Journal Title
XX(X):1–13

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Hierarchical and Robust Bayesian Meta-Models for Estimating Cadmium Dose-Response Curves

Jeongin Lee[1], Inyoung Beak[1], Jeongseon Kim[2], Jaeoh Kim[1], and Seongil Jo[1]

Abstract

Cadmium is a toxic heavy metal that accumulates in human tissues and poses long-term health risks, particularly nephrotoxicity. Benchmark dose (BMD) modeling provides a scientific basis for establishing health-based guidance values; however, conventional dose-response models often fail to address heterogeneity and outliers across studies. This study conducts a Bayesian meta-analysis of 48 studies and 250 exposure groups published between 1983 and 2022, focusing on the relationship between urinary cadmium (UCd) and urinary 2-microglobulin (2MG), a biomarker of early renal dysfunction. We propose a robust hierarchical Bayesian framework based on a Hierarchical Mixture of Experts (HME) model with a piecewise linear structure and Student's t-distribution to capture latent subgroup structure and improve robustness against outliers. Compared with standard models, the HME-PL model achieved superior performance in terms of Deviance Information Criterion (DIC) and Normalized Root Mean Squared Error (NRMSE). Age- and sex-specific BMD and BMDL estimates indicate that older adults, particularly females, exhibit greater sensitivity to cadmium exposure. These findings highlight the importance of incorporating heterogeneity in dose-response assessment and provide a practical methodological contribution for deriving regulatory reference points in toxicological risk assessment.

Keywords

Cadmium; Benchmark dose; Bayesian meta-analysis; Hierarchical Mixture of Experts; Robust modeling; Risk assessment

1 Introduction

Cadmium is a toxic trace metal found in both natural environments and as a byproduct of industrial processes such as non-ferrous metal smelting, nickel-cadmium battery production, and phosphate fertilizer manufacturing. Its chemical stability and low biological excretion rate contribute to high environmental persistence and bioaccumulation, particularly in aquatic systems and agricultural soils. Dietary intake is the primary route of human exposure, as cadmium is readily absorbed by crops like *Oryza sativa* (rice) and leafy vegetables. Following absorption, cadmium accumulates in renal and hepatic tissues, where it induces chronic toxicity with a biological half-life ranging from 10 to 30 years (Nordberg and Nordberg 2022). It is associated with nephrotoxic, osteotoxic, and carcinogenic effects (Prozialeck and Edwards 2012), and is classified as a Group 1 carcinogen by the International Agency for Research on Cancer (IARC). In response, regulatory agencies—including the EU's REACH, the U.S. OSHA, and Korea's Ministry of Employment and Labor—have designated cadmium as a priority hazardous substance. Major international bodies such as WHO, EFSA, and ATSDR have established health-based guidance values (HbGVs) to mitigate risk from long-term exposure (EFSA Panel on Contaminants in the Food Chain (CONTAM) 2011).

Chemical risk assessment generally proceeds through four key stages: hazard identification, dose-response assessment, exposure assessment, and risk characterization. Of these, the dose-response assessment is essential for quantifying the relationship between exposure levels and adverse health outcomes, thereby providing the scientific basis for deriving reference points (RPs) used in regulatory frameworks Lee et al. (2022). In cadmium risk assessment, particular attention is given to the relationship between internal exposure and early biological effects on the kidney. Urinary cadmium (UCd) is widely accepted as a long-term internal dose indicator, while urinary β2-microglobulin (β2MG), a low molecular weight protein reabsorbed by the renal proximal tubule, serves as a sensitive biomarker of cadmium-induced nephrotoxicity (Prozialeck and Edwards 2012). Numerous studies have reported a nonlinear dose-response relationship between UCd and urinary β2MG, especially at low to moderate exposure levels. Accurately modeling this nonlinear association is crucial for determining benchmark dose (BMD) values and their lower confidence bounds (BMDLs), which are increasingly used as reference points in health-based guidance value derivation.

Various dose-response models have been employed in cadmium risk assessment, including linear regression, Poisson models, penalized splines, and Cox proportional hazards models, to estimate benchmark doses (BMDs) (Stayner et al. 1992; Steenland and Deddens 2004; Chen et al. 2024). In addition, hormesis models have been applied to describe non-monotonic dose-response patterns, particularly those exhibiting low-dose stimulation and high-dose inhibition (Nweke and

[1]Department of Statistics and Data Science, Inha University, South Korea
[2]Department of Cancer AI and Digital Health, Graduate School of Cancer Science and Policy, National Cancer Center, South Korea

Prepared using sagej.cls *[Version: 2017/01/17 v1.20]*

Ogbonna 2017; Yang et al. 2021; Pruvost-Couvreur et al. 2020). These models are generally implemented under the frequentist framework, where inference is based on point estimates and confidence intervals derived from sample data. Frequentist approaches, however, are limited in their ability to quantify the full range of uncertainty associated with model parameters—especially when dealing with small sample sizes, measurement error, or inter-study heterogeneity. They also do not yield probability-based interpretations for quantities such as the BMD and its lower confidence bound (BMDL), which are critical for setting regulatory thresholds.

Bayesian methods address these limitations by incorporating prior information and generating full posterior distributions for model parameters. This facilitates formal uncertainty quantification and supports probabilistic interpretation of quantities such as the BMD and BMDL (Lee et al. 2013; Moran et al. 2021; Du et al. 2024; Billoir et al. 2011). Compared to conventional approaches, Bayesian modeling offers enhanced interpretability and robustness, particularly when data are sparse or heterogeneity is substantial. Meanwhile, to account for structural variability across studies—such as differences in exposure metrics, population characteristics, and biomarker definitions—a hierarchical modeling framework is often necessary. The Bayesian paradigm naturally accommodates this structure through multilevel models that jointly represent within-study and between-study variation. This allows for partial pooling of information and enables stable estimation even when individual studies lack statistical power Gelman et al. (1995). Furthermore, it permits posterior inference at both the population level and study-specific level, supporting generalizable as well as context-sensitive risk characterization.

To address the limitations of existing cadmium dose-response analyses, this study conducts a Bayesian meta-analysis using an expanded dataset comprising 48 studies and 250 exposure groups published between 1983 and 2022. We analyze the relationship between UCd and urinary β2MG, a biomarker indicative of early renal dysfunction, and demonstrate that conventional piecewise linear (PL) models often fail to accommodate the heterogeneity and outlier patterns inherent in large-scale aggregated data.

Relevant literature was systematically retrieved from the PubMed and Embase databases using the search terms `("benchmark dose")` and `("cadmium")` up to September 30, 2023. Additional references were identified via manual review of bibliographies from selected publications. Following criteria similar to those adopted in previous studies (Prankel et al. 2004; Gallagher and Meliker 2010; Woo et al. 2015; Filippini et al. 2020), we included studies that: (1) reported BMDs derived from urinary cadmium concentrations and renal biomarkers such as β2MG or NAG; (2) excluded cadmium intake based on dietary lifetime exposure; and (3) provided both BMD and its lower confidence limit (BMDL). Extracted data included publication metadata, exposure and biomarker levels (e.g., UCd, β2MG, NAG), threshold criteria, and dose-response estimates. Where only group-level statistics were available, pooled values were computed via weighted averaging for use in meta-regression. To overcome the limitations of standard PL models under such heterogeneity, we propose a robust Bayesian framework based on a Hierarchical Mixture of Experts (HME) model (Bishop and Svensén 2012). This model uses latent indicator variables to separate the dataset into 'mainstream' and 'non-mainstream' subgroups and fits distinct PL models to each. A shifted Student's t-distribution is employed to enhance robustness to heavy-tailed noise and outlier contamination. The primary contributions of this study are as follows:

- Construction of a comprehensive UCd–β2MG meta-analytic dataset spanning four decades
- Development of a robust Bayesian HME model that captures latent subgroup structure and provides stable BMD estimates
- Presentation of a practical framework for health-based guidance value derivation under data heterogeneity and uncertainty

The remainder of this paper is organized as follows. Section 2 provides a detailed overview of dose-response modeling approaches and BMD methodology. Section 3 describes the data collection and curation process, including study selection criteria, and introduces the proposed modeling framework based on PL functions and the HME structure. In Section 4, we compare the performance of the standard PL model (Standard-PL) and the HME-PL model under both normal and shifted Student's t-distribution assumptions, using deviance information criterion (DIC) and normalized root mean squared error (NRMSE) as model selection criteria. Based on the optimal model, we derive age- and sex-specific BMD estimates for cadmium-induced nephrotoxicity. Finally, Section 5 summarizes the main findings, discusses the implications and limitations of the study, and outlines directions for future research.

