 (5,5)	ZnO	5 (5,5)	SiO2	5 (5,5)	QD	5 (5,5)

Table B.2: **Rankings within outcomes based on upper** 25% **loss functions.** Aggregated ranks across each each outcome and aggregated across cell-lines. Posterior expected ranks and 95% posterior intervals computed by minimizing squared error loss. Each cell-line is given equal weight.

	MSF		EIC		CMD		MMP	
Cell-line	ENM	Rank	ENM	Rank	ENM	Rank	ENM	Rank
RAW 264.7	QD	1 (1,1)	Pt	1 (1,1)	QD	1.5 (1.5,1.5)	Fe3O4	1.5 (1.5,1.5)
	Pt	2 (2,5)	Fe3O4	2 (2,5)	ZnO	1.5 (1.5,1.5)	SiO2	1.5 (1.5,1.5)
	Fe_3O_4	3 (2,5)	Ag	3 (2.5,5)	Ag	5.5 (5.5,5.5)	Ag	5.5 (5.5,5.5)
	SiO_2	4 (4,5)	ZnO	4 (2.5,5)	Au	5.5 (5.5,5.5)	Au	5.5 (5.5,5.5)
	Ag	6.5 (6,6.5)	Au	6.5 (6.5,6.5)	Pt	5.5 (5.5,5.5)	Pt	5.5 (5.5,5.5)
	Au	6.5 (6,6.5)	Al2O3	6.5 (6.5,6.5)	Al2O3	5.5 (5.5,5.5)	Al2O3	5.5 (5.5,5.5)
	Al_2O_3	6.5 (6,6.5)	SiO2	6.5 (6.5,6.5)	Fe3O4	5.5 (5.5,5.5)	QD	5.5 (5.5,5.5)
	ZnO	6.5 (6,6.5)	QD	6.5 (6.5,6.5)	SiO2	5.5 (5.5,5.5)	ZnO	5.5 (5.5,5.5)
BEAS-2B	QD	1 (1,1)	Fe3O4	1.5 (1.5,1.5)	QD	1 (1,1)	ZnO	1 (1,1)
	Au	2 (2,5)	QD	1.5 (1.5,1.5)	Ag	2 (2,5)	QD	2 (2,5)
	Fe_3O_4	3 (2.5,5)	Ag	5.5 (5.5,5.5)	ZnO	3 (2,5)	Au	3 (2.5,5)
	Al_2O_3	4 (2.5,5)	Au	5.5 (5.5,5.5)	Au	6 (6,6)	Fe3O4	4 (2.5,5)
	Ag	6.5 (6.5,6.5)	Pt	5.5 (5.5,5.5)	Pt	6 (6,6)	SiO2	5 (5,5)
	Pt	6.5 (6.5,6.5)	Al2O3	5.5 (5.5,5.5)	Al2O3	6 (6,6)	Ag	7 (6.5,7)
	SiO_2	6.5 (6.5,6.5)	SiO2	5.5 (5.5,5.5)	Fe3O4	6 (6,6)	Pt	7 (6.5,7)
	ZnO	6.5 (6.5,6.5)	ZnO	5.5 (5.5,5.5)	SiO2	6 (6,6)	Al2O3	7 (6.5,7)

Table B.3: **Individual rankings based on upper** 25% **loss functions.** Posterior expected ranks and 95% posterior intervals computed by minimizing squared error loss. Each cell-line is given weight and each outcome (CMD,MSF,MMP,EIC) is given weights: $(.37, .27, .21, .15)$.

APPENDIX C

Appendix C: Relating ENM Properties to Toxicity

C.1 Full Conditional Distributions

In this appendix, we describe some of the full conditional distributions for the model described in the Chapter 3. Let $y_{ij}(d,t)$ denote a multivariate response corresponding to ENM i $(i = 1, ..., I)$ and replicate j $(j = 1, ..., J)$, at some dose d in $\boldsymbol{d} = (d_1, ...d_D)'$ and some time t in $\boldsymbol{t} = (t_1, ...t_T)'$.

Also let $\boldsymbol{\omega} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_\epsilon, \boldsymbol{\sigma_\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\rho})$ denote the full parameter vector, and let $\boldsymbol{\omega}_{\backslash \boldsymbol{\delta}}$ denote the vector containing all components of $\boldsymbol{\omega}$ except for some parameter δ in $\boldsymbol{\omega}$. Moreover, we denote with $\boldsymbol{Y_i}$ the complete set of response values for particle i. Finally, let $\boldsymbol{H}$ denote a $(M_d M_t) \times N$ dimensional design matrix, which can be defined as $(\mathcal{B}_1(d)\mathcal{B}_1(t), \ldots,$ $\mathcal{B}_{m_d}(d)\mathcal{B}_{m_t}(t), \ldots \mathcal{B}_{M_d}(d)\mathcal{B}_{M_t}(t))$, and X an $I \times (P+1)$ dimensional design matrix, which includes the P covariates. Using the notation above we define the full conditional distributions for all available parameters as follows.

C.1.1: Full conditional distributions for the random effect parameters α_i and $\boldsymbol{\beta_i}$.

$$\alpha_i \mid \boldsymbol{Y_i}, \boldsymbol{\omega}_{\backslash \alpha_i} \sim N\left(\frac{\frac{\sum_{d,t,k} \tilde{y}_{\alpha_{ij}}(d,t)}{\sigma_\epsilon^2/\tau_i} + \frac{m_{alphai}}{s_{\alpha_i}}}{\frac{D\times T\times J}{\sigma_\epsilon^2/\tau_i} + \frac{1}{\sigma_{\alpha_i}}}, \frac{1}{\frac{D\times T\times J}{\sigma_\epsilon^2/\tau_i} + \frac{1}{\sigma_{\alpha_i}}} \right),$$

where

$$\tilde{y}_{\alpha_{ij}}(d,t) = \begin{cases} y_{ij}(d,t) & \text{if } \gamma_i = 0 \\ y_{ij}(d,t) - \boldsymbol{H\beta} & \text{if } \gamma_i = 1 \end{cases}$$

$$\boldsymbol{\beta_i} \mid \boldsymbol{Y_i}, \boldsymbol{\omega}_{\setminus \boldsymbol{\beta_i}} \sim N_{M_d M_t}$$
$$\left(\left(\boldsymbol{\Sigma}_{\boldsymbol{\beta_i}}^{-1} + \sum_{d,t,k} \frac{\boldsymbol{H'H}}{\sigma^2_{\epsilon_j}/\tau_i} \right)^{-1} \left(\boldsymbol{\Sigma}_{\boldsymbol{\beta_i}}^{-1} \boldsymbol{m}_{\beta_i} + \sum_{d,t,j} \frac{\boldsymbol{H}\tilde{y}_{\beta_{ij}}(d,t)}{\sigma^2_\epsilon/\tau_i} \right), \left(\boldsymbol{\Sigma}_{\boldsymbol{\beta_i}}^{-1} + \sum_{d,t,j} \frac{\boldsymbol{H'H}}{\sigma^2_\epsilon/\tau_i} \right)^{-1} \right),$$

where $\tilde{y}_{\beta_{ij}}(d,t) = y_{ij}(d,t) - \alpha_i$ and $\Sigma_{\beta_i} = \sigma^2_{\beta_i}(K_d \otimes K_t)$.

