Journal of
Toxicology and Risk Assessment

RESEARCH ARTICLE

Combining Hazard Information from Multiple Dose-Response Experiments for Benchmark Dose Estimation in Risk Assessment

Lucy Kerns*

Department of Mathematics and Statistics, Youngstown State University, Tressel Way, Youngstown, OH, USA

*Corresponding author: *Dr. Lucy Kerns, Department of Mathematics and Statistics, Youngstown State University, Tressel Way, Youngstown, OH 44555, USA*

Abstract

An important goal of risk assessment is the determination of the minimum dose levels (benchmark doses or BMDs) of a hazardous chemical at which a specified bench mark risk (BMR) is attained. In practice, more than one experiment is often conducted for a hazard to determine the BMD values. In such cases, synthesizing all available hazard information to produce an average BMD value becomes an important task for risk analysts, which can be challenging when there is significant between-experiment heterogeneity in toxicological effects of the hazard. Statistical methods for combining information from different sources that can account for data-source heterogeneity play a critical role in accomplishing such a task. In this paper, we illustrate the use of three statistical techniques for estimating BMD values in risk studies involving multiple dose-response experiments: The Generalized Estimation Equation (GEE), Generalized Linear Mixed Model (GLMM), and meta-analysis approaches. Monte Carlo simulation studies are conducted to compare the performance of these statistical methods. An example is given to illustrate the use of the methods.

Keywords

Bench mark dose, Risk assessment, Generalized estimation equation, Generalized linear mixed model, Meta-analysis approach

Introduction

Quantitative risk assessment is mainly concerned about the identification and quantification of adverse effects after exposures to hazardous chemicals. The adverse effects could be death, birth defect, weight loss, cancer, or mutation caused by chemicals. Risk in this context is the probability of a specified adverse response in a subject or object from exposure to a chemical. A major effort in risk analysis involves the statistical characterization of the dose-response relationship and the derivation of an estimate of a point of departure (POD), which is defined as the lowest dose/concentration level corresponding to an estimated low effect level or no effect level. The bench mark dose (BMD) the dose level that causes a specified bench mark risk (BMR) is commonly used by risk analysts for setting the POD when assessing hazardous exposures [1].

The BMD methodology is a widely used statistical technique for assessing toxicological potency of a chemical and identifying the BMD in risk assessment. It was first introduced by Crump for quantal response data [2], but has been subsequently extended with definitions for continuous response data [3,4]. In this approach a dose-response model describing the relationship between exposure and response is fit to the data, and this mode list hen used to identify the dose level (i.e, the BMD) associated with a specified risk level (i.e., the BMR). If the exposure is measured as a concentration, benchmark concentration (BMC) is used for the exposure point at which a specified BMR is attained. To account for the uncertainty in the estimation of BMD, it is standard practice to determine the one-sided lower $100(1-\alpha)\%$ confidence limit of BMD, called the benchmark dose lower limit, or BMDL [4].

In practice, more than one experiment is often carried out to assess the toxicity of a chemical and determine BMD estimates of the chemical; for example,

these experiments could be conducted in the same lab on different days, or in different labs. The estimated dose-response curves are likely to vary from experiment to experiment due to various sources of variability in the experimental design; this includes biological variability (i.e., variability due to testing subjects/objects, organisms, etc.) and technical variability (i.e., variability due to testing methods, equipment, measurements, sample preparation, etc.).

In such cases, a simple pooled analysis - that is, an analysis based on a single dataset resulted from combining data sets from different sources-fails to take into account this between experiment/lab variability. Consequently, this practice of data-pooling may result in mischaracterized dose-response relationships and hence biased BMD estimates (i.e., either too high or low estimates having confidence intervals with coverage far from the usual nominal level of 95%). As such, it is important for risk analysts to employ appropriate strategies for combining the hazard information from different sources that can account for such data-source heterogeneity and provide an accurate potency assessment to better inform decision-making process.

In the case where dose response data are compiled from different experiments, one can calculate the individual-level BMD estimates based on the estimated dose-response relationship for each experiment as well as the overall or population-averaged BMD estimates obtained by summarizing BMD estimates from all experiments. When tasked with determining the potential for a chemical to cause adverse effects to humans or the environment, risk assessors are often more interested in obtaining a general population effect of the chemical, instead of the effect on a single subject/object. In this sense, the population-averaged BMD estimate is useful, since it can be interpreted naturally in a population context, making it easier to generalize the results to a much broader population than the representative sample. Here, we illustrate the use of three statistical techniques for determining the population level BMD estimates from multiple dose response experiments.

One strategy is to use the Generalized Estimation Equation (GEE) approach, which is a popular statistical approach to fit a marginal model for correlated data analysis [5-7]. It is a population-level approach based on a quasi-likelihood function and provides the population-level estimates of the parameters (i.e., parameters of the model as well as derived parameters such as the BMDs). An alternative approach is to fit the Generalized Linear Mixed-Effects Model (GLMM) for the correlated dose-response data. The GLMM method [8] is a conditional approach by adopting random effects to capture the correlation between the observations of the same lab/experiment (or the same subjects as often the case in longitudinal studies). It takes into account the variabilities

within and among experiments simultaneously, and can be used to provide both individual level (or experiment-specific) and population averaged estimates of the parameters. The third approach is to estimate individual BMDs from each experiment by fitting a dose response model to the data for each experiment, and combine these estimates using a meta-analytic approach [9,10].

Research studies have appeared in literature that focused on the estimation of specific potency end points in multi-experiment risk assessments. Simmons, et al. [11] applied a hierarchical model or GLMM in the Bayesian setting to combine mutagenic potency estimates from multiple bioassays used for environmental risk assessment. Coull, et al. [12] also used a Bayesian hierarchical model to study the adverse effects of methylmercury by combining BMD estimates from multiple studies. Wheeler and Bailer [13] proposed a general hierarchical model in a Bayesian frame work that can be used to estimate the average BMD for response data assumed to be generated from the exponential family of distributions. Jiang and Kopp-Schneider [14] investigated two strategies, the GLMM and meta-analysis methods, for summarizing the half maximal effective concentration (EC50) estimates from multiple experiments. They later expanded their earlier work by further examining these two strategies in two cases of multiple dose-response experiments: Complete dose response relationships are observed in all experiments, and complete dose response relationships are observed only in a subset of experiments [15].

Quantal response data are commonly encountered in dose-response studies where binary outcomes are observed and the observations are often in the form of proportions associated with each dose level. Most of the works mentioned above have focused on continuous response data, with less emphasis on nonnormal data such as quantal data. Our focus herein concerns the estimation of BMD and BMDL in risk studies with quantal response data coming from multiple experiments. Section 2 describes the three strategies for determining the population-level BMDs and BMDLs in risk studies involving multiple experiments: The GEE, GLMM, and meta-analysis approaches. Section 3 presents simulation studies to compare the three strategies, and Sections 4 provides an illustrative example. A brief discussion is given in Section 5.

