 DOI: 10.1111/j.1541-0420.2005.00506.x

A Potential Outcomes Approach to Developmental Toxicity Analyses

Michael R. Elliott,[1,2,*] Marshall M. Joffe,[3] and Zhen Chen[3]

[1]Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, Michigan 48109, U.S.A.

[2]Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, Michigan 48106, U.S.A.

[3]Department of Biostatistics and Epidemiology, University of Pennsylvania, 423 Guardian Drive, Philadelphia, Pennsylvania 19104, U.S.A.

[*]*email:* mrelliot@umich.edu

Summary. Estimating the effects of a toxin on fetal development in animal models such as mice can be problematic, because the number of pups that develop and survive until birth may simultaneously affect developmental outcomes such as birth weight and be affected by the introduction of a toxin into the fetal environment. Also, comparing pups that survived until birth at a high dose of the toxin with pups that survived at low doses may underestimate the effect of the toxin, because the lower dose means include the less healthy pups that would not survive if exposed to a higher level of toxin. We consider this problem in a potential outcomes framework that defines the effect of the dose on the outcome as the difference between what the outcome would have been for a pup had the dam in which the pup develops been exposed to dose level $Z = z^*$ rather than dose level $Z = z$. To disentangle the direct effect of dose from the effect of litter size, we focus on effects defined within principal strata that are a function of the survival status of the pups at each of the possible dose levels. A unique contribution to the potential outcomes literature is that we allow the outcome for a subject to be dependent on the principal stratum to which other subjects within a cluster belong.

Key words: Causal model; Clustering; Litter size; Principal effects; Principal strata.

1. Introduction

In developmental toxicology, the number of pups that develop and survive until birth may simultaneously affect developmental outcomes such as birth weight and be affected by the introduction of a toxin into the fetal environment. Catalano and Ryan (1992) attempted to adjust for this effect by including litter size as a covariate in their model, which considered a joint outcome of birth weight and fetal malformation. Dunson, Chen, and Harry (2003) considered a latent variable model that modeled birth weight, fetal malformation, and live pup litter size using a random effect common to both model components to induce dependence. However, these approaches do not directly address the fact that, if experiments are being conducted in placental mammals where multiple births are common (e.g., mice), had the developmental environment been different, a different number of pups might have survived to birth, affecting the amount of nutrients and space available for growth in the womb. Further, pups that survive to birth at a high dose level of a toxin might be healthier on average, and thus any comparison between the observed outcome under high doses with the observed outcome under low doses might underestimate the effect of the toxin, because the lower dose means include the less healthy pups that would not survive if exposed to a higher level of toxin.

To deal with these issues we consider an alternative "counterfactual" or "potential outcomes" approach (Neyman, 1923; Rubin, 1974, 1978) in which the effect of the dose on the outcome is defined as the difference between what the outcome would have been for the mice in a dam had the dam been exposed to dose level $Z = z^*$ rather than dose level $Z = z$. A particular problem that we encounter in making an inference in this context is censoring by death (Robins, 1986; Frangakis and Rubin, 2002; Gilbert, Bosch, and Hudgens, 2003): mice pups might survive under one dose assignment and not another, which implies that we are comparing different subsets of mice under the different dosing regimes, as well as failing to account for the effect of the differing numbers of surviving pups under the different dosing regimes. Hence, we develop principal strata (Imbens and Rubin, 1997; Frangakis et al., 2004) that are a function of the survival status of the pups at each of the possible dose levels, and obtain estimates of "principal effects" as intent-to-treat (ITT) contrasts within principal strata (Frangakis and Rubin, 2002). As in Frangakis, Rubin, and Zhou (2002), we accommodate correlation within a litter "cluster" of pups for both the outcome and principal strata. Our primary contribution to the potential outcomes literature is that, in order to account for the potentially different number of surviving littermates under the different dosing

regimes, we allow the outcome for a subject to be dependent on the principal stratum (PS) to which other subjects within a cluster belong.

Section 2 discusses developmental toxicology analyses, defines principal strata in terms of the potential survival status at each dose level, and describes models for the principal strata and the potential outcomes at each dose level in terms of the principal strata of both the index pup and the other pups in the litter. Section 2 also discusses the within-principal-strata contrasts or "principal effects," which might be direct effects relevant to contrasts in which the number of surviving pups is unchanged at different potential doses, or litter-size-adjusted effects that account for the differing number of surviving pups under different potential doses. Section 3 considers an application of the method to data from a developmental toxicology study of ethylene glycol in mice conducted by the National Toxicology Program (NTP) (Price et al., 1985). Section 4 discusses the results of the NTP analysis in relation to previous analyses of these data, and considers limitations of the approach as well as future extensions.

2. Methodology

2.1 *Developmental Toxicology Data*

In a typical type II toxicology experiment, each of $j = 1, \ldots, N$ dams are impregnated by male mice. Shortly after impregnation, a toxin of quantity $x(Z_j)$ at level $Z_j = 0, \ldots, L$ is administered at random to the jth dam. Just before birth, the dam is sacrificed and the individual pups $i = 1, \ldots, n_j$ in its litter are examined. Survival status D_{ij} for each pup is determined, usually as a three-level variable: alive, dead, and resorbed, where the latter are ascertained from implantation sites on the uterus. Outcomes Y_{ij} such as birth weight and developmental deformities are also determined, which typically are defined only if the pup is alive at the time of sacrifice.

In our application to the National Toxicology Program data on exposure to ethylene glycol (Price et al., 1985), we make a "monotonicity" assumption that is critical to the identifiability of the model presented: that pups that die under a given dose of toxin would also die if exposed to a higher dose. We make three additional simplifications in order to not overly complicate the presentation of key ideas. First, we restrict the outcome Y to birth weight only. Second, as only 12 of the 1192 pups were considered dead at sacrifice, we combine these with the 152 resorbed pups into a single category, denoted as "dead" below. Third, we assume that the litter size n_j, which includes both the live and dead/resorbed pups, is the same at the time of exposure as at the time of sacrifice.

