Flexible Bayesian modelling for clustered categorical responses in developmental toxicology

Athanasios Kottas

University of California at Santa Cruz, U.S.A.

and

Kassandra Fronczyk

The University of Texas MD Anderson Cancer Center, U.S.A.

Summary: We present a Bayesian nonparametric approach to modelling and risk assessment in developmental toxicity studies. The primary objective for this dose-response setting is to quantify the relationship between the level of exposure of pregnant laboratory animals to a toxic chemical and the probability of prenatal death or a physiological response for viable fetuses. Hence, the data involve clustered categorical responses, and typically suggest response distributions and dose-response relationships that can not be captured well by standard parametric approaches. The focus of our modelling approach is on the dose-dependent response distributions, which are represented through a nonparametric trinomial mixture model. The nonparametric mixing is built from a dependent Dirichlet process prior with the dependence of the mixing distributions governed by the dose level. The key implication of this modelling strategy is flexible inference for both the response distribution as well as for the dose-response curves associated with the different endpoints, including the capacity of the model to capture non-monotonic dose-response relationships. The practical utility of the methodology is illustrated with data from an experiment on the effects of diethylhexalphthalate, a commonly used plasticizing agent.

1 Introduction

Developmental toxicity studies investigate birth defects induced by toxic chemicals. In particular, under the standard Segment II developmental toxicity experiment, at each experimental dose

level, x_i, $i = 1, ..., N$, a number, n_i, of pregnant laboratory animals (dams) are exposed to the toxin. Dam j at dose x_i has m_{ij} implants, of which the number of resorptions, that is, undeveloped embryos or very early fetal deaths, and prenatal deaths are recorded as R_{ij}, and the number of live pups at birth with a certain malformation are recorded as y_{ij}. Consequently, the number of viable fetuses for dam j at dose x_i is $m_{ij} - R_{ij}$. Additional continuous outcomes measured on each of the live pups may include body weight and length.

The main objective of developmental toxicity studies is to examine the relationship between the level of exposure to the toxin, which we generically refer to as the dose level, and the probability of the various responses of interest. We focus on the clustered categorical endpoints of embryolethality, that is, non-viable fetuses, and fetal malformation for live pups; thus, the data structure comprises $\{(m_{ij}, R_{ij}, y_{ij}) : \ i = 1, \ldots, N; \ j = 1, \ldots, n_i\}$. The corresponding dose-response curves are defined by the probability of the endpoints across dose levels. Also of interest is quantitative risk assessment, which evaluates the probability that adverse effects may occur as a result of the exposure to the substance.

Plotted in Figure 1 is a motivating data set, available from the National Toxicology Program database, from an experiment that explored the effects of diethylhexalphthalate (DEHP), a commonly used plasticizing agent. The left and middle panels correspond to the endpoints of non-viable fetus, that is, resorption or prenatal death, and malformation, that is, external, visceral or skeletal malformation of a live fetus. The right panel plots the proportions of combined negative outcomes, that is, adding the number of non-viable fetuses and malformations.

The number of dams is 30 for the control group, and 26, 26, 17, and 9 for doses 25, 50, 100, and 150 mg/kg $\times$1000. The number of implants across all dams and dose levels ranges from 4 to 18, with 25th, 50th, and 75th percentiles equal to 11, 13, and 14, respectively. Particularly noteworthy is the drop in the proportions of malformations, and combined negative outcomes, from dose 0 to 25 mg/kg $\times$1000, which may correspond to a hormetic dose-response relationship. Hormesis refers to a dose-response phenomenon characterized by favorable biological responses to low exposures to toxins, and thus by opposite effects in small and large doses. For endpoints involving disease incidence, such as mutation, birth defects or cancer, hormesis results in a J-shaped dose-response curve. Although the possibility of different low dose effects is accepted, the suggestion of positive low dose effect is debated, hence, hormesis remains a controversial

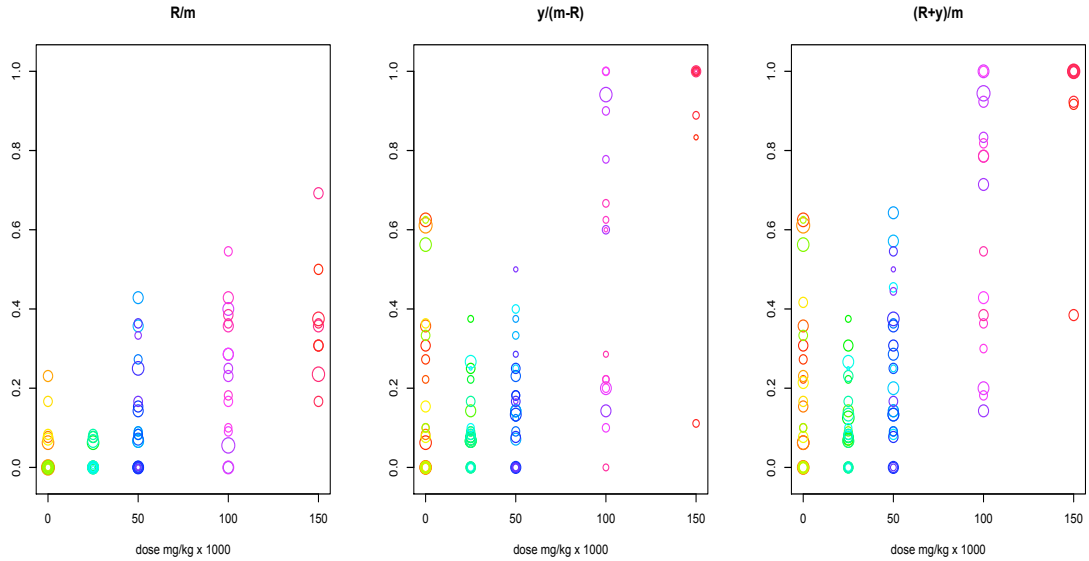

Figure 1: DEHP data. In each panel, a circle corresponds to a particular dam and the size of the circle is proportional to the number of implants. The coordinates of the circle are the dose level and the proportion of the specific endpoint: non-viable fetuses among implants (left panel); malformations among live pups (middle panel); combined negative outcomes among implants (right panel).

concept in the toxicological sciences [e.g., 1].

Notwithstanding the ultimate scientific conclusions, data such as the ones from the DEHP study motivate our modelling framework for the dose-dependent response distributions which enables rich inference for the implied, possibly non-monotonic, dose-response curves. Building flexible modelling for the response distribution is easy to justify for developmental toxicology data, which typically indicate vast departures from parametric models. This can be attributed to the inherent heterogeneity in the data due to the clustering of individuals within a group and the variability of the reaction of the individuals to the toxin. Note also that the typical toxicity experiment discussed above provides information on potentially different dose-response relationships for the distinct endpoints of embryolethality and fetal malformation, and it is thus biologically relevant to jointly analyze the clustered responses. This stands in contrast with the prevailing data structure found in the statistical literature, where the variables involved are the number of implants and the sum of all negative outcomes, as in the right panel of Figure 1.