2 Background

Cadmium can be absorbed not only through industrial activities causing pollution, but also occurs naturally, easily entering aquatic and terrestrial environments, and subsequently accumulating in foods such as shellfish, crops, and leafy vegetables. As a heavy metal that is difficult to fully control, cadmium is classified under hazardous substances regulated by industrial safety and health regulations and is treated as a specially controlled substance due to its carcinogenic properties. Therefore, its concentrations must be maintained below permissible limits. Various countries and international organizations establish Health-based Guidance Values (HbGV) for such hazardous substances to ensure food safety through hazard assessment (Yoon et al. 2011; EFSA Panel on Contaminants in the Food Chain (CONTAM) 2011).

Risk assessment comprises four main stages: hazard identification, dose-response assessment, exposure analysis, and risk characterization. Dose-response assessment, a critical component, involves quantitatively evaluating human health impacts using toxicity data from human or animal studies to establish a reference point (RP) for hazard thresholds used in determining occupational risk (Lee et al. 2022). This process utilizes dose-response curves to quantitatively assess harmful effects on humans.

In dose-response assessment, a key indicator for internal accumulation is the concentration of cadmium, which sensitively responds to long-term exposure. The biomarker used to assess responses to exposure is $\beta 2MG$, which indicates renal damage due to chronic cadmium exposure. Reduced protein $\beta 2MG$ reflects impaired renal function where it fails to be reabsorbed in the tubules and is instead excreted. The process of establishing threshold values for cadmium toxicity based on dose-response data involves two main steps. Initially, the dose-response curve is modeled, followed by calculation of the toxicity threshold using the modeled curve. Detailed descriptions of each step are provided in the following sections.

2.1 Modeling Dose-Response Curve

In this section, we discuss the modeling of dose-response curves for establishing threshold values (Reference Point, RP) for cadmium toxicity. A meta-analytic approach is used to synthesize evidence across multiple studies that have reported exposure and response information. Meta-analysis allows for the statistical integration of quantitative results from different studies when they are sufficiently comparable (National Evidence-based Healthcare Collaborating Agency (NECA) 2023).

We collected relevant literature that reported geometric means of UCd and $\beta 2MG$, sample sizes, and corresponding response indicators. Based on these data, the dose-response relationship was modeled to capture the average trend in cadmium exposure and renal biomarker levels across studies.

Specifically, we consider the Piecewise Linear (PL) model, which has been used in prior research to reflect the nonlinear association between UCd concentration and $\beta 2MG$ (EFSA; Lee et al. 2013).

The PL model, represented by Equation (1), describes a piecewise linear function in which the slope (β) changes at a breakpoint (δ), illustrating the nonlinear relationship between exposure and response.

$$\mu(d) = \begin{cases} \alpha_1 + \beta_1 d & \text{if } d < \delta \\ \alpha_2 + \beta_2 (d - \delta) & \text{if } d \geq \delta, \end{cases} \tag{1}$$

where α_2 denotes the response effect at the breakpoint δ.

2.2 Benchmark Dose

In this section, we describe the process of estimating the RP for cadmium toxicity based on the dose-response curve modeled in the previous section. In traditional risk assessment, RPs have typically been derived from observed dose levels such as the No Observed Adverse Effect Level (NOAEL) or the Lowest Observed Adverse Effect Level (LOAEL) (Lee et al. 2022).

NOAEL refers to the highest dose at which no statistically or biologically significant adverse effects are observed, whereas LOAEL is the lowest dose at which such effects are observed. These values are determined based on the specific dose levels tested in experiments. As such, they are limited in that they depend heavily on the choice of dose levels, sample size, and experimental design. Furthermore, they do not fully utilize the shape of the dose-response curve (Lee et al. 2022).

To address these limitations, Crump (Crump 1984) proposed the concept of the BMD. The BMD corresponds to a dose associated with a specified change in response, referred to as the benchmark response (BMR), relative to the background dose (bd). Here, the background dose is defined as the minimum dose level included in the dose-response data. Unlike NOAEL or LOAEL, the BMD is not an observed value, but is estimated from the modeled dose-response curve.

In recent years, a growing body of research has emphasized that BMD is a more scientifically robust basis for deriving toxicological reference values compared to NOAEL (Zarn et al. 2020; Baralić et al. 2020). The EFSA 2022 guidance document (Committee et al. 2022) explicitly recommends the BMD approach over the NOAEL approach for deriving RPs, highlighting its methodological advantages. Moreover, it encourages a shift from the frequentist quantile-based framework in the 2017 EFSA guidance (Committee et al. 2017) toward a Bayesian paradigm. In line with these recommendations, the present study adopts a Bayesian approach to estimate BMD and derive the corresponding RP for cadmium exposure.

To calculate the BMD, it is first necessary to define the BMR, which represents a specific magnitude of response considered biologically or toxicologically meaningful. Two common definitions of risk are used in BMD modeling: additional risk and extra risk. These are defined as follows:

$$\begin{aligned} \text{BMR} &= P(BMD) - P(bd), \quad \text{(additional risk)} \\ \text{BMR} &= \frac{P(BMD) - P(bd)}{1 - P(bd)}, \quad \text{(extra risk)} \end{aligned} \tag{2}$$

In Equation (2), $P(d)$ denotes the probability that the response variable Y_d at dose d exceeds a predefined cut-off value c^*, which reflects the threshold at which a toxic or adverse effect is considered to occur. Assuming Y_d follows a normal distribution with mean $\mu(d)$ and constant variance σ^2, i.e., $Y_d \sim N(\mu(d), \sigma^2)$, the probability $P(d)$ can be expressed as:

$$P(d) = P(Y_d > c^*) = 1 - \Phi\left(\frac{c^* - \mu(d)}{\sigma}\right) \tag{3}$$

In this study, we adopt the extra risk definition of BMR. Under this framework, the BMD can be

derived using Equations (4) and (5). As is common in toxicological data analysis, we apply a logarithmic transformation to the dose variable to account for the typically skewed distribution of dose values.

$$1 - P(\log(\text{BMD})) = (1 - \text{BMR}) \cdot \Phi\left(\frac{\log c^* - \mu(\log(bd))}{\sigma}\right) \equiv p^* \quad (4)$$

$$\text{BMD} = \exp\left\{\mu^{-1}\left[\log c^* - \sigma \cdot \Phi^{-1}(p^*)\right]\right\} \quad (5)$$

Based on the estimated BMD, we derive the lower bound of the 95% credible interval, referred to as BMDL. In the BMD approach, this lower bound is commonly used as the RP for risk assessment.

- **BMDL**x: the dose below which the change in response is likely to be less than x%, where 'likely' is defined by the statistical credible level (usually 95%) (Committee et al. 2022).