2.2 *Defining Principal Strata and Principal Effects*

The dose level of the administered toxin is considered to be a potential "controllable factor" with levels $z = 0, \ldots, L$. Denote the potential survival status of the ith pup in the jth litter by $D_{ij}^{(z)}$: $D_{ij}^{(z)} = 0$ if the pup would have died at dose level z and $D_{ij}^{(z)} = 1$ if it would have survived. Denote the observed survival status by $D_{ij}^{(Z_j)}$. Similarly denote the potential outcome at dose level z by $Y_{ij}^{(z)}$ and the observed outcome at the actual dose level assignment Z_j by $Y_{ij}^{(Z_j)}$. Our models assume $Y_{ij}^{(z)}$ is defined only if $D_{ij}^{(z)} = 1$. We then define a PS by the vector of "survival" statuses $D_{ij}^{(z)}$ at each of the potential dose levels: $(D_{ij}^{(0)}, \ldots, D_{ij}^{(L)})'$, and denote the PS to which the pup belongs by the scalar random variable S_{ij} that represents a particular pattern of $(D_{ij}^{(0)}, \ldots, D_{ij}^{(L)})'$. Denote the vector of potential outcomes at dose z for all pups in the jth litter by $Y_j^{(z)} = \{Y_{1j}^{(z)}, \ldots, Y_{n_j j}^{(z)}\}$, and similarly denote S_j. The goal of our analysis is to compare potential outcomes under the different dose levels z^* within a PS level s, which Frangakis and Rubin (2002) term "principal effects":

$$\left\{Y_{ij}^{(z^*)} : S_{ij} = s\right\} \quad \text{and} \quad \left\{Y_{ij}^{(z)} : S_{ij} = s\right\}.$$

These principal effects may be viewed as ITT contrasts within principal strata. We consider ITT contrasts of functions of $Y_{ij}^{(z)}$ under a specific parametric model that accounts for the PS memberships of the other pups in the litter.

We make three standard assumptions to assist in the identification of the potential outcomes model: randomization, stable unit treatment value assumption or SUTVA, and monotonicity. Because of the randomization mechanism, we assume that the potential birth weight and survival outcome at each dose level are independent of the actual dose: $Y_{ij}^{(0)}, \ldots, Y_{ij}^{(L)}$, $S_{ij} \perp Z_j$. Thus, the PS to which a pup belongs is independent of its dose assignment. Also, because the assignment of dose $Z_{ij} \equiv Z_j$ is the same for all of the pups within a litter, and we assume that the potential birth weight outcomes are independent across litters, the SUTVA (Rubin, 1990) that the birth weight of the ith pup in the jth litter is not affected by the assignment of dose to the i'th pup in the j'th litter is preserved: if $j' = j$, then the dose assigned to the ith and i'th pup must be the same; if $j' \neq j$, then the dose assigned to the ith pup will not affect the outcome for the i' pup. Finally, if we make the monotonicity assumption that a pup that dies at level z will also die at $z^* > z$, then $D_{ij}^{(z^*)} \leq D_{ij}^{(z)}$ for all $z^* > z$, and Table 1 illustrates that there are $L + 2$ possible patterns of survival statuses that define the values of S_{ij}. $S = 1$, then, consists of the pups that would survive at all dose levels, or the "healthiest" mice pups. Less "healthy" mice belonging to higher principal strata: $S = 2$ consists of the pups that survive at all but the highest dose level, $S = 3$ consists of those that survive only at dose levels below the highest two, and so on, up to $S = L + 1$, which consists of pups that survive only at zero dose level, and $S = L + 2$, which consists of pups that die at all dose levels. Even though we observe $D_{ij}^{(z)}$ only for the observed dose Z_j, under the assumption of monotonicity we might observe the PS membership for some pups when the jth dam is assigned to either $Z_j = 0$—dead pups at the lowest dose belong to $S_{ij} = L + 2$, or $Z_j = L$—surviving pups at the

Table 1
Definition of principal strata under monotonicity assumption. $D^{(z)}$ indicates survival status (0 = dead, 1 = alive) at dose $Z = z$.

S	$Z = 0$	$Z = 1$	$\cdots$	$Z = L - 1$	$Z = L$
1	$D^{(0)} = 1$	$D^{(1)} = 1$	$\cdots$	$D^{(L-1)} = 1$	$D^{(L)} = 1$
2	$D^{(0)} = 1$	$D^{(1)} = 1$	$\cdots$	$D^{(L-1)} = 1$	$D^{(L)} = 0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$L + 1$	$D^{(0)} = 1$	$D^{(1)} = 0$	$\cdots$	$D^{(L-1)} = 0$	$D^{(L)} = 0$
$L + 2$	$D^{(0)} = 0$	$D^{(1)} = 0$	$\cdots$	$D^{(L-1)} = 0$	$D^{(L)} = 0$

highest dose belong to $S_{ij} = 1$. Parametric assumptions about the potential outcome also assist in model identification, because the resulting mixture models will have mixing fractions that are independent of dose under randomization.

2.3 *Modeling Principal Strata and Potential Outcomes*

We consider the joint distribution of potential outcomes and survival statuses by factoring $f(Y^{(z)}, S)$ into $f(S)$ and $f(Y^{(z)} \mid S)$. The PS to which the ith pup in the jth litter belongs is modeled using a "proportional odds random effect" model that assumes that the log odds of a pup belonging to PS level l versus $l-1$ changes by b_j for all pups within the jth litter, relative to the log odds for the overall population

$$\text{logit}(P(S_{ij} \le s) \mid b_j) = \theta^s + b_j, \quad s = 1, \ldots, L+2 \qquad (1)$$
$$b_j \sim N\big(0, \tau_b^2\big),$$

where by definition $\theta^0 = -\infty$ and $\theta^{L+2} = \infty$. The random effect b_j is used to account for any within-litter correlations of PS memberships. Assuming that the vector of PS memberships S_j is known for all pups i in litter j, the potential outcome for the birth weight of the ith pup in the jth litter at the zth dose level is modeled as

$$Y_{ij}^{(z)} \,\Big|\, \mu^s, \beta^s, \gamma_0^s, \gamma_1^s, \sigma_s^2, a_j, S_j \overset{\text{ind}}{\sim}$$
$$\begin{cases} N\big(\mu^s + \beta^s x(z) + \gamma_0^s n_{0ij}^{(z)} + \gamma_1^s n_{1ij}^{(z)} + a_j, \sigma_s^2\big) & \text{if } D_{ij}^{(z)} = 1 \text{ for } S_{ij} = s \\ \text{does not exist} & \text{if } D_{ij}^{(z)} = 0 \text{ for } S_{ij} = s \end{cases} \qquad (2)$$
$$a_j \sim N(0, \tau^2),$$

where $x(z)$ is the quantity of the toxin received at dose level z, $n_{dij}^{(z)} = \sum_{k,k\neq i} I(D_{kj}^{(z)} = d)$ is the number of pups in litter j other than pup i with survival status d at dose level z (which would be known if S_j was known), and s is the PS membership of the ijth pup. This model assumes that there is a PS-specific intercept μ^s, a PS-specific linear dose effect β^s, and PS-specific linear litter size effects that may differ by whether the littermates are dead (γ_0^s) or alive (γ_1^s). The random effect a_j common to all pups in the litter is used to account for any within-litter correlations of birth weights. Under the monotonicity assumption the number of potential birth weight outcomes for a pup in the sth PS is restricted to $L + 2 - s$, because $Y_{ij}^{(z)}$ exists if and only if $D_{ij}^{(z)} = 1$.