Methods

Dose response models

Suppose that a series of k experiments is performed in a dose response study for a specific hazardous chemical. For experiment $i(i = 1,...,k)$, suppose that there are n_i observations. Let Y_{ij} denote the number of subjects/objects exhibiting an adverse effect in the ith experiment at the jth dose level d_{ij}, where $i = 1,...,k$ and $j = 1,...,n_i$. We assume that Y_{ij}s are binomial variates

with parameters N_{ij} and $R(d_{ij})$, where N_{ij} is the number of subjects/objects tested and $R(d_{ij})$ is the risk function representing the probability that an individual subject/object will respond adversely at dose level d_{ij}. It should be noted that a typical design for a quantal response risk study employs a series of pre-specified doses which are assumed increasing, and there are usually several replicates at each dose level in dose response experiments.

We write $\eta_{ij} = \beta_0 + \beta_1 d_{ij}$ as the first-order linear predictor, where β_0 and β_1 are unknown parameters. In the quantal data setting, we use a link function g to associate the linear predictor η_{ij} with the response probability P_{ij}, $g(P_{ij}) = \eta_{ij}$. The inverse link is $g^{-1}(\eta_{ij}) = P_{ij}$. In our notation, this translates to the risk function, $R_{ij} = g^{-1}(\eta_{ij})$. Some common link functions for quantal response data include the logit link, the probit link, and the complementary log link. In this study, we use the logit link and the resulting two-parameter logistic model has the following form:

$$\log \frac{P_{ij}}{1-P_{ij}} = \beta_0 + \beta_1 d_{ij} \qquad (1)$$

Under the model in (1), the risk function is given as follows:

$$R(d_{ij}) = [1 + \exp(-\beta_0 - \beta_1 d_{ij})]^{-1} \qquad (2)$$

In many risk studies, it is common to further refine the risk function $R(d_{ij})$ into the extra risk function, $R_E(d_{ij})$, for the purpose of assessing risks above back ground. The extra risk is defined as $R_E(d_{ij}) = [R(d_{ij})-R(0)]/[1-R(0)]$, where $R(0)$ is an independent background risk and $0 \leq R(0) \leq 1$; that is, it is defined as the risk solely due to dose and independent of background risk. The BMD is then determined by setting $R_E(d_{ij})$ = BMR and solve for the dose level. For the two-parameter logistic model in equation (1), the BMD is given as

$$BMD = \frac{1}{\beta_1} \log\left[\frac{BMR \times e^{-\beta_0} + 1}{1 - BMR} \right] \qquad (3)$$

Generalized estimating equations

We begin with an over view of the GEE method for correlated or clustered data. Let $Y_i = (Y_{i1},...,Y_{ij},...,Y_{in}i)$ represent the response vector for the ith cluster ($i=1,...,k$), where we assume the observations from the same cluster are correlated. The GEE approach relaxes distributional assumptions and only requires the specification of the first and second moments. The marginal expectations $E(Y_{ij}) = \mu_{ij}$ are related to the linear predictor $x'_{ij}\beta$ by a link function as

$$g(\mu_{ij}) = x'_{ij}\beta \qquad (4)$$

where β is the unknown parameter vector and $x_{ij} = (x_{ij,1},...,x_{ij,p})$ is the p-dimensional vector of covariates for the jth member of the ith cluster. The variance is specified as

$$Van(Y_{ij}) = \phi v(\mu_{ij}) \qquad (5)$$

where ϕ is a scale or over-dispersion parameter which may need to be estimated and v is a known variance function. The GEE incorporates a working correlation structure to account for the correlation of repeated measures. Let $R_i(\alpha)$ be the working correlation matrix of Y_i, which is completely described by the parameter vector α, and let $V_i = \phi A_i^{1/2}R_i(\alpha)A_i^{1/2}$ be the corresponding working covariance matrix of Y_i, where A_i is the $n_i \times n_i$ diagonal, matrix with the variance of Y_i as diagonal element.

GEE yield a consistent estimate, $\hat{\beta}$, for the coefficients β, which are the solution of the equation

$$\sum_{i=1}^{k} D'_i V_i^{-1}(Y_i - \mu_i) = 0, \qquad (6)$$

Where $D_i = \partial\mu/\partial\beta'$. Moreover, $\hat{\beta}$ is asymptotically normally distributed with a mean β and a covariance matrix $\sum = \sum_0^{-1} \sum_1 \sum_0^{-1}$, where

$$\sum_0 = \sum_{i=1}^{k} D'_i V_i^{-1} D_i, and, \sum_1 = \sum_{i=1}^{k} D'_i V_i^{-1} Cov(Y_i)V_i^{-1} D_i \qquad (7)$$

A so called sandwich estimate of $\hat{\sum}$ of $\sum$ can be obtained by replacing β, ϕ and α with their consistent estimates and the covariance matrix Cov (Y_i) with $(Y_i - \hat{\mu}_i)(Y_i - \hat{\mu})'$ in equation (7).

Within the context of risk assessment for quantal data, the risk, or the response probability of a subject/object exhibiting an adverse effect is the expectation of the binary outcome, and if we use the logit link function to connect the mean and the covariates (i.e, the dose level of the agent in our case), then the marginal mean model in equation (4) with the first-order linear predictor can be written as the two-parameter logistic model given in equation (1). The variance function in equation (5) is given by $v(\mu_{ij}) = P_{ij}(1-P_{ij})$. There are a number of popular choices for the correlation structure, $R_i(\alpha)$, including independence, exchangeable, first-order Autoregressive (AR-1), and unstructured. We will fit the GEE within dependence and exchangeable correlation structures to illustrate the use of the method in the current setting.

We denote the estimator of the regression coefficients using $\beta^{GEE} = (\beta_0^{GEE}, \beta_1^{GEE})$. Given these estimates of coefficients, the risk at a specific dose level and the BMD at a given BMR can be estimated using equation (2) and equation (3), respectively, by substituting the estimates of the regression coefficients for β_0 and β_1. Inference on the BMD can be drawn from standard asymptotic theory for the GEE regression coefficients estimates. That is, we can construct the Wald-type confidence intervals to produce a 100 (1-α)% BMDL as $B\hat{M}D - z_\alpha se\left(B\hat{M}D\right)$, where z_α is the upper-α standard normal critical point and $se\left(B\hat{M}D\right)$ is the standard error of $B\hat{M}D$ found via the delta method.

Logistic random intercept models

The GLMM is an extension of the Generalized Linear Model (GLM) for clustered categorical data in which the linear predictor contains random effects in addition to the fixed effects. It is a conditional model that provides an estimate of individual-level (e.g., subject/experiment-specific) or conditional effects, but can also be used to estimate population-averaged effects by summarizing individual-level effects from multiple sources. A commonly used GLMM for modeling correlated binary outcomes is the Binomial GLMM, in which each Y_{ij}, given random effects, is assumed to have a binomial distribution. In this study, we consider a specific type of the Binomial GLMM, a two-parameter logistic random intercept models:

$$\log\left[\frac{P(Y_{ij}=1\,|\,u_i)}{1-P(Y_{ij}=1\,|\,u_i)}\right] = \beta_0 + u_i + \beta_1 d_{ij} \qquad (8)$$

Where u_i is the random effect for the ith experiment, $i = 1,...,k$, which is usually assumed to be normally distributed with mean 0 and variance σ^2_{μ}.

In the context of risk studies involving multiple experiments, the regression coefficients in this logistic random intercept model are experiment-specific or conditional coefficients, and have the experiment-specific interpretation conditional on the random effect for each experiment. Maximum likelihood methods or restricted maximum likelihood methods are often used to estimate the coefficients, and they are implemented in popular R packages, such as "lme4". We denote the estimates of the coefficients as $\hat{\beta}_0^{es}$ and $\hat{\beta}_1^{es}$, and the estimate of variance for the random effects as $\hat{\sigma}_u^2$, where the subscript "es" is used to reflect the fact that they represent effects of dose levels conditioning on the random experiment effect.