We develop a Bayesian nonparametric mixture model for the joint distribution of the number of non-viable fetuses and malformations. We seek mixture modelling for response distributions

that are related across doses with the level of dependence driven by the distance between the dose values. To this end, we consider a dependent Dirichlet process (DDP) prior [18] for the dose-dependent mixing distributions. Particular emphasis is placed on the choice of the mixture kernel and the DDP prior formulation to ensure an increasing trend in prior expectation for the implied dose-response curves, but without restricting prior realizations to be necessarily monotonic. This is key for the model's capacity to capture non-standard dose-response relationships. The nonparametric mixture model structure enables flexible inference for the response distributions at any observed dose level. Moreover, the dependence of the DDP prior across dose levels allows inference for the induced dose-response relationships through interpolation and extrapolation over any range of dose values of interest.

The modelling approach developed here extends work for the simpler data setting with combined negative outcomes [10]. To our knowledge, the literature does not include any Bayesian nonparametric approaches to modelling developmental toxicology data with a multicategory response classification. A Bayesian semiparametric model for the combined negative endpoints case, based on a product of Dirichlet process prior, was proposed by [5], and more recently extended in [19]. Examples of parametric Bayesian hierarchical models for toxicology data with discrete-continuous outcomes include [6] and [8]. Regarding the classical literature, a Dirichlet-trinomial model is presented in [2]; [24] develop an extended Dirichlet-multinomial model with Weibull dose-response functions; and [21] and [17] use quasi-likelihood and generalized estimating equations, respectively, to fit multinomial models which incorporate overdispersion.

The outline of the paper is as follows. Section 2 develops the DDP mixture model, including model properties and methods for prior specification and Markov chain Monte Carlo posterior inference. In Section 3, we present the application to the analysis of the DEHP data. Section 4 concludes with a summary.

2 Methods

2.1 The modelling approach

Under the Segment II toxicity study design, exposure occurs after implantation, and thus it is natural to treat the number of implants as a random quantity containing no information

about the dose-response relationship. Hence, the modelling for the number of implants (m), the number of non-viable fetuses (R), and the number of malformations (y) is decomposed to $f(m, R, y) = f(m)f(R, y \mid m)$, where only the conditional distribution $f(R, y \mid m)$ will depend on dose level x. We assume a shifted Poisson distribution for $f(m)$ such that $m \geq 1$, although more general distributions can be readily utilized. Inference for the implant distribution is carried out separately from inference for $f(R, y \mid m)$, and is not discussed further.

To develop a flexible inference framework for risk assessment, we propose a nonparametric mixture model for the dose-dependent conditional distribution of the number of non-viable fetuses and malformations given the number of implants. Specifically, for a generic dose x,

$$f(R, y \mid m) \equiv f(R, y \mid m; G_x) = \int k(R, y \mid m; \varphi)\, \mathrm{d}G_x(\varphi)$$

where $k(R, y \mid m; \varphi)$ is a parametric kernel, with parameters φ, and G_x the dose-dependent mixing distribution. Placing a nonparametric prior on G_x results in a nonparametric mixture prior for $f(R, y \mid m; G_x)$. Nonparametric Bayesian mixture priors offer flexible modelling tools that, with the appropriate structure, can capture the complexity inherent in the data. These models can be viewed as extensions of finite mixture or continuous mixture models, where the random mixing distribution is not defined with a particular parametric family of distributions.

Regarding the mixture kernel, we take $k(R, y \mid m; \gamma, \theta) = \mathrm{Bin}(R; m, \pi(\gamma))\mathrm{Bin}(y; m-R, \pi(\theta))$, where $\pi(u) = \exp(u)/\{1+\exp(u)\}$, $u \in \mathbb{R}$, will be used to denote the logistic function. Therefore, $\pi(\gamma)$ is the kernel probability of a non-viable fetus, and $\pi(\theta)$ the conditional probability of a malformation for a live pup. This formulation of the trinomial kernel distribution is natural as it highlights the nested nature of the count responses. Moreover, the logistic transformation for the Binomial probabilities is used to facilitate the formulation of the nonparametric prior model for the collection of mixing distributions, $G_{\mathcal{X}} = \{G_x : x \in \mathcal{X}\}$, where $\mathcal{X} \subseteq \mathbb{R}^+$.

As discussed in the Introduction, we seek modelling for the response distributions that allows nonparametric dependence structure across dose levels. We achieve such modelling by placing a DDP prior on the dose-dependent mixing distributions $\{G_x : x \in \mathcal{X}\}$. The DDP prior arises through extension of the Dirichlet process (DP) prior [9], the most widely used prior for mixing distributions in nonparametric or semiparametric mixture models. We use $\mathrm{DP}(\alpha, G_0)$ to denote

the DP prior defined in terms of a parametric centering distribution G_0, and precision parameter $\alpha > 0$.

To define the form of the DDP prior we use for $G_{\mathcal{X}}$, the almost sure discrete representation for the regular DP [22] is extended to

$$G_{\mathcal{X}}(\cdot) = \sum_{l=1}^{\infty} \omega_l \delta_{\eta_{l\mathcal{X}}}(\cdot) \tag{1}$$

where δ_a denotes a point mass at a, the $\eta_{l\mathcal{X}} = \{\eta_l(x) : x \in \mathcal{X}\}$ are independent realizations from a stochastic process $G_{0\mathcal{X}}$ over $\mathcal{X}$, and the weights are defined by a stick-breaking process: $\omega_1 = \zeta_1$, and $\omega_l = \zeta_l \prod_{r=1}^{l-1}(1-\zeta_r)$ for $l \geq 2$, with ζ_l independent from a Beta(1,α) distribution, independently of the $\eta_{l\mathcal{X}}$. A key feature of the DDP prior is that for any finite collection of dose levels $(x_1, ..., x_k)$ it induces a multivariate DP prior for the corresponding collection of mixing distributions $(G_{x_1}, ..., G_{x_k})$. Therefore, the DDP prior model involves a countable mixture of realizations from stochastic process $G_{0\mathcal{X}}$ with weights matching those from the standard DP; this prior structure is referred to as single-p DDP prior. Single-p DDP mixture models have been applied to analysis of variance settings [4], spatial modelling [11; 15], dynamic density estimation [20], quantile regression [16], and survival regression [3].