Litter size is associated with birth weight (see Figure 1 in Dunson et al., 2003), and it is included in the birth weight model through $n_{dij}^{(z)}$, the littermate survival status in (2). It is important to understand that littermate survival status $n_{dij}^{(z)}$ is a potential outcome like $Y_{ij}^{(z)}$—at each dose level z the survival status of each pup in the litter could be observed and $n_{dij}^{(z)}$ computed. Because $n_{dij}^{(z)}$ is a deterministic function of the PS of the other pups in a litter and of z, the dose regime under the control of the investigator, it meets the requirements of a potential outcome in the sense of Rubin (1978). Hence, across all pups in the jth litter, there are $(L+2)^{n_j}$ patterns of PS memberships to be considered before any data are collected; the actual dose received and the observed survival status of the pups restrict the PS membership possibilities to a subset of these. To make this clearer, consider a litter with two pups, and assume that the experiment consists of the dam being exposed to one of two dose levels. Thus, there are three PS: $S = 1$ is defined by pups that survive at both dose levels, $S = 2$ by pups that survive only at the lowest dose level, and $S = 3$ by pups that do not survive at any dose level. Table 2 illustrates the littermate survival status potential outcomes when $Z_j = 1$ (dam received high dose). If both pups survive, then $S_{1j} = S_{2j} = 1$, and we have that $n_{0ij}^{(0)} = n_{0ij}^{(1)} = 0$ and $n_{1ij}^{(0)} = n_{1ij}^{(1)} = 1$ for $i = 1, 2$. If only the first pup survives, then we know that $S_{1j} = 1$ and that $S_{2j} = 2$ or $S_{2j} = 3$. If $S_{2j} = 2$, then the second pup would survive at the low dose

Table 2

Potential littermate status for a two-pup litter, under a two-dose regime, where dam receives high dose, under different observed survival patterns for the pups. NA indicates that a principal stratum membership pattern is impossible, given the observed survival statuses and dose assignment.

				Potential littermate survival status			
				$n_{0ij}^{(z)}$		$n_{1ij}^{(z)}$	
	Pup	S_{1j}	S_{2j}	$z=0$	$z=1$	$z=0$	$z=1$
$D_{1j}^{(Z_j=1)} = 1, D_{2j}^{(Z_j=1)} = 1$	$i=1$	1	1	0	0	1	1
		1	2,3	NA	NA	NA	NA
		2,3	1,2,3	NA	NA	NA	NA
	$i=2$	1	1	0	0	1	1
		2,3	1	NA	NA	NA	NA
		1,2,3	2,3	NA	NA	NA	NA
$D_{1j}^{(Z_j=1)} = 1, D_{2j}^{(Z_j=1)} = 0$	$i=1$	1	1	NA	NA	NA	NA
		1	2	0	1	1	0
		1	3	1	1	0	0
		2,3	1,2,3	NA	NA	NA	NA
$D_{1j}^{(Z_j=1)} = 0, D_{2j}^{(Z_j=1)} = 1$	$i=2$	1	1	NA	NA	NA	NA
		2	1	0	1	1	0
		3	1	1	1	0	0
		1,2,3	2,3	NA	NA	NA	NA

level, and thus $n_{01j}^{(0)} = 0$, $n_{01j}^{(1)} = 1$ and $n_{11j}^{(0)} = 1$, $n_{11j}^{(1)} = 0$. If $S_{2j} = 3$, then the second pup does not survive at either dose level, and thus $n_{01j}^{(0)} = n_{01j}^{(1)} = 1$ and $n_{11j}^{(0)} = n_{11j}^{(1)} = 0$. A similar pattern is observed when only the second pup survives, with the first and second pup indices permuted. When neither pup survives, then determination of $n_{dij}^{(z)}$ is irrelevant, because neither pup will be used to estimate the effect of dose on birth weight.

2.4 *Estimation*

Consider the complete data to consist of the dose assignments, the potential survival statuses for each pup at each dose level that define the principal strata, and the potential outcomes at each dose level that, in the analysis below, are defined only if the pup survives at the given dose level. For each litter, then, the complete data distribution is given by

$$\begin{aligned}
&f\big(Z_1,\ldots,Z_N,Y_{11}^{(0)},\ldots,Y_{11}^{(L)},Y_{21}^{(0)},\ldots,Y_{21}^{(L)},\ldots,\\
&\quad Y_{n_1,1}^{(0)},\ldots,Y_{n_1,1}^{(L)},\ldots,\ldots,Y_{n_N,N}^{(0)},\ldots,Y_{n_N,N}^{(L)},\\
&\quad D_{11}^{(0)},\ldots,D_{11}^{(L)},D_{21}^{(0)},\ldots,D_{21}^{(L)},\ldots,D_{n_1,1}^{(0)},\ldots,\\
&\quad D_{n_1,1}^{(L)},\ldots,\ldots,D_{n_N,N}^{(0)},\ldots,D_{n_1,N}^{(L)}\big)\\
&=\prod_j f(Z_j=z)f\big(Y_{1j}^{(0)},\ldots,Y_{1j}^{(L)},Y_{2j}^{(0)},\ldots,Y_{2j}^{(L)},\ldots,\\
&\qquad\qquad Y_{n_jj}^{(0)},\ldots,Y_{n_jj}^{(L)},D_{1j}^{(0)},\ldots,D_{1j}^{(L)},\\
&\qquad\qquad D_{2j}^{(0)},\ldots,D_{2j}^{(L)},\ldots,D_{n_jj}^{(0)},\ldots,D_{n_jj}^{(L)}\big), \qquad (3)
\end{aligned}$$

where equality in (3) follows from the SUTVA and randomization assumptions. As our inference of interest is with respect to Y and D, we integrate out the unobserved potential outcomes from each of the j elements in the right factor in (3). Thus