3 Bayesian Meta-Analysis Models for Cadmium Dose-Response Evaluation

3.1 Robust Bayesian Meta-Analysis Model based on Piecewise Linear Model

Let B_{ijk} and U_{ijl} denote the β2MG and UCd concentration for individual k in group i of study j, respectively, and assume that both follow a log-normal distribution. The collected data Y_{ij} and X_{ij} represent the geometric means of B_{ijk} and U_{ijl}, and thus $log(Y_{ij})$ and $log(X_{ij})$ follow a normal distribution:

$$log(Y_{ij}) = \frac{1}{n_i}\sum_{k=1}^{n_i} log(B_{ijk}) \sim N(\mu_{ij}, \frac{\sigma^2}{n_i}),$$

$$log(X_{ij}) = \frac{1}{m_i}\sum_{l=1}^{m_i} log(U_{ijl}) \sim N(d_{ij}, \frac{\theta^2}{m_i}),$$

where j is the index for the $J = 49$ studies, i indexes the subject groups within each study, and k denotes the individual subjects ($\sum_{j=1}^{J} t_j = 250$). The variable X_{ij} is modeled as the actual cadmium dose absorbed, denoted d_i, while Y_{ij} is modeled through a piecewise linear (PL) dose-response function μ_{ij}.

Although the initial model assumes normality for simplicity, model fitting revealed potential sensitivity to outliers, especially due to the inclusion of extended data. To address this limitation, we considered a more robust distributional assumption based on the Student's t-distribution. Specifically, we assume that both $\log(B_{ijk})$ and $\log(U_{ijl})$ follow a Student's t-distribution. As a result, their sample means, $\log(Y_{ij})$ and $\log(X_{ij})$, also follow a scaled version of the t-distribution due to the additive property of independent t-like variables. This relationship is formalized in Proposition 1:

Proposition 1. *Let* $Z_1, \ldots, Z_n \overset{iid}{\sim} \mathcal{N}(0,1)$ *and* $V \sim \chi^2_\nu$, *all mutually independent. Define*

$$T_n = \frac{Z_1 + \cdots + Z_n}{\sqrt{V/\nu}}.$$

Then $T_n \sim \sqrt{n} \cdot t_\nu$, *i.e.,*

$$\frac{T_n}{\sqrt{n}} \sim t_\nu.$$

Using this result, we model the log-transformed geometric means as:

$$\log(Y_{ij}) = \frac{1}{n_i}\sum_{k=1}^{n_i} \log(B_{ijk}) \sim t\left(\mu_{ij}, \frac{\sigma^2}{n_i}, \nu\right), \quad (6)$$

where the distribution is understood as a scaled and shifted Student's t-distribution with location parameter μ_{ij}, scale parameter σ^2/n_i, and degrees of freedom ν. The probability density function (pdf) of the shifted and scaled Student's t-distribution is given by:

$$f(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{1}{\nu\pi\tau^2}\right)^{1/2}\left\{1 + \frac{(y-\mu)^2}{\nu\tau^2}\right\}^{-\frac{\nu+1}{2}}.$$

Note this formulation enhances robustness against outliers, a property we later confirm in the data analysis section by comparing model performance under both normal and t-distributional assumptions. The results demonstrate that the t-distribution provides improved stability in the presence of extreme values.

The PL model used in this study is defined in Equation (7), incorporating the effects of age (a_i) and sex (g_i), which were found to be significant covariates (see Figure 2). To account for study-level heterogeneity across studies (j), we included a random effect s_j modeled as a normally distributed term.

$$u_{ij} = \begin{cases} \alpha_1 + \beta_1 d_i + r_a a_i + r_g g_i + s_j, & \text{if } d_i \le \delta \\ \alpha_2 + \beta_2(d_i - \delta), & \text{if } d_i > \delta, \end{cases} \quad (7)$$

where $s_j \sim \mathcal{N}(0, \tau_{st})$. Descriptions of the model parameters are provided in Table 1.

Table 1. Parameters of the PL Model

Parameter	Description
β_1	Slope before the breakpoint
β_2	Slope after the breakpoint
α_1	Background response at zero dose
α_2	Response level at the breakpoint δ
δ	Breakpoint dose where the slope changes
d_i	Latent urinary cadmium dose
r_a	Regression coefficient for age effect (a_i)
r_g	Regression coefficient for sex effect (g_i)
s_j	Study-specific random intercept for study j

If available, observed variances of UCd and β2MG (denoted S^2_{UCd} and $S^2_{\beta 2MG}$, respectively) would follow chi-squared distributions and be modeled as:

$$S^2_{i,Ucd} \sim \frac{1}{m_i}\chi^2(m_i - 1), \quad i = 1, 2, \ldots, 250$$

$$S^2_{i,\beta 2MG} \sim \frac{1}{n_i}\chi^2(n_i - 1), \quad i = 1, 2, \ldots, 250$$

However, due to data unavailability, these variance components were not incorporated in the present

analysis. To estimate the model parameters, we employed a Bayesian approach using the Markov Chain Monte Carlo (MCMC) method. The prior distributions used are as follows:

$$\begin{aligned}
&\alpha_1 \sim \mathcal{N}(0.001, 1000)\mathbb{I}(0, 100), && \delta \sim \mathcal{N}(0.001, 1000),\\
&\beta_1 \sim \mathcal{N}(0.001, 1000)\mathbb{I}(0, 100), && \beta_2 \sim \mathcal{N}(0.001, 1000)\mathbb{I}(0, 100),\\
&r_g \sim \mathcal{N}(0.001, 1000), && r_e \sim \mathcal{N}(0.001, 1000),\\
&r_a \sim \mathcal{N}(0.001, 1000), && \frac{1}{\sigma^2} \sim \text{Gamma}(0.01, 0.01),\\
&\frac{1}{\theta^2} \sim \text{Gamma}(0.01, 0.01), && \frac{1}{\tau_{st}} \sim \text{Gamma}(0.01, 0.01),\\
&d_i \sim \mathcal{N}(0.001, 1000), && \nu \sim \text{Gamma}(2, 1)\mathbb{I}(2, 30),
\end{aligned}$$

where d_i denotes the latent true cadmium dose and is modeled as a normally distributed latent variable due to its unobserved nature.

Figure 1 displays the fitted PL model under both normal and Student's t-distribution assumptions. Compared to previous studies Lee et al. (2013), both models exhibit a reduction in slope after the breakpoint δ. Notably, the robust Student's t model estimates a smaller slope than the normal model, highlighting the model's sensitivity to outlier classification. Since the newly incorporated data may contain atypical subgroups, we further extend the framework in the following section by introducing a Hierarchical Mixture of Experts (HME) model to better capture both dominant and atypical data patterns.

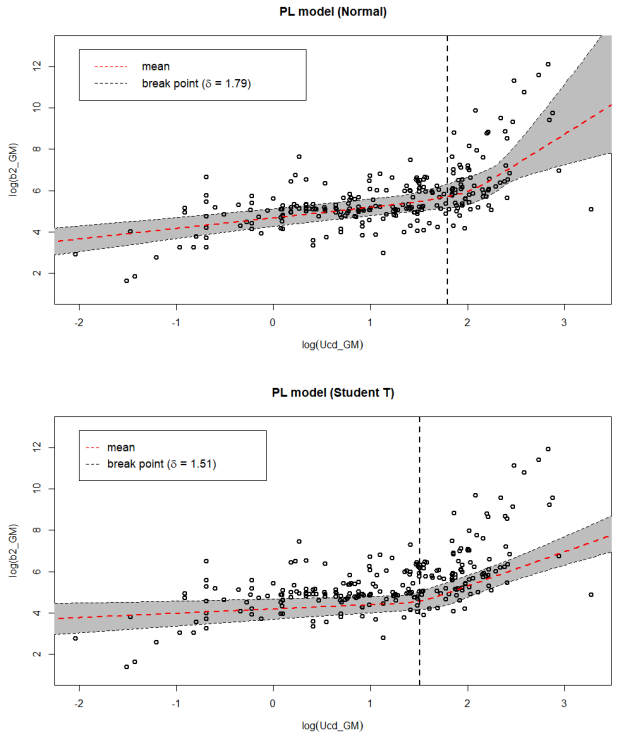

Figure 1. Standard Piecewise Linear (Standard-PL) model fitted to the complete dataset without the HME structure. The model assumes a normal distribution (upper) and a Student's t-distribution (below).