C.1.2: Full conditional distributions for the error variance parameter σ^2_ϵ and variance inflation parameter τ_i.

$$1/\sigma^2_\epsilon \mid \boldsymbol{Y_i}, \boldsymbol{\omega}_{\setminus \sigma_\epsilon} \sim Gamma\left(a_\epsilon + \tfrac{I \times D \times T \times J}{2}, \tfrac{1}{2} \sum_{d,t,j,i} (y_{ij}(d,t) - m_i(d,t))^2 \tau_i + b_\epsilon \right),$$

where

$$m_i(d,t) = \begin{cases} \alpha_i & \text{if } \gamma_i = 0 \\ \boldsymbol{H\beta} + \alpha_i & \text{if } \gamma_i = 1. \end{cases}$$

$$\tau_i \mid \boldsymbol{Y_i}, \boldsymbol{\omega}_{\setminus \tau_i} \sim Gamma\left(\tfrac{\nu}{2} + \tfrac{J \times D \times T}{2}, \tfrac{1}{2} \sum_{d,t,j} \tfrac{(y_{ij}(d,t) - m_i(d,t))^2}{\sigma^2_\epsilon} + \tfrac{\nu}{2} \right),$$

where

$$m_i(d,t) = \begin{cases} \alpha_i & \text{if } \gamma_i = 0 \\ \boldsymbol{H\beta} + \alpha_i & \text{if } \gamma_i = 1. \end{cases}$$

C.1.3: Full conditional distributions for other variance parameters.

$$1/\sigma^2_{\beta_{i\ell}} \mid \boldsymbol{\omega}_{\setminus \sigma_{\beta_{i\ell}}} \sim Gamma\left(a_{\beta_i} + \tfrac{M_d M_t}{2}, b_{\beta_i} + \tfrac{1}{2}(\boldsymbol{\beta}_i - m_\beta)'(K_d \otimes K_t)(\boldsymbol{\beta}_{ij} - m_\beta) \right).$$

C.1.4: Full conditional distributions for covariate parameters θ_p and Z_i.

$$Z_i \mid \boldsymbol{\theta}, \gamma_i = 1 \sim N\left(x_i'\boldsymbol{\theta}, 1\right) I(Z_i \leq 0),$$

$$Z_i \mid \boldsymbol{\theta}, \gamma_i = 0 \sim N\left(x_i'\boldsymbol{\theta}, 1\right) I(Z_i > 0),$$

$$\boldsymbol{\theta} \mid \boldsymbol{Z}, \boldsymbol{\gamma} \sim N_{P+1}\left((X_\rho' X_\rho)^{-1} X_\rho' Z, (X_\rho' X_\rho)^{-1} \right).$$

C.1.5: Derivation of $P(\boldsymbol{\gamma} \mid \boldsymbol{Y}, \boldsymbol{\omega}_{\setminus\boldsymbol{\gamma}})$.