Finally, the DDP prior mixture model is given by

$$f(R, y \mid m; G_{\mathcal{X}}) = \int \text{Bin}(R; m, \pi(\gamma))\text{Bin}(y; m-R, \pi(\theta))\, \text{d}G_{\mathcal{X}}(\gamma, \theta), \;\; G_{\mathcal{X}} \sim \text{DDP}(\alpha, G_{0\mathcal{X}}). \tag{2}$$

Here, $\text{DDP}(\alpha, G_{0\mathcal{X}})$ denotes the DDP prior for $G_{\mathcal{X}} = \sum_{l=1}^{\infty} \omega_l \delta_{\eta_{l\mathcal{X}}}$, where $\eta_l(x) = (\gamma_l(x), \theta_l(x))$, for $x \in \mathcal{X}$, with precision parameter α and base stochastic process $G_{0\mathcal{X}}$. We define $G_{0\mathcal{X}}$ through two independent Gaussian processes, one driving each probability of response, each with a linear mean function, constant variance, and isotropic exponential correlation function. Hence, to introduce notation, we assume for all l, $\text{E}(\gamma_l(x) \mid \xi_0, \xi_1) = \xi_0 + \xi_1 x$, and $\text{E}(\theta_l(x) \mid \beta_0, \beta_1) = \beta_0 + \beta_1 x$; $\text{var}(\gamma_l(x) \mid \tau^2) = \tau^2$, and $\text{var}(\theta_l(x) \mid \sigma^2) = \sigma^2$; $\text{corr}(\gamma_l(x), \gamma_l(x') \mid \rho) = \exp(-\rho|x - x'|)$, and $\text{corr}(\theta_l(x), \theta_l(x') \mid \phi) = \exp(-\phi|x - x'|)$, with $\rho > 0$ and $\phi > 0$. As discussed in the next section, this specification for $G_{0\mathcal{X}}$ and, in particular, the linear mean functions, are key for flexible inference about the dose-response relationships implied by model (2). The full Bayesian model is implemented with priors on α and on the $G_{0\mathcal{X}}$ hyperparameters, $\psi = (\xi_0, \xi_1, \tau^2, \rho, \beta_0, \beta_1, \sigma^2, \phi)$.

2.2 Dose-response relationships

Here, we study the dose-response curves implied by DDP mixture model (2), including the probability of a non-viable fetus, the conditional probability of a malformation for a live pup, and a risk function that combines the two endpoints.

To develop the dose-response curves, it is useful to note a connection of the mixture model with the clustered Binomial kernels with the model based on products of Bernoullis kernel for the underlying binary responses. That is, for a generic dam at dose level x with m implants, let $R^* = \{R_k^* : k = 1, \ldots, m\}$ be the individual non-viable fetus indicators and denote by $y^* = \{y_s^* : s = 1, \ldots, m - \sum_{k=1}^m R_k^*\}$ the malformation indicators for the viable fetuses. Therefore, $R = \sum_{k=1}^m R_k^*$ and $y = \sum_{s=1}^{m-\sum_k R_k^*} y_s^*$. Then, a DDP mixture model for the clustered binary responses can be formulated as

$$f^*\left(R^*, y^* \mid m; G_{\mathcal{X}}\right) = \int \prod_{k=1}^{m} \text{Bern}(R_k^*; \pi(\gamma)) \prod_{s=1}^{m-\sum_k R_k^*} \text{Bern}(y_s^*; \pi(\theta))\, \mathrm{d}G_{\mathcal{X}}(\gamma, \theta), \tag{3}$$

where $G_{\mathcal{X}}$ is assigned the same DDP prior as the one for model (2). It is straightforward to show that mixture models (2) and (3) are equivalent with regard to the distribution for (R, y) conditional on m; in particular, the joint moment generating function for $(\sum_{k=1}^m R_k^*, \sum_{s=1}^{m-\sum_k R_k^*} y_s^*)$ under model (3) is equal to the joint moment generating function for (R, y) under model (2), in both cases, conditioning on m. Hence, we can define the dose-response curves under the DDP mixture model (2) working with probabilities of the two endpoints for a generic implant; this involves implicit conditioning on $m = 1$, which we suppress in the notation below.

The first risk assessment quantity of interest is the probability of embryolethality across effective dose levels, which is defined by

$$D(x) \equiv \text{pr}(R^* = 1; G_x) = \int \pi(\gamma)\, \mathrm{d}G_x(\gamma, \theta), \quad x \in \mathcal{X}.$$

Risk assessment for the malformation endpoint is based on the conditional probability that a generic pup has a malformation given that it is a viable fetus, i.e.,

$$M(x) \equiv \text{pr}(y^* = 1 \mid R^* = 0; G_x) = \frac{\text{pr}(R^* = 0, y^* = 1; G_x)}{\text{pr}(R^* = 0; G_x)} = \frac{\int \{1 - \pi(\gamma)\}\pi(\theta)\mathrm{d}G_x(\gamma, \theta)}{\int \{1 - \pi(\gamma)\}\mathrm{d}G_x(\gamma, \theta)}, \quad x \in \mathcal{X}.$$

Moreover, a full risk function at any given dose level can be defined through the combination of the probability of a non-viable fetus and the probability of a live, malformed pup; that is, the combined risk at dose level x is given by

$$\begin{aligned} r(x) &\equiv \text{pr}(R^* = 1 \text{ or } y^* = 1; G_x) = \text{pr}(R^* = 0, y^* = 1; G_x) + \text{pr}(R^* = 1; G_x) \\ &= \int \{1 - \pi(\gamma)\}\pi(\theta)\, \text{d}G_x(\gamma, \theta) + \int \pi(\gamma)\, \text{d}G_x(\gamma, \theta) \\ &= 1 - \int \{1 - \pi(\gamma)\}\{1 - \pi(\theta)\}\, \text{d}G_x(\gamma, \theta), \quad x \in \mathcal{X}. \end{aligned}$$

A key aspect of the modelling approach is that it does not force a non-decreasing shape restriction to the dose-response functions, which is the traditional assumption for more standard quantal bioassay experiments. As discussed in the Introduction and illustrated in Section 3 with the DEHP data, the model's capacity to capture non-standard, possibly non-monotonic, dose-response relationships is an asset of the proposed methodology. At the same time, given the relatively small number of observed dose levels in developmental toxicity studies, some structure is needed in the prior model in order to obtain meaningful interpolation and extrapolation posterior inference results for the dose-response curves. Under the specific formulation of the DDP prior for mixture model (2), such structure can be incorporated in the form of a non-decreasing trend in prior expectation for the dose-response curves.

Consider first the prior expectation for the dose-dependent probability of a non-viable fetus,

$$\text{E}\{D(x)\} = \text{E}\left\{\int \pi(\gamma)\text{d}G_x(\gamma, \theta)\right\} = \int \pi(\gamma)\text{d}G_{0x}(\gamma, \theta) = \int \pi(\gamma)\text{dN}(\gamma; \xi_0 + \xi_1 x, \tau^2),$$

that is, $\text{E}\{D(x)\}$ is the expectation of the (increasing) logistic function with respect to the $\text{N}(\xi_0 + \xi_1 x, \tau^2)$ distribution, which is stochastically ordered in x when $\xi_1 > 0$. Hence, $D(x)$ is a non-decreasing function of x in prior expectation, under the $\xi_1 > 0$ prior restriction. Similarly,

$$\begin{aligned} \text{E}\{r(x)\} &= 1 - \textstyle\int\{1 - \pi(\gamma)\}\{1 - \pi(\theta)\}\text{d}G_{0x}(\gamma, \theta) \\ &= 1 - \left[\textstyle\int\{1 - \pi(\gamma)\}\text{dN}(\gamma; \xi_0 + \xi_1 x, \tau^2)\right]\left[\textstyle\int\{1 - \pi(\theta)\}\text{dN}(\theta; \beta_0 + \beta_1 x, \sigma^2)\right]. \end{aligned}$$

