Journal of Environmental Statistics
November 2015, Volume 7, Issue 6. http://www.jenvstat.org

A Bayesian Model-Based Approach for Determining Multivariate Tolerable Regions

Edward L. Boone
Department of Statistical Sciences
and Operations Research
Virginia Commonwealth University

J. Paul Brooks
Department of Statistical Sciences
and Operations Research
Virginia Commonwealth University

Epiphanie Nyirabahizi
Epidemiology and Public Health
School of Public Health
Yale University

Xi Chen
Industrial & Systems Engineering
Virginia Tech

Abstract

A crucial concern of toxicologists is to determine an acceptable exposure level(s) to a hazardous substance(s). Often lab experiments produce data featuring multiple hazards and multiple outcome measures. The current practice evaluates each hazard and outcome combination separately, which leads to multiple statistical tests that suffer from inflated Type I error rates. This paper introduces a Bayesian model-based approach for analyzing data of similar nature. This approach is dimension-preserving in that it permits simultaneous quantification of an acceptable exposure level among multiple hazards. Furthermore, we introduce the concept of significance probabilities to assess the importance of the outcomes in determining an acceptable exposure level. The proposed methodology is motivated and illustrated through analyzing the dataset from a rodent study of pesticides on neurotoxicity conducted by Moser *et al.* (2005).

Keywords: environmental health; Bayesian modeling and analysis; Markov chain Monte Carlo; simultaneous inference.

1. Introduction

Toxicologists are concerned with determining and reducing the risk of an *adverse effect* on an organism that is exposed to potentially toxic materials. In laboratory settings toxicolo-

gists consider the effect of a single toxic substance (stressor) on a single outcome (endpoint) via rodent studies, even though they may be making measurements on multiple endpoints. This practice is clearly a poor use of resources as the collected data are not fully utilized. In fact, an exposure to the stressor of interest often affects multiple endpoints, all of which could potentially signal an adverse effect. Thus determining a single acceptable dosage of an exposure can prove difficult. As an illustration, consider the work of Moser *et al.* (2005), who attempted to determine the effect of exposures to pesticides on neurotoxicity using laboratory experiments with rats. The problem is that neurotoxicity is ill-defined as it can be evidenced by several distinct endpoints. For rats, neurotoxicity can be measured in terms of endpoints such as blood cholinesterase, brain cholinesterase, motor activity, and tail pinch response, and a decrease in any measurement is considered adverse. The issue can be further complicated in more general settings by the fact that organisms are typically exposed to multiple stressors in a form of a mixture of chemicals/compounds or even under diverse environmental conditions (Monosson 2005). For example, commercial pesticides often come as a combination of several pesticides or different pesticides may be applied at different stages of the plant growth, resulting in an application of a combination of pesticides. In either case, any organism consuming the plant will be exposed to more than a single pesticide.

The US Environmental Protection Agency (USEPA) is required to establish guidelines for exposures to stressors. These guidelines lead to the establishment of the Point Of Departure (POD), which is the level of an exposure at which an adverse effect is detected. There are two common methods of determining POD's: the No Observable Adverse Effect Level (NOAEL) and the Benchmark Dose (BMD). The NOAEL approach makes the statistical comparison of the mean response of a variety of dose groups against that of the control group. Specifically, the NOAEL is considered to be the highest dose level for which the mean response of the dose groups is not significantly different from the mean of the control group. One flaw of NOAEL is that the responses may be highly sensitive to the stressors, and hence any dose given could result in an observable adverse effect. Consequently, from such experiments one may end up concluding that no level of exposure is tolerable (e.g., Crump 1984; Leisenring and Ryan 1992; Slob 1999). Follow up experiments could be conducted at lower and lower dose levels to determine the NOAEL, but each successive experiment would result in an even lower NOAEL (Crump 2002; Scholze and Kortenkamp 2007). Hence the main deficiency of NOAEL is that it may not be easily found for the stressors of interest.

1.1. The Benchmark Dose Approach

The use of BMD was developed due to the aforementioned issue with NOAEL. Instead of looking directly at differences from the control, BMD, typically, fits a monotone invertible function f to explain the relationship between the dose of a stressor and the measured endpoint (Crump 1984). To use BMD one first specifies the value of the response that is considered unacceptably adverse, known as the Benchmark Response (BMR). For simplicity the BMR is typically indexed by an η% change from the control group, denoted by BMR_η. For example, suppose one deems a 10% change from the control group to be the acceptable upper bound for an adverse effect; then for a particular endpoint the BMR_{10} is the value of that endpoint which corresponds to the 10% change. Once BMR_η is specified we use f to find the dosage that corresponds to this value of BMR_η, which is known as the Benchmark Dose, BMD_η. Considering the uncertainty associated with the estimated parameters in f and the unavoidable experimental errors, a C% confidence interval about BMD_η can be further calculated.

The lower bound of this confidence interval is the Benchmark Dose Lower bound, BMDL_η, and it is the greatest dosage at which one is $C\%$ confident that an unacceptable adverse effect on the endpoint under consideration will not occur. A common POD adopted in practice is BMDL_η. In fact, the USEPA accepts the replacement of NOAEL with BMD whenever appropriate quantitative data are available (USEPA 1995, 2002). Yu and Catalano (2008) assessed BMD's behavior relative to BMR and recommended to use higher BMR (e.g., 5% and 10%) given the variability of the BMD distribution due to background noise. A good overview of the BMD methodology and its application is given by Piegorsch and Bailer (2005).

For a single response variable (endpoint), y_i, resulting from a single predictor (stressor), x_i, one can determine the BMD as follows. One first specifies the acceptable percentage change from the control/placebo, η, and then translates that into, BMR_η, the response value that is considered the maximum acceptable level of an adverse effect. For instance, if an adverse effect of no more than 5% is claimed acceptable, then we set $\text{BMR}_\eta = 5\%$; if the response scale is from 0 to 100 and higher values correspond to a more adverse effect, then set $\eta = 5$. Once η is determined, a link function that relates the measurement on the endpoint y_i and the dose levels of the stressor x_i can be established

$$y_i = f(\beta_0 + \beta_1 x_i), \tag{1}$$

where $\boldsymbol{\beta}$ denotes the vector of governing parameters. One then solves for $\text{BMD}_\eta = x^*$ through

$$f^{-1}(1 - \eta/100) = \beta_0 + \beta_1 x^*. \tag{2}$$

To account for parameter uncertainty, one can determine a BMDL on x^* by constructing a $C\%$ confidence interval for x^* (USEPA 1995). The parameter vector $\boldsymbol{\beta}$, the standard errors for $\boldsymbol{\beta}$, and x^* can be estimated via a quasi-Newton iterative algorithm (Proc NLMIXED in SAS), profile likelihood (Pawitan 2001), or Markov Chain Monte Carlo methods based on Bayesian analysis (Tao *et al.* 2004).