$$
\begin{aligned}
P(\boldsymbol{\gamma} \mid \tilde{\boldsymbol{Y}}, \boldsymbol{\omega}_{\setminus\gamma}) &\propto P(\tilde{\boldsymbol{Y}} \mid \boldsymbol{\omega}_{\setminus\boldsymbol{\beta}}) P(\boldsymbol{\gamma} \mid \boldsymbol{\theta}) \\
P(\tilde{\boldsymbol{Y}} \mid \boldsymbol{\omega}_{\setminus\boldsymbol{\beta}}) &= \int P(\tilde{\boldsymbol{Y}} \mid, \boldsymbol{\omega}) P(\boldsymbol{\beta} \mid \boldsymbol{\omega}_{\setminus\boldsymbol{\beta}})\, d\boldsymbol{\beta}_{\gamma i} \\
&= \int \prod_{i=1}^{I} (2\pi)^{\frac{-DTKJ}{2}} (2\pi)^{\frac{-M_t M_d J_\gamma}{2}} \left(\frac{\sigma_\epsilon^2}{\tau_i}\right)^{-\frac{1}{2}} \mid \boldsymbol{\Sigma}_{\boldsymbol{\beta i}} \mid^{-\frac{1}{2}} \\
&\quad exp\left\{-\tfrac{1}{2}\left(\tfrac{\tau_i}{\sigma_\epsilon^2}(\tilde{\boldsymbol{Y}}_{\boldsymbol{i}} - \boldsymbol{X}_{\boldsymbol{\gamma i}})'(\tilde{\boldsymbol{Y}}_{\boldsymbol{i}} - \boldsymbol{X}_{\boldsymbol{\gamma i}}) + (\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}} \boldsymbol{\beta}_{\boldsymbol{\gamma}_i})\right)\right\} \\
&= C_1 \int exp\left\{-\frac{1}{2}\left(\frac{\tau_i}{\sigma_\epsilon^2}(\tilde{\boldsymbol{Y}}'_{\boldsymbol{i}} \tilde{\boldsymbol{Y}}'_{\boldsymbol{i}} + \boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} \boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - 2\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})\right.\right. \\
&\quad \left.\left. + \boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}} \boldsymbol{\beta}_{\boldsymbol{\gamma}_i}\right)\right\} d\boldsymbol{\beta}_{\gamma i} \\
&= C_1 exp\left\{-\tfrac{1}{2}\, \tfrac{\tau_i}{\sigma_\epsilon^2}(\tilde{\boldsymbol{Y}}'_{\boldsymbol{i}} \tilde{\boldsymbol{Y}}'_{\boldsymbol{i}})\right\} \\
&\quad \int exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i}(\frac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} + \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}}) \boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - 2\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} \frac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}}\right)\right\} d\boldsymbol{\beta}_{\gamma i} \\
&= C_1 C_2 \int exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} V^{\text{-}1} \boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - 2\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} D\right)\right\} d\boldsymbol{\beta}_{\gamma i} \\
&= C_1 C_2 exp\left\{\tfrac{1}{2}\,(D' A A^{\text{-}1} A D)\right\} \\
&\quad exp\left\{-\tfrac{1}{2}\left(\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} V^{\text{-}1} \boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - 2\boldsymbol{\beta}'_{\boldsymbol{\gamma}_i} V^{\text{-}1} V D + D' V V^{\text{-}1} V D\right)\right\} \\
&= C_1 C_2 exp\left\{\tfrac{1}{2}((\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})'(\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} + \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}})^{\text{-}1}(\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}}))\right\} \\
&\quad \int exp\left\{-\frac{1}{2}\left((\boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - VD)' V^{\text{-}1} (\boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - VD)\right)\right\} d\boldsymbol{\beta}_{\gamma i} \\
&= C_1 C_2 C_3 \int exp\left\{-\frac{1}{2}\left((\boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - VD)' V^{\text{-}1} (\boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - VD)\right)\right\} d\boldsymbol{\beta}_{\gamma i} \\
&= C_1 C_2 C_3 (2\pi)^{\frac{M_t M_d J}{2}} \mid V \mid^{\frac{1}{2}} \\
&\quad \int (2\pi)^{\frac{-M_t M_d J}{2}} \mid V \mid^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\left((\boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - VD)' V^{\text{-}1} (\boldsymbol{\beta}_{\boldsymbol{\gamma}_i} - VD)\right)\right\} d\boldsymbol{\beta}_{\gamma i} \\
&= \prod_{i=1}^{I} (2\pi)^{\frac{-DTKJ}{2}} \left(\frac{\sigma_\epsilon^2}{\tau_i}\right)^{-\frac{1}{2}} \mid \boldsymbol{\Sigma}_{\boldsymbol{\beta i}} \mid^{-\frac{1}{2}} \mid (\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} + \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}}) \mid^{-\frac{1}{2}} \\
&\quad exp\left\{-\tfrac{1}{2}\left(\tfrac{\tau_i}{\sigma_\epsilon^2} \tilde{\boldsymbol{Y}}'_{\boldsymbol{i}} \tilde{\boldsymbol{Y}}'_{\boldsymbol{i}}\right)\right\} \\
&\quad exp\left\{\tfrac{1}{2}\left((\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})'(\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} + \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}})^{\text{-}1}(\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})\right)\right\} \\
&\propto \prod_{i=1}^{I} \mid \boldsymbol{\Sigma}_{\boldsymbol{\beta i}} \mid^{-\frac{1}{2}} \mid (\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} + \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}}) \mid^{-\frac{1}{2}} \\
&\quad exp\left\{\tfrac{1}{2}\left((\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})'(\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \boldsymbol{X}_{\boldsymbol{\gamma i}} + \boldsymbol{\Sigma}^{\text{-}1}_{\boldsymbol{\beta i}})^{\text{-}1}(\tfrac{\tau_i}{\sigma_\epsilon^2} \boldsymbol{X}'_{\boldsymbol{\gamma i}} \tilde{\boldsymbol{Y}}_{\boldsymbol{i}})\right)\right\}.
\end{aligned}
$$

C.2 Model Assessment

In this appendix we discuss model assessment. First we assess goodness of fit using the conditional predictive ordinate (cpo), as described by Geisser (1980). Next we plot the probability integral transform (PIT) histogram, as a measure of predictive performance (Gneiting et al. (2007)). Finally, we present some graphical posterior predictive checks.

C.2.1: Conditional Predictive Ordinate (CPO)

The conditional predictive ordinate (CPO) is a diagnostic tool for detecting observations with poor model fit. If we let $\boldsymbol{Y}$ denote the complete set of responses, let $\boldsymbol{Y}_{-k}$ denote observation $\boldsymbol{Y}$ with the k-th component omitted, and let $\boldsymbol{Y}_k^{obs}$ denote the kth component of observation $\boldsymbol{Y}$, then CPO_k can be defined as follows:

$$\begin{aligned} CPO_i &= \pi(\boldsymbol{Y}_k^{obs} \mid \boldsymbol{Y}_k) = \int \pi(\boldsymbol{Y}_k^{obs} \mid \boldsymbol{Y}_{-k}, \boldsymbol{\omega})\pi(\boldsymbol{\omega} \mid \boldsymbol{Y}_{-k})\, \mathrm{d}\boldsymbol{\omega}, \\ \pi(\boldsymbol{Y}_k^{obs} \mid \boldsymbol{Y}_{-k}, \boldsymbol{\omega}) &= \frac{1}{\int \frac{\pi(\boldsymbol{\omega} \mid \boldsymbol{Y})}{\pi(\boldsymbol{Y_k^{obs}} \mid \boldsymbol{\omega})}\, \mathrm{d}\boldsymbol{\omega}}. \end{aligned} \tag{C.1}$$

Here, $\boldsymbol{\omega} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_\epsilon, \boldsymbol{\sigma}_\beta, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\rho})$ denotes the full parameter vector. Given N MCMC samples, $n = 1, ..., N$, from the posterior distribution $P(\boldsymbol{\omega} \mid \boldsymbol{Y})$, we can obtain the harmonic mean estimate of CPO_k as follows:

$$\hat{CPO}_i = \frac{N}{\sum_{n=1}^{N} 1/\pi\left(\boldsymbol{Y_k^{obs}} | \boldsymbol{\omega}_k^{(n)}\right)}. \tag{C.2}$$

The expression above is evaluated at posterior samples $\boldsymbol{\omega}_k^{(1)}, \ldots, \boldsymbol{\omega}_k^{(N)}$.

A plot of -log(CPO_k) can be used to diagnose poor model fit. Large values of -log(CPO_k) indicate observations that are not consistent with the model. The *top* panel of Figure C.1, provides a plot of -log($\hat{CPO}_i(d,t)$) for the model and data described in Chapter 4. Overalll values of -log($\hat{CPO}_i(d,t)$) are relatively low, indicating good model fit. The *middle* panel indicates that the largest values of -log($\hat{CPO}_i(d,t)$) tend to be observations with large exposure times, This is to be expected, as cell death is followed after sometime by the

dissolution of cell nuclei, hindering the measurement of cellular responses.