Provided $\xi_1 > 0$ and $\beta_1 > 0$, distributions $\text{N}(\xi_0 + \xi_1 x, \tau^2)$ and $\text{N}(\beta_0 + \beta_1 x, \sigma^2)$ are stochastically ordered in x, which implies that both $\int\{1 - \pi(\gamma)\}\text{dN}(\gamma; \xi_0 + \xi_1 x, \tau^2)$ and $\int\{1 - \pi(\theta)\}\text{dN}(\theta; \beta_0 +$

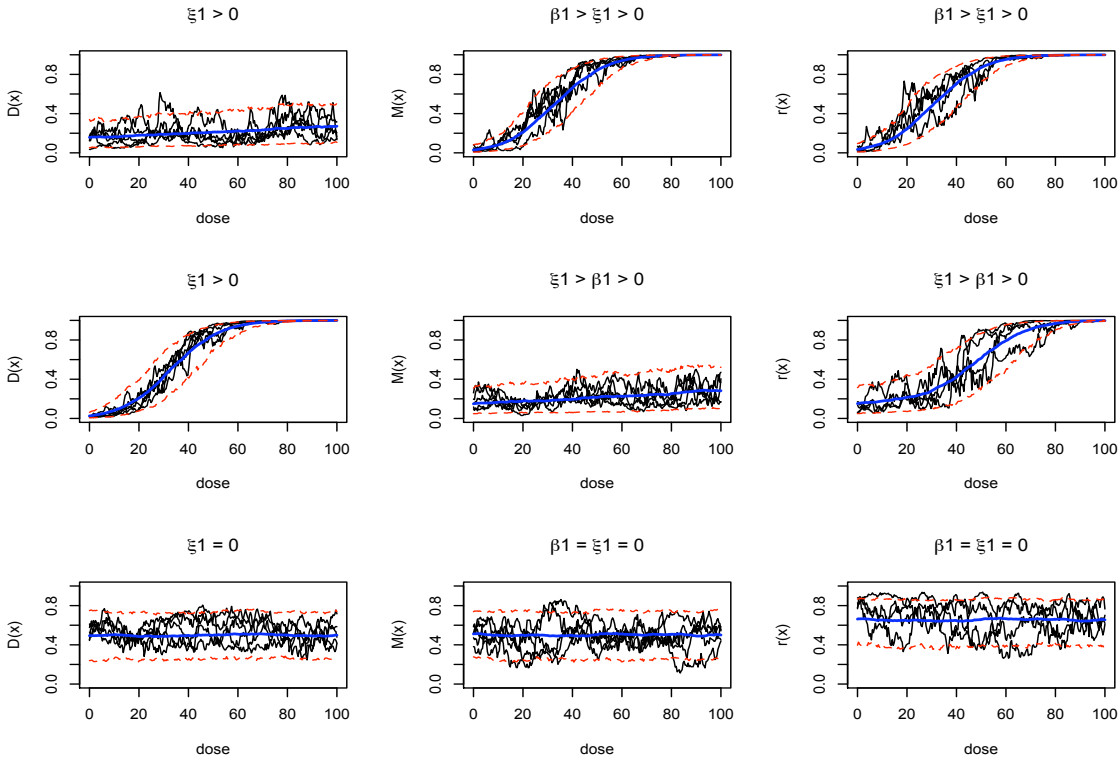

Figure 2: Prior mean and 90% interval estimates, along with 5 individual prior realizations, for the three dose-response curves. See Section 2.2 for details.

$\beta_1 x, \sigma^2$) are decreasing functions of x, with values in the unit interval. Thus, E$\{r(x)\}$ is non-decreasing in x when $\xi_1 > 0$ and $\beta_1 > 0$.

Therefore, with the $\xi_1 > 0$ and $\beta_1 > 0$ prior restrictions, we can build to both the probability of a non-viable fetus and to the combined risk function the non-decreasing trend in prior expectation. Although the same argument does not extend to the conditional probability of malformation, $M(x)$, the restriction $\xi_1 > 0$ and $\beta_1 > 0$ appears sufficient to provide the prior expectation non-decreasing trend for all three dose-response curves. In this respect, it is useful to note that, even though we develop inference about three dose-response relationships, there are only two endpoints and, consequently, the model is driven at any specific dose level by a bivariate random mixing distribution.

Note that the argument above relies on both the constant Gaussian process variances for the two components of $G_{0\mathcal{X}}$ – which ensures the stochastic ordering of the induced normal distributions – and on the linear Gaussian process mean functions – which enables the non-

decreasing trend through the restriction on the slope parameters. Indeed, the linear mean functions are crucial for practicable posterior inference. As suggested by Figure 2, if the model is applied using constant mean functions for the DDP prior centering Gaussian processes, that is, setting $\xi_1 = \beta_1 = 0$, we should not expect practically useful results outside the observed dose levels. For illustration, Figure 2 plots results from prior simulation for the embryolethality and malformation dose-response curves, and for the combined risk function, using fixed values for α $(= 1)$ and ψ. In particular, $(\xi_1 = 0.0085,\ \beta_1 = 0.12)$ and $(\xi_1 = 0.12,\ \beta_1 = 0.01)$ in the top and middle row, respectively. Although the relative magnitude of ξ_1 and β_1 affects the rate of increase for the different curves, in all cases with $\xi_1 > 0$ and $\beta_1 > 0$, the non-decreasing trend in prior expectation is preserved.

Finally, smoothness properties of prior realizations for the dose-response curves relate directly to the respective properties of the centering process $G_{0\mathcal{X}}$. For details, we refer to the arguments in [18] and [11], extended and formalized by [12], but note briefly that the continuity of the realizations from the two Gaussian processes that define $G_{0\mathcal{X}}$ implies that as the distance between x and x' gets smaller, the difference between G_x and $G_{x'}$ gets smaller; moreover, it yields continuous prior realizations for the three dose-response functions defined above. The practical implication is that in prediction for the probability mass function $f(R, y \mid m; G_x)$ and for the corresponding dose-response curves, we learn more from dose levels x' nearby x than from more distant doses, a desirable property for distributions that are expected to evolve relatively smoothly with the dose level.

2.3 Implementation details

2.3.1 Markov chain Monte Carlo posterior simulation

Regarding the hierarchical model formulation for the data $= \{(m_{ij}, R_{ij}, y_{ij}) :\ i = 1, \ldots, N;\ j = 1, \ldots, n_i\}$, we observe that for the DEHP data (discussed in the Introduction) the dams are labeled and recorded in ascending numerical order across dose levels; that is, the smallest identification number corresponds to data from the first dam at the first dose level, the first dam at the second dose level has the next identification number, and so on. This is also the case for other data sets available from the database of the National Toxicology Program. Therefore, to write the model for the data, the animals are linked as a response vector across the dose levels

with the conditional independence assumption built for the replicated response vectors. Hence, the data structure and corresponding hierarchical model is along the lines of the spatial DP [11] rather than, for instance, the ANOVA DDP [4].