The aforementioned approach works for a single pair of stressor and endpoint only. However, in practice multiple endpoints are often considered for risk assessment (Ryan 1992; D. W. Gaylor *et al.* 1998). Extending BMD to situations that concern multiple endpoints is nontrivial, and it poses challenges for data analysis and interpretation due to the non-comparability aspect of BMDs for different stressors (Englehardt 2004). For a single stressor risk assessment, the analysis should focus on the most sensitive endpoint. Establishing the most sensitive endpoint typically requires the use of multiple statistical tests that result in inflated Type I error rates. Correcting methods such as Bonferroni correction if adopted typically reduce the power of the tests to detect the effect of interest. Furthermore, such results lack an overall interpretation on the effect of the stressor under consideration.

To avoid solely considering the most sensitive endpoint, several approaches have been proposed such as using multiple endpoints in the same model or combining multiple endpoints into a single composite endpoint. Sammel *et al.* (1997) and Regan and Catalano (1999b) used a latent-variable model for mixed discrete and continuous correlated endpoints. Coffey *et al.* (2007) proposed an overall score based on a desirability function for analyzing multiple-endpoint data. The overall score is an outcome of a dimension reduction approach that converts the multiple-endpoint dataset into a single composite endpoint dataset. Regan and Catalano (1999a) evaluated the developmental effects of Ethylene Glycol (EG) on fetus malformation and fetal weight by considering separate BMRs for the two endpoints,

and they employed generalized estimating equations to incorporate correlations between the endpoints. Yu and Catalano (2005) developed a likelihood model that allows separate dose-response models for all endpoints while accounting for the bivariate correlation to obtain an overall characterization of the risk associated with an exposure.

Several Bayesian approaches have been proposed for dealing with multiple-endpoint experiments. Choi *et al.* (2010) and Faes *et al.* (2004) used Bayesian hierarchical models for analyzing multiple-endpoint data that result from the exposure to Perchlorate and data from the exposure to EG on mice, respectively. Faes *et al.* (2004) proposed a two-stage Bayesian hierarchical structure in which the first-stage models the probability that a fetus is non-viable and the second stage models the probability that a survivor fetus has a malformation. Choi *et al.* (2010) and Faes *et al.* (2004), however, only considered models for exposures to a single chemical/stressor and each endpoint was modeled separately. Geys *et al.* (1999) proposed a pseudolikelihood approach for modeling multivariate endpoints. Coull *et al.* (2003) used Bayesian hierarchical models to synthesize information for several endpoints across several studies. Li et al. (2008) used parametric and semiparametric logistic methods in a Bayesian framework; however, they only modeled for multi-binary data. Budtz-Jørgensen (2007) considered using structural equation models to determine the BMD for multiple endpoints and multiple stressors. Our formulation is similar to Dunson (2000), who considered combining multiple endpoints via random effects. We combine BMD with the random effects under a Bayesian framework to determine which dosages and combination of dosages are tolerable.

The remainder of this paper is organized as follows. Section 2 presents the organophosphate pesticide (OP) dataset collected by Moser *et al.* (2005), which is used to demonstrate the methodology we propose. Traditional BMD analysis is conducted assuming that the stressors are independent of each other. Section 3 develops the modeling approach and establishes the definitions of BenchMark Dose Tolerable Region and BenchMark Dose Tolerable Set, respectively, for stressors and endpoints. Section 4 demonstrates the proposed method by analyzing the OP dataset. Section 5 concludes with a discussion and suggests potential enhancements to the proposed method.

2. The Organophosphate Pesticide (OP) Dataset

Moser *et al.* (2005) investigated the neurotoxic effects of organophosphates pesticides (OP) which are commonly used in agriculture. OP pesticides are suspected of producing neurotoxic effects, however, the outcome of neurotoxicity has been ill-defined and is often measured as a combination of several endpoints. Moser *et al.* (2005) considered the following neurotoxicity endpoints: Brain cholinesterase (BrainCHE) (y_1), Blood cholinesterase (BloodCHE) (y_2), motor activity (y_3), and tail pinch (y_4). These endpoints were considered in conjunction with exposures to the following pesticides (i.e., stressors): acephate (ACE), diazinon (DIA), chlorpyriphos (CPF), dimethoate (DTO) and malathion (MAL). The pesticides were given in ranging doses to a total of 349 rats. Specifically, ACE doses ranged from 0 to 120 (mg/kg), DIA doses ranged from 0 to 250 (mg/kg), DTO doses ranged from 0 to 75 (mg/kg) and MAL doses ranged from 0 to 500 (mg/kg). Each of the pesticides was applied when all other pesticides had dose 0; in addition, a variety of pesticide mixtures was also given. The stressors were given orally to Long/Evans rats and the endpoints above were measured on each rat. Since the measurement on BrainCHE required the animal be sacrificed, only one dose combination and one set of endpoint measurements was available for each rat. See Moser

et al. (2005) for details of the data collection process. One interesting complexity with this dataset is that tail pinch is a binomial outcome as it counts the total number of positive reactions to a pre-specified number of pinches. This complexity does not allow modeling the data as a multivariate normal random vector. Due to differences in tail pinch protocols across different experimental sites the number of trials associated with tail pinch varied and values of 7, 8, 12 or 16 were observed.

Analysis is conducted for each stressor and endpoint combination *separately* and the BMDL_{50} for all stressors are calculated using a method consistent with that used by Moser *et al.* (2005). For the continuous endpoints, motor activity, BrainCHE and BloodCHE, the following dose-response model is used to describe the relationship between a stressor and a given endpoint:

$$y_{ji} \sim N\left(e^{\beta_{0j}+x_i\beta_{1j}}, \sigma^2\right), \ j = 1, 2, 3,$$

where the index i denotes the ith subject, β_{0j} and β_{1j} are the regression parameters and x_i is the dose of the stressor given to subject i. As mentioned above, the tail pinch endpoint is not continuous as it arose from a binomial experiment in which the number of positive reactions to a set of n_i tail pinches was counted. Recall that the value of n_i was not consistent across all experimental sites. Therefore, we adopt the following model for this endpoint:

$$y_{4i} \sim Binom\left(n_i, \frac{1}{1+\exp(\beta_{04}+x_i\beta_{14})}\right).$$

Table 1 shows the resulting BMDL_{50} for all endpoint-stressor combinations.

Table 1: BMDL_{50} resulting from an independent analysis of each endpoint considered separately from exposures to ACE, DIA, CPF, DTO and MAL, measured in mg/kg.

Endpoint	ACE	DIA	CPF	DTO	MAL
BrainChE	21.73	111.51	36.77	23.76	500
BloodChE	23.79	7.92	1.85	17.61	500
Motor activity	24.33	131.23	32.45	51.57	500
Tail pinch	56.44	160.44	37.73	48.23	500
$\min(\text{Endpoint}_j)$	21.73	7.92	1.85	17.61	500

The goal of the analysis is not only to determine the BMDL_{50} for each endpoint-stressor combination but also the BMDL_{50} for a stressor across all endpoints. Notice from Table 1 that for stressor ACE the BMDL_{50} associated with BrainChE is 21.73 mg/kg. This is the smallest BMDL_{50} for ACE across all endpoints. Hence any exposure to ACE that is below this level would be deemed tolerable across all endpoints. Since BrainChE is the endpoint that defines the BMDL_{50} across all endpoints we say that it is the most sensitive endpoint for ACE. Similarly for stressor DIA, the endpoint BloodChE is most sensitive. Notice that motor activity and tail pinch do not contribute to defining the BMDL_{50} for any of the stressors. Therefore one would conclude from Table 1 that BrainChE and BloodChE are considered as the most sensitive endpoints among the four as they determine the BMDL_{50} for all stressors. While the analysis above gives very important information it ignores any additive effects from combining various stressors and it fails to determine whether a combination of two stressors will lead to an unacceptable adverse effect.