C.2.2: Probability Integral Transform (PIT)

The probability integral transform (PIT), as described by Gneiting et al. (2007), is frequently used as a measure of posterior predictive calibration. Here calibration is defined as the statistical consistency between the posterior predictive distribution and the observed responses $\boldsymbol{Y}$. The PIT is described as the value of the observed response $\boldsymbol{Y_k}$ attained under the predictive cumulative distribution function. Using the same notation as above, the PIT can be defined as follows:

$$PIT_k = \int P(\boldsymbol{Y}_k \leq \boldsymbol{\omega})\pi(\boldsymbol{\omega})\mathrm{d}\boldsymbol{\omega} = P_k(\boldsymbol{Y}_k). \tag{C.3}$$

Given N MCMC samples, $n = 1, ..., N$, from the posterior distribution $P(\boldsymbol{\omega} \mid \boldsymbol{Y})$, we can estimated PIT as follows:

$$\hat{PIT}_k = \tfrac{1}{N} \sum_{n=1}^{N} I(\boldsymbol{Y_k} \leq \tilde{\boldsymbol{Y}}_k^{(n)}). \tag{C.4}$$

where $\boldsymbol{Y}_{rep\,k}^{(n)}$ is a sample from the posterior predictive distribution.

A plot of the PIT histogram can be used to visually assess the calibration of the model. Under good predictive performance of the model, the PIT histogram has a uniform distribution (see Diebold et al. (1997) for a formal proof). Inspection of the PIT histogram can also indicate reasons for poor predictive performance. A hump-shaped PIT histogram indicates prediction intervals that are, on average, to wide due to over dispersion of the predicative distribution. A U-shaped PIT histogram indicates that the predictive distribution is too narrow. Finally, a triangle shaped PIT histogram corresponds to biased predictive distributions (Gneiting et al. 2007).

The *bottom* panel of Figure C.1 provides a plot of the PIT histogram for the entire model, including all doses, times, and particles. Visual assessment indicates that the plot does tend toward uniformity, indicating good overall predictive performance.

C.2.2: Posterior Predictive Checking

A common tool for model checking in Bayesian inference involves posterior predictive checks. The basic idea behind posterior predictive checking is that if the model is a good fit to

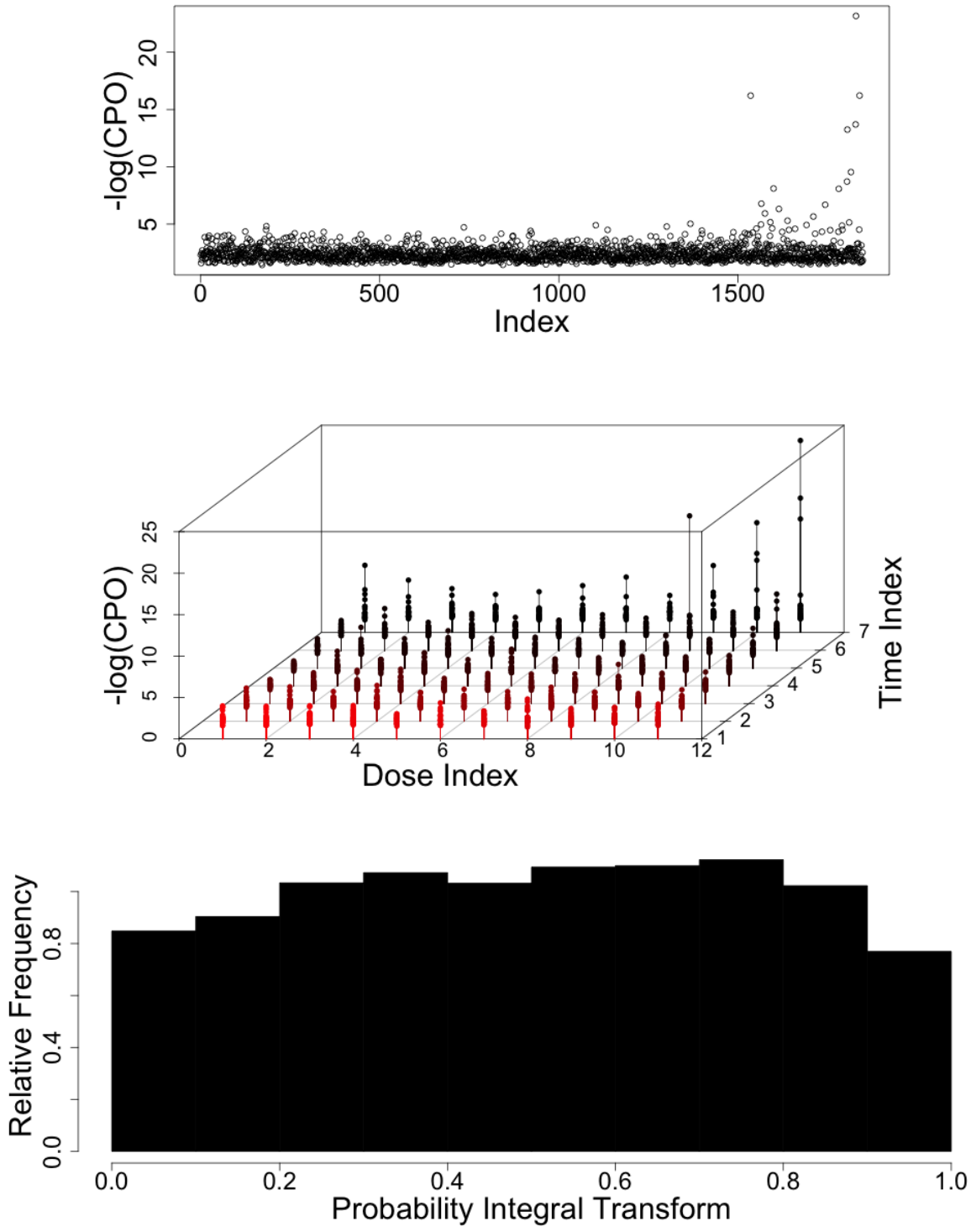

Figure C.1: **Graphical model diagnostics.** *(*Top) Estimate of $-\log(cpo_i(d,t))$ for detecting observations with poor model fit. *(*Middle) Plot of $-log(cpo_i(d,t))$ as a function of dose and time, indicating any relationship between outlying observations and the administered dose or duration of exposure. *(*Bottom) Probability Integral Transform assessing empirical calibration of the posterior predictive distribution.