Multiple methods for estimating the population-averaged (or marginal) BMD could be considered. Hedeker, et al. [16] proposed a marginalization method to produce population-averaged or marginal estimates for the regression coefficients from their conditional counterparts in GLMMs for longitudinal binary outcomes. Here, we employ their method to find marginal estimates for β_0 and β_1 in the logistic random intercept model, from which the population-averaged BMD estimate is determined. Specifically, we write the marginal model in equation (1) in a matrix form as $\lambda^{pa} = X\beta^{PA}$, where λ^{pa} is a n × 1 vector $(n = \sum_{i=1}^{k} n_i)$ containing elements $\log[P_{ij}^{pa}/(1-P_{ij}^{pa})]$ with P_{ij}^{pa} being the marginal probability, X is the n × 2 covariate matrix including a column of ones for the intercept and a column of dose levels d_{ij}, and $\beta^{pa} = (\beta_0^{pa}, \beta_1^{pa})$ is the marginal regression coefficients. This operation yields the following equation for describing the relationship between the marginal regression coefficients and the marginal probabilities:

$$\beta^{pa} = (X'X)^{-1} X' \lambda^{pa} \qquad (9)$$

Note that the subscript "pa" for λ and β indicates that these parameters are population averaged or marginal parameters. Given the estimates of conditional coefficients, $\hat{\beta}_0^{es}$ and $\hat{\beta}_1^{es}$, the marginal probabilities, P_{ij}^{pa}, can be calculated by integrating over the distribution of the random effect in the logistic random intercept model:

$$\hat{P}_{ij}^{pa} = \int_{-\infty}^{\infty} \frac{1}{1+\exp(-(\hat{\beta}_0^{es}+u+\hat{\beta}_1^{es}d_{ij}))} f(u)du, \qquad (10)$$

Where $f(u)$ is the density function for a Normal distribution with mean 0 and variance σ_u^2, which can be substituted by its estimate, $\hat{\sigma}_u^2$. We may approximate the integral using Gauss-Hermite quadrature:

$$\hat{P}_{ij}^{pa} \approx \sum_{q=1}^{n} \frac{1}{1+\exp(-(\hat{\beta}_0^{es}+b_q+\hat{\beta}_1^{es}d_{ij}))} \omega_q, \qquad (11)$$

where $\{b_q\}_{q=1}^n$ are the set of quadrature points and $\{w_q\}_{q=1}^n$ are the related weights.

Replacing the marginal probability P_{ij}^{pa} in λ^{pa} by its estimates given in equation (11) yields an estimate of λ^{pa}, $\hat{\lambda}^{pa}$, from which marginal regression coefficients, $\hat{\beta}^{pa} = (\hat{\beta}_0^{pa}, \hat{\beta}_1^{pa})$, can be obtained using equation (9). The standard errors of the estimates $\hat{\beta}^{pa}$ can be derived using the delta method. Then, using these estimates of marginal regression coefficients and their standard errors, we can determine the marginal BMD estimate and the Ward-type BMDL for a specific BMR in the same way as we did with the GEE method. It should be pointed out that although we discussed Hedeker, et al.'s method in the logistic random intercept model with logit link function and one random effect, their method can be applied to the more general cases of multiple random effects as well as other link functions; see Hedeker, et al. [16] for more details.

Another method to obtain the marginal BMD that we consider here also relies on the estimation of marginal regression coefficients. For the logistic random intercept model, we may use the following approximation to obtain the estimates of marginal regression coefficients from their conditional counterparts [17]: $\hat{\beta}_i^{pa} \approx \hat{\beta}_i^{es}/\sqrt{(c^2\sigma_u^2+1)}$, where i = 0, 1, and c = 16√3/(15π). As with the Hedeker, et al.'s method, the marginal BMD and BMDL can be determined based on these approximate estimates for the marginal regression coefficients.

Meta-analysis models

Meta-analysis is a statistical method for integrating the results of multiple studies addressing the same basic question, with the aim of averaging the estimates of a comparable parameter from each study. It is typically a two-step process. In the first step, each study produces

an estimate of the parameter and its corresponding standard error. These results are then used to form the data input of a standard fixed or random effects meta-analysis model in the second step, to find the average of parameter estimates. A fixed-effects meta-analysis assumes that the true effect is the same in all studies, i.e., there is no statistical heterogeneity between studies, while for a random-effects meta-analysis model, the effects being estimated in the different studies are assumed to be different and follow some distribution. The latter is the focus of this paper. Here, we describe the meta-analysis random-effects model in the context of risk assessment, with a specific focus on combining BMD estimates from multiple experiments to generate a population-level BMD estimate.

For notational convenience, let D_i denote the BMD estimate obtained by fitting a two-parameter logistic model to the data from the ith experiment (i = 1, . . . , k) for a specific BMR, and let ϑ_i denote the corresponding (unknown) true BMD value. Then, D_i can be models as

$$D_i = \theta_i + e_i, \tag{12}$$

where e_i represents the sampling error, $e_i \sim N(0, \sigma_i^2)$. For random-effects models, the true BMD values ϑ_i are assumed as random and modeled as:

$$\theta_i = \theta + b_i \tag{13}$$

Where ϑ is the true average BMD value and $b_i \sim N(0, \tau^2)$ represents the random effect term by which the ϑ_i deviates from the true average BMD value. The variance τ^2 represents the amount of heterogeneity among the true BMD values of each experiment. Combining equations (12) and (13) yields the following meta-analytic random effects model:

$$D_i = \theta + b_i + e_i = \theta + \in_i \tag{14}$$

If b_i and e_i are assumed to be independent, then, $\in_i \sim N(0, \sigma_i^2 + \tau^2)$, and $D_i \sim N(0, \sigma_i^2 + \tau^2)$. With the meta-analysis approach, the task of estimating the population averaged BMD and BMDL from a series of dose-response experiments can be accomplished by finding an estimate of the parameter θ in equation (14) and its standard error.

To fit the random-effects model in equation (14), a two-step approach can be used [18], in which τ^2 is estimated in the first step and ϑ is estimated via weighted least squares with weights equal to $w_i = 1/(\sigma_i^2 + \tau^2)$ in the second step. The estimate of weight is $\hat{w}_i = 1/(\hat{\sigma}_i^2 + \hat{\tau}^2)$, where $\hat{\tau}^2$ is the estimate of τ^2 obtained in the first step, and $\hat{\sigma}_i^2$ is the estimated within-experiment variance which can be obtained by fitting the two-parameter logistic model to the dose-response data of each experiment. Note that various methods are available for estimating the amount of heterogeneity, τ^2, some of which are implemented in the "metafor" R package [19]. Jiang and Kopp-schneider

[14] conducted simulation studies to investigate the impact of these heterogeneity estimators for τ^2 on the estimation of an overall average EC50 value (i.e., concentration level of a chemical that gives half-maximal response), and they found that the choice of heterogeneity estimators did not influence the estimation of the overall average EC50 value. For a description and comparison of these estimators, see Jiang and Kopp-schneider [14]. In this study, the Hunter-Schmidt estimator [20,21] is used.