In the next section we propose a novel unified Bayesian model-based approach to evaluating dose-response relationships among multiple endpoints that result from exposures to multiple stressors. This approach will equip toxicology researchers with a powerful systematic tool to find a dosage *region* (in the stressor space) that controls for an acceptable level analogous to the BMD. Furthermore, we define the notions of *hypersensitive*, *cosensitive* and *hyposensitive* which reflect the importance of each endpoint in defining the tolerable region of dosages. In addition, we illustrate how to quantify the importance of an endpoint. The aforementioned issues are of vital importance to toxicology researchers; however, to the best of our knowledge, none of these have been addressed by the existing methods. We hope the proposed methodology will help fill the research gap.

3. The Bayesian Model-Based Approach

In this section we present in detail our Bayesian model-based approach to evaluating dose-response relationships. The detailed steps are provided and a general algorithm is presented at the end of this section.

Suppose that we are interested in learning the dose-response relationships among J endpoints and K stressors through an experiment conducted on n subjects. Let y_{ij} be the response of the ith subject on the jth endpoint. For the ith subject the J endpoint measurements can be put into the $J \times 1$ vector $Y_i = (y_{i1}, y_{i2}, ..., y_{iJ})^\top$, $i = 1, 2, \ldots, n$. The $K \times 1$ vector $X_i = (x_{i1}, x_{i2}, ..., x_{iK})^\top$ gives doses on the K stressors applied to the ith subject. The relationships between the K doses and the J endpoints can be described by the $J \times 1$ vector of functions $\mathbf{f} = (f_1, f_2, ..., f_J)^\top$, where the f_j's are assumed to be monotone and invertible. Let $\boldsymbol{\beta}_j$ be the $K \times 1$ vector of parameters that corresponds to the jth endpoint and let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_J)$. We model the response of the ith subject on the J endpoints due to an exposure to the K doses as

$$Y_i = \mathbf{f}(X_i^\top \boldsymbol{\beta}) = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iJ} \end{pmatrix} = \begin{pmatrix} f_1(X_i^\top \boldsymbol{\beta}_1 + r_i) \\ f_2(X_i^\top \boldsymbol{\beta}_2 + r_i) \\ \vdots \\ f_J(X_i^\top \boldsymbol{\beta}_J + r_i) \end{pmatrix}, \; i = 1, 2, ..., n. \tag{3}$$

This formulation allows each endpoint to have its own model while using a subject specific random effect term r_i to capture the within subject effect across endpoints. For consistency, in formulation (3) all the K stressors in vector X_i are included in each of the J endpoint models. For a given $\text{BMR}_\eta = (\eta_1, \eta_2, ..., \eta_J)^\top$ the invertibility of the f_j's permits the BMD or the tolerable region, T, to be determined for all endpoints simultaneously by solving the following system of inequalities,

$$\mathbf{f}^{-1}(\eta) = \begin{pmatrix} f_1^{-1}\left((1-\eta_1/100)\widehat{f}_1(\mathbf{0})\right) \\ f_2^{-1}\left((1-\eta_2/100)\widehat{f}_2(\mathbf{0})\right) \\ \vdots \\ f_J^{-1}\left((1-\eta_J/100)\widehat{f}_J(\mathbf{0})\right) \end{pmatrix} \leq \begin{pmatrix} \boldsymbol{\beta}_1^\top \\ \boldsymbol{\beta}_2^\top \\ \vdots \\ \boldsymbol{\beta}_J^\top \end{pmatrix} \cdot \; X^* \;\; , \; X^* \geq 0, \tag{4}$$

where $\widehat{f}_j(\mathbf{0})$ is the estimated function value on endpoint j when zero dosage is given for

all stressors. Notice that any combination of dosage of stressors, X^*, that satisfies (4) is considered tolerable with respect to the given BMR_η.

3.1. Prior and Posterior Distributions

In this subsection we give the likelihood for the data, the prior distributions associated with the parameters, and details on how to obtain samples from the posterior distribution. Due to the possible complexity of handling both continuous and discrete endpoints, care should be exercised in constructing the likelihood as a multivariate normal distribution is no longer appropriate. Let $g_j(y_{ij}|f_j(X_i^\top\boldsymbol{\beta}_j+r_i),\gamma_j)$ be the probability density for the response measurement of the ith subject on the jth endpoint, $j = 1, 2, ..., J$; the γ_js are additional parameters that may be required to define the distribution. Given the $\boldsymbol{\beta}_j$'s and γ_j's the endpoints are conditionally independent and hence the likelihood can be constructed as follows,

$$L(Y|\{X_i\}_{i=1}^n,\boldsymbol{\beta}) = \prod_{i=1}^{n}\prod_{j=1}^{J} g_j(y_{ij}|f_j(X_i^\top\boldsymbol{\beta}_j + r_i),\gamma_j). \tag{5}$$

We specify the following proper prior distributions:

$$\begin{aligned}
\boldsymbol{\beta}|\boldsymbol{\mu},\boldsymbol{\Omega} &\sim N(\boldsymbol{\mu},\boldsymbol{\Omega})\\
\boldsymbol{\mu}|a,A &\sim N(a,A)\\
\boldsymbol{\Omega}|R,\rho &\sim Wishart(R,\rho)\\
\gamma_j &\sim p(\gamma_j)\\
r_i|\varpi^2 &\sim N(0,\varpi^2)\\
\varpi^2 &\sim Gamma(1,1)\ ,
\end{aligned} \tag{6}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top,\boldsymbol{\beta}_2^\top,\ldots,\boldsymbol{\beta}_J^\top)^\top$ denotes the vector of regression coefficients (a slight abuse of notation without causing difficulty in understanding); $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ represent the mean vector and the precision matrix, respectively, for $\boldsymbol{\beta}$. Here R is a specified correlation matrix and ρ gives the associated degrees of freedom. The prior distribution for γ_j, $p(\gamma_j)$, is chosen appropriately; and a and A are hyperparameters governing $\boldsymbol{\mu}$.

Samples from the posterior distribution are obtained by Markov Chain Monte Carlo (MCMC) techniques via standard open source MCMC sampling software such as WinBUGS, OpenBUGS or JAGS, etc. These samples should be analyzed to ensure convergence of the algorithm and satisfactory quality of the samples. For more details on MCMC methods and diagnostics, see Gelman *et al.* (2005).