the data, then data replicated under the model should resemble the observed response $\boldsymbol{Y}$. In posterior predictive checking, replicate samples $\boldsymbol{Y_{rep}}$, are simulated from the posterior predictive distribution and compared to the observed data $\boldsymbol{Y}$. Potential problems with the model can be detected by looking for systematic differences between the simulated posterior predictive samples and the observed response. Using the same notation described above, the posterior predictive distribution can be described as follows:

$$p(\boldsymbol{Y}_{rep} \mid \boldsymbol{Y}) = \int P(\boldsymbol{Y}_{rep} \mid \boldsymbol{\omega}) P(\boldsymbol{\omega} \mid \boldsymbol{Y}) \mathrm{d}\boldsymbol{\omega}. \tag{C.5}$$

Given N MCMC samples, $n = 1, ..., N$, from the posterior distribution $P(\boldsymbol{\omega} \mid \boldsymbol{Y})$, we can draw samples $\boldsymbol{Y}_{rep}^{(n)}$, $n = 1, ..., N$, from the posterior predictive distribution

Diagnostics of posterior predictive performance are obtained by comparing draws from the posterior predictive distribution to the observed data, using both formal tests and graphical checks. Graphical model checking involves the display of the simulated data from the posterior predictive distribution alongside the observed data $\boldsymbol{Y}$, and visually looking for large discrepancies such as lack of coverage (Gelman et al. 2004).

Figures C.2 and C.3 provide plots of the distribution of the posterior predictive mean response averaged across all doses and times of exposure (black), for each particle. The mean and associated 95% posterior intervals for the posterior predictive distribution are marked using vertical lines (black). Also included is the empirical mean response across all doses and times of exposure (red). Figure C.4, summarizes these results by plotting the mean and 95% posterior intervals of the posterior predictive mean response (black), along with the the empirical mean response across all doses and times (red), for each particle. In all cases the empirical mean response is contained within the 95% posterior intervals of the posterior predictive mean distribution, indicating relatively good posterior coverage across all doses and times of exposure.

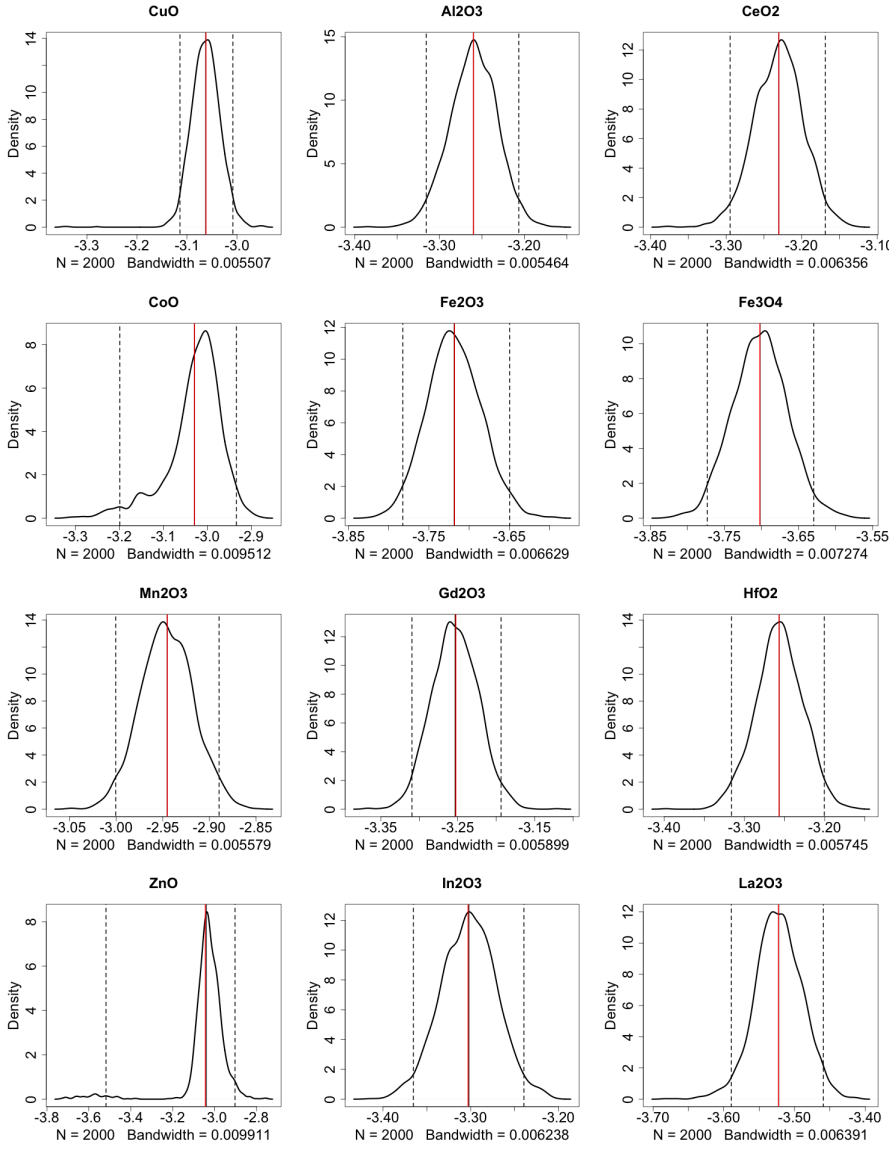

Figure C.2: **Posterior predictive mean distributions for CuO, Al_2O_3, CeO_2, CoO, Fe_2O_3, Fe_3O_4, Mn_2O_3, Gd_2O_3, HfO_2, ZnO, In_2O_3, and La_2O_3 ENMs.** For each particle we plot the distribution of the posterior predictive mean response across all doses and times of exposure (black), along with the mean (solid black line) and associated 95% posterior intervals (dotted black lines) for this distribution. Also included is the empirical mean response across all doses and times of exposure (red).