3.2 Extension to Hierarchical Mixtures of Experts

Mixture of Experts (MoE) models provide a flexible framework that combines the predictions of multiple expert models using a gating network. This structure allows MoE models to effectively capture nonlinear relationships, making them well-suited for modeling complex dose-response curves. For example, Patel et al. (2014) employed MoE to probabilistically model the effects of physicochemical properties of nanoparticles on toxicity using a gating mechanism. Canale et al. (2018) proposed a convex mixture regression approach combining two specialized experts for low- and high-dose extremes using a dose-dependent gating structure, and applied it to estimate benchmark doses (BMD) for arsenic exposure in drinking water.

Extending MoE, the Hierarchical Mixture of Experts (HME) model introduces a tree-structured conditional probability framework composed of multiple gating nodes and expert nodes Bishop and Svensén (2012). At each gate node, branching probabilities are computed, enabling HME to flexibly model complex and heterogeneous data.

In our study, we constructed a binary HME tree consisting of one gating node and two expert nodes. The latent variable z_i indicates which expert node individual i is routed to:

$$z_i \sim \text{Bernoulli}(p_i), \quad p_i = \frac{\exp(\alpha_0 + \beta_0 a_i)}{1 + \exp(\alpha_0 + \beta_0 a_i)},$$

where, p_i is defined by a logistic (sigmoid) function of age a_i, reflecting the observed separation in UCd and β2MG levels by age.

Each expert node follows a separate PL model. For $z_i = 1$, the outcome model is:

$$h_{ij} = \begin{cases} \alpha_1 + \beta_1 d_i + r_a a_i + r_g g_i + s_j, & \text{if } d_i \le \delta_1 \\ \alpha_2 + \beta_2 (d_i - \delta_1), & \text{if } d_i > \delta_1 \end{cases} \tag{8}$$

For $z_i = 0$, a second PL model is used:

$$h'_{ij} = \begin{cases} \alpha_3 + \beta_3 d_i + r_a a_i + r_g g_i + s_j, & \text{if } d_i \le \delta_2 \\ \alpha_4 + \beta_4 (d_i - \delta_2), & \text{if } d_i > \delta_2 \end{cases} \tag{9}$$

In both models, $s_j \sim \mathcal{N}(0, \tau_{st})$ is a study-specific random effect. The final mean model is defined as a mixture:

$$\mu_{ij} = z_i h_{ij} + (1 - z_i) h'_{ij} \tag{10}$$

To estimate the model parameters, we assign the following prior distributions:

$$
\begin{aligned}
\alpha_1, \alpha_2, \alpha_3, \alpha_4 &\sim \mathcal{N}(0.001, 1000)\, \mathbb{I}(0, 5) \\
\beta_1, \beta_2, \beta_3, \beta_4 &\sim \mathcal{N}(0.001, 1000)\, \mathbb{I}(0, 100) \\
\delta_1 &\sim \mathcal{N}(0.001, 0.001) \\
\delta_2 &\sim \mathcal{N}(0.001, 0.001)\, \mathbb{I}(\delta_1, \infty) \\
r_g, r_a &\sim \mathcal{N}(0.001, 1000) \\
\frac{1}{\sigma^2}, \frac{1}{\theta^2}, \frac{1}{\tau_{st}} &\sim \text{Gamma}(0.01, 0.01) \\
d_i &\sim \mathcal{N}(0.001, 1000) \\
\nu &\sim \text{Gamma}(2, 1)\, \mathbb{I}(2, 30).
\end{aligned}
$$

4 Cadmium Data Analysis

4.1 Data Description and Exploratory Analysis

To conduct a Bayesian meta-analysis of cadmium dose-response relationships, we compiled a dataset from 250 articles derived from 48 independent studies published between 1983 and 2022. The dataset contains study-level summaries, including the study author, year of publication, country of study, type and level of environmental cadmium exposure, threshold levels for β2MG (μg/g creatinine) and NAG (U/g creatinine), geometric means of β2MG and NAG, geometric mean UCd concentration (μg/g creatinine), sample sizes, and reported BMD and BMDL values. Studies were included based on the following criteria: (1) BMD must be estimated using UCd concentration and renal biomarkers; (2) Studies based on lifetime exposure or dietary cadmium intake (e.g., from cereals) were excluded; (3) Both BMD and BMDL values must be explicitly reported.

Figure 2 displays exploratory visualizations of the geometric means of UCd and β2MG concentrations, as well as their distributions stratified by age and sex. The following key observations were made: First, a nonlinear association is observed between UCd and urinary β2MG concentrations. Second, the distribution of urinary β2MG differs significantly across sex and age groups. Third, individuals aged 50 or older generally show higher urinary β2MG levels than those under 50. Fourth, among those aged 50 and older, females tend to have higher urinary β2MG levels than males.

5 Results

This section presents the results of Bayesian modeling using the PL model and corresponding prior distributions introduced in the previous Sections. The posterior samples were obtained via the MCMC method implemented through the `rjags` package in `R`, which utilizes the JAGS library. Model fitting was performed using two parallel Markov chains. For each chain, 100,000 burn-in iterations were discarded, after which 10,000 posterior samples were retained by thinning every 30 iterations. The Piecewise Linear (PL) model was fitted within the Hierarchical Mixture of Experts (HME) framework, hereafter referred to as the HME-PL model. In the HME-PL model, the log-transformed response variable, $\log(Y_{ij})$, is assumed to follow either a normal distribution or a Student's t-distribution. The following subsections summarize the model fitting results, clustering behavior, and performance comparison across distributional assumptions and modeling structures.

Table 2. Comparison of Gelman-Rubin statistics (GR-stats) for the HME-PL model under Normal versus Student's t-distribution assumptions

Normal	GR-stats	Student's t	GR-stats
r_a (Age)	1.59	r_a (Age)	1.00
r_m (Male)	1.07	r_m (Male)	1.00
r_f (Female)	1.06	r_f (Female)	1.00
α_1 (Background)	1.56	α_1 (Background)	1.01
δ_1 (Breakpoint1)	1.00	δ_1 (Breakpoint1)	1.00
δ_2 (Breakpoint2)	1.00	δ_2 (Breakpoint2)	1.00
β_1 (Slope1)	1.03	β_1 (Slope1)	1.02
β_2 (Slope2)	1.00	β_2 (Slope2)	1.00
β_3 (Slope3)	1.04	β_3 (Slope3)	1.01
β_4 (Slope4)	1.00	β_4 (Slope4)	1.00
τ_{st} (Study RE)*	1.85	τ_{st} (Study RE)*	1.00
α_0 (Gate Intercept)	1.43	α_0 (Gate Intercept)	1.00
β_0 (Gate Effect)	1.00	β_0 (Gate Effect)	1.01

* Study RE: Study Random Effect.

5.1 Estimation Results of Dose-Response Curves Using HME

We first examine the convergence of the posterior samples obtained via MCMC for the PL models incorporating the HME framework, where the distribution of $\log(Y_{ij})$ is assumed to follow either a Normal or a Student's t-distribution. To assess the convergence of each model, the Gelman–Rubin statistic was calculated. This diagnostic evaluates between- and within-chain variances and indicates satisfactory convergence when the statistic is close to 1.

As shown in Table 2, the Student's t-distribution model exhibited excellent convergence properties, with all parameters having Gelman–Rubin statistics very close to 1. This suggests that the posterior samples from two parallel chains, after 100,000 burn-in iterations, converged well. Trace plots further confirmed that the posterior draws were well-mixed and displayed no apparent trend or periodicity.