3.2. Evaluation of the Benchmark Dose Tolerable Area

At a specified BMR_η each inequality in (4) defines a tolerable region T_j associated with the jth endpoint, $j = 1, 2, \ldots, J$. The intersection of all J tolerable regions, $T = \cap_{j=1}^J T_j$, gives the BenchMark Dose Tolerable Region (BMDTR), which is analogous to BMD. For a given $\boldsymbol{\beta}$, the dosages X^*, that satisfy the inequalities in (4) are considered to present an acceptable risk of adverse effect. Furthermore, each MCMC sample of $\boldsymbol{\beta}^{(m)}$ leads to a BMDTR, $T^{(m)}$. Define $A_{T^{(m)}}$ as the area of $T^{(m)}$ corresponding to the mth MCMC sample,

$$A_{T^{(m)}} = \int_{T^{(m)}} dX\ , m = 1, 2, \ldots, M,$$

where M denotes the total MCMC samples generated. The tolerable areas $\{A_{T^{(m)}}\}_{m=1}^{M}$ can then be ranked and a lower bound on the area of the tolerable region is defined by the MCMC sample of $\boldsymbol{\beta}$ that corresponds to the qth quantile of the $\{A_{T^{(m)}}\}_{m=1}^{M}$, e.g., $q = 5$. This gives a $100 \times q\%$ credible region of the BMDTR, parallel to the traditional BMDL.

3.3. Hyposensitive, Hypersensitive and Cosensitive Endpoints

Endpoints are not necessarily of equal importance, hence one needs to determine which endpoints are most sensitive to the stressors. Recall that the endpoint setting the minimum dose for a particular stressor is deemed the most sensitive endpoint for that stressor. Consider the OP data analysis given in Table 1. It has been identified that for ACE the most sensitive endpoint is BrainChE and for DIA it is BloodChE. There the most sensitive endpoint is found by intersecting the intervals reaching from 0 to the respective BMDLs across all the endpoints under investigation. Here we introduce some key concepts to define a sensitive endpoint more rigorously which is central to our proposed method. Let T_j be the tolerable region corresponding to endpoint y_j. Recall that the BMDTR is given by intersecting the tolerable regions corresponding to all endpoints, i.e., BMDTR $= \cap_{j=1}^{J} T_j$.

Definition: A given endpoint y_j is said to be *hypersensitive* if that endpoint is the only one necessary to define the BMDTR, i.e. the tolerable region associated with this endpoint is the intersection of all tolerable regions. Mathematically speaking,

$$\bigcap_{l=1}^{J} T_l = T_j, \quad j \in \{1, 2, \ldots, J\}. \tag{7}$$

Definition: An endpoint y_j is said to be *hyposensitive* if that endpoint is not necessary to define the BMDTR, i.e. the intersection of all the tolerable regions can be determined without considering T_j. Mathematically speaking,

$$\bigcap_{\ell=1, \ell \neq j}^{J} T_l = \bigcap_{l=1}^{J} T_l, \quad j \in \{1, 2, \ldots, J\}. \tag{8}$$

Definition: An endpoint that is neither hypersensitive nor hyposensitive is considered to be *cosensitive*.

The BenchMark Dose Tolerable Set (BMDTS) is the set of endpoints that are hyper or cosensitive, i.e. the minimum set of endpoints necessary to define the BMDTR. (The BMDTS is the minimum cardinality set that consists of hyper or cosensitive endpoints.) Formally, the BMDTS can be written as

$$\text{BMDTS} = argmin_{y_j}\{y_j | \bigcap_{j=1}^{J} T_j = \text{BMDTR}\}.$$

The BMDTS tells researchers which endpoints are critical in defining an acceptable level of exposures to stressors. In particular, we note that hyposensitive endpoints may be omitted

from future studies if the same set of endpoints and stressors is of interest. Doing so could potentially generate significant savings in resource investment in the first place.

Let us apply the concepts above to analyze the OP dataset. In the single dimension (one-stressor) case one is simply taking intersections of intervals to find the BMDTR and the BMDTS. Take the stressor DIA for an example, the tolerable regions due to different endpoints are, respectively, $T_1 = [0, 111.51)$ for BriainChE, $T_2 = [0, 7.92)$ for BloodChE, $T_3 = [0, 131.23)$ for motor activity and $T_4 = [0, 160.44)$ for tail pinch. Since the BMDTR for DIA is given by $\cap_{j=1}^{4} T_j = T_2 = [0, 7.92)$, which corresponds to the second endpoint BloodChE, hence BloodChE would be deemed hypersensitive for DIA. In a similar fashion, BrainChE is found to be hypersensitive for ACE. Again we find that neither of the endpoints motor activity and tail pinch contributes to defining the BMDTR for any of the stressors under study, therefore they are hyposensitive.

3.4. Determine the BMDTR and Endpoint Probabilities

In addition to determining a BMDTR with a probability guarantee as explained in Subsection 3.2, the MCMC samples can be used to determine which endpoints play a role in defining the BMDTRs with corresponding probabilities. This can be achieved via the GNU Linear Programming Kit (GLPK) which is a library of routines for solving linear programming problems of the following form (written as a minimization problem):

$$\begin{aligned} &\min\ C^\top x, \\ &\text{subject to } \mathbf{G}x \leq b, \end{aligned}$$

where C denotes the vector of coefficients that specifies the linear objective function, and vector b and matrix $\mathbf{G}$ together define a system of linear constraints.

In the context of determining the BMDTS as defined in Subsection 3.3, each endpoint corresponds to a linear constraint in the linear program and the polytope defining the set of feasible solutions in fact gives the BMDTR. Notice that the BMDTS is the minimum set of endpoints whose corresponding constraints define the BMDTR. The BMDTS can be calculated by testing the feasibility of a series of linear programs as follows:

- The ℓth endpoint defines an inequality,

$$X^\top \boldsymbol{\beta}_\ell \leq f_\ell^{-1}(\eta_\ell),\ \ell = 1, 2, ..., J. \tag{9}$$

- For the jth endpoint, construct linear program LP^j by making the jth constraint active and including the remaining $J-1$ inequality constraints given by (9), i.e.,

$$X^\top \boldsymbol{\beta}_\ell = f_\ell^{-1}(\eta_\ell),\ \ell = 1, 2, ..., J,\ \ell \neq j, \tag{10}$$

  and setting $C = 0$, for $j = 1, 2, ..., J$. If LP^j has a feasible solution, then we say that the jth endpoint contributes to defining the BMDTR. This is a very computationally efficient method for determining if an endpoint is important.

Now we are in a good position to calculate endpoint *significance probabilities* by applying the aforementioned procedure to the generated MCMC samples. Recall that M gives the MCMC

sample size. We estimate the significance probability $P(y_j \in \text{BMDTS}_\eta)$ by $\widehat{P}_j$, defined as follows,

$$\widehat{P}_j = \frac{1}{M}\sum_{m=1}^{M} I\{y_j \in Z^{(m)}\},\ j = 1, 2, \ldots, J, \tag{11}$$

where $Z^{(m)}$ represents the BMDTS_η for the mth MCMC sample; $I\{U\}$ is an indicator function that takes value 1 if event U is true and 0 otherwise. The estimated significance probability for the jth endpoint, $\widehat{P}_j$, reflects the importance of that endpoint in defining the BMDTR. If the estimated significance probability of a given endpoint is close to 1, then corresponding endpoint is considered crucial in determining the BMDTR. In contrast, if the estimated significance probability is close to 0, then that endpoint is expected to play little role in defining the BMDTR; as a result, we can exclude that endpoint from further consideration with great confidence. Notice that the aforementioned approach to determining the endpoint probability heavily relies on the dose-response model specified being a generalized linear model with only first-order terms.