Therefore, let $R_j = (R_{1j}, \ldots, R_{Nj})$, and $y_j = (y_{1j}, \ldots, y_{Nj})$ be the j-th response replicates with corresponding number of implants vector $m_j = (m_{1j}, \ldots, m_{Nj})$, for $j = 1, \ldots, n$, where $n = \max_i n_i$. Moreover, denote by $\gamma_j \equiv \gamma_j(x) = (\gamma_j(x_1), \ldots, \gamma_j(x_N))$, and $\theta_j \equiv \theta_j(x) = (\theta_j(x_1), \ldots, \theta_j(x_N))$ the latent mixing vector for R_j and y_j, respectively, where $x = (x_1, \ldots, x_N)$. We introduce missing value indicators, s_{ij}, such that $s_{ij} = 1$ if the j-th response replicates at dose level i are present and $s_{ij} = 0$ otherwise. Note that the s_{ij} are fixed for any particular data set. Then, the first stage of the hierarchical model for the data can be written as

$$\{(R_{ij}, y_{ij})\} \mid \{m_{ij}\}, \{(\gamma_j, \theta_j)\} \sim \prod_{j=1}^{n} \prod_{i=1}^{N} \{\text{Bin}(R_{ij}; m_{ij}, \pi(\gamma_j(x_i)))\text{Bin}(y_{ij}; m_{ij} - R_{ij}, \pi(\theta_j(x_i)))\}^{s_{ij}}$$

where the (γ_j, θ_j), given G_x, are i.i.d. from G_x, which follows a $\text{DP}(\alpha, G_{0x})$ prior implied by the DDP prior for $G_{\mathcal{X}}$. In particular, G_{0x} comprises two independent N-variate normal distributions, induced by the two Gaussian processes that define $G_{0\mathcal{X}}$; the first normal distribution has mean vector $(\xi_0+\xi_1 x_1, \ldots, \xi_0+\xi_1 x_N)'$ and covariance matrix with (i, j)-th element $T_{ij} = \tau^2 \exp(-\rho|x_i - x_j|)$; the mean of the second normal distribution is $(\beta_0+\beta_1 x_1, \ldots, \beta_0+\beta_1 x_N)'$ and its covariance matrix has (i, j)-th element $\Sigma_{ij} = \sigma^2 \exp(-\phi|x_i - x_j|)$.

Hence, the hierarchical model for the data is a DP mixture model induced by the DDP mixture prior. For Markov chain Monte Carlo posterior simulation, we use blocked Gibbs sampling [e.g., 13], which offers relatively ready implementation and also, in our context, can easily handle unbalanced response replicates. The approach is based on a finite truncation approximation of G_x such that $G_x \approx G_x^L = \sum_{l=1}^{L} p_l \delta_{(U_l(x), Z_l(x))}$, where the weights p_l arise from a truncated version of the stick-breaking construction: $p_1 = V_1$, $p_l = V_l \prod_{r=1}^{l-1}(1 - V_r)$, $l = 2, \ldots, L-1$, and $p_L = 1 - \sum_{l=1}^{L-1} p_l$, with the V_l i.i.d., given α, from $\text{Beta}(1, \alpha)$. Moreover, $U_l(x) = (U_l(x_1), \ldots, U_l(x_N)) \equiv U_l$ and $Z_l(x) = (Z_l(x_1), \ldots, Z_l(x_N)) \equiv Z_l$, with the (U_l, Z_l) i.i.d., given ψ, from G_{0x}, for $l = 1, ..., L$. Hence, under the truncated version of mixing distribution G_x, $(\gamma_j, \theta_j) = (U_l, Z_l)$ with probability p_l, and $G_x^L \equiv (p, U, Z)$, where $p = (p_1, \ldots, p_L)$, $U = (U_1, \ldots, U_L)$ and $Z = (Z_1, \ldots, Z_L)$.

To represent the hierarchical model for the data under the DP truncation approximation, configuration variables $w = (w_1, \ldots, w_n)$ are introduced, such that $w_j = l$ if and only if $(\gamma_j, \theta_j) = (U_l, Z_l)$, for $l = 1, \ldots, L$ and $j = 1, \ldots, n$. Then, the model for the data can be expressed as

$$\begin{aligned}
\{(R_j, y_j)\} \mid \{m_j\}, w, (U, Z) &\sim \prod_{j=1}^{n} \prod_{i=1}^{N} \left\{ \mathrm{Bin}(R_{ij}; m_{ij}, \pi(U_{w_j}(x_i))) \mathrm{Bin}(y_{ij}; m_{ij} - R_{ij}, \pi(Z_{w_j}(x_i))) \right\}^{s_{ij}} \\
w_j \mid p &\sim \prod_{j=1}^{n} \sum_{l=1}^{L} p_l \delta_l(w_j) \\
p, (U, Z) \mid \alpha, \psi &\sim f(p \mid \alpha) \times \prod_{l=1}^{L} G_{0x}(U_l, Z_l \mid \psi) \qquad (4)
\end{aligned}$$

where $f(p \mid \alpha) = \alpha^{L-1} p_L^{\alpha-1} (1 - p_1)^{-1} \{1 - (p_1 + p_2)\}^{-1} \times \cdots \times (1 - \sum_{l=1}^{L-2} p_l)^{-1}$, a special case of the generalized Dirichlet distribution, is the prior for p, given α, induced by the truncated stick-breaking construction. The full Bayesian model is completed with independent hyperpriors for the DDP precision parameter α and the parameters ψ of the centering Gaussian processes. Specifically, we place a gamma(a_α, b_α) prior on α; normal priors $\mathrm{N}(m_\xi, s_\xi^2)$ and $\mathrm{N}(m_\beta, s_\beta^2)$ on ξ_0 and β_0; exponential priors $\mathrm{Exp}(b_\xi)$ and $\mathrm{Exp}(b_\beta)$ on ξ_1 and β_1 to promote the non-decreasing trend in prior expectation for the dose-response functions; inverse gamma priors inv-gamma(a_τ, b_τ) and inv-gamma(a_σ, b_σ) on the variance terms τ^2 and σ^2; and uniform priors $\mathrm{Unif}(0, b_\rho)$ and $\mathrm{Unif}(0, b_\phi)$ on the range parameters ρ and ϕ. Prior specification is discussed in Section 2.3.3.

Denote the n^* distinct values of vector w by $w_1^*, \ldots, w_{n^*}^*$, and let $M_k^* = |\{j : w_j = w_k^*\}|$, for $k = 1, \ldots, n^*$, and $M_l = |\{j : w_j = l\}|$, for $l = 1, \ldots, L$. Then, sampling from the posterior distribution $p(U, Z, w, p, \alpha, \psi \mid \text{data})$ corresponding to model (4) is based on simulation from the following posterior full conditional distributions.

The (U_l, Z_l) that correspond to $l \notin \{w_k^* : k = 1, \ldots, n^*\}$ are sampled from G_{0x} given its currently imputed parameters ψ. For $l = w_k^*$, $k = 1, \ldots, n^*$, the posterior full conditional for $U_{w_k^*}$ is proportional to $G_{0x}^{\gamma}(U_{w_k^*} \mid \psi) \prod_{\{j: w_j = w_k^*\}} \prod_{i=1}^{N} \{\mathrm{Bin}(R_{ij}; m_{ij}, \pi(U_{w_k^*}(x_i)))\}^{s_{ij}}$, and the posterior full conditional for $Z_{w_k^*}$ to $G_{0x}^{\theta}(Z_{w_k^*} \mid \psi) \prod_{\{j: w_j = w_k^*\}} \prod_{i=1}^{N} \{\mathrm{Bin}(y_{ij}; m_{ij} - R_{ij}, \pi(Z_{w_k^*}(x_i)))\}^{s_{ij}}$. Here, G_{0x}^{γ} and G_{0x}^{θ} denote the respective N-variate normal distributions arising from G_{0x}. Each of $U_{w_k^*}$ and $Z_{w_k^*}$ is updated using a random-walk Metropolis-Hastings step with an N-variate normal distribution as the proposal. The proposal covariance matrices were estimated dynamically, using initial runs based on normal proposals with scaled identity covariance matrices.