Next, to address the component allocation of observations in the HME model, we employed Dahl (2006)'s method to summarize the clustering structure implied by the posterior samples of the latent allocation variables Z_i. At each MCMC iteration b, a clustering configuration $c_b = (z[1], \ldots, z[250])_b$ was recorded, yielding a collection $\{c_1, \ldots, c_B\}$ across B total iterations.

The procedure for Dahl's method is as follows:

- **Posterior Samples of Component Allocations Z:** MCMC samples of how each observation is assigned to mixture components in each iteration.
- **Association Matrix $\delta(c_b)$:** For each clustering c_b, construct a binary matrix indicating whether each pair of observations share the same cluster.
- **Pairwise Probability Matrix $\hat{\pi}$:** Compute the average of $\delta(c_b)$ over all iterations, representing

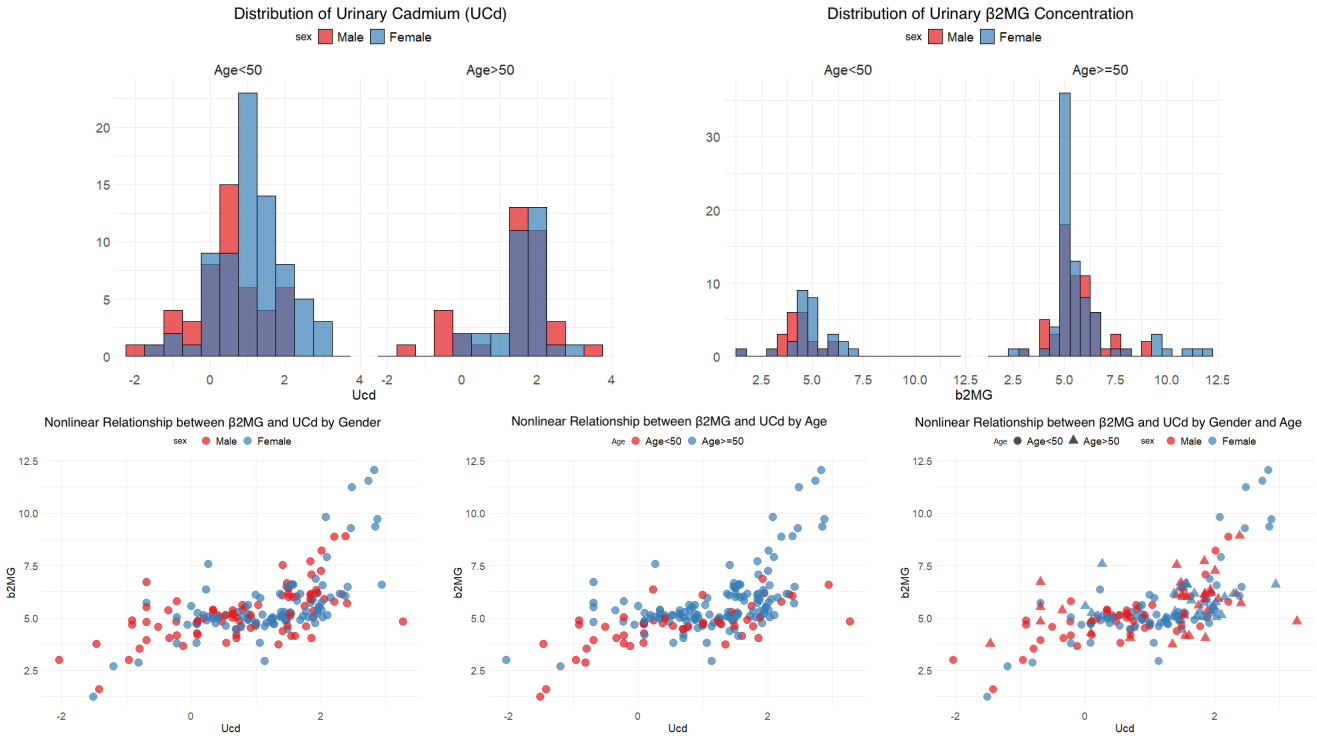

Figure 2. Distributions and nonlinear associations between UCd and β_2MG, stratified by sex and age

the empirical posterior probability that a pair of observations co-cluster.

- **Least-Squares Clustering c_{LS}:** Select the clustering c_b that minimizes the squared deviation from $\hat{\pi}$:

$$c_{LS} = \arg \min_{c \in \{c_1, \ldots, c_B\}} \sum_{i=1}^{G} \sum_{j=1}^{G} \left(\delta_{i,j}(c) - \hat{\pi}_{i,j}\right)^2$$

Based on this clustering approach, the final allocation results indicated that under the normal distribution assumption, posterior samples of Z_i were distributed as 167 to Component 1 and 83 to Component 2. Under the Student's t-distribution model, the allocation was 176 to Component 1 and 74 to Component 2.

Table 3 presents the posterior summary statistics for the model parameters, and Figure 3 visualizes the mean dose-response curves estimated from the posterior samples based on Equations (8) and (9). In both distributional assumptions, the estimated breakpoint δ_2 corresponding to Component 2 is substantially larger than the breakpoint δ_1 of Component 1. This result implies that the data assigned to Component 2 tend to follow a relatively linear pattern on the log-scale, which is also visually evident in Figure 3, where the green line representing Component 2 shows a much flatter slope.

In addition, the logistic coefficient β_0 associated with the binary indicator a_i (indicating whether the subject is over 50 years old) was estimated to be positive under both distributional assumptions. This suggests that individuals aged 50 or older are more likely to be classified into Component 1 by the gating network.

A notable difference between the two models lies in the estimated slope β_2 after the breakpoint δ_1 in Component 1. Under the normal distribution, the posterior mean of β_2 was estimated to be

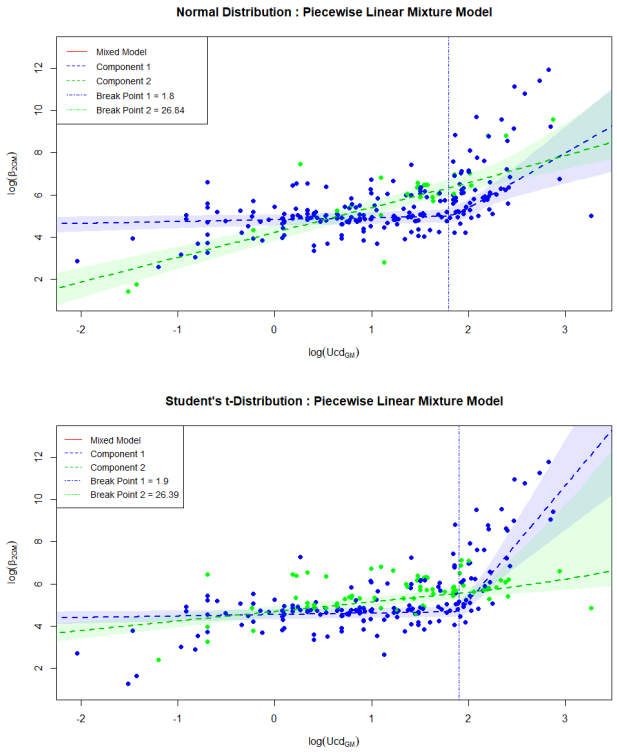

Figure 3. Fitted HME-PL dose-response curves under Normal (upper) and Student's t-distribution (below) assumptions. The model visualizes two distinct components, with Component 1 (blue) representing the mainstream data and Component 2 (green) capturing atypical patterns.

approximately 1.77, whereas under the Student's t-distribution, it was significantly larger at around 4.39. This indicates that, for the main component (Component 1) where most data points are assigned, the normal model—being sensitive to extreme values—may have yielded a downward-biased slope due to a few influential outliers. In contrast, the Student's t model provided a more robust estimation that better captures the true nonlinear trend of the data. This supports the benefit of the Student's t-distribution in mitigating the influence of outliers in dose-response modeling using HME structures.