We give a generic algorithm to summarize the steps of the Bayesian model-based approach introduced in this section in Algorithm 1.

Algorithm 1 A generic algorithm

For $m = 1$ to M

1. Draw a sample $\boldsymbol{\beta}^m \sim p(\boldsymbol{\beta}|D)$.
2. Determine the tolerable region associated with $\boldsymbol{\beta}^m$, denoted by $T^{(m)}$.
3. Determine the area of $T^{(m)}$, $A_{T^{(m)}}$.
4. Determine the smallest set of endpoints necessary that defines $T^{(m)}$, $\text{BMDTS}_\eta^{(m)}$.

End

Rank the elements in $\{A_{T^{(m)}}\}_{m=1}^{M}$.

BMDTS_η is the $T^{(m^*)}$ that corresponds to the ηth percentile of $\{A_{T^{(m)}}\}_{m=1}^{M}$

The significance probability for the jth endpoint is estimated by $\widehat{P}_j, j = 1, 2, \ldots, J$.

4. Example

We detail an application of the proposed methodology presented in Section 3 to analyzing the OP dataset. Two examples at different levels of complexity have been studied. The first example concerns two of the five stressors and all four endpoints and the second example extends the analysis to all five stressors present in the OP dataset. For the sake of brevity, only the steps for analyzing the two-stressor case is presented. We refer the interested reader to Tables 5—7 in the Appendix for results of the five-stressor case.

4.1. Analysis of OP Data with Two Stressors

Recall the OP dataset introduced in Section 2. This subsection considers analyzing dose-response relationships of two stressors ACE (x_1) and DIA (x_2) and four endpoints BrainChE

(y_1), BloodChE (y_2), motor activity (y_3) and tail pinch (y_4). Recall that BrainChE, BloodChE and motor activity are continuous measurements which are scaled to percentage to control and tail pinch counts the number of positive reactions to n_i tail pinches.

The dose-response model adopted to describe the data is as follows:

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

where Y_i denotes the vector of responses of the ith subject on the four endpoints due to exposures to the two stressors mentioned; the other parameters can be understood without difficulty following the notation established for (3) in Section 3. The formulation given in (12) allows the three continuous endpoints and the binomial endpoint to be modeled simultaneously. Here r_i denotes a random effect that accounts for within subject correlations across the endpoints for the ith subject. From the model given in (12) the likelihood can be easily constructed.

Given that the dose-response model has been defined for each endpoint and likelihood has been specified, prior distributions are needed for obtaining the posterior distribution of the parameters. To this end, the following prior distributions are assigned:

$$\begin{aligned} \boldsymbol{\beta} &\sim N(\boldsymbol{\mu}, \boldsymbol{\Omega}) \\ \boldsymbol{\Omega} &\sim Wishart(\mathbf{I}_{12}, 12) \\ \boldsymbol{\mu} &\sim N(0, 100\mathbf{I}_{12}) \\ \sigma_j &\sim Gamma(1,1) \\ r_i &\sim N(0, \varpi^2) \\ \varpi^2 &\sim Gamma(1,1) \end{aligned} \tag{13}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \boldsymbol{\beta}_3^\top, \boldsymbol{\beta}_4^\top)^\top$ and $\boldsymbol{\beta}_i = (\beta_{i0}, \beta_{i1}, \beta_{i2})^\top, i = 1, 2, 3, 4$; $\mathbf{I}_{12}$ denotes the 12×12 identity matrix and other parameters are as specified for (6) in Subsection 3.1. The posterior distribution follows easily from (12) and (13),

$$\begin{aligned} P(\boldsymbol{\beta}, \boldsymbol{\Omega}, \boldsymbol{\mu}|D) \;\propto\; & \exp\left\{-\sum_{j=1}^{3}\sigma_j^2\right\} \times \exp\left\{-\frac{1}{200}\sum_{j=1}^{4}\mu_j^2\right\} \times \varpi^{-1}\exp\left\{-\frac{1}{2\varpi^2}\sum_{i=1}^{n} r_i^2\right\} \\ & \times \exp\{-\varpi^2\} \times |\boldsymbol{\Omega}|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu})^\top\boldsymbol{\Omega}^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu})\right\} \\ & \times \prod_{i=1}^{n}\begin{pmatrix} n_i \\ y_{4i} \end{pmatrix}\left(\left(1+\exp\{x_i^\top\boldsymbol{\beta}_4 + r_i\}\right)^{-1}\right)^{y_{4i}}\left(1-\left(1+\exp\{x_i^\top\boldsymbol{\beta}_4 + r_i\}\right)^{-1}\right)^{n_i - y_{4i}} \\ & \times \left(2^{72}\pi^{44}\prod_{i=1}^{12}\Gamma\left(\frac{13-i}{2}\right)\right)^{-1}|I_{12}|^{-6}|\boldsymbol{\Omega}|^{-1/2}\exp\left\{-\frac{1}{2}tr(\boldsymbol{\Omega})\right\} \\ & \times \left(\prod_{j=1}^{3}\sigma_j^2\right)^{-n/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{3}\sigma_j^{-2}\left(y_{ji}-\exp\{x_i^\top\boldsymbol{\beta}_j + r_i\}\right)^2\right\}. \end{aligned} \tag{14}$$

Table 2: The quantiles of the posterior samples and $\widehat{R}$ for the regression parameters for the OP data with exposures to stressors ACE and DIA. Credible intervals in **bold** do not contain 0. The calculation is based on 10,000 posterior samples.

Endpoint	Parameter	2.5%	50%	97.5%	$\widehat{R}$
Brain ChE	β_{10}	**-0.100**	**-0.049**	**-0.001**	1.000
	β_{11}	**-0.005**	**-0.004**	**-0.003**	1.001
	β_{12}	**-0.085**	**-0.064**	**-0.046**	1.001
Blood ChE	β_{20}	-0.085	-0.012	0.059	1.002
	β_{21}	**-0.060**	**-0.039**	**-0.026**	1.001
	β_{22}	**-0.072**	**-0.043**	**-0.025**	1.001
Motor activity	β_{30}	-0.079	-0.028	0.022	1.005
	β_{31}	**-0.003**	**-0.002**	**-0.001**	1.001
	β_{32}	**-0.033**	**-0.024**	**-0.018**	1.004
Tail pinch	β_{40}	**0.986**	**0.894**	**0.802**	1.003
	β_{41}	**0.004**	**0.002**	**0.0001**	1.000
	β_{42}	**0.011**	**0.006**	**0.001**	1.001

Computation.

OpenBUGS is used to generate 10 chains of 101,000 MCMC samples from the posterior distribution given in (14). To verify convergence of the chains traceplots are examined to guarantee good mixing and the estimated potential scale reduction $\widehat{R}$ is verified to be less than 1.006 for all parameters. The first 1,000 samples from each chain are discarded as burn-in samples. The remaining 100,000 samples are thinned by 100 to minimize autocorrelations in the samples, resulting in 1,000 samples from each chain. The 10,000 samples from the combined chains are used to draw all inferences. The sampling process takes 115 seconds on a computer with 2.3 GHz Intel Core i7 Processor with 8GB RAM.