The posterior full conditional for each w_j, $j = 1, ..., n$, is given by a discrete distribution with values $l = 1, ..., L$ and corresponding probabilities

$$\tilde{p}_{lj} \propto p_l \prod_{i=1}^{N} \{\text{Bin}(R_{ij}; m_{ij}, \pi(U_l(x_i)))\text{Bin}(y_{ij}; m_{ij} - R_{ij}, \pi(Z_l(x_i)))\}^{s_{ij}}, \quad l = 1, ..., L.$$

The updates for parameters α and p are the same with a generic DP mixture model [14]. Finally, the joint posterior full conditional for the hyperparameters ψ of the DDP prior centering Gaussian processes is proportional to

$$p(\xi_0)p(\xi_1)p(\tau^2)p(\rho)p(\beta_0)p(\beta_1)p(\sigma^2)p(\phi) \times \prod_{k=1}^{n^*} G_{0x}(U_{w_k^*}, Z_{w_k^*} \mid \psi)$$

where $p(\cdot)$ denotes the prior for each parameter. The form of G_{0x} and the parametric priors for the components for ψ result in normal posterior full conditionals for ξ_0 and β_0, and inverse gamma full conditionals for τ^2 and σ^2. We use Metropolis-Hastings updates for ξ_1 and β_1, and sample ρ and ϕ by discretizing their bounded support.

2.3.2 Inference for risk assessment

The samples from the posterior distribution of model (4) yield the mixing distribution G_x^L at all the observed dose levels through the posterior samples for (p, U, Z). To expand the inference over any range of doses of interest, we augment the N observed dose levels with M new doses, $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_M)$. Now, in the prior model, the $(U_l(x), U_l(\tilde{x}))$ and the $(Z_l(x), Z_l(\tilde{x}))$, for $l = 1, ..., L$, are independent realizations from two independent $(N+M)$-variate normal distributions, induced by the Gaussian processes that define $G_{0\mathcal{X}}$, with mean vectors and covariance matrices that are of the same form as above extending x to $(x, \tilde{x})$. But then, to sample from the conditional posterior distributions for each of the $U_l(\tilde{x})$ and $Z_l(\tilde{x})$, the additional sampling needed is from conditional M-variate normal distributions given the currently imputed (U_l, Z_l), $l = 1, ..., L$, and the parameters ψ.

Using the posterior samples from model (4), augmented with the posterior samples for $(U_l(\tilde{x}), Z_l(\tilde{x}))$, $l = 1, ..., L$, full inference for the response distributions and for risk assessment through the dose-response curves can be obtained by evaluating the relevant expressions developed in Sections 2.1 and 2.2. Under the DP truncation approximation used for posterior

simulation, the integrals are replaced with sums. For instance, for any generic dose x_0 in $(x, \tilde{x})$, the posterior distribution for the probability of a non-viable fetus arises from $\sum_{l=1}^{L} p_l \pi(U_l(x_0))$, and for the combined risk function through $1-\sum_{l=1}^{L} p_l\{1-\pi(U_l(x_0))\}\{1-\pi(Z_l(x_0))\}$. Moreover, for a specified number of implants m_0, the conditional probability mass function for the number of malformations given R_0 non-viable fetuses, is evaluated through $\sum_{l=1}^{L} q_l(x_0)\text{Bin}(y; m_0 - R_0, \pi(Z_l(x_0)))$, where $q_l(x_0) = p_l\text{Bin}(R_0; m_0, \pi(U_l(x_0)))/\{\sum_t^L p_t\text{Bin}(R_0; m_0, \pi(U_t(x_0)))\}$. The posterior samples can be summarized with means and percentiles to provide posterior mean estimates and uncertainty bands for dose-response curves and probability mass functions for the response distributions; Section 3 reports such inferences for the DEHP data.

2.3.3 Prior specification

To specify the uniform priors for the range parameters ρ and ϕ, we consider the limiting case of the DDP model with $\alpha \to 0^+$, which yields the kernel of the mixture in (2) as the model's first stage with $G_{0\mathcal{X}}$ defining Gaussian process priors for the Binomial probabilities on the logistic scale. Then, under the exponential correlation function, $3/\rho$ is the range of dependence, that is, the distance between dose levels that yields correlation 0.05 for the Gaussian process realizations that define the probability of a non-viable fetus, and, analogously, for $3/\phi$. The range is usually assumed to be a fraction of the maximum interpoint distance over the index space. Let $D_{\max}$ be the maximum distance between observed doses. Since $3/b_\rho < 3/\rho$, we specify b_ρ such that $3/b_\rho = rD_{\max}$ for a small r; $r = 0.002$ was used for the DEHP data analysis in Section 3 leading to a Unif$(0, 10)$ prior for ρ; the same uniform prior was used for ϕ. This approach to prior specification for ρ and ϕ is conservative, in particular, the posterior distributions for ρ and ϕ are concentrated on values substantially smaller than b_ρ and b_ϕ.

We set the prior means for ξ_0 and β_0 to 0, and the shape parameters of the inverse gamma priors for τ^2 and σ^2 to 2, implying infinite prior variance. The prior variances for ξ_0 and β_0, and the prior means for ξ_1, β_1, τ^2 and σ^2 are chosen by studying the induced prior distribution for the dose-response curves defined in Section 2.2. For the DEHP data, we placed a N$(0, 10)$ prior on ξ_0 and β_0, an exponential prior with mean $b_\xi^{-1} = b_\beta^{-1} = 0.1$ on ξ_1 and β_1, and an inv-gamma$(2, 10)$ on τ^2 and σ^2. Under this prior choice, the prior means for functions $D(x)$ and $M(x)$ have a relatively weak increasing trend starting around 0.5, with 90% uncertainty bands