An approximate 100(1-α)% Wald-type confidence interval for ϑ can be obtained under the assumption of asymptotic normality for $\hat{\theta}$, which yields a 100(1-α)% BMDL as $\hat{\theta} - z_\alpha se(\hat{\theta})$, where $se(\hat{\theta})$ is the standard error of $\hat{\theta}$, and can be estimated by $s\hat{e}(\hat{\theta}) = \sqrt{1/\sum_{i=1}^{k} \hat{w}_i}$. However, the estimation of the standard error of $\hat{\theta}$ depends on $\hat{w}_i$, which depends on the unknown τ^2. Considering the uncertainty in the estimate of τ^2, Hartung and Knapp [22,23] proposed a modified Wald-type confidence interval to account for such uncertainty:

$$\hat{\theta} - t_{k-1,\alpha} \sqrt{\frac{\sum_{i=1}^{k} \hat{w}_i (D_i - \hat{\theta})^2}{(k-1)\sum_{i=1}^{k} \hat{w}_i}} \tag{15}$$

where $t_{k-1,\alpha}$ is the upper α quantile of the t-distribution with k-1 degrees of freedom. Jiang and Kopp-Schneider [14] compared the performance of the Wald-type confidence interval and its modified version in terms of coverage probability in simulation studies, along with two other methods for adjusting the standard error of $\hat{\theta}$. Their results indicated that the Hartung-Knapp method outperformed the other methods by producing confidence intervals with coverage probabilities being closest to the nominal level, and hence it was recommended for use in the construction of confidence intervals when using the meta-analysis strategy. Therefore, for the purpose of estimating BMDL here, we use the modified Wald-type confidence intervals for the overall average BMD.

Simulation Study

Simulation design

To compare the performance of the methods discussed in Section 2, simulation studies were conducted. The specifications for the simulation studies were based on the example introduced in the next section. We simulated from a two-parameter random intercept logistic model given in equation (1) and the values of the fixed effects parameters were set as: β₀ = -2.75 and β₁ = 7.50. The dose levels were identical for each experiment i (i = 1, . . . , k) at d = 0, 0.075, 0.15, 0.3, 0.6, 1.2, 1.5, with 8 replicates of response measurements at the control level (i.e., d = 0) and 4 replicates for the other dose levels. An equal number of trials, n = 5 was

used for each dose level in all experiments. To study the effect of number of experiments performed, three values of k were selected: k = 3, 9, 15. Additionally, four different values for the standard deviation of the random intercept were considered, σ_u = 0.1, 0.5, 1, 2, to examine the influence of different variation of the random intercept.

We proceeded in simulating binomial data by first generating a random normal variate for each experiment i with mean 0 and a pre-defined standard deviation σ_u and then calculating the risk P_{ij} at each dose level for each experiment using these random variates and the specified fixed effects parameters according to equation (2). Binomial data were then generated with n = 5 and probability of success P_{ij}, and this simulation procedure was repeated 1000 times for each of the parameter settings (i.e., for each of three values of k: 3, 9, 15 and for each of four values of σ_u: 0.1, 0.5, 1, 2).

We considered a BMR of 0.05, and the "true" marginal BMDs corresponding to the chosen BMR were obtained by Monte Carto integration; specifically, for each parameter setting, we sampled 100,000 numbers from a normal distribution conditioning on the pre-defined, true variance of the random intercept, for calculating the "true" marginal risks at specific dose levels, which were then used to find the "true" marginal BMD for the given BMR value. The point estimates of the BMD value obtained by the different methods were evaluated with respect to the mean, bias, and root mean square error (RMSE) of the estimates. Confidence intervals (CIs) for the BMD value constructed using the methods discussed in Section 2 were compared in terms of coverage probability (CP) of the CIs, which was calculated by determining the number of times out of 1000 simulations that the lower limits on the BMD

values (i.e, the BMDL) are below the true values of marginal BMD. The error probability α of the CIs was set at 0.05 throughout the simulation.

We used the statistical program R (version 4.2.2, R Core Team, 2022) for all simulations and computations. The **GEE** marginal regression models were fitted using the **geeglm** function of the R package **geepack** [24], and the logistic random intercept models were fitted using the **mixed_model** function of the **R** package **GLMMadaptive** [25]. For the meta-analysis approach, we fit the logistic dose-response model to the data from sub-experiment through the use of the model-fitting function **drm** in the **R** package **drc** [26] and then fit the random effects meta-analytic models using **rma** function of **R** package **metafor** [19].

Simulation results

Table 1 and Table 2 present the simulation results for the BMD values estimated based on the following models: The GEE with independent correlation structure (GEE-ind), GEE with exchangeable correlation structure (GEE-exc), GLMM using marginalized regression coefficients (GLMM-marg), GLMM with approximated marginal coefficients (GLMM-app), and meta-analysis random-effects model (META), when BMR was 0.05 and σ_u was set to be 0.1 and 0.5 (Table 1), and (Table 2). Note that for notational convenience, the general term GLMM is used in Sections 3 and 4 to represent the two-parameter logistic model discussed in Section 2.

The GEE and GLMM methods gave similar simulation results in terms of mean, bias, and RMSE of the estimates when the variation of the random intercept is small (i.e, σ_u = 0.1, 0.5); however, the GLMM methods yielded smaller bias and RMSE for larger variances of the random intercept. Between the two GEE methods,

Table 1: Simulation results for the population-averaged BMD when σ_u is 0.1 and 0.5

k	Method	σ_u = 0.1				σ_u = 0.5			
		Mean	Bias	RMSE	Coverage	Mean	Bias	RMSE	Coverage
3	GEE-ind	0.0849	0.0012	0.0102	0.8560	0.0834	0.0039	0.0195	0.8360
	GEE-exc	0.0847	0.0010	0.0105	0.8540	0.0839	0.0044	0.0199	0.8290
	GLMM-marg	0.0847	0.0010	0.0102	0.9990	0.0828	0.0033	0.0194	0.9700
	GLMM-app	0.0849	0.0012	0.0102	0.9990	0.0835	0.0040	0.0195	0.9670
	META	0.0831	-0.0006	0.0101	0.9540	0.0818	0.0023	0.0192	0.9610
9	GEE-ind	0.0842	0.0005	0.0060	0.9180	0.0819	0.0024	0.0114	0.8840
	GEE-exc	0.0842	0.0005	0.0060	0.9100	0.0824	0.0029	0.0117	0.8840
	GLMM-marg	0.0842	0.0005	0.0060	0.9940	0.0818	0.0023	0.0124	0.9050
	GLMM-app	0.0842	0.0005	0.0060	0.9940	0.0820	0.0025	0.0114	0.8890
	META	0.0816	-0.0021	0.0062	0.9720	0.0813	0.0018	0.0115	0.9510
15	GEE-ind	0.0840	0.0003	0.0045	0.9290	0.0812	0.0017	0.0088	0.9150
	GEE-exc	0.0840	0.0003	0.0046	0.9300	0.0817	0.0022	0.0092	0.9090
	GLMM-marg	0.0840	0.0003	0.0045	0.9850	0.0811	0.0016	0.0088	0.9150
	GLMM-app	0.0840	0.0003	0.0045	0.9850	0.0814	0.0019	0.0088	0.9070
	META	0.0811	-0.0026	0.0052	0.9810	0.0812	0.0017	0.0092	0.9360