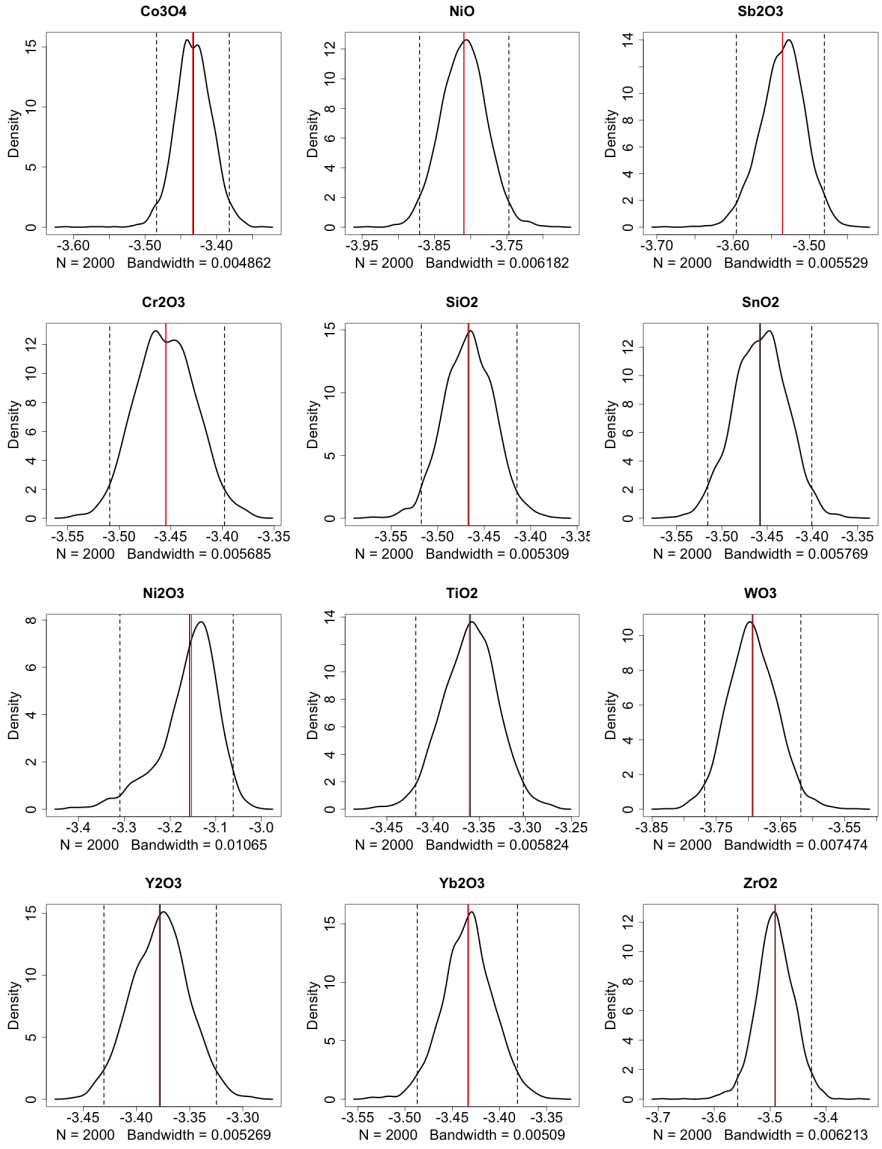

Figure C.3: **Posterior predictive mean distributions for Co_3O_4, NiO, Sb_2O_3, Cr_2O_3, SiO_2, SnO_2, Ni_2O_3, TiO_2, WO_3, Y_2O_3, Yb_2O_3, and ZrO_2 ENMs.**] For each particle we plot the distribution of the posterior predictive mean response across all doses and times of exposure (black), along with the mean (solid black line) and associated 95% posterior intervals (dotted black lines) for this distribution. Also included is the empirical mean response across all doses and times of exposure (red).

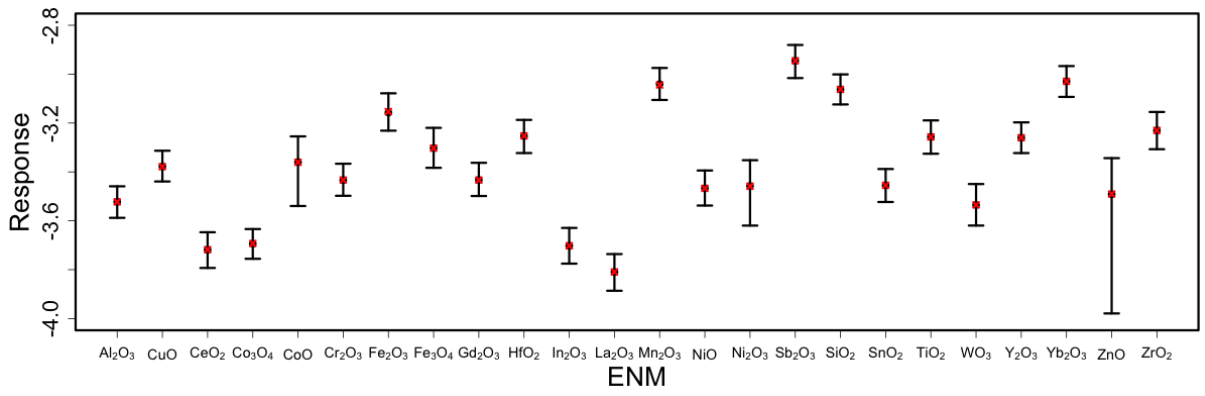

Figure C.4: **Summary of posterior predictive mean coverage.** Mean and 95% posterior intervals of the posterior predictive mean response across all doses and times of exposure, for all 24 particles.) Also included are the empirical mean responses across all doses and times of exposure (red).

Bibliography

Albert, J. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of American Statistical Association 88*(422), 669–679.

Baladandayuthapani, V., B. Mallick, and R. Carroll (2005). Spatially adaptive *b*ayesian p-splines with heteroscedastic errors. *Journal of Computational & Graphical Statistics 14*, 378–622.

Barlow, S., J. Greig, J. Bridges, A. Carere, A. Carpy, C. Galli, J. Kleiner, I. Knudsen, H. Koeĺ ter, L. Levy, C. Madsen, S. Mayer, J.-F. Narbonne, F. Pfannkuch, M. Prodanchuk, M. Smith, and P. Steinberg (2002). Hazard identification by methods of animal-based toxicology. *Food and Chemical Toxicology 40*, 145–191.

Besag, J. and D. Higdon (1999). Bayesian analysis of agricultural field experiments. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 61*(4), 691–746.