5.2 Model Comparison of HME-PL Dose-Response Models

In this section, we compare the fitted PL models under two distributional assumptions for $\log(Y_{ij})$: the normal distribution and the Student's t-distribution. We evaluate model performance using two widely accepted Bayesian model comparison criteria: the Deviance Information Criterion (DIC, Spiegelhalter et al. (2002)) and the Normalized Root Mean Square Error (NRMSE).

The DIC is defined as follows:

$$\text{DIC} = \bar{D} + 2p_D, \tag{11}$$

where the deviance is given by $D(\theta) = -2\log(p(y \mid \theta))$, and $\bar{D} = \mathbb{E}_\theta[D(\theta)]$ is its posterior mean. The effective number of parameters p_D is defined as $p_D = \bar{D} - D(\bar{\theta})$, where $\bar{\theta}$ is the posterior mean of the parameters. A smaller DIC indicates a better trade-off between model fit and complexity.

The NRMSE is defined as:

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\sigma}, \tag{12}$$

where $\hat{y}_i$ denotes the predicted value for observation y_i, and σ is the standard deviation of the observed data. Smaller values of NRMSE indicate better predictive accuracy, normalized by the data scale.

The `rjags` package automatically reports DIC as part of the model fitting process. We use these outputs to compare the normal and Student's t-based HME-PL models. Table 4 summarizefs the comparison results in terms of both DIC and NRMSE.

As shown in Table 4, both the DIC and NRMSE values were lower under the Student's t-distribution assumption compared to the normal distribution. These results suggest that the PL model based on the Student's t-distribution provides a better fit to the data and achieves greater predictive accuracy.

This improvement can be attributed to the robustness of the Student's t-distribution to outliers, which enables the model to more accurately capture the underlying structure of the dose-response relationship while mitigating the influence of extreme observations. Based on this comparison, we conclude that the HME-PL model with the Student's t-distribution is the most appropriate model for estimating the cadmium dose-response relationship and adopt it as the final model in this study.

5.3 BMD Estimation Based on the Student's t HME-PL Model

In this section, we estimate the BMD, a reference point for cadmium risk assessment, using the Student's t-based HME-PL model, which demonstrated superior model fit in the previous analysis. Although the HME structure involves two components, we compute the BMD using the posterior samples assigned to the major component the one with the majority of data points—based on posterior clustering results. Specifically, we utilize the 176 samples classified into the major component under the Student's t assumption.

The BMD estimation follows the *extra risk* approach described in Section 2.2, in which BMD is defined as the dose corresponding to a specified BMR, typically set at 5% or 10%. In our implementation, we extracted 20,000 posterior samples from the fitted HME-PL model using `rjags`, then applied Equation (5) to calculate 20,000 posterior BMD samples. From these samples, we obtained the posterior mean and the lower bound of the 95% credible interval (BMDL), which together provide reference values for cadmium risk.

A threshold value (c^*), representing the critical response level indicating toxic or adverse effects, is also required for BMD estimation. Based on prior toxicological studies (EFSA); Gamo et al. (2006), we adopted a clinical cut-off value of 1000 μg/g creatinine. For further sensitivity analysis, an alternative threshold of 300 μg/g creatinine was also considered. Table 5 summarizes the estimated BMD and BMDL values for each subgroup (gender and age: $<$50 and $\geq$50), based on both cut-off levels and BMR values (5% and 10%).

The results in Table 5 indicate that, for both BMR levels (5% and 10%), older adults ($\geq$50 years) had consistently lower BMD and BMDL values compared to younger adults ($<$50 years). Similarly, females exhibited lower benchmark values than males across both age groups. This suggests that older females are the most sensitive subgroup to cadmium exposure, highlighting the need for stricter exposure control and monitoring for this population. These findings align with prior toxicological evidence and reinforce the necessity for age- and sex-specific risk management strategies.

5.4 Comparison of Model Fit With and Without the HME Structure

To assess the utility of incorporating a latent allocation variable Z_i and mixture distribution structure via the HME model, we compare the performance of HME-based models with their non-mixture counterparts. Specifically, we fit a standard piecewise linear (Standard-PL) model assuming either a Normal or a Student's t distribution, without introducing any mixture components or latent class indicators.

Posterior summary statistics for the standard PL models under Normal and Student's t assumptions are reported in Table 7, while convergence diagnostics based on the Gelman-Rubin statistic are summarized in Table 6. As shown, all parameters exhibit Gelman-Rubin values very close to 1, indicating

Table 3. Posterior summary statistics for the HME-PL model parameters, comparing results under the Normal and Student's t-distribution assumptions.

Parameter (Normal)	Mean	Standard Error	Parameter (Student's t)	Mean	Standard Error
r_a (Age)	2.8852	0.3956	r_a (Age)	0.2340	0.0995
r_m (Male)	-0.1908	0.2049	r_m (Male)	0.0118	0.1005
r_f (Female)	-0.2415	0.2022	r_f (Female)	-0.0651	0.0951
α_1 (Background)	1.9819	0.4810	α_1 (Background)	3.7573	0.3690
δ_1 (breakpoint1)	1.2545	0.1084	δ_1 (breakpoint1)	1.9202	0.1001
δ_2 (breakpoint2)	27.9862	18.4512	δ_2 (breakpoint2)	26.3445	18.8010
β_1 (Slope1)	0.0676	0.0599	β_1 (Slope1)	0.3406	0.1225
β_2 (Slope2)	1.7712	0.2581	β_2 (Slope2)	4.3913	1.0792
β_3 (Slope3)	0.2669	0.0679	β_3 (Slope3)	0.7569	1.6290
β_4 (Slope4)	24.6955	18.7945	β_4 (Slope4)	24.4152	18.6976
τ_{st} (Study Random effect)	1.7325	0.5448	τ_{st} (Study Random effect)	2.1715	0.5529
α_0 (Gate intercept)	-3.2797	0.8002	α_0 (Gate intercept)	4.7183	13.5327
β_0 (Gate Effect)	30.4911	17.6171	β_0 (Gate Effect)	3.8459	11.5803

Table 4. Model comparison results of the HME-PL model under Normal versus Student's t-distribution assumptions using DIC and NRMSE.

	HME-PL Model	
	Normal	Student's t
DIC	809.0	520.4
NRMSE	1.000	0.854

good convergence across chains after 100,000 burn-in iterations. Additionally, trace plots confirm that posterior samples were appropriately mixed, with no apparent trends or cyclic patterns.

Based on the posterior summary statistics in Table 7, the dose-response curves for the Standard-PL models were estimated under both Normal and Student's t distributional assumptions. As described in Section 2.1, the dose-response curve is derived from the posterior predictive mean response function μ_{ij} defined in Equation (7). For each of the 20,000 MCMC posterior samples, we computed μ_{ij} and visualized the resulting predictive dose-response relationships for both distributional assumptions.

Figure 1 in Section 3.1 displays the predictive dose-response curves estimated under both the Normal and Student's t assumptions. In both cases, the slope parameter β_2 after the breakpoint tends to be underestimated, indicating a limited ability to capture the increasing trend in the post-breakpoint region. This suggests that the Standard-PL model is insufficient in explaining the nonlinear behavior of the observed data, particularly when compared to the HME-PL model visualized in Figure 3.

To quantify model fit more formally, we compared models using the DIC and NRMSE. Table 8 presents DIC and NRMSE values for both the Standard-PL and HME-PL models under Normal and Student's t assumptions.

As shown in Table 8, the Student's t assumption consistently outperformed the Normal assumption in terms of both DIC and NRMSE. Moreover, incorporating the HME structure led to even greater improvements in fit, particularly under the Student's t distribution, which achieved the lowest DIC and NRMSE values overall. These results demonstrate that the HME-PL model provides a more flexible and robust fit to the data, particularly in the presence of potential outliers.