Table 2 shows the 2.5%, 50% and 97.5% quantiles for each regression parameter. Notice that the coefficients indicate a decreasing relationship between both stressors ACE and DIA and all endpoints. Using the rule that if zero is in the credible interval then the particular parameter is not considered statistically significant from zero, it is found that except for the intercepts β_{20} and β_{30} all parameters are statistically significant. Nevertheless, we can conclude that the stressors ACE and DIA indeed have a significant effect on the endpoints.

Evaluating BMDTR.

Using the MCMC samples of $\boldsymbol{\beta}$ obtained from the posterior distribution $p(\boldsymbol{\beta}|D)$, we can determine the BMDTR_η for the stressors. Recall that BMDTR_η is the dosage region (in the stressor space) which elicits responses on all endpoints that differ no more than η% from those under the no exposure condition. In order to determine the BMDTR_η, we need to first find the corresponding BMR_η associated with each endpoint (in the response space), which can

be calculated as follows:

$$\text{BMR}_\eta = \begin{pmatrix} \ln\left((1-\eta/100)\widehat{f}_1(\mathbf{0})\right) \\ \ln\left((1-\eta/100)\widehat{f}_2(\mathbf{0})\right) \\ \ln\left((1-\eta/100)\widehat{f}_3(\mathbf{0})\right) \\ \ln\left(\frac{(1-\eta/100)\widehat{p}_4(\mathbf{0})}{1-(1-\eta/100)\widehat{p}_4(\mathbf{0})}\right) \end{pmatrix}, \tag{15}$$

where $\widehat{f}_j(\mathbf{0})$ corresponds to the estimated function value of the model for jth endpoint when no stressor is applied. This is obtained by setting the dosage for both ACE and DIA to zero and using the median values for the model parameters and evaluating f_j; Similarly, $\widehat{p}_4(\mathbf{0})$ is the calculated probability of a positive reaction to a tail pinch when using zero doses for stressors ACE and DIA and evaluating the function with the median values of the parameters. Notice that (15) provides a vector of exposure limits that defines a η% tolerable region of stressor dosages.

Using each of the MCMC samples, $\boldsymbol{\beta}^{(m)}$, we determine $A_{T^{(m)}}$ as explained in Subsection 3.2 and calculate the 5th quantile of $A_{T^{(m)}}$ and the corresponding $\boldsymbol{\beta}$. This results in a BMDTR_η for a given value of η. For different values of η, the corresponding BMDTR_η can be determined by following the approach presented in Subsection 3.4.

Consider the OP dataset. The BMDTR_{50} (i.e., η=50) is defined as the dosages (x_1^*, x_2^*) that satisfy the following constraints:

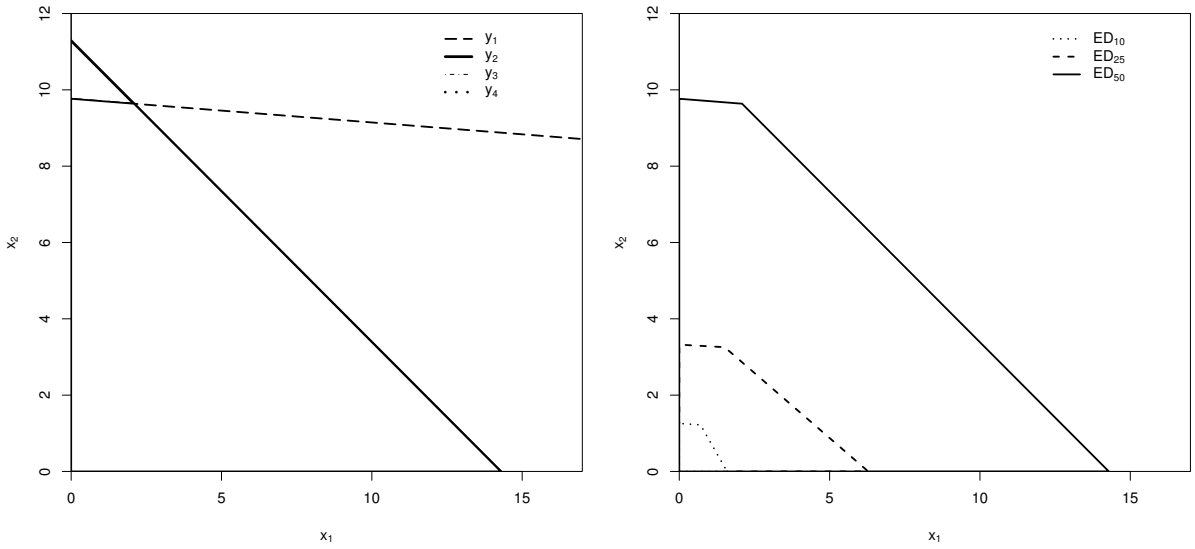
$$\begin{pmatrix} -0.742 \\ -0.705 \\ -0.720 \\ 0.597 \end{pmatrix} \leq \begin{pmatrix} -0.041 & -0.004 & -0.072 \\ -0.009 & -0.050 & -0.063 \\ -0.018 & -0.002 & -0.034 \\ 0.860 & 0.001 & 0.007 \end{pmatrix} \begin{pmatrix} 1 \\ x_1^* \\ x_2^* \end{pmatrix}, \tag{16}$$

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

Notice that the left-hand side of (16) is given by evaluating (15) at $\eta = 50$. Figure 1 shows the plot of the BMDTR_{50} associated with exposures to ACE and DIA as well as the contours of the BMDTR_η for $\eta = 10, 25$ and 50. Notice that y_1 (BrainChE) and y_2 (BloodChE) are cosensitive endpoints as these are the only two endpoints that contribute to defining the BMDTR_{50} . Furthermore, from visual inspection of Figure 1(a) it appears that y_2 is the endpoint that mostly defines the BMDTR_{50} as the edge associated with y_1 is very short. Hence one may be tempted to conclude that y_1 could be omitted from consideration along with y_3 and y_4. From Figure 1(b) observe that the shape of the tolerable region is consistent across different levels of η. More importantly, y_1 contributes to the definition of the BMDTR_η across different levels of η thus omitting y_1 would be not advised. This observation demonstrates the need to determine which endpoints the BMDTR_η is sensitive to and their corresponding sensitivity levels. To proceed we next determine the significance probability $P(y_j \in \text{BMDTS}_\eta)$ for each of endpoint under consideration.

Endpoint significance probability.

Using Figure 1 alone one can readily tell that the BMDTR_η is defined by y_1 and y_2. In higher dimensions, however, visualization of the BMDTR_η is not as straightforward or even becomes impossible. Equation (11) given in Subsection 3.4 provides a rigorous way to determine the



(a) (b)

Figure 1: (a) The BMDTR_{50} associated with exposures to ACE (x_1) and DIA (x_2) and (b) the contours of BMDTR_η for $\eta = 10, 25$ and 50.

Table 3: Estimated endpoint significance probabilities, $\widehat{P}(y_j \in \text{BMDTS})$, for the OP data with exposures to ACE and DIA. The calculation is based on 10,000 posterior samples.