Table 2: Simulation results for the population-averaged BMD when σ_u is 1 and 2.

k	Method	$\sigma_u = 1$				$\sigma_u = 2$			
		Mean	Bias	RMSE	Coverage	Mean	Bias	RMSE	Coverage
3	GEE-ind	0.0816	0.0109	0.0348	0.8190	0.0806	0.0206	0.0593	0.8380
	GEE-exc	0.0836	0.0130	0.0364	0.8240	0.0873	0.0273	0.0627	0.8488
	GLMM-marg	0.0801	0.0095	0.0341	0.8870	0.0784	0.0184	0.0585	0.8879
	GLMM-app	0.0819	0.0113	0.0344	0.8690	0.0816	0.0216	0.0583	0.8478
	META	0.0815	0.0109	0.0364	0.9469	0.0815	0.0216	0.0677	0.9546
9	GEE-ind	0.0767	0.0061	0.0189	0.8860	0.0685	0.0085	0.0260	0.9250
	GEE-exc	0.0784	0.0078	0.0200	0.8970	0.0746	0.0146	0.0305	0.9290
	GLMM-marg	0.0761	0.0054	0.0188	0.8990	0.0667	0.0067	0.0257	0.9510
	GLMM-app	0.0773	0.0067	0.0189	0.8670	0.0700	0.0100	0.0266	0.8320
	META	0.0800	0.0094	0.0222	0.8920	0.0718	0.0118	0.0361	0.9310
15	GEE-ind	0.0750	0.0044	0.0142	0.9040	0.0655	0.0055	0.0189	0.9470
	GEE-exc	0.0767	0.0061	0.0154	0.9300	0.0716	0.0116	0.0233	0.9350
	GLMM-marg	0.0743	0.0036	0.0141	0.9160	0.0633	0.0033	0.0185	0.9530
	GLMM-app	0.0757	0.0051	0.0143	0.8740	0.0670	0.0071	0.0196	0.8340
	META	0.0792	0.0086	0.0179	0.8758	0.0681	0.0081	0.0258	0.9220

the GEE method with independent correlation structure generally produced estimates with smaller bias and RMSE compared with the ones resulted from the GEE method with exchangeable correlation structure. This could be the result of binomial data for sub-experiments in each experiment being independently generated. In comparison to the GLMM approximation method, the GLMM marginalized approach produced better estimates with smaller bias and RMSE, and this became more apparent with larger variation of the random intercept. The bias in the BMD estimates was small (0.0003, 0.0044) for a small value of σ_u (i.e., σ_u = 0.1, .5), even for k = 3, but it became larger as σ_u increased. We observed that there was consistent reduction in both bias and RMSE of the estimates based on the GEE and GLMM methods with increasing k in all simulation conditions, which was expected.

The simulation results based on the meta-analytic approach were comparable to those given by the other methods in terms of mean, bias, and RMSE of the BMD estimates. For a small σ_u (i.e, 0.1 and 0.5), the bias was small (-0.0026, 0.0023), but it increased with increasing σ_u. However, unlike the GEE and GLMM methods, the bias of the estimates resulted from the meta-analytic approach did not display a consistent decreasing trend as the number of experiments k increased; but a consistent decreasing tendency was found with the RMSE of the estimates.

The coverage probabilities of the 95% CIs for the BMD values based on the GEE methods were generally low (CP < 0.93) except under one simulation condition: k = 15 and σ_u = 2 for the GEE with independent correlation structure. The GLMM methods produced CIs that had a wide range of coverage probabilities; for instance, the CPs were high when σ_u = 0.1 (CP ≥ 0.985) for all values

of k, but below the nominal level (i.e, CP < 0.95) when σ_u = 0.5. This might be explained by the fact that the parameter estimates using the GLMM are conditioned on the estimated variance of the random effects, which are subject to uncertainty, but this uncertainty is not accounted in the estimation process. It is noted that in general, the GLMM marginalized method produced CIs with the coverage probability closer to the nominal level than the GLMM approximation method. The CPs of the 95% CIs using the meta-analysis approach were generally close to the nominal level when σ_u is small; however, the CPs became smaller (CP < 0.93) as σ_u increased for k = 9, 15, while remaining close to the nominal level for k = 3.

Overall, the three approaches displayed similar performance in terms of mean, bias and RMSE of the BMD estimates when there was a small variation in the random intercept, but as the random intercept variance increased, the GLMM approach showed a slight advantage, particularly the GLMM marginalized approach. This could be attributed to the GLMM's ability to borrow strength across experiments by shrinking fixed effects (i.e., the regression coefficients) towards a common mean, which results in improved parameter estimates with smaller bias. It appears that the meta analysis approach can be recommended when the random intercept exhibited small variability, as it produces estimates as accurate as the other methods but has CPs that are closer to the nominal level. However, as the variance of the random intercept increased, it became harder to find a clear winner if we take into consideration all of the evaluation metrics considered (i.e., bias, RMSE, and coverage probabilities), since the simulation results varied according to the variance of the intercept as well as the number of experiments considered.

Example

We illustrate the estimation and inference methods for the population-averaged BMD using the data from a toxicology experiment reported by Gottardi and Cedergreen [27], in which the acute toxicity of the parathyroid insecticide α-cypermethrin on the fresh water crustacean *Daphnia magna* was investigated. Nine experiments were performed independently at different times, and in each experiment, *D. magna* neonates (< 1-day-old) were exposed to different concentrations of α-cypermethrin (4 replicates, 5 organisms in each) or acetone controls (8 replicates, 5 organisms in each) for 48 hours in M7 medium. Six different concentrations (5 non-zero concentrations and control) were used in experiments 1-7, and seven concentrations (6 non-zero concentrations and control) were used in experiments 8 and 9. Immobility of *D. magna* was visually determined after 48h of exposure. These data have also been used by Jensen, et al. [28] in their **R** package **bmd** to demonstrate the estimation of BMDs for binomial data obtained from a hierarchical design.

We first fit a GEE model in equation (1) with exchangeable correlation structure to the data, which yielded the following marginal estimates for the regression coefficients: $\beta_0^{GEE}(exc) = -2.25$ and $\beta_1^{GEE}(exc) = 7.25$. We also fit a GEE model with independent correlation structure, and the estimated marginal regression coefficients were $\beta_0^{GEE}(ind) = -2.76$ and $\beta_1^{GEE}(ind) = 7.45$. Three common values of BMR were considered, BMR = 0.01, 0.05, 0.10, and the corresponding estimates of BMC and 95% BMCL using the GEE with both independent and exchangeable correlation are reported in Table 3. Note that we used BMC and BMCL in this case, since the exposure level of α-cypermethrin was measured as a concentration.

A number of R packages can be used to fit the two parameter logistic random intercept model. For our purposes here, we used the R package **GLMMadaptive**, in which Hedeker, et al.'s marginalization method of the regression parameters in GLMMs is implemented. The conditional regression coefficients were estimated to be $\hat{\beta}_0^{ss} = -3.19$ and $\hat{\beta}_1^{ss} = 7.79$, and the estimated variance for the random effect among experiments was $\hat{\sigma}_u^2 = 1.02$. Hedeker et al.'s method yielded the following estimates for the marginal regression coefficients: $\hat{\beta}_0^{pa}(m\arg) = -2.79$ and $\hat{\beta}_1^{pa}(m\arg) = 7.79$. We also calculated the estimated marginal regression coefficients using the approximation method given by Zeger, et al. [17]: $\hat{\beta}_0^{pa}(appox) = -2.75$ and $\hat{\beta}_1^{pa}(appox) = 7.42$. These estimates of marginal regression coefficients were then used to estimate the BMCs and BMCLs corresponding to BMR = 0.01, 0.05, 0.10. The results are shown in Table 3.