Besag, J. and C. Kooperberg (1995). On conditional and intrinsic autoregression. *Biometrika 82*, 733–746.

Burello, E. and P. Worth (2011). Qsar modeling of nanomaterials. *WIREs Nanomedicine and Nanobiotechnology 3*, 298–306.

Calabrese, E. and L. Baldwin (2003). Toxicology rethinks its central belief. *Nature 421*, 691–692.

Clayton, D. (1996). *Markov Chain Monte Carlo in Practice, eds.* London: Chapman and Hall.

DECHEMA/VCI (2011). *10 Years of Research: Risk Assessment, Human and Environmental Toxicology of Nanomaterials.* Germany: DECHEMA/VCI working group: Responsible Production and Use of Nanomaterials.

DeSemet, Y., J. Sprinagael, and P. Kunsch (2002). Towards statistical multicriteria decision modeling: A first approach. *Journal of Multi-Criteria Decision Analysis 11*(6), 305–313.

Diebold, F., T. Gunther, and A. Tay (1997). Evaluating density forecasts. *International Economic Review 39*, 863–883.

Dwork, C., R. Kumar, M. Naor, and D. Sivakumar (2001). Rank aggregation methods for the web. *Proceedings of the 10th International Conference on World Wide Web 10*, 613–622.

Edler, L., K. Poirier, M. Dourson, J. Kleiner, B. Mileson, H. Nordmann, A. Renwick, W. Slob, K. Walton, and G. Würtzen (2002). Mathematical modelling and quantitative methods. *Food and Chemical Toxicology 40*, 283–326.

Emmens, C. (1940). The dose-response relation for certain principles of the pituitary gland, and of the serum and urine of pregnancy. *Journal Of Endocrinology 2*, 194–225.

Fagin, R., R. Kumar, and D. Sivakumar (2003). Comparing top-k lists. *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms 17*(1), 134–160.

Finney, D. (1979). Bioassay and the practice of statistical inference. *International Statistical Review 47*, 1–12.

Geisser, S. (1980). Discussion on sampling and bayes inference in scientific modeling and robustness. *Journal of the Royal Statistical Society: Series A 143*, 416–417.

Gelfand, A. and A. Smith (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association 85*, 398–409.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis 4*, 515–533.

Gelman, A., J. Carlin, H. Stern, and D. Rubin (2004). *Bayesian Data Analysis*. Boca Raton, London, New York, Washington, D.C.: Chapman and Hall/CRC.

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

George, S., S. Pokhrel, T. Xia, B. Gilbert, Z. Ji, M. Schowalter, A. Rosenauer, R. Damoiseaux, K. Bradley, L. Madler, and A. Nel (2009). Use of a rapid cytotoxicity screening approach to engineer a safer zinc oxide nanoparticle through iron doping. *ACS Nano 4*(1), 15–29.

George, S., T. Xia, R. Rallo, Y. Zhao, Z. Ji, S. Lin, X. Wang, H. Zhang, B. France, D. Schoenfeld, R. Damoiseaux, R. Liu, S. Lin, K. Bradley, Y. Cohen, and A. Nel (2011). Use of a high-throughput screening approach coupled with in vivo zebrafish embryo screening to develop hazard ranking for engineered nanomaterials. *ACS Nano 5*(3), 1805–1817.

Geys, H., M. Regan, P. Catalano, and G. Molenberghs (2001). Two latent variable risk assessment approaches or mixed continuous and discrete outcomes from developmental toxicity data. *Journal of Agricultural, Biological, and Environmental Statistics 6*(3), 340–355.

Gneiting, T., F. Balabdaoui, and A. Raftery (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69*(2), 243–268.

Green, P. (1995). Reversible jump markov chain monte carlo computation and Bayesian model determination. *Biometrika 82*(4), 711–732.

Hastie, T. and R. Tibshirani (1986). Generalized additive models. *Statistical Science 1*(3), 297–318.

Hill, A. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *The Journal of Physiology 40*, iv–vii.

Hoheisel, J. (2006). Microarray technology: Beyond transcript profiling and genotype analysis. *Nature Review Genetics 7*, 200–210.

Johnson, V. and J. Albert (1999). *Ordinal Data Modeling*. New York: Springer-Verlag.

Kong, M. and R. Eubank (2006). Monotone smoothing with application to dose-response curve. *Communications in Statistics–Simulation and Computation 35*, 991–1004.

Lang, S. and A. Brezger (2004). Bayesian p-splines. *Journal of Computational and Graphical Statistics 13*(1), 183–212.

Lang, S. and A. Brezger (2006). Generalized structured additive regression based on Bayesian p-splines. *Computational Statistics and Data Analysis 50*, 967–991.

Lee, K., N. Sha, E. Dougherty, V. M., and B. Mallick (2003). Gene selection: a Bayesian variable selection approach. *Bioinformatics 19*(1), 90–97.

Lerche, D., R. Brüggemann, P. SØrensen, L. Carlsen, and O. Nielsen (2002). A comparison of partial order technique with three methods of multi-criteria analysis for ranking of chemical substances. *Journal of Chemical Information and Computer Sciences 42*, 1086–1098.

Lerche, D., S. Matsuzaki, P. SØrensen, L. Carlsen, and O. Nielsen (2004). Ranking of chemical substances based on the japanese pollutant release and transfer register using partial order theory and random linear extensions. *Chemosphere 55*, 1005– 1025.

Lerche, D. and P. SØrensen (2003). Evaluation of the ranking probabilities for partial orders based on random linear extensions. *Chemosphere 53*, 981– 992.

Li, C. and D. Hunt (2004). Regression splines for threshold selection with applications to a random-effects logistic dose-response model. *Computational Statistics and Data Analysis 46*(1), 1–9.

Lilienblum, W., W. Dekant, H. Foth, T. Gebel, J. Hengstler, R. Kahl, P. Kramer, H. Schweinfurth, and K. Wollin (2008). Alternative methods to safety studies in experimental animals: Role in the risk assessment of chemicals under the new european chemicals legislation (reach). *Archives of Toxicology 82*(4), 211– 236.

Lin, R., T. Louis, S. Paddock, and G. Ridegeway (2006). Loss function based ranking in two-stage, hierarchical models. *Bayesian Analysis 1*(4), 915–946.