$$\begin{aligned}
&f\big(Y_j^{(Z_j)},D_j^{(Z_j)}\,\big|\,a_j,b_j,\zeta^Y,\zeta^S\big)\\
&=\int\!\!\int f\big(Y_{1j}^{(0)},\ldots,Y_{1j}^{(L)},Y_{2j}^{(0)},\ldots,Y_{2j}^{(L)},\ldots,Y_{n_jj}^{(0)},\ldots,\\
&\qquad Y_{n_jj}^{(L)},D_{1j}^{(0)},\ldots,D_{1j}^{(L)},D_{2j}^{(0)},\ldots,D_{2j}^{(L)},\ldots,\\
&\qquad D_{n_jj}^{(0)},\ldots,D_{n_jj}^{(L)}\,\big|\,a_j,b_j,\zeta^Y,\zeta^S\big)\,dY_j^{\text{mis}}\,dD_j^{\text{mis}}\\
&=\prod_i\int\!\!\int f\big(Y_{ij}^{(0)},\ldots,Y_{ij}^{(L)}\,\big|\,D_{1j}^{(0)},\ldots,D_{1j}^{(L)},D_{2j}^{(0)},\ldots,\\
&\qquad D_{2j}^{(L)},\ldots,D_{n_jj}^{(0)},\ldots,D_{n_jj}^{(L)},a_j,\zeta^Y\big)\\
&\quad\times f\big(D_{ij}^{(0)},\ldots,D_{ij}^{(L)}\,\big|\,b_j,\zeta^S\big)\,dY_j^{\text{mis}}\,dD_j^{\text{mis}}\\
&=\prod_i\int\!\!\int f\big(Y_{ij}^{(0)},\ldots,Y_{ij}^{(L)}\,\big|\,S_{1j},\ldots,S_{n_jj},a_j,\zeta^Y\big)\\
&\quad\times f\big(S_{ij}\mid b_j,\zeta^S\big)\,dY_j^{\text{mis}}\,dD_j^{\text{mis}}\\
&=\prod_i\sum_{s=1}^{L-Z_j+1}p\big(S_{ij}=s\,\big|\,b_j,\zeta^S\big)\\
&\quad\times p\big(Y_{ij}^{(Z_j)}\,\big|\,S_j,a_j,\zeta^Y\big)I\big(D_{ij}(Z_j)=1\big)\big)\\
&\quad+\sum_{L-Z_j+2}^{L+2}p\big(S_{ij}=s\,\big|\,b_j,\zeta^S\big)I\big(D_{ij}(Z_j)=0\big), \qquad (4)
\end{aligned}$$

where Y_j^{mis} and D_j^{mis} contain the elements of $\{Y_j^{(0)},\ldots,Y_j^{(L)}\}$ and $\{D_j^{(0)},\ldots,D_j^{(L)}\}$ for which $z \neq Z_j$, $\zeta^Y = (\mu^1,\ldots,\mu^{L+2}, \beta^1,\ldots,\beta^{L+2}, \gamma_0^1,\ldots,\gamma_0^{L+2}, \gamma_1^1,\ldots,\gamma_1^{L+2}, \sigma_1^2,\ldots,\sigma_{L+2}^2, \tau^2)$ parameterizes the outcome of interest, and $\zeta^S = (\theta^1,\ldots,\theta^{L+1}, \tau_b^2)$ parameterizes PS membership. The second equality in (4) follows from the PS model (1), which assumes that the potential survival statuses of the pups within a litter are independently conditional on a common random effect b_j; the birth weight model (2), which implies that $Y_{ij}^{(z)} \perp Y_{ij'}^{(z)} \mid a_j, S_j$, that is, only the survival status of the other pups in the litter affects the outcome of a given pup, conditional on the common outcome random effect a_j; and the independence assumption of a_j and b_j. The third equality in (4) follows from the principal strata notation. The fourth equality in (4) follows from the structure of the missing potential observations—living pups at observed dose Z_j must belong to principal strata 1 through $L - Z_j + 1$, while those dead at dose Z_j must belong to principal strata $L - Z_j + 2$ through $L + 2$.

Taking a fully hierarchical Bayesian approach, then, the posterior distribution of ζ^Y and ζ^Z is given by

$$\begin{aligned}
&f(\zeta^Y,\zeta^S\mid Y^{(Z)},D^{(Z)})\\
&\quad\propto\left[\prod_j\int\!\!\int f\big(Y_j^{(Z_j)},D_j^{(Z_j)}\,\big|\,a_j,b_j\zeta^Y,\zeta^S\big)\right.\\
&\qquad\left.\times p\big(a_j\big|\tau^2\big)p\big(b_j\big|\tau_b^2\big)\,da_j\,db_j\right]p(\zeta^Y,\zeta^S), \qquad (5)
\end{aligned}$$

where $p(a_j \mid \tau^2) \overset{\text{ind}}{\sim} N(0,\tau^2), p(b_j \mid \tau_b^2) \overset{\text{ind}}{\sim} N(0,\tau_b^2)$, and $p(\zeta^Y,\ \zeta^S) = \prod_s[p(\mu^s,\ \beta^s,\ \gamma_0^s,\ \gamma_1^s)p(\theta^s)p(\sigma_s^2)]p(\tau^2)p(\tau_b^2)$ for $p(\mu^s,\beta^s,\gamma_0^s,\gamma_1^s) \overset{\text{ind}}{\sim} N(\beta_0,\Omega), p(\theta^s \overset{\text{ind}}{\sim} N(\theta^0,\Sigma), p(\sigma_s^2) \sim \text{Inv}-\chi_a^2(b), p(\tau^2) \sim \text{Inv}-\chi_c^2(d), p(\tau_b^2) \sim \text{Inv}-\chi_e^2(f)$. Hyperpriors β_0, Ω, θ^0, Σ, a, b, c, d, e, and f are assumed known. Simulations from the posterior distribution (5) are obtained using a Markov chain Monte Carlo (MCMC) approach that combines Gibbs sampling (Gelfand and Smith, 1990) with Metropolis–Hastings draws. To facilitate computation, we obtain draws of a_j, b_j, and S_{ij} as part of the MCMC sampling routine. The full conditional draws are as follows:

1. $f(S_{ij} \mid \text{rest}) \sim \text{Multi}(1, \pi_{ij}^{s^*})$ where

$$\begin{cases}
\pi_{ij}^{s^*} = \dfrac{\pi_{s_j}\phi\big(Y_{ij}^*;\eta_{ij}^s,\sigma_s^2\big)}{\sum_{s=1}^{L-Z_j+1}\pi_{s_j}\phi\big(Y_{ij}^*;\eta_{ij}^s,\sigma_s^2\big)} & \text{if } D_{ij}(Z_j)=1,\\
\pi_{ij}^{s^*} = \dfrac{\pi_{s_j}}{\sum_{s=L-Z_j+2}^{L+2}\pi_{s_j}} & \text{if } D_{ij}(Z_j)=0,
\end{cases}$$

   where $\pi_{s_j} = \frac{\exp(\theta^s+b_j)}{1+\exp(\theta^s+b_j)} - \frac{\exp(\theta^{s-1}+b_j)}{1+\exp(\theta^{s-1}+b_j)}$, $Y_{ij}^* = Y_{ij}^{(Z_j)} - a_j$, $\phi(z;\mu,\sigma^2)$ is the probability density function (PDF) of a normal distribution evaluated at z with mean μ and variance σ^2, and $\eta_{ij}^s = \mu^s + \beta^s x(Z_j) + \gamma_0^s n_{0ij}^{(Z_j)} + \gamma_1^s n_{1ij}^{(Z_j)}$.
2. $f(\theta \mid \text{rest})$ is obtained via sampling importance resampling (Smith and Gelfand, 1992): proposal draws of $f(\theta \mid \text{rest})$ are obtained by first fitting a multinomial

model to the latest PS membership draws with offsets given by the litter effects b_j, and then drawing from a multivariate normal model centered at the multinomial logit intercept parameter estimates with the associated parameter estimate covariance matrix, inflated by the factor c to ensure domination of the target distribution.