With the meta-analysis method, the individual BMCs and their corresponding standard errors were first estimated by fitting the univariate two-parameter logistic model to each dose-response data separately. The individual experiment-specific fitted dose response curves are shown in Figure 1 along with the observed proportions of immobile *D. magna* neonates for each replicate in the study. It is evident that there is a large amount of heterogeneity between experiments, since analyses based on individual experiments yielded considerably different dose-response curves. The individual BMC estimates and their standards error were obtained for each of the BMR values (i.e., BMR = 0.01, 0.05, 0.10). We then fit the meta-analytic random-effects model in equation (14) to combine the individual BMC estimates for determining an estimate of overall average BMC, and the results are given in Table 3. Figure 2 is a forest plot displaying the individual BMC estimates and their 95% confidence intervals for a BMR of 5%. Also included in the forest plot is the population-averaged BMC along with the corresponding 95% CI.

Figure 3 displays the fitted dose-response curves based on the GEE model within dependence and exchangeable correlation structures as well as the fitted marginal dose response curves for the GLMM constructed using the marginalization and approximation approaches. The estimated overall average BMC value is also shown in the figure. In this case, these approaches produced similar results - i.e., the curves agreed well - except for the GEE with exchangeable correlation structure, which produced a dose-response curve that falls above the other curves, particularly noticeable for low concentrations. It is also noted that the GEE method with independent correlation structure and the GLMM method based on approximation produced virtually identical fits.

Table 3: Estimated BMC and 95% BMCL for a specific BMR.

Method		BMR = 0.01		BMR = 0.05		BMR = 0.1	
		$\hat{BMC}$	$\hat{BMCL}$	$\hat{BMC}$	$\hat{BMCL}$	$\hat{BMC}$	$\hat{BMCL}$
GEE	independent	0.021	0.014	0.085	0.061	0.141	0.105
	exchangeable	0.014	0.008	0.061	0.036	0.106	0.067
GLMM	marginalization	0.021	0.011	0.083	0.052	0.138	0.096
	approximation	0.021	0.012	0.085	0.057	0.141	0.104
Meta-analysis		0.018	0.011	0.082	0.054	0.140	0.094
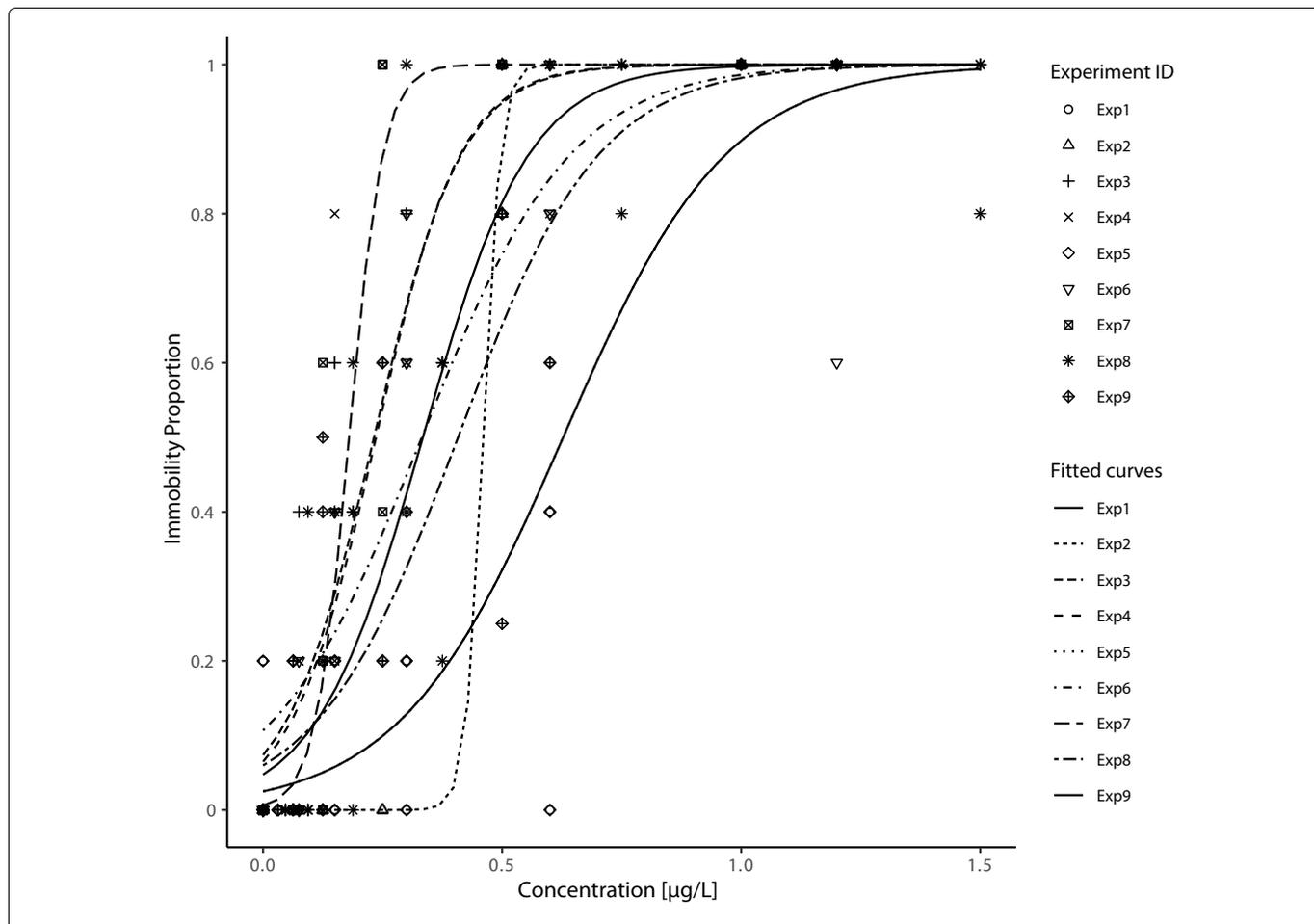
Figure 1: Individual experiment-specific fitted dose-response curves assuming a two parameter logistic regression relationship.
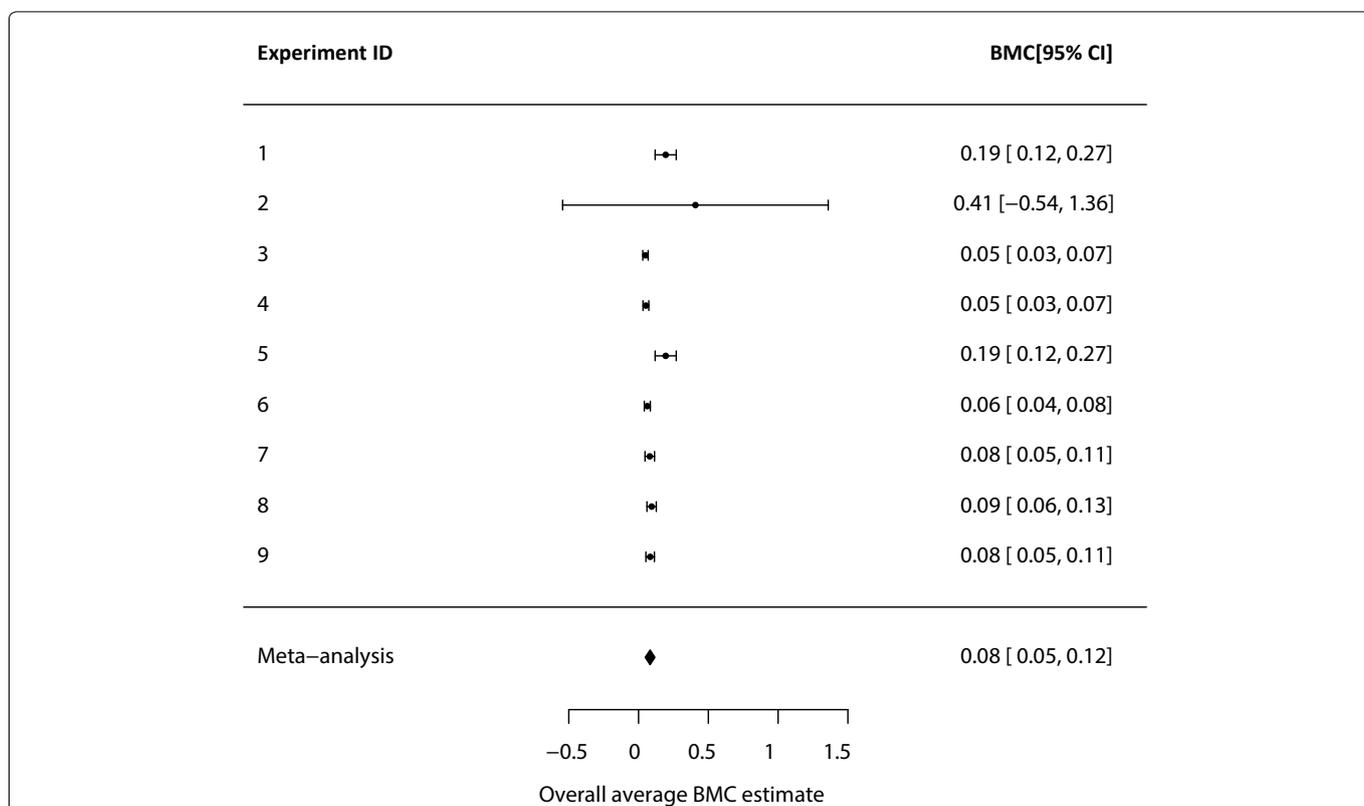
Figure 2: Forest plot showing the individual BMC estimates with 95% CI of 9 experiments investigating the toxicity of α-cypermethrin. The four-sided polygon at the bottom of the figure shows the overall average BMC (the center of the polygon) with corresponding 95% CI (the left and right edges) estimated based on the random-effects model by the meta-analysis approach.

