

基于贝叶斯模型平均的基准剂量估计及其在白银市人群镉暴露评估中的应用

贺娜娜¹, 孙金芳¹, 何杰宇², 余小金¹

1. 东南大学公共卫生学院, 江苏 南京 210009

2. 昆明医科大学公共卫生学院, 云南 昆明 650500



DOI 10.11836/JEOM22055

摘要:

[背景] 最优模型法估计基准剂量(BMD)时未考虑模型选择的不确定性, 国内暂无使用贝叶斯模型平均(BMA)估计 BMD 的研究。

[目的] 将 BMA 应用于我国镉污染的暴露评估研究, 讨论 BMA 在基于剂量-反应模型的 BMD 估计中的作用, 为有害物质健康风险评估提供方法学支持。

[方法] 模拟研究基于甘肃省白银市某镉污染区数据获得的不同剂量-反应模型(Gamma、Log-logistic、Log-probit、Two-stage 和 Weibull 模型)的参数和中国 5 个镉污染区的尿镉范围, 模拟不同正确模型、不同剂量组数(5、8)、不同样本量(50、100、200)的模拟数据, 比较 BMA 和最优模型法的优劣。实例分析使用甘肃省白银市某镉污染区的镉暴露数据。模拟研究和实例分析均以尿镉为镉暴露指标, β_2 -微球蛋白异常率作为效应指标, 基准反应设定为 10%, 分别使用正确模型(模拟数据时使用的模型)、最优模型法[赤池信息准则(AIC)最小的模型]和 BMA 估计 BMD 及其 95%置信区间下限(BMDL); 比较不同方法估计的 BMD、BMDL 及相对偏差。

[结果] 对于模拟研究, 随着样本量或剂量组数增加, BMD 的第 5 百分位数和第 90 百分位数区间大致呈现变窄趋势; 当正确模型为单一模型时, BMA 估计的 BMD 相对偏差大于最优模型法; 当正确模型为等权重混合模型时, BMA 估计的 BMD 相对偏差小于最优模型法。基于实例数据, 最优模型为 Log-probit 模型(AIC=1814.46), 其次为 Log-logistic 模型(AIC=1814.57); Log-probit 模型、Log-logistic 模型和 BMA 估计的 BMD(BMDL)分别为 3.46(2.68)、3.16(2.33)和 2.92(2.07) $\mu\text{g}\cdot\text{g}^{-1}$ 。

[结论] 在正确模型已知时, 最优模型法依然是值得推荐的方法; 但在有害物质剂量-反应关系高度不确定或来源人群、暴露分组等不同时, BMA 与最优模型法相比, 理论上更充分地考虑了可能存在的多个备选模型, 能提供更为稳定的 BMD 和 BMDL 估计。

关键词: 基准剂量; 剂量-反应模型; 贝叶斯模型平均; 尿镉; 暴露评估

Benchmark dose estimation based on Bayesian model averaging and its application to evaluation of cadmium exposure of population in Baiyin HE Nana¹, SUN Jinfang¹, HE Jieyu², YU Xiaojin¹ (1. School of Public Health, Southeast University, Nanjing, Jiangsu 210009, China; 2. School of Public Health, Kunming Medical University, Kunming, Yunnan 650500, China)

Abstract:

[Background] The optimal model method for estimation of benchmark dose (BMD) does not consider the uncertainty of model selection. There is a lack of studies on using Bayesian model averaging (BMA) to estimate BMD.

[Objective] To apply BMA to the exposure assessment of cadmium pollution in China, discuss the role of BMA in estimating BMD based on dose-response models, and to provide methodological support for health risk assessment of hazardous substances.

[Methods] The parameters of five dose-response models (Gamma, Log-logistic, Log-probit, Two-stage, and Weibull models) estimated from the data from a cadmium-contaminated area in Baiyin City of Gansu Province and the urinary cadmium ranges in five cadmium-contaminated areas in China were used to simulate the data of varied correct models with different numbers of dosage groups (5 and 8) and different sample sizes (50, 100, and 200), then the performance of BMA and traditional optimal model were compared. The case analysis used the cadmium exposure data in Baiyin, Gansu Province. All analyses set urinary cadmium as the indicator of cadmium exposure, the abnormal rate of β_2 -microglobulin as the effect indicator, and the benchmark response to 10%. The correct model (the model used when simulating data), optimal model [the model with

基金项目

国家自然科学基金面上项目(81573159); 江苏省研究生科研与实践创新计划项目(SJCX20_0064); 中央高校基本科研业务费专项资金(3225002110D)

作者简介

贺娜娜(1997—), 女, 硕士生;
E-mail: 19834514119@163.com

通信作者

余小金, E-mail: xiaojinyu@seu.edu.cn

伦理审批 不需要

利益冲突 无申报

收稿日期 2022-02-18

录用日期 2022-08-18

文章编号 2095-9982(2022)09-1004-07

中图分类号 R12

文献标志码 A

引用

贺娜娜, 孙金芳, 何杰宇, 等. 基于贝叶斯模型平均的基准剂量估计及其在白银市人群镉暴露评估中的应用[J]. 环境与职业医学, 2022, 39(9): 1004-1010.

本文链接

www.jeom.org/article/cn/10.11836/JEOM22055

Funding

This study was funded.

Correspondence to

YU Xiaojin, E-mail: xiaojinyu@seu.edu.cn

Ethics approval Not required

Competing interests None declared

Received 2022-02-18

Accepted 2022-08-18

To cite

HE Nana, SUN Jinfang, HE Jieyu, et al. Benchmark dose estimation based on Bayesian model averaging and its application to evaluation of cadmium exposure of population in Baiyin[J]. Journal of Environmental and Occupational Medicine, 2022, 39(9): 1004-1010.

Link to this article

www.jeom.org/article/en/10.11836/JEOM22055

smallest Akaike information criterion (AIC)], and BMA were used to estimate BMD and lower confidence limit of benchmark dose (BMDL); the BMDs, BMDLs, and relative deviations from different methods were compared.

[Results] In the simulation study, with increasing sample size or the number of dosage groups, the intervals of the 5th percentile and the 90th percentile of BMD tended to be narrower; when the correct model was a single model, the relative deviation of BMD estimation by BMA was greater than that of the traditional optimal model; when the correct model was an equal weight mixed model, the relative deviation of BMD estimation by BMA was less than that by the traditional optimal model. For the data of cadmium-contaminated areas, the optimal model was a Log-probit model (AIC=1814.46), followed by a Log-logistic model (AIC=1814.57); the BMDs (BMDLs) estimated by the Log-probit model, the Log-logistic model, and BMA were 3.46 (2.68), 3.16 (2.33), and 2.92 (2.07) $\mu\text{g}\cdot\text{g}^{-1}$, respectively.

[Conclusion] The traditional optimal model is still recommended when the correct model is known. However, when the dose-response relationship of a hazardous substance is uncertain or with different sources or exposure grouping, compared with the traditional optimal model, BMA theoretically provides more stable estimation of BMD and BMDL by considering multiple possible alternative models.

Keywords: benchmark dose; dose-response model; Bayesian model averaging; urinary