3. $f(b_j \,|\, \text{rest})$ is obtained via sampling importance resampling: proposal draws of $f(b_j \,|\, \text{rest})$ are obtained from an $N(0,\, c\tau_b^2)$ distribution where c again ensures domination of the target distribution.
4. $f(\tau_b^2 \,|\, \text{rest}) \sim \text{Inv} - \chi^2_{e+N}(f + Q_b^*)$, where $Q_b^* = N^{-1}\sum_j b_j^2$.
5. $f(\mu^s, \beta^s, \gamma_0^s, \gamma_1^s \,|\, \text{rest}) \sim N((D^s)^{-1}d_s, (D^s)^{-1})$ where $D^s = X^T W_s X + \Omega^{-1}$ and $d_s = X^T W_s Y^* + \Omega^{-1}\beta_0$, X is a design matrix consisting of an intercept, observed dose $x(Z)$, and observed dead and living littermate sizes $n_0^{(Z)}$ and $n_1^{(Z)}$, and W_s is a diagonal matrix with diagonal elements equal to σ_s^{-2} if $S_{ij} = s$ and $D_{ij} = 1$ and 0 otherwise.
6. $f(\sigma_s^2 \,|\, \text{rest}) \sim \text{Inv} - \chi^2_{a+N_s}(b + Q_s)$ where $N_s = \sum_i \sum_j \times I(S_{ij} = s)I(D_{ij}(Z_j) = 1)$ and $Q_s = N_s^{-1}\sum_i \sum_j I(S_{ij} = s)(Y_{ij}^* - \eta_{ij}^s)^2$.
7. $f(a_j \,|\, \text{rest}) \sim N(\hat{a}_j, V_{a_j})$ where $V_{a_j} = (\sum_i \sum_s I(S_{ij} = s)\sigma_s^{-2} + \tau^{-2})^{-1}$ and $\hat{a}_j = (\sum_i \sum_s I(S_{ij} = s)\tilde{Y}_{ij}^s/\sigma_s^2)/V_{a_j}$ for $\tilde{Y}_{ij}^s = Y_{ij}^{(Z_j)} - \eta_{ij}^s$.
8. $f(\tau^2 \,|\, \text{rest}) \sim \text{Inv} - \chi^2_{c+N}(d + Q^*)$ where $Q^* = N^{-1}\sum_j a_j^2$.

2.5 *Principal Effects of Interest*

Our focus is on estimating principal effects, that is, the posterior predictive distribution of ITT contrasts among pups that belong to a given PS and that have a common survival pattern across the potential dosing regimes. Alternatively, we might estimate the posterior predictive distribution of ITT contrasts among all pups for whom the potential outcomes exist at both dose levels $Z = z$ and $Z = z^*$, which may require averaging over several PS.

The posterior predictive distribution of the potential outcomes and survival statuses $Y^{(z)}$, $D^{(z)} \equiv S$ is given by

$$f(Y^{(z)}, S \mid Y^{(Z)}, D^{(Z)}) = \iint f(Y^{(z)} \mid \zeta^Y, S) f(S \mid \zeta^S) \times f(\zeta^Y, \zeta^S \mid Y^{(Z)}, D^{(Z)})\, d\zeta^Y d\zeta^Z.$$

Our MCMC sampling algorithm outlined above yields draws from $f(\zeta^Y,\, \zeta^S \,|\, Y^{(Z)},\, D^{(Z)})$ and from $f(S \,|\, \zeta^S)$. To obtain a draw $\tilde{Y}_{ij}^{(z)}$ from $f(Y^{(z)} \,|\, \zeta^Y,\, S)$, let

$$\tilde{Y}_{ij}^{(z)} = \mu^s + \beta^s x(z) + \gamma_0^s n_{0ij}^{(z)} + \gamma_1^s n_{1ij}^{(z)} + a_j + \epsilon_{ij}^{(z)},$$

where $\epsilon_{ij}^{(z)} \sim N(0,\, \sigma_s^2)$ and $\text{Corr}(\epsilon_{ij}^{(z^*)}, \epsilon_{ij}^{(z)}) = \rho_P$. The observed data tell us nothing about the within-pup correlation structure among potential outcomes $Y_{ij}^{(0)}, \ldots, Y_{ij}^{(L)}$ (Rubin, 1990; Imbens and Rubin, 1997); for ease of presentation and analysis, we consider $\rho_P = 1$, and thus a draw of $\tilde{Y}_{ij}^{(z)}$ is given by $Y_{ij}^{(Z_j)} + \beta^s(x(z) - x(Z_j)) + \gamma_0^s(n_{0ij}^{(z)} - n_{0ij}^{(Z_j)}) + \gamma_1^s(n_{1ij}^{(z)} - n_{1ij}^{(Z_j)})$. But by the law of large numbers the results are very insensitive to the choice of ρ_P if N is large, because the principal effects we consider below are sums of linear contrasts of potential outcomes.