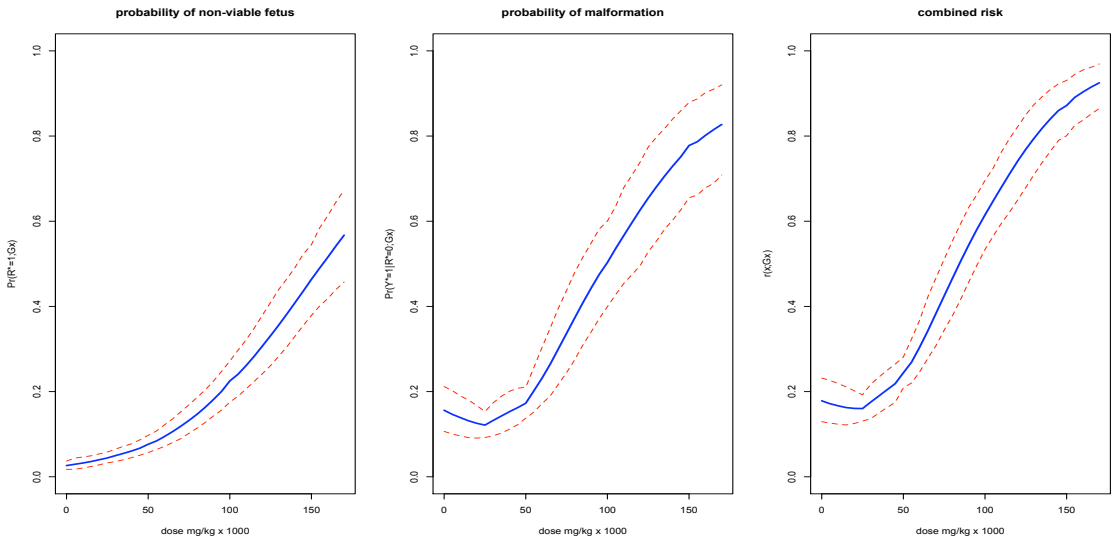

Figure 3: For the DEHP data, the posterior mean (solid lines) and 90% interval bands (dashed lines) for the risk assessment functions: probability of a non-viable fetus (left panel); conditional probability of malformation (middle panel); combined risk (right panel).

that cover almost the entire unit interval.

The DDP prior precision parameter, α, controls the number, n^*, of distinct mixture components [e.g., 7]. In particular, for moderate to large sample sizes, a useful approximation to the prior expectation $\mathrm{E}(n^* \mid \alpha)$ is given by $\alpha \log\{(\alpha + n)\alpha^{-1}\}$. This expression can be averaged over the prior for α to obtain $\mathrm{E}(n^*)$, thus selecting the gamma prior parameters to agree with a guess at the expected number of distinct mixture components. A gamma$(2, 1)$ prior was used for the DEHP data example corresponding to $\mathrm{E}(n^*) \approx 5$. Prior sensitivity analysis revealed robust posterior inference under more dispersed priors.

Finally, the level L for the DP truncation approximation can be chosen using standard distributional properties for the weights arising from the stick-breaking structure in (1). For instance, $\mathrm{E}(\sum_{l=1}^{L} \omega_l \mid \alpha) = 1 - \{\alpha/(\alpha + 1)\}^L$, which can be averaged over the prior for α to estimate $\mathrm{E}(\sum_{l=1}^{L} \omega_l)$. Given a tolerance level for the approximation, this expression is solved numerically to obtain the corresponding value L. For the analysis of the DEHP data, we used $L = 50$, which yields $\mathrm{E}(\sum_{l=1}^{L} \omega_l) \approx 0.9999593$ under the gamma$(2, 1)$ prior for α.

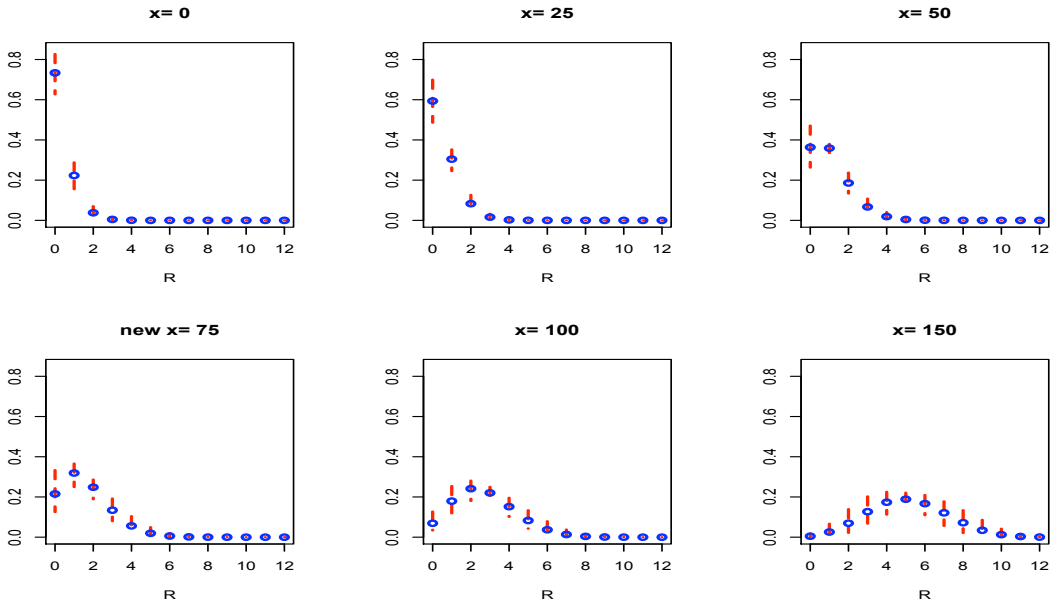

Figure 4: The posterior mean ("o") and 90% probability bands (dashed lines) of the probability mass functions for the number of non-viable fetuses given $m = 12$ implants.

3 Data example

We illustrate the proposed DDP mixture modelling approach with the DEHP data set discussed in the Introduction (Figure 1). It is known that plasticizers, such as the DEHP plasticizing agent, may leak in small quantities from plastic containers with various solvents such as food or milk. The possibility of toxic effects from these agents have been recognized and tested in developmental toxicity studies such as the one described in [23]. Recall that the two endpoints are non-viable fetus corresponding to resorption or actual prenatal death, and malformation involving external, visceral or skeletal malformation of a live fetus.

Figure 3 plots the posterior mean and 90% interval estimates for the three dose-response curves developed in Section 2.2 for risk assessment. The probability of a non-viable fetus across dose levels is a monotonically increasing function, with uncertainty bands around the posterior mean estimate that increase with increasing dose values, consistent with the decreasing number of dams for larger dose levels. The conditional probability of malformation, however, reveals a non-monotonic behavior at the low dose levels, and this J-shaped pattern carries over to the combined risk which also exhibits the dip in the probability from the control through dose 25