Lin, R., T. Louis, S. Paddock, and G. Ridegeway (2009). Ranking usrds provider specific smrs from 1998Ð2001. *Health Services Outcomes Reseach Methodology 4*, 515–533.

Liu, R., R. Rallo, S. George, Z. Ji, S. Nair, A. Nel, and Y. Cohen (2011). Classification nanosar development for cytotoxicity of metal oxide nanoparticles. *Small 7*(8), 1118–1126.

Lockwood, J., T. Louis, and D. McCaffrey (2002). Uncertainty in rank estimation: Implications for value-added modeling accountability systems. *Journal of Educational and Behavioral Statistics 27*(3), 255–270.

Louis, T. and W. Shen (1999). Innovations in bayes and empirical bayes methods: estimating parameters, populations and ranks. *Statistics in Medicine 18*, 2493Ű2505.

Maynard, A., R. Aitken, T. Butz, V. Colvin, K. Donaldson, G. Oberdörster, M. Philbert, J. Ryan, A. Seaton, V. Stone, S. Tinkle, L. Tran, N. Walker, and D. Warheit (2006). Safe handling of nanotechnology. *Nature Biotechnology 444*, 267–268.

Meng, H., M. Liong, T. Xia, Z. Li, Z. Ji, J. Zink, and N. A. E (2010). Engineered design of mesoporous silica nanoparticles to deliver doxorubicin and p-glycoprotein sirna to overcome drug resistance in a cancer cell line. *ACS nano 4*(8), 4539Ű4550.

Nel, A., L. Mädler, D. Velegol, T. Xia, E. Hoek, P. Somasundaran, F. Klaessig, V. Castranova, and M. Thompson (2009). Understanding biophysicochemical interactions at the nano−bio interface. *Nature Materials 8*, 543–557.

Nel, A., T. Xia, L. Mädler, and N. Li (2006). Toxic potential of materials at the nanolevel. *Science 311*(5761), 622–627.

Noma, H., M. Matsui, T. Omori, and T. Sato (2010). Bayesian ranking and selection methods using hierarchical mixture models in microarray studies. *Biostatistics 11*(2), 281–289.

Plummerm, M., N. Best, K. Cowles, and K. Vines (2006, March). CODA: Convergence diagnosis and output analysis for MCMC. *R News 6*(1), 7–11.

Puzyn, T., D. Leszczynska, and J. Leszczynski (2009). Toward the development of "nano-qsars": Advances and challenges. *Small 5*(22), 2494Ű2509.

Puzyn, T., B. Rasulev, A. Gajewicz, X. Hu, T. Dasari, A. Michalkova, H. Hwang, A. Toropov, D. Leszczynska, and J. Leszczynski (2011). Using nano-qsar to predict the cytotoxicity of metal oxide nanoparticles. *Nature Nanotechnology 6*, 175–178.

Ramsay, J. (1988). Monotone regression splines in action. *Statistical Science 3*(4), 425–461.

Ramsey, J. and B. Silverman (2005). *Functional Data Analysis*. New York: Springer.

Regan, M. and P. Catalano (1999). Bivariate dose-response modeling and risk estimation in developmental toxicology. *Journal of Agricultural, Biological, and Environmental Statistics 4*(3), 217–237.

Ritz, C. (2010). Toward a unified approach to dose-response modeling in ecotoxicology. *Environmental Toxicology and Chemistry 29*(1), 220–229.

Roberts, O. and J. Rosenthal (2001). Optimal scaling for various metropolis hastings algorithms. *Statistical Science 16*(4), 351–367.

Schultz, W., M. Cronin, and T. Netzevab (2003). The present status of qsar in toxicology. *Journal of Molecular Structure: THEOCHEM 622*(1-2), 23–38.

Scott, J. and J. Berger (2006). An exploration of aspects of Bayesian multiple testing. *Journal of Statistical Planning and Inference 136*(7), 2144–2162.

Severini, T. and J. Stainwalis (1994). Quasi-likelihood estimation in semiparametric models. *Journal of the American Statistical Association 89*, 501–512.

Shen, W. and T. Louis (1998). Triple-goal estimates in two-stage, hierarchical models. *Journal of the Royal Statistical Society 60*, 455–471.

Society, R. (2004). Nanoscience and nanotechnologies: Opportunities and uncertainties. *The Royal Society, Science Policy Section, London, England*, 2494–2509.

Stanley, S., E. Westly, M. Pittet, A. Subramanian, S. Schreiber, and R. Weissleder (2008). Pertubational profiling of nanomaterial biologic activity. *Proceedings of the National Academy of Sciences 105*(21), 7387–7392.

Stern, S. and S. McNeil (2008). Nanotechnology safely concerns revisited. *Toxicological Sciences 101*(1), 4–21.

Telesca, D. and L. Inoue (2009). Differential expression and network inferences through functional data modeling. *Biometrics 65*(3), 793–804.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics 22*(4), 1701–1729.

Tsuji, J., A. Maynard, P. Howard, J. James, C. Lam, D. Warheit, and A. Santamaria (2005). Research strategies for safety evaluation of nanomaterials, part iv: Risk assessment of nanoparticles. *Toxicological Sciences 89*(1), 42Ű50.

West, M. (1984). Outlier models and prior distributions in Bayesian linear regression. *Journal of the Royal Statistical Association. Series B (Methodological) 46*(3), 431–439.

White, R. (2000). High-throughput screening in drug metabolism and pharmacokinetic support of drug discovery. *Annual Review of Pharmacology and Toxicology 40*, 133–157.

Xia, T., M. Kovochich, J. Brant, M. Hotze, J. Sempf, T. Oberley, C. Sioutas, J. Yeh, M. Wiesner, and N. AE (2006). Comparison of the abilities of ambient and manufactured nanoparticles to induce cellular toxicity according to an oxidative stress paradigm. *Nano Letters 6*(8), 1794–1807.

Yu, Z. and P. Catalano (2005). Quantitative risk assessment for multivariate continuous outcomes with application to neurotoxicology: The bivariate case. *Biometrics 61*, 757–766.

Zhang, H., J. Zhaoxia, T. Xia, H. Meng, C. Low-Kam, L. Rong, P. Suman, S. Lin, X. Wang, L. Yu-Pei, M. Wang, L. Linjiang, R. Rallo, R. Damoiseaux, D. Telesca, L. Madler, Y. Cohen, J. Zink, and A. Nel (2012). Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. *ACS Nano*, 1–74.