We consider the posterior predictive distribution of two principal effects

$$\begin{aligned}\Psi_1^s(z^*, z) &= (N^s)^{-1}\sum_i \sum_j I(S_{ij} = s) \\ &\quad \times \big(\tilde{Y}_{ij}^{(z^*)} - \tilde{Y}_{ij}^{(z)} \,\big|\, S_{ij} = s, n_{dij}^{(z^*)} = n_{dij}^{(z)}\big) \\ &= \beta^s(x(z^*) - x(z))\end{aligned}$$

$$\begin{aligned}\Psi_2^s(z^*, z) &= (N^s)^{-1}\sum_i \sum_j I(S_{ij} = s) \\ &\quad \times \big(\tilde{Y}_{ij}^{(z^*)} - \tilde{Y}_{ij}^{(z)} \,\big|\, S_{ij} = s, n_{dij}^{(z^*)}, n_{dij}^{(z)}\big) \\ &= \beta^s(x(z^*) - x(z)) + \gamma_0\big(\bar{n}_{0^{(z^*)}}^s - \bar{n}_{0^{(z)}}^s\big) \\ &\quad + \gamma_1\big(\bar{n}_{1^{(z^*)}}^s - \bar{n}_{1^{(z)}}^s\big),\end{aligned}$$

where $N_s = \sum_i \sum_j I(S_{ij} = s)$ and $\bar{n}_{d^{(z)}}^s = N_s^{-1}\sum_i \sum_j \times I(S_{ij} = s)n_{dij}^{(z)}$. We restrict the set of z^*, z for which contrasts are made to those for which $D^{(z^*)} = D^{(z)} = 1$ when $S = s$. Contrast Ψ_1^s, which we term as the "direct" effect of dose of birth weight, considers the effect of the change in dose from z to z^* within a PS s among pups for whom $n_{dij}^{(z^*)} = n_{dij}^{(z)}$ for all d, that is, among pups in the subset of litters where the survival status of all pups in the litter happens to be unchanged between dose level z and z^*. This may be an important measure in extrapolating results to human effects, where singleton births are the norm. Alternatively, we might interpret the ITT effect of Ψ_1^s as the direct effect of the change in dose from z to z^* assuming that the survival status of pups in the litter could somehow be fixed to be the same under each dose level (Robins and Greenland, 1992; Pearl, 2001). Contrast Ψ_2^s, which we term the "litter-size-adjusted" effect of dose on birth weight, considers the difference in the sample average of the expected birth weights at level z^* and z among all pups belonging to PS level s. We also consider the posterior predictive distribution of the average of the principal effects over the set of principal strata for which contrasts between $Z = z^*$ and $Z = z$ are defined, assuming $z^* > z$:

$$\Psi_3(z^*, z) = \sum_{s=1}^{L-z^*+1} N_s \Psi_1^s(z^*, z) \Big/ \sum_{s=1}^{L-z^*+1} N_s$$

$$\Psi_4(z^*, z) = \sum_{s=1}^{L-z^*+1} N_s \Psi_2^s(z^*, z) \Big/ \sum_{s=1}^{L-z^*+1} N_s.$$

All contrasts considered assume that the total number of pups n_j in a litter is not directly affected by the dose. A marginal association between number of pups surviving and dose is induced by the assumption of monotonicity and the specific values of θ^s.

3. Application to National Toxicology Program Data

We consider an application of this method using data from a developmental toxicology study of ethylene glycol in mice

Table 3
Total number of litters (N_Z), total number of observed pups (n_Z), percentage surviving, and mean birth weight, by dose $x(Z)$ of ethylene glycol

Dose level Z	Dose $x(Z)$ (mg/kg)	No. litters N_Z	No. pups n_Z	Percentage surviving	Mean birth weight (g)
3	3.0	23	283	79.9	0.704
2	1.5	22	266	86.1	0.764
1	0.75	24	310	89.0	0.877
0	0	25	333	89.2	0.972

conducted by the National Toxicology Program (Price et al., 1985). Pregnant mice (dams) were exposed to ethylene glycol during organogenesis at one of four different dose levels: 0, 0.75, 1.5, and 3 mg/kg. These data have also been considered in Catalano and Ryan (1992), Molenberghs and Ryan (1999), Gueorguieva and Agresti (2001), and Dunson et al. (2003), among others.

We obtain our prior on the birth weight regression parameters β_0 as $\hat{\beta}_0$, the maximum likelihood regression estimates from linear regression of the observed birth weights on an intercept, the observed dose, the observed number of dead pups in the litter, and the observed number of surviving pups in the litter minus one (i.e., the number of surviving pups other than the index pup), and set Ω to $nV_{\hat{\beta}_0}$, where $V_{\hat{\beta}_0}$ is the associated maximum likelihood covariance matrix of $\hat{\beta}_0$. Also, in the same very weak data-driven prior spirit, we set the prior inverse chi-square parameters for the birth weight mixture variance as $a = 1$, $b = 0.012$, equivalent to one observation with a residual variance equal to the estimated residual variance obtained from the same linear regression. Although both are "data-based" priors, they are extremely weak, yet ensure that posterior is proper and help stabilize the mixture model estimation. For the hyperprior inverse chi-square parameters for the between-litter variances, we chose the nearly noninformative values $c = d = e = f = 0.01$. For the prior on the proportional odds intercept parameters, we fit a multinomial logit model with 10 outcomes in each of $L + 2$ categories, using the parameter estimates and associated covariance matrix for θ_0 and Σ, again a relatively weak prior given that the total sample size is 1192. We initially considered a model that allows PS-specific estimates of litter survival status effects but found their estimation to be very unstable, and not inconsistent with the assumption that the effect of the other pups on birth weight is independent of the PS to which they belong. Hence, we proceeded under a model that assumed common litter survival status effects ($\gamma_0^s = \gamma_0$ and $\gamma_1^s = \gamma_1$ for all s).

Table 3 presents the total number of litters, total number of pups across all litters, percentage of pups surviving until birth, and mean birth weight among surviving pups, by dose level. Under the simplifying assumption that $\tau_b^2 = 0$ (independence of PS within litter), the fraction of pups belonging to PS 1, $\pi^1 = \frac{\exp(\theta^1)}{1+\exp(\theta^1)}$, could be estimated as the fraction of live pups in the highest dose group; similarly $\pi^1 + \pi^2$, the fraction of pups belonging to PS 1 or PS 2, could be estimated as the fraction of surviving pups in the second-highest dose groups, and so forth. Hence, as in Frangakis et al. (2004), we can estimate the fraction of pups belonging to PS 2 by the difference between the survival fractions in the second-highest and highest dose group, the fraction of pups belonging to PS 3 by the difference between the survival fractions in the third-highest and second-highest dose group, and so on. Estimates of the five PS sampling fractions under the assumption of within-litter independence are given by $\hat{\pi}^1 = 0.799, \hat{\pi}^2 = 0.062, \hat{\pi}^3 = 0.029, \hat{\pi}^4 = 0.002, \hat{\pi}^5 = 0.108$. Because $\hat{\pi}^s \geq 0$ for all principal strata, the assumption of monotonicity appears satisfied; however, because the proportion of the population that belongs to PS 4 is so small, we will collapse it with PS 3, thus assuming that any mouse pup that survives at dose level 0 would also have survived had it received dose level 1 (0.75 mg/kg). All further analyses assume four principal strata.