While the HME model relies on major component membership to estimate BMD values—thus excluding minor component observations—this exclusion can be justified when such observations represent heterogeneous or anomalous behavior that could otherwise bias the overall dose-response estimation. Therefore, the use of the HME-PL model with a Student's t distribution is both statistically and practically advantageous for robust dose-response modeling and benchmark dose estimation.

6 Conclusion

This study presents a robust Bayesian meta-analytic framework for evaluating the dose-response relationship of cadmium exposure by integrating recent data beyond previous studies such as Committee et al. (2022) and Lee et al. (2013). Based on a Bayesian hierarchical model, we modeled the nonlinear association between UCd and urinary β2MG using a piecewise linear (PL) structure. Unlike previous approaches, the proposed model explicitly accounts for data heterogeneity and outliers arising from the expansion of recent datasets by incorporating a Hierarchical Mixture of Experts (HME) framework. Furthermore, to improve robustness against outliers, we employed the Student's t-distribution as a prior, instead of the conventional normal distribution.

The model fitting results demonstrate that the HME-PL model with a Student's t-distribution outperformed the normal-based model in terms of Deviance Information Criterion (DIC) and Normalized Root Mean Square Error (NRMSE), and better captured the nonlinear characteristics observed in the empirical data. Using the optimal model, we estimated cadmium RPs through the BMD approach and found that risk levels varied significantly across age and sex groups. In particular, the BMD and BMDL values estimated for individuals aged 50 and older were consistently

Table 5. Estimates of BMD and BMDL from the Student's t HME-PL model using a cut-off of 1000 μg/g creatinine based on the major component

Cut-off: 1000 μg/g creatinine				
	Male, age < 50	Male, age ≥ 50	Female, age < 50	Female, age ≥ 50
BMD5 (BMDL5)	7.2967 (6.8926)	7.0753 (6.6875)	7.2727 (6.6875)	7.0500 (6.6804)
BMD10 (BMDL10)	7.9084 (7.5293)	7.6145 (7.1641)	7.8749 (7.5372)	7.5732 (7.2267)

Table 6. Comparison of Gelman-Rubin statistics (GR-stats) for the standard PL model under Normal versus Student's t-distribution assumptions

Normal	GR-stats	Student's t	GR-stats
r_a (Age)	1.00	r_a (Age)	1.00
r_m (Male)	1.00	r_m (Male)	1.01
r_f (Female)	1.00	r_f (Female)	1.01
α_1 (Background)	1.00	α_1 (Background)	1.03
δ (Breakpoint)	1.00	δ (Breakpoint)	1.01
β_1 (Slope1)	1.00	β_1 (Slope1)	1.03
β_2 (Slope2)	1.00	β_2 (Slope2)	1.00
τ_{st} (Study RE)*	1.00	τ_{st} (Study RE)*	1.00

* Study RE: Study Random Effect.

lower, indicating the need for stricter cadmium exposure control in this demographic. These findings align with earlier conclusions drawn by EFSA (2009) and others(EFSA; Lee et al. 2013), while also offering updated insights based on recent data.

The advantages of our approach can be summarized as follows. First, by incorporating data from 1983 to 2022, we captured temporal variations in cadmium exposure effects and a broader range of heterogeneity across study populations. Unlike previous studies that relied solely on older data, our approach allows for more relevant and timely estimation of cadmium reference levels using recent observations. Second, we addressed challenges associated with outliers and atypical (non-mainstream) data resulting from dataset expansion by implementing the HME structure and a shifted Student's t-distribution. This provided greater robustness compared to normality assumptions, as evidenced by superior model fit metrics such as DIC and NRMSE.

Nonetheless, there are limitations to this study. The covariates used were restricted to age and sex, and we did not account for spatiotemporal variations that may influence the observed relationships between UCd and β2MG. Future research could improve estimation accuracy by incorporating a wider set of covariates and employing dynamic models that reflect both spatial and temporal patterns in cadmium exposure data.

In conclusion, this study demonstrates that the proposed HME-PL model within a Bayesian meta-analysis framework can effectively capture the nonlinear dose-response characteristics of cadmium exposure and provide robust estimates in the presence of outliers. By incorporating recent data, the model enables more accurate estimation of cadmium reference levels and offers a meaningful methodological contribution to the assessment of heavy metal toxicity and the establishment of health-based safety standards.

Acknowledgements

This research was supported by grant (22191MFDS071) from the Ministry of Food and Drug Safety Korea. Seongil Jo was also supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Korea government(MSIT) (RS-2023-00209229) and INHA UNIVERSITY research grant.

References

Baralić K, Javorac D, Antonijević E, Buha-Đorđević A, Ćurčić M, Đukić Ćosić D, Bulat Z and Antonijević B (2020) Relevance and evaluation of the benchmark dose in toxicology. *Arhiv za farmaciju* 70(3): 130–141. DOI: 10.5937/ARHFARM2003130B. URL https://doi.org/10.5937/ARHFARM2003130B.

Billoir E, Delhaye H, Clément B, Delignette-Muller ML and Charles S (2011) Bayesian modelling of daphnid responses to time-varying cadmium exposure in laboratory aquatic microcosms. *Ecotoxicology and environmental safety* 74(4): 693–702.

Bishop CM and Svensén M (2012) Bayesian hierarchical mixtures of experts. *arXiv preprint arXiv:1212.2447* .

Canale A, Durante D and Dunson DB (2018) Convex mixture regression for quantitative risk assessment. *Biometrics* 74(4): 1331–1340. DOI:10.1111/biom.12917. URL https://doi.org/10.1111/biom.12917.

Chen F, Lin H, Zhang Y, Zhang Y and Chen S (2024) Investigating how blood cadmium levels influence cardiovascular health scores across sexes and dose responses. *Frontiers in Public Health* 12: 1427905.

Committee ES, Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen KH, More S, Mortensen A, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Silano V, Solecki R, Turck D, Aerts M, Bodin L, Davis A, Edler L, Gundert-Remy U, Sand S, Slob W, Bottex B, Abrahantes JC, Marques DC, Kass G and Schlatter JR (2017) Update: use of the benchmark dose approach in risk assessment. *EFSA Journal* 15(1): e04658. DOI:https://doi.org/10.2903/j.efsa.2017.4658. URL https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2017.4658.

Committee ES, More SJ, Bampidis V, Benford D, Bragard C, Halldorsson TI, Hernández-Jerez AF, Bennekou SH, Koutsoumanis K, Lambré C et al. (2022) Guidance on the use of the benchmark dose approach in risk assessment. *Efsa Journal* 20(10): e07584.

Crump KS (1984) A new method for determining allowable daily intakes. *Fundamental and Applied Toxicology* 4(5): 854–871. DOI:https://doi.org/10.1016/0272-0590(84)90107-6. URL https://www.sciencedirect.com/science/article/pii/0272059084901076.

Table 7. Posterior summary statistics for the Standard-PL model parameters, comparing results under the Normal and Student's t-distribution assumptions.

Parameter (Normal)	Mean	Standard Error	Parameter (Student's t)	Mean	Standard Error
r_a (Age)	0.3236	0.13317	r_a (Age)	0.28017	0.06205
r_m (Male)	-0.2119	0.20807	r_m (Male)	-0.05206	0.15065
r_f (Female)	-0.3270	0.19258	r_f (Female)	-0.15100	0.15331
α_1 (Background)	3.2224	0.34072	α_1 (Background)	4.09834	0.25142
δ (Breakpoint)	1.8464	0.28578	δ (Breakpoint)	1.50671	0.13044
β_1 (Slope1)	0.5161	0.08115	β_1 (Slope1)	0.20829	0.06559
β_2 (Slope2)	3.1942	1.59342	β_2 (Slope2)	1.65780	0.26785
τ_{st} (Study Random Effect)	2.9789	0.86172	τ_{st} (Study Random Effect)	2.15676	0.54091

Table 8. Model comparison results of the Standard-PL and HME-PL models under Normal versus Student's t-distribution assumptions using DIC and NRMSE.