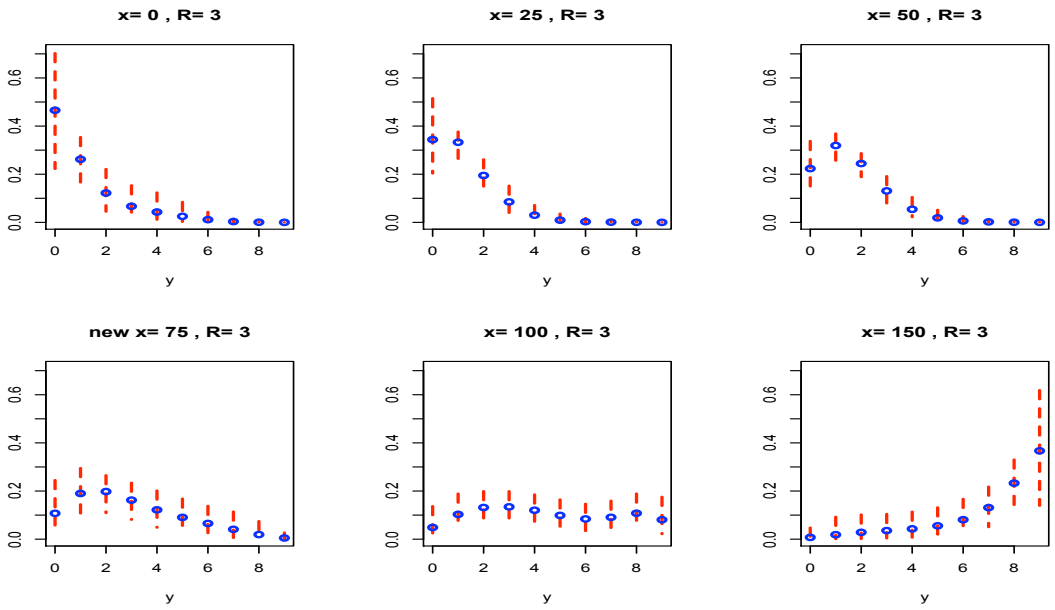

Figure 5: The posterior mean ("o") and 90% interval bands (dashed lines) of the probability mass functions for the number of malformations given $m = 12$ implants and $R = 3$ non-viable fetuses.

mg/kg $\times$ 1000. The inference for the combined risk function agrees with the estimated dose-response curve for the combined negative outcomes version of the DEHP data, as obtained in [10] based on a DDP Binomial mixture model. The modelling approach presented in this paper is key to uncovering the malformation endpoint as the one that contributes to the non-monotonic, possibly hormetic, combined dose-response relationship.

Inference for response distributions is illustrated with posterior mean and 90% interval estimates for the probability mass function of the number of non-viable fetuses given $m = 12$ implants (Figure 4) and the number of malformations given $m = 12$ implants and $R = 3$ non-viable fetuses (Figure 5). Results are reported for the control group, the four effective dose levels, and a new dose at $x = 75$ mg/kg $\times$ 1000. As expected, there is more uncertainty in the estimation of the conditional response distributions for malformation. The interpolation at the new dose level appears to be influenced more by the distribution at dose 50, which can be attributed to the larger sample size relative to dose 100. The estimated distributions for the number of non-viable fetuses have relatively standard shapes, whereas there is some evidence of a bimodal shape at dose 100, and skewness in the estimated malformation distributions.

4 Summary

We have developed a Bayesian nonparametric modelling approach for risk assessment in developmental toxicity studies. The motivation for the proposed methodology is that it is critical to model flexibly the dose-dependent dam specific response distribution associated with the clustered categorical outcomes of a non-viable fetus and of malformation for a live pup. The model is built from a mixture with a product Binomial kernel, to capture the nested structure of the responses, and a dependent Dirichlet process prior for the dose-dependent mixing distributions. The resulting nonparametric DDP mixture model provides rich inference for the response distributions as well as for the dose-response curves. Data from a toxicity experiment involving a plasticizing agent were used to illustrate the scientifically relevant feature of the DDP mixture model with regard to estimation of different dose-response relationships for different endpoints, including non-monotonic dose-response curves.

Acknowledgements

This research is part of the Ph.D. dissertation of Kassandra Fronczyk, completed at University of California, Santa Cruz, and was supported in part by the National Science Foundation under award DEB 0727543, and by a Special Research Grant awarded by the Committee on Research, University of California, Santa Cruz.

References

[1] Calabrese, E. J. (2005). Paradigm lost, paradigm found: The re-emergence of hormesis as a fundamental dose response model in the toxicological sciences. *Environmental Pollution*, **138**, 378–411.

[2] Chen, J.J., Kodell, R.L., Howe, R.B., and Gaylor, D.W. (1991). Analysis of trinomial responses from reproductive and developmental toxicity experiments. *Biometrics*, **47**, 1049–1058.

[3] DeIorio, M., Johnson, W. O., Müller, P., and Rosner, G. L. (2009). Bayesian nonparametric nonproportional hazards survival modeling. *Biometrics*, **65**, 762–771.

[4] DeIorio, M., Müller, P., Rosner, G. L., and MacEachern, S. N. (2004). An ANOVA model for dependent random measures. *Journal of the American Statistical Association*, **99**, 205–215.

[5] Dominici, F. and Parmigiani, G. (2001). Bayesian semiparametric analysis of developmental toxicology data. *Biometrics*, **57**, 150–157.

[6] Dunson, D., Chen, Z., and Harry, J. (2003). A Bayesian approach for joint modeling of cluster size and subunit-specific outcomes. *Biometrics*, **59**, 521–530.

[7] Escobar, M.D. and West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, **90**, 577–588.

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

[9] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, **1**, 209–230.

[10] Fronczyk, K. and Kottas, A. (2010). A Bayesian nonparametric modeling framework for developmental toxicity studies. Technical Report UCSC-SOE-10-11, University of California Santa Cruz, Department of Applied Mathematics and Statisics.

[11] Gelfand, A. E., Kottas, A., and MacEachern, S. (2005). Bayesian nonparametric spatial modeling with dirichlet process mixing. *Journal of the American Statistical Association*, **100**, 1021–1035.

[12] Guindani, M. and Gelfand, A. E. (2006). Smoothness properties and gradient analysis under spatial Dirichlet process models. *Methodology and Computing in Applied Probability*, **8**, 159–189.

[13] Ishwaran, H. and James, L. (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, **96**(453), 161–173.

[14] Ishwaran, H. and Zarepour, M. (2000). Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika*, **87**(2), 371–390.

[15] Kottas, A., Duan, J., and Gelfand, A. E. (2008). Modeling disease incidence data with spatial and spatio-temporal Dirichlet process mixtures. *Biometrical Journal*, **50**, 29–42.

[16] Kottas, A. and Krnjajić, M. (2009). Bayesian semiparametric modelling in quantile regression. *Scandinavian Journal of Statistics*, **36**, 297–319.

[17] Krewski, D. and Zhu, Y. (1994). Applications of multinomial dose-response models in developmental toxicity risk assessment. *Risk Analysis*, **14**, 613–627.

[18] MacEachern, S.N. (2000). Dependent Dirichlet processes. Technical report, Ohio State University, Department of Statistics.

[19] Nott, D. J. and Kuk, A. Y. C. (2009). Analysis of clustered binary data with unequal cluster sizes: A semiparametric Bayesian approach. *Journal of Agricultural, Biological, and Environmental Statistics*, **15**, 101–118.

[20] Rodriguez, A. and ter Horst, E. (2008). Bayesian dynamic density estimation. *Bayesian Analysis*, **3**, 339–366.

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

[22] Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica*, **4**, 639–650.

[23] Tyl, R.W., Jones-Price, C., Marr, M.C., and Kimmel, C.A. (1983). Teratologic evaluation of diethylhexyl phthalate (cas no. 111-81-7). Final Study Report for NCTR/NTP contract 222-80-2031(c), NITS PB85105674, National Technical Information Service, Springfield, Virginia.

[24] Zhu, Y., Krewski, D., and Ross, W.H. (1994). Dose-response models for correlated multinomial data from developmental toxicity studies. *Applied Statistics*, **43**, 583–598.