Kerns. J Toxicol Risk Assess 2024, 10:058

• Page 9 of 12 •

Figure 3: Fitted dose-response curves based on the GEE method with independent (the dashed curve) and exchangeable (the solid curve) correlation structures and the GLMM method with marginalized (the dotted curve) and approximated (the dot-dash curve) regression coefficients assuming a two-parameter logistic regression relationship. The vertical line indicates the overall average BMC value estimated from all 9 experiments corresponding to a BMR of 0.05. The different symbols in the plot represent the observed proportion of immobility for every dose level in each experiment.

Discussion

When a series of experiments is performed to determine the BMD values for a hazardous chemical, risk analysts are often confronted with the challenge of synthesizing data from multiple experiments. In such cases, appropriate and powerful statistical methods are required to perform valid and reliable risk assessment. In this paper, we have discussed applications of three statistical methods -i.e., the GEE, GLMM, and meta-analysis approaches- to estimate population-averaged BMD values in multi-study risk assessment. More specifically, we focused on the two-parameter logistic random intercept model to demonstrate the use of the GLMM approach, and we used the random-effects model as an illustration of the meta-analysis approach.

The GEE is a marginal model popularly applied for longitudinal or grouped/clustered data. It only requires the specification of the marginal mean model and the variance function as well as the link function which connects the covariates of interest and marginal means. The GEE incorporates a working correlation structure to account for the within-group correlation (or the within-subject correlation in longitudinal study). Although the GEE yields consistent parameter estimators even when the correlation structure is mis-specified, the estimators can be inefficient if the specified structure

does not incorporate all of the information on the within-subject/group correlation. So, it is important to select an appropriate correlation structure to improve the efficiency of parameter estimation, particularly in small samples. The GEE is a semi-parametric method because it does not rely on a fully specified probability model, and thus computationally simple relative to the maximum likelihood estimation (MLE). The parameter estimates from the GEE can be directly interpreted as population parameters, but the information about the individual/group-specific effect (or the experiment-specific effect of a hazard in the current study) is lost. Therefore, it is most useful when the overall effect of a covariate is of primary interest such as, in our case, the population-averaged BMD in multi-experiment risk studies.

The GLMM approach produces the population-averaged BMD by simultaneously taking within- and between-experiment variability into consideration, and unlike the GEE method, it explicitly models the within-experiment correlation by using random effects. The GLMM is a conditional model since it parameterizes the conditional distribution of the response given the specified random effects. The method is flexible, in the sense that it can be used to provide both conditional and marginal interpretation of results, both of which

Kerns. J Toxicol Risk Assess 2024, 10:058

• Page 10 of 12 •

can be of importance depending on the specific research questions. Another advantage of the GLMM is its ability to provide improved parameter estimates with lower bias than the other two methods -as demonstrated by the simulation results- because the random effects allow for borrowing strength across multiple experiments. The extent of this is determined by experiment similarity; in general, this borrowing is more apparent for the experiments that produce extremely different results and those that have less data. However, the GLMM method has limitations. One is that the estimation of the population-averaged BMD is conditioning on estimated variance of the random effects, but the uncertainty involved in using estimated variance components was ignored in the estimation process. This may lead to underestimation of the uncertainty of the BMD estimates, resulting in narrower confidence interval with coverage probabilities smaller than the desired nominal level. Another limitation of the GLMM relates to its high computational complexity. The estimation of the regression parameters of the GLMM is more complicated compared with other methods such as the GEE. This is because with the GLMM, the marginal likelihood is formed by integrating over the distribution of random effects so that the fixed effects (i.e., the regression parameters) can be estimated. Unfortunately, the integral cannot be solved analytically in a closed form for binary data, and therefore numerical approximation is needed to evaluate the marginal likelihood. Moreover, in order to determine the population-averaged BMD, marginal parameter estimates are generated from its conditional counterparts based on the marginal probabilities, which are obtained by integrating the probability function with respect to the random effects. This integral is again intractable and must be approximated, adding more computational cost to the method.