	Standard-PL		HME-PL	
	Normal	Student's t	Normal	Student's t
DIC	1093.0	815.9	809.0	520.4
NRMSE	1.000	0.899	1.000	0.854

Dahl DB (2006) Model-based clustering for expression data via a dirichlet process mixture model. *Bayesian inference for gene expression and proteomics* 4: 201–218.

Du Y, Chen Y, Cao A, Pu Y, Zhang K, Ai S and Dang Y (2024) Estimation of urinary cadmium benchmark dose thresholds for preschool children in a cadmium-polluted area based on bayesian model averaging. *Environmental Geochemistry and Health* 46(7): 253.

(EFSA) EFSA (2009a) Meta-analysis of dose-effect relationship of cadmium for benchmark dose evaluation. *EFSA Journal* 7(3): 254r. DOI:https://doi.org/10.2903/j.efsa.2009.254r. URL https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2009.254r.

(EFSA) EFSA (2009b) Meta-analysis of dose-effect relationship of cadmium for benchmark dose evaluation. *EFSA Journal* 7(3): 254r. DOI:https://doi.org/10.2903/j.efsa.2009.254r. URL https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2009.254r.

EFSA Panel on Contaminants in the Food Chain (CONTAM) (2011) Statement on tolerable weekly intake for cadmium. *EFSA Journal* 9(2): 1975. DOI:https://doi.org/10.2903/j.efsa.2011.1975. URL https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2011.1975.

Filippini T, Torres D, Lopes C, Carvalho C, Moreira P, Naska A, Kasdagli MI, Malavolti M, Orsini N and Vinceti M (2020) Cadmium exposure and risk of breast cancer: a dose-response meta-analysis of cohort studies. *Environment International* 142: 105879.

Gallagher CM and Meliker JR (2010) Blood and urine cadmium, blood pressure, and hypertension: a systematic review and meta-analysis. *Environmental health perspectives* 118(12): 1676–1684.

Gamo M, Ono K and Nakanishi J (2006) Meta-analysis for deriving age- and gender-specific dose–response relationships between urinary cadmium concentration and 2-microglobulinuria under environmental exposure. *Environmental Research* 101(1): 104–112. DOI:https://doi.org/10.1016/j.envres.2005.09.004.

Gelman A, Carlin JB, Stern HS and Rubin DB (1995) *Bayesian data analysis*. Chapman and Hall/CRC.

Lee M, Choi T, Kim J and Woo HD (2013) Bayesian analysis of dose-effect relationship of cadmium for benchmark dose evaluation. *The Korean Journal of Applied Statistics* 26(3): 453–470.

Lee M, Kim M and Jang J (2022) A study on the introduction of benchmark dose for improving the utility of inhalation toxicity test results. Technical report, Korea Occupational Safety and Health Agency.

Moran KR, Dunson D, Wheeler MW and Herring AH (2021) Bayesian joint modeling of chemical structure and dose response curves. *The annals of applied statistics* 15(3): 1405.

National Evidence-based Healthcare Collaborating Agency (NECA) (2023) Bayesian meta-analysis methodology. Technical report, National Evidence-based Healthcare Collaborating Agency (NECA). URL https://www.neca.re.kr/lay1/bbs/S1T11C102/F/39/view.do?article_seq=5331&cpage=1&rows=10&condition=A.TITLE&keyword=&show=&cat=.

Nordberg M and Nordberg GF (2022) Metallothionein and cadmium toxicology—historical review and commentary. *Biomolecules* 12(3): 360.

Nweke CO and Ogbonna CJ (2017) Statistical models for biphasic dose-response relationships (hormesis) in toxicological studies. *Ecotoxicology and Environmental contamination* 12(1): 39–55.

Patel T, Telesca D, Low-Kam C, Ji Z, Zhang H, Xia T, Zinc JI and Nel AE (2014) Relating nanoparticle properties to biological outcomes in exposure escalation experiments. *Environmetrics* 25(1): 57–68. DOI:10.1002/env.2246. URL https://doi.org/10.1002/env.2246.

Prankel S, Nixon R and Phillips C (2004) Meta-analysis of feeding trials investigating cadmium accumulation in the livers and kidneys of sheep. *Environmental Research* 94(2): 171–183.

Prozialeck WC and Edwards JR (2012) Mechanisms of cadmium-induced proximal tubule injury: new insights with implications for biomonitoring and therapeutic interventions. *The Journal of pharmacology and experimental therapeutics* 343(1): 2–12.

Pruvost-Couvreur M, Le Bizec B, Béchaux C and Rivière G (2020) A method to assess lifetime dietary risk: Example of cadmium exposure. *Food and Chemical Toxicology* 137: 111130.

Spiegelhalter DJ, Best NG, Carlin BP and Van Der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology)* 64(4): 583–639.

Stayner L, Smith R, Thun M, Schnorr T and Lemen R (1992) A dose-response analysis and quantitative assessment of lung cancer risk and occupational cadmium exposure. *Annals of epidemiology* 2(3): 177–194.

Steenland K and Deddens JA (2004) A practical guide to dose-response analyses and risk assessment in occupational epidemiology. *Epidemiology* 15(1): 63–70.

Woo HD, Chiu WA, Jo S and Kim J (2015) Benchmark dose for urinary cadmium based on a marker of renal dysfunction: a meta-analysis. *PLoS One* 10(5): e0126680.

Yang L, Wang J, Cheke RA and Tang S (2021) A universal delayed difference model fitting dose-response curves. *Dose-Response* 19(4): 15593258211062785.

Yoon H, Jeong K, Hwang M, Eom J, Jeong J, Kim H and Kwon S (2011) Study for establishment of health based guidance value on hazardous material. Final Report TRKO201200007066, National Institute of Food and Drug Safety Evaluation (NIFDS), Cheongju, Republic of Korea. NTIS Project ID 1475006100.

Zarn JA, Zürcher UA and Geiser HC (2020) Toxic responses induced at high doses may affect benchmark doses. *Dose-Response* 18(2): 1559325820919605. DOI:10.1177/1559325820919605. URL https://doi.org/10.1177/1559325820919605. PMID: 32341684.

A Sensitivity Analysis with an Alternative Cut-off

Table 9. Estimates of BMD and BMDL from the Student's t HME-PL model using a cut-off of 300 μg/g creatinine

	Male, age < 50	Male, age ≥ 50	Female, age < 50	Female, age ≥ 50
BMD5 (BMDL5)	6.1523 (0.8820)	1.7911 (0.4463)	5.2537 (0.9020)	1.7030 (0.4522)
BMD10 (BMDL10)	6.8964 (5.7371)	6.6156 (2.0791)	6.8754 (5.9972)	6.6093 (2.0587)

To evaluate the robustness of our findings, we conducted a sensitivity analysis by additionally estimating the BMD and BMDL values using a more conservative clinical cut-off value of 300 μg/g creatinine. The results are presented in Table 9.

As noted in the EFSA technical report (EFSA), the BMD and BMDL estimates obtained with the 300 μg/g creatinine cut-off under the PL model were substantially unstable. Compared with the values derived from the 1000 μg/g cut-off used in the main analysis, the estimates for most subgroups were considerably lower. While it is expected that a stricter (lower) definition of the critical effect level would naturally yield lower dose estimates for the same benchmark response, the large discrepancy observed between the BMD and BMDL suggests instability in the reference estimates. This instability is a typical artifact of the PL model, which produces a "broken" dose-effect curve due to the presence of a breakpoint (EFSA).

Although the estimates were unstable, the key pattern remained consistent despite the absolute differences between the BMD and BMDL values. Specifically, adults aged 50 years and older and females consistently exhibited lower BMD and BMDL estimates. This indicates that even under more conservative assumptions, older females are identified as the most sensitive subgroup. Such consistency across thresholds strengthens the primary conclusion of this study regarding the need for age- and sex-specific risk management for cadmium exposure.