To obtain draws from the joint posterior distribution (5), we ran three chains starting with widely divergent starting values, obtaining 2500 iterations after a 2500 iteration burn-in. The maximum Gelman–Rubin statistic (Gelman et al., 2004, p. 296–297) was 1.030, suggesting a high degree of convergence. The results are given in Table 4. The model-based estimates suggest that a greater fraction of pups belongs to $S = 3$ (pups that survive only up to 0.75 mg/kg) and a smaller fraction of pups belongs to $S = 2$ (pups that survive up to 1.5 mg/kg of ethylene glycol) than the nonparametric estimates derived from Table 3 in the previous paragraph. There is evidence that the dose effect is greater in pups belonging to $S = 2$ and $S = 3$ than $S = 1$: $(P(\beta^2 < \beta^1 \,|\, \text{data}) = 0.69)$; $(P(\beta^3 < \beta^1 \,|\, \text{data}) = 0.81)$, consistent with their more "fragile" status.

The "dead pup" effect on birth weight is -0.007 (95% PPI = -0.015, 0.002) g/pup, while the "live pup" effect is -0.016 (95% PPI = -0.022, -0.010) g/pup, consistent with increasing resource draws from the dam for pups born alive

Table 4
Posterior means and 95% posterior predictive intervals (PPI) in subscript from birth weight model (2). $\text{expit}(x) = \exp(x)/(1 + \exp(x))$.

	Principal stratum s			
Parameter	1	2	3	4
PS fraction $(\text{expit}(\theta^s) - \text{expit}(\theta^{s-1}))$	$0.791_{(0.734,0.846)}$	$0.024_{(0.007,0.056)}$	$0.097_{(0.025,0.166)}$	$0.088_{(0.062,0.122)}$
Intercept (μ^s)	$1.12_{(1.05,1.20)}$	$0.95_{(0.76,1.15)}$	$1.16_{(0.94,1.26)}$	NA
Dose effect (β^s)	$-0.092_{(-0.107,-0.078)}$	$-0.123_{(-0.263,-0.000)}$	$-0.125_{(-0.224,0.037)}$	NA
Intra-litter correlation $(\tau^2/(\tau^2 + \sigma_s^2))$	$0.52_{(0.40,0.65)}$	$0.35_{(0.13,0.67)}$	$0.72_{(0.27,0.88)}$	NA

Table 5

Posterior means and 95% *PPI in subscript of "direct" dose effect* $\Psi_1^s(z^*, z)$ *and "litter-size-adjusted" dose effect* $\Psi_2^s(z^*, z)$*, by principal stratum, and averaged over the principal stratum for which the contrast exists:* $\Psi_3(z^*, z)$ *and* $\Psi_4(z^*, z)$

	Principal stratum s			
Effect	1	2	3	Average
Direct				
0.75 vs. 0	$-0.069_{(-0.080,-0.059)}$	$-0.092_{(-0.198,-0.000)}$	$-0.094_{(-0.168,0.028)}$	$-0.073_{(-0.084,-0.061)}$
1.5 vs. 0	$-0.137_{(-0.160,-0.117)}$	$-0.185_{(-0.395,-0.000)}$	NA	$-0.139_{(-0.162,-0.118)}$
3.0 vs. 0	$-0.275_{(-0.320,-0.235)}$	NA	NA	$-0.275_{(-0.320,-0.235)}$
Litter-size-adjusted				
0.75 vs. 0	$-0.069_{(-0.080,-0.059)}$	$-0.092_{(-0.198,-0.000)}$	$-0.094_{(-0.168,0.028)}$	$-0.073_{(-0.083,-0.061)}$
1.5 vs. 0	$-0.128_{(-0.154,-0.107)}$	$-0.173_{(-0.386,0.010)}$	NA	$-0.129_{(-0.156,-0.108)}$
3.0 vs. 0	$-0.263_{(-0.311,-0.224)}$	NA	NA	$-0.263_{(-0.311,-0.224)}$

rather than dead or resorbed. The within-litter birth weight correlation was lower for PS 2 and greater for PS 3 as compared to PS 1. The posterior between-litter variance for the proportional odds of PS membership was 0.82 (95% PPI = 0.38, 1.47), suggesting a moderate degree of within-litter correlation.

Table 5 shows the "litter-size-adjusted" and the "direct" principal effects of dose of birth weight. We see that the "litter-size-adjusted" effect at the higher dose levels is attenuated somewhat over the unadjusted "direct" effect that estimates the effect of dose on pups in litters whose number of live births is unaffected by dose. This is consistent with $\gamma_1 < \gamma_0$, i.e., the negative effect of live pups on birth weight is stronger than the effect of dead pups, meaning that birth weights in the observed high dosage group include a positive effect of fewer live pups competing for resources over what would have been the case had the dam been assigned to a placebo group, while the birth weight for the $S = 1$ pups in the $Z = 0$ dosage group includes a negative effect of more live pups for resources over what would have been the case had the dam been assigned to a high dosage group. The direct and litter-size-adjusted effects are the same when estimating the effect of the 0.75 mg/kg dose, because we have assumed that all pups that survive at the zero dose level also survive at the 0.75 mg/kg dose level. The results obtained by averaging over the PS for which the contrast is defined are also given in Table 5.

Figure 1a plots the litter mean of the posterior means of the PS membership for each pup, by the litter total size (both living and dead pups) and the litter dose level. Very small and very large litter sizes were associated with more pups belonging to "weaker" PS. This is consistent with small litters being born to dams that are somehow "weaker," either because of genetic or environmental variability, while large litters may depress the survival capability of the individual pups even in the "stronger" dams. This result is sensitive to the $n_j = 3$ and $n_j = 5$ litters; no statistically significant association between litter size and mean PS membership remains after removing these litters. Figure 1b plots the posterior mean of the litter level random effect of PS memberships b_j against the proportion of the pups observed to survive, indicating the expected negative relationship between the proportional odds random effect and the observed survival rate within a litter.

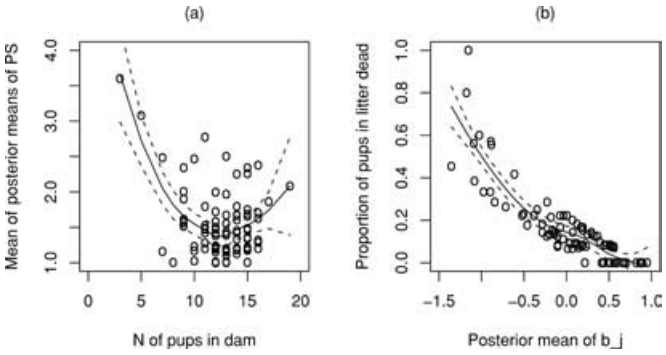

Figure 1. (a) Posterior means of principal stratum memberships in each litter, by total litter size; (b) observed survival rate in a litter, by posterior mean of random effect b_j on principal stratum membership of pups in the jth litter. (—) shows (a) smoothed estimate of mean of posterior means of principal stratum membership by total litter size and (b) smoothed estimate of the proportion of pups that die for a given litter effect b_j; (- - -) gives 95% CI.