	η		
Endpoint (y_j)	10	25	50
BrainChE (y_1)	0.856	0.945	0.938
BloodChE (y_2)	0.998	1	1
Motor activity (y_3)	0.001	0	0
Tail pinch (y_4)	0	0	0

significance of each endpoint by calculating the estimate $\widehat{P}(y_j \in \text{BMDTS}_\eta)$. Table 3 shows the estimated significance probabilities, $\widehat{P}(y_j \in \text{BMDTS}_\eta)$, for different levels of η. Notice that the values of $\widehat{P}(y_j \in \text{BMDTS}_\eta)$ vary across the levels of η. The magnitudes of the estimated significance probabilities of y_1 (BrainChE) and y_2 (BloodChE) imply that both endpoints play a critical role in determining the BMDTR_η and hence should be deemed cosensitive. On the other hand, the significance probability of y_4 (tail pinch) is estimated to be consistently 0, which manifests that this endpoint is hyposensitive; a similar conclusion can be reached for y_3 (motor activity) as well. Since motor activity and tail pinch are considered hyposensitive, one could omit them from future consideration.

Robustness analysis.

To determine the robustness of the analysis relative to prior distribution specification, a sensitivity study is conducted by varying the prior distribution variance parameters for $\boldsymbol{\mu}$. Recall

Table 4: Posterior endpoint significance probabilities, $\widehat{P}(y_j \in \text{BMDTS})$, for the OP data with exposures to ACE and DIA with varying σ_μ^2 values. The calculation is based on 10,000 posterior samples.

		$\eta = 10$			$\eta = 25$		
Endpoint (y_j)	σ_μ^2	100	50	10	100	50	10
BrainChE (y_1)		0.856	0.851	0.849	0.945	0.938	0.941
BloodChE (y_2)		0.998	0.998	0.998	1	1	1
Motor Activity (y_3)		0.001	0.001	0.001	0	0	0
Tail pinch (y_4)		0	0	0	0	0	0

$\boldsymbol{\mu}$ is the mean of the $\boldsymbol{\beta}$ vector and is normally distributed with mean $\mathbf{0}$ and variance matrix $100\mathbf{I}_{12}$. Let σ_μ^2 represent the variance of $\boldsymbol{\mu}$ (in this case 100). As σ_μ^2 decreases the prior distribution becomes more informative and the posterior estimates will be shrunk towards the prior mean $\mathbf{0}$. For σ_μ^2 =10, 50 and 100 the model is refit and the endpoint significance probabilities $\widehat{P}(y_j \in \text{BMDTS}_\eta)$'s are computed. Parameter quantiles are obtained from the posterior distributions for each value of σ_μ^2 and are found to be sufficiently similar to those given in Table 2, hence we conclude that the inferences drawn for the model parameters are not sensitive to prior distribution specification. Table 4 shows the results of the sensitivity study with respect to endpoint significance probabilities. Notice that the estimated endpoint significance probabilities do not change much as the prior distribution on $\boldsymbol{\mu}$ becomes increasingly informative. Hence we conclude that the specification of the prior distribution used in the analysis does not affect the final result and is appropriate with lack of strong prior information.

5. Discussion and Future Research

This paper presents a methodology to analyze datasets with multiple endpoints and multiple stressors in a dose-response setting. The proposed model is easy to fit using WinBUGS or OpenBUGS with reasonable computational resources invested. The method utilizes the notion of area or volume to determine the BMDTR for multiple stressors. The idea allows endpoints that do not contribute to defining the BMDTR to be systematically omitted from future consideration. Furthermore, the endpoint significance probabilities permit researchers to determine the importance of endpoints in determining the BMDTR and hence provide a rigorously justified tool to screen endpoints for future studies. This could lead to significant amount of savings in computational resources and would prove to be much more valuable in the lab experiment setting. The proposed method was demonstrated through two examples on the OP pesticides dataset.

In Section 4 the proposed method gives a different tolerable region from that obtained by treating each endpoint independently. Consider the dosage combination of two chemicals 12 (mg/kg) ACE and 7 (mg/kg) of DIA. Recall that the independent analysis based on Table 1 would conclude that this dosage combination is tolerable as it is lower than the BMDL_{50} for both stressors. However, examining Figure 1 suggests that this combination should be deemed unacceptable when considering the aggregated effects of the two chemicals, which is taken

into account by the proposed method. This could be an indication that there may be some interaction between the two chemicals that is pulling the tolerable region in a conservative manner. The proposed method does not account for possible interactive effects.

Since the OP dataset did not have a wide range of dose combinations, the method proposed is based on the assumption that the effects of stressors are additive. Taking into account of interaction effects appears to be straightforward to achieve by simply adding interaction terms into each of the dose-response model expressions. Nevertheless, potential issues could arise and need to be addressed. While using the area or volume to define the BMDTR is still viable, the approach of employing a linear program to determine the significance probability $P(y_j \in \text{BMDTS})$ no longer applies. Therefore, a different strategy will need to be proposed to determine the significance probabilities. Another issue that needs to be addressed is how sensitive the method is to the size of the dataset. In many data collection efforts it is difficult to obtain large samples such as the one provided in our example. This could be done via a simulation study to determine how well the method performs under different sample sizes and how it performs when larger numbers of endpoints are considered.

In this paper, we use simple dose-response models that include all the stressors under investigation as predictors for the convenience of analysis without introducing undue complications. Nevertheless, one should consider alternative choices of dose-response models to better fit their data and produce more accurate tolerable regions. Moreover, if parsimony is of interest one should also consider selecting a subset of stressors to use for each endpoint model. This is potentially quite involved as simultaneously choosing the appropriate endpoint model and the corresponding set of stressors is not straightforward. The aforementioned issues among others are considered as future research topics and we expect an in-depth study of these to yield fruitful results.

Acknowledgement

This work was partially supported by Grant No. T32 MH014235 from the National Institute of Health.

References

Budtz-Jørgensen E (2007). "Estimation of the benchmark dose by structural equation models." *Biostatistics*, **8**, 675–688.

Choi T, Schervish MJ, Schmitt K, Small MJ (2010). "Bayesian Hierarchical Analysis for Multiple Endpoints in a Toxicity Study." *Journal of Agricultural, Biological, and Environmental Statistics*, **15**, 290–307.

Coffey JT, Gennings C, Moser VM (2007). "The simultaneous analysis of discrete and continuous outcomes in a dose-response study: Using desirability functions." *Regulatory Toxicology and Pharmacology*, **48**, 51–58.

Coull BA, Mezzetti M, Ryan LM (2003). "A Bayesian hierarchical model for risk assessment of methylmercury." *Journal of Agricultural, Biological, and Environmental Statistics*, **8**, 253–270.

Crump KS (1984). "A new method for determining allowable daily intakes." *Fundamental Applied Toxicology*, **4**, 854–871.

Crump KS (2002). "Critical issues in benchmark calculations from continuous data." *Critical Reviews Toxicology*, **32**, 133–153.

D W Gaylor LR, Krewski D, Zhu YL (1998). "Procedures for calculating benchmark doses for health risk assessment." *Regulatory Toxicology and Pharmacology*, **28**, 150–164.

Dunson DB (2000). "Bayesian latent variable models for clustered mixed outcomes." *Journal of the Royal Statistical Society, Series B*, **62**, 355–366.