Unlike the GLMM approach, the random-effects meta-analysis estimates the within and between-experiment variance separately in two steps. In the first step, individual BMD values and within-experiment variance are estimated by fitting a dose-response model for each experiment. The second step involves using a simple linear random effects model to estimate a mean value of the individual BMD estimates, taking into account the between-experiment variance. The random-effects meta-analysis approach is an attractive strategy to summarize BMD estimates from multiple experiments, due to its simplicity and high computational efficiency. It allows for heterogeneity across studies that cannot be readily explained by assuming the effects underlying different studies follow some distribution. The conventional choice of distribution is a normal distribution. Unfortunately, in practice it is often hard to justify the validity of this distributional assumption, and this is a common criticism of the random-effects meta-analysis. Another concern with this approach is

that when there is little information -either because there are a small number of studies or if the studies are small- it may not be able to provide reliable estimates of the standard error of the estimated overall average effect that describes the extent of heterogeneity, which may lead to unreliable statistical inferences about the overall average effect.

There are a number of issues that warrant further research. For example, to quantify the uncertainty in the BMD estimates, we used the Wald-type method to construct confidence intervals and derive the BMDL. The Wald-type construction relies on large sample approximation for valid inference, since it is based on the assumption of asymptotic normality. Therefore, while it is a simple and widely used method for building confidence intervals, it may show poor performance (that is, failing to achieve the intended nominal coverage probability) if the assumption is not met. As such, it is important to consider other methods to obtain confidence intervals with better performance in terms of coverage probability as well as the interval width, particularly in cases where the number of experiments or the experiment sample size is small in multi-experiment risk assessment. One possible alternative is the bootstrap method, which is a popular approach for producing confidence intervals due to its generality; however, it is known to be more computationally intensive than its competitors. We also realize that the two-parameter logistic random intercept model considered in the current study is a simple form of the GLMM with only one random effect. In general, the GLMM approach can be extended to allow for multiple random effects. This is useful when we have data with more than one source of random variability, but the model then becomes more complex due to additional parameters. The added random effects result in higher dimensional integrals formed when computing marginal likelihood and marginal probabilities, causing the model to be even more computationally expensive and sometimes may fail to converge. We should further point out that various numerical methods can be used to approximate the integrals appeared in the GLMM, and some of the most commonly used methods include: 1) Gaussian quadrature (GQ), which is a classical method for approximating the marginal likelihood in GLMMs; 2) Adaptive Gauss-Hermite quadrature (AQ), which is the default method for approximating the marginal likelihood in the R package GLMM adaptive; 3) Laplace approximation, which is a special case of AQ with only a single quadrature point and the default estimation method for all models in the popular R package lme4; and 4) Monte Carlo integration, which is used for calculating the marginal probabilities in the GLMM adaptive to produce marginal estimates of regression coefficients. Each method has its advantages and disadvantages, and the choice of method depends on the type of data, the specific integral to be evaluated (e.g., the formula of

the integrand, the dimension of the integrals, etc.), and amount of computational expense can be tolerated, among others. This, along with deciding the optimal number of quadrature points to use in the GQ and AGQ, remains as an active research area.

References

1. Kodell RL (2005) Managing uncertainty in health risk assessment. Int J Risk Assessm Managem 5: 193-205.

2. Crump KS (1984) A new method for determining allowable daily intake. Fundam Appl Toxicolo 4: 854-871.

3. Gaylor DW, Slikker W (1990) Risk assessment for neurotoxic effects. Neurotoxicol 11: 211-218.

4. Crump KS (1995) Calculation of benchmark doses from continuous data. Risk Analysis 15: 79-89.

5. Wang M (2014) Generalized estimating equations in longitudinal data analysis: A review and recent developments. Adv Statist 1: 1-11.

6. Dahmen G, Ziegler A (2004) Generalized estimating equations in controlled clinical trials: Hypotheses testing. Biometric J 46: 214-232.

7. Liang KY, Zeger SL (1986) Longitudinal data-analysis using generalized linear models. Biomet 73: 13-22.

8. Clayton DG (1996) Generalized linear mixed models. In: Gilks WR, Richardson S, Spiegelhalter DJ, Markov Chain Monte Carloin Practice. London: Chapman and Hall.

9. Ritz C, Jensen SM, Gerhard D, Streibig JC (2019) Dose-response analysis using R. Boca Raton: Chapman and Hall.

10. Jensen SM, Kluxen, Ritz C (2019) A review of recent advances in benchmark dose methodology. Risk Analysis 39: 2295-2315.

11. Simmons SJ, Piegorsch WW, Nitcheva D, Zeiger E (2003) Combining environmental information via hierarchical modeling: An example using mutagenic potencies. Environ metrics 14: 159-168.

12. Coull BA, Mezzetti M, Ryan LM (2003) A Bayesian hierarchical model for risk assessment of methyl mercury. J Agricult Biolo Environm Statist 8: 253-270.

13. Wheeler MW, Bailer AJ (2009) Benchmark dose estimation incorporating multiple data sources. Risk Analysis 29: 249-256.

14. Jiang X, Kopp-Schneider A (2014) Summarizing EC50 estimates from multiple dose-response experiments: A comparison of meta-analysis strategy to a mixed-effects model approach. Biom J 56: 493-512.

15. Jiang X, Kopp-Schneider A (2015) Statistical strategies for averaging EC50 from multiple dose-response experiments. Arch Toxicol 89: 2119-2127.

16. Hedeker D, Toit S, Demirtas H, Gibbon RD (2018) A note on marginalization of regression parameters from mixed models of binary outcomes. Biometrics 74: 354-361.

17. Zeger SL, Liang KY, Albert PS (1988) Models for longitudinal data: A generalized estimating equation approach. Biometrics 44: 1049-1060.

18. Raudenbush SW (2009) Analyzing effect sizes: Random effects models. In: H Cooper LV Hedges, JC Valetines, (2nd edn), The handbook of research synthesis and meta-analysis, Russell Sage Foundation, New York, 295-315.

19. Viechtbauer W (2010) Conducting meta-analyses in R with the meta for package. Journal of Statistical Software 36: 1-48.

20. Hunter JE, Schmidt FL (1990) Methods of meta-analysis: Correcting error and bias in research findings. Thousand Oaks, CA: Sage.

21. Viechtbauer W (2005) Bias and efficiency of meta-analytic variance estimators in the random-effects model. Journal of Educational and Behavioral Statistics 30: 261-293.

22. Hartung J, Knapp G (2001) On tests of the overall treatment effect in meta- analysis with normally distributed responses. Stat Med 20: 1771-1782.

23. Hartung J, Knapp G (2001) A refined method for the meta-analysis of controlled clinical trials with binary outcome. Stat Med 20: 3875-3889.

24. Halekoh U, Højsgaard S, Yan J (2006) The R package geepack for Generalized Estimating Equations. Journal of Statistical Software 15: 1-11.

25. Rizopoulos D (2023) GLM Madaptive: Generalized linear mixed models using adaptive gaussian quadrature.

26. Ritz C, Baty F, Streibig JC, Daniel G (2015) Dose-response analysis using R. PLoS One 10: e0146021.

27. Gottardi M, Cedergreen N (2019) The synergistic potential of azole fungicides does not directly correlate to the inhibition of cytochrome P450 activity in aquatic invertebrates. Aquat Toxicol 207: 187-196.

28. Jensen SM, Kluxen FM, Streibig JC, Cedergreen N, Ritz C (2020) bmd: An R package for benchmark dose estimation. Peer J 8: e10557.