4. Discussion

This manuscript considers a potential outcomes approach to estimating the effect of a toxin on birth weight, accommodating the effect of changes in litter survival status that might occur in a given litter had the litter been exposed to differing levels of the toxin. This method also restricts the estimates of birth weight change between given dose levels to the set of pups estimated to survive at both dose levels under the monotonicity assumption that a pup that dies at a given dose level will die at a higher dose level. This approach parallels the ITT contrasts within compliance classes in the Rubin causal model (Holland, 1986; Imbens and Rubin, 1997), where causal effect of a treatment is estimated within the set of "compliers" that would have received the treatment if and only if randomized to treatment. Our method can accommodate both the fact that only "healthier," and more likely heavier, pups

survive at high dose levels, which in turn frees up additional space and nutrients for these survivors in comparison to what would have happened had the dam been exposed to lower dose levels.

Catalano and Ryan (1992), who adjusted for surviving pup litter size as a fixed covariate in a joint outcomes model, estimated a dose effect of −0.095 g/mg/kg (95% CI = −0.111, −0.078) g/mg/kg, yielding estimated birth weight changes of −0.071, −0.142, and −0.285 g/mg/kg for 0.75, 1.5, and 3.0 mg/kg of ethylene glycol, respectively. This is an underestimate of the principal dose effect at lower levels and an overestimate of the principal dose effect at higher levels. This difference occurs because the potential outcomes approach untangles a mixture of slopes and intercepts that suggest the toxin has a stronger effect on pups that can tolerate only low doses than it does on pups that can tolerate higher doses. Dunson et al. (2003), who estimated fetal outcome and litter size via a shared latent variable model (Sammel, Ryan, and Legler, 1997), estimated a dose effect of −0.088 g/mg/kg (95% HPD = −0.105, −0.071) g/mg/kg, yielding estimated birth weight changes of −0.066, −0.132, and −0.264 g/mg/kg for 0.75, 1.5, and 3.0 mg/kg of ethylene glycol, respectively. The Dunson et al. estimate is closer to what we obtained in the potential outcomes framework at higher doses, although it underestimates the principal effect at low doses because no direct adjustment was made for the effect of other pups on birth weight.

One limitation of this approach is the monotonicity assumption that pups that would die at a given dose level would also die if the dose level were higher. This is admittedly a strong assumption, although not contradicted by the data, because survival rates at higher dose levels are monotonically lower. Unfortunately, this assumption is required to have an identifiable model; otherwise it would be impossible to determine the pups that belong to the lowest and highest stratum, and thus to identify the mixture components for the regression model.

Another limitation of the approach considered here is that the litter size n_j is assumed to be the same at the time of sacrifice as it was at the time of exposure to the toxin. However, if this assumption is incorrect, and if unobserved implantation sites are more common at higher doses, then our approach may not fully account for the effect of litter size at different doses. While this assumption is plausible because the exposure to the toxin occurs after mating and embryo implantation, Table 3 hints that it may not have held—the mean total litter size declines from 13.3 pups/litter at $Z = 0$ to 12.3 pups/litter at $Z = 3$ ($p = 0.102$). To accommodate these "missing pups," it may be possible to add a third outcome ("missing") to the dead/alive pup status and extend the PS membership table under the monotonicity assumption that pups that die at a given dose do not survive at a higher dose, and pups that are missing at a given dose are also missing at a higher dose.

Acknowledgements

The authors would like to thank Dylan Small and Thomas Ten Have, along with the three reviewers, the associate editor, and the editor, whose helpful comments improved the article. This research was supported in part by National Institute of Mental Health grant P30-MH066270.

References

Catalano, P. and Ryan, L. M. (1992). Bivariate latent variable models for clustered discrete and continuous outcomes. *Journal of the American Statistical Association* **87,** 651–658.

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

Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics* **58,** 21–29.

Frangakis, C. E., Rubin, D. B., and Zhou, X.-H. (2002). Clustered encouragement design with individual noncompliance: Bayesian inference and application to advance directive forms. *Biostatistics* **3,** 147–164.

Frangakis, C. E., Brookmeyer, R. S., Varadhan, R., Safaeian, M., Vlahov, D., and Strathdee, S. A. (2004). Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a needle exchange program. *Journal of the American Statistical Association* **99,** 239–249.

Gelfand, A. E. and Smith, A. M. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85,** 389–409.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*, 2nd edition. New York: Chapman and Hall.

Gilbert, P. B., Bosch, R. J., and Hudgens, M. G. (2003). Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. *Biometrics* **59,** 531–541.

Gueorguieva, R. V. and Agresti, A. (2001). A correlated probit model for joint modeling of cluster binary and continuous responses. *Journal of the American Statistical Association* **96,** 1102–1112.

Imbens, G. W. and Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Annals of Statistics* **25,** 305–327.

Molenberghs, G. and Ryan, L. M. (1999). An exponential family model for clustered multivariate binary data. *Environmetrics* **10,** 279–300.

Neyman, J. (1923). On the application of probability theory to agricultural experiments: Essay on principles, Sec. 9, translated in *Statistical Science* (1990) **5,** 465–480.

Pearl, J. (2001). Direct and indirect effects. *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, 411–420. San Francisco: Morgan Kaufmann Publishers.

Price, C. J., Kimmel, C. A., Tyl, R. W., and Marr, M. C. (1985). The developmental toxicity of ethylene glycol in rats and mice. *Toxicology and Applied Pharmacology* **81,** 113–127.

Robins, J. M. (1986). A new approach to causal inference in mortality studies with sustained exposure

periods—Application to control of the healthy worker survivor effect. *Mathematical Modelling* **7,** 1393–1512.

Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology* **3,** 143–155.

Rubin, D. B. (1974). Estimating causal effects in randomized and non-randomized studies. *Journal of Educational Psychology* **66,** 688–701.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics* **6,** 34–58.

Rubin, D. B. (1990). Comment on Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science* **5,** 472–480.

Sammel, M. D., Ryan, L. M., and Legler, J. M. (1997). Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society, Series B* **59,** 667–678.

Received December 2004. *Revised September* 2005.
Accepted September 2005.