Englehardt JD (2004). "Predictive Bayesian dose-response assessment for appraising absolute health risk from available health risk from available information." *Human and Ecological Risk Assessment*, **10**, 69–78.

Faes C, Geys H, Aerts M, Molenberghs G, Catalano PJ (2004). "Modeling combined continuous and ordinal outcomes in a clustered setting." *Journal of Agricultural, Biological, and Environmental Statistics*, **9**, 515–530.

Gelman A, Carlin JB, Stern HS, Rubin DB (2005). *Bayesian Data Analysis*. Chapman Hall and CRC Press, Boca Raton, FL.

Geys H, Molenberghs G, Ryan ML (1999). "Pseudolikelihood modeling of multivariate outcomes in developmental toxicology." *Journal of the American Statistical Association*, **94**, 734–745.

Leisenring W, Ryan L (1992). "Statistical properties of the NOAEL." *Regulatory Toxicology and Pharmacology*, **15**, 161–171.

Monosson E (2005). "Chemical mixtures: Considering evolution of toxicology and chemical assessment." *Environmental Health Perspectives*, **113**, 383–390.

Moser VC, Casey M, Hamm A, Carter WHJ, Simmons JE, Gennings C (2005). "Neurotoxicological and statistical analyses of a mixture of five organophosphorus pesticides using a ray design." *Toxicological Sciences*, **86**, 101–115.

Pawitan Y (2001). *In All Likelihood: Statistical modeling and Inference Using Likelihood*. Oxford University Press, NY.

Piegorsch WW, Bailer AJ (2005). *Analyzing Environmental Data*. John Wiley & Sons, Chichester, England.

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

Regan MM, Catalano PJ (1999b). "Likelihood models for clustered binary and continuous outcomes: Application to developmental toxicology." *Biometrics*, **55**, 760–768.

Ryan L (1992). "Quantitative risk assessment for developmental toxicity." *Biometrics*, **48**, 163–174.

Sammel MD, Ryan ML, Legler JM (1997). "Latent variable models for mixed discrete and continuous outcomes." *Journal of the Royal Statistical Society Series B*, **59**, 667–678.

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

Slob W (1999). "Thresholds in toxicology and risk assessment." *International Journal of Toxicology*, **18**, 259–268.

Tao J, Shi NZ, Lee SY (2004). "Drug risk assessment with determining the number of subpopulations under finite mixture normal models." *Computational Statistics & Data Analysis*, **46**, 661–676.

USEPA (1995). "The use of the benchmark dose method in health risk assessment." *Technical Report EPA/630/R-94/007*, Risk Assessment Forum, U. S. Environmental Protection AgencyËIJ, Washington D. C.

USEPA (2002). "Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity." *Technical report*, Office of Pesticide Programs, U. S. Environmental Protection AgencyËIJ, Washington D. C.

Yu ZF, Catalano PJ (2005). "Quantitative risk assessment for bivariate continuous outcomes with application to neurotoxicology: the bivariate case." *Biometrics*, **61**, 757–766.

Yu ZF, Catalano PJ (2008). "A simulation study of quantitative risk assessment for bivariate continous outcomes." *Risk Analysis*, **28**, 1415–1430.

Appendix

Table 5: The 2.5%, 50% and 97.5% quantiles and $\widehat{R}$ of the posterior samples for the regression parameters for the OP data with exposures to ACE, DIA, DTO, CPF and MAL. Credible intervals in **bold** do not contain 0. The calculation is based on 10,000 posterior samples.

Endpoint	Parameter		2.5%	50%	97.5%	$\widehat{R}$
BrainCHE	β_{10}	Intercept	**-0.126**	**-0.076**	**-0.037**	1.003
	β_{11}	ACE	**-0.007**	**-0.006**	**-0.004**	1.005
	β_{12}	DIA	**-0.082**	**-0.052**	**-0.035**	1.000
	β_{13}	MAL	**-0.037**	**-0.029**	**-0.022**	1.002
	β_{14}	CPF	-0.0006	-0.0003	0.0001	1.001
	β_{15}	DTO	**-0.022**	**-0.018**	**-0.013**	1.002
BloodChE	β_{20}	Intercept	**-0.135**	**-0.085**	**-0.041**	1.002
	β_{21}	ACE	**-0.074**	**-0.059**	**-0.046**	1.003
	β_{22}	DIA	**-0.038**	**-0.029**	**-0.021**	1.001
	β_{23}	MAL	**-0.039**	**-0.031**	**-0.024**	1.004
	β_{24}	CPF	**-0.0022**	**-0.0016**	**-0.0011**	1.003
	β_{25}	DTO	**-0.670**	**-0.564**	**-0.455**	1.003
Motor Activity	β_{30}	Intercept	**-0.097**	**-0.085**	**-0.027**	1.002
	β_{31}	ACE	**-0.004**	**-0.003**	**-0.002**	1.002
	β_{32}	DIA	**-0.027**	**-0.020**	**-0.015**	1.000
	β_{33}	MAL	**-0.015**	**-0.012**	**-0.009**	1.001
	β_{34}	CPF	-0.0003	-0.0000	0.0003	1.006
	β_{35}	DTO	**-0.031**	**-0.025**	**-0.018**	1.003
Tail pinch	β_{40}	Intercept	**0.887**	**0.785**	**0.686**	1.003
	β_{41}	ACE	**0.004**	**0.003**	**0.002**	1.002
	β_{42}	DIA	**0.011**	**0.006**	**0.002**	1.003
	β_{43}	MAL	**0.010**	**0.004**	**0.002**	1.001
	β_{44}	CPF	-0.0015	0.0006	0.0003	1.001
	β_{45}	DTO	**0.019**	**0.009**	**0.000**	1.003

Table 6: Posterior endpoint significance probabilities, $\widehat{P}(y_j \in \text{BMDTS})$, for the OP data with exposures to ACE, DIA, DTO, CPF and MAL. The calculation is based on 10,000 posterior samples.

	η		
Endpoint (y_j)	10	25	50
BrainChE (y_1)	0.9728	0.9986	1
BloodChE (y_2)	0.9982	1	1
Motor activity (y_3)	0.0014	0	0
Tail pinch (y_4)	0.024	0.002	0.0026

Table 7: Posterior endpoint significance probabilities, $\widehat{P}(y_j \in \text{BMDTS})$, for the OP data with exposures to ACE, DIA, DTO, CPF and MAL with varying values of σ^2_μ. The calculation is based on 10,000 posterior samples.

		$\eta = 10$			$\eta = 25$		
Endpoint (y_j)	σ^2_μ	100	50	10	100	50	10
BrainChE (y_1)		0.9728	0.9788	0.9751	0.9986	0.9994	0.9987
BloodChE (y_2)		0.9982	0.9982	0.9992	1	1	1
Motor activity (y_3)		0.0014	0	0.0041	0	0	0
Tail pinch (y_4)		0.0240	0.0197	0.0250	0.0020	0.0008	0.0013

Affiliation:

Edward L. Boone
Department of Statistical Sciences and Operations Research
Virginia Commonwealth University
Richmond, VA 23284
E-mail: elboone@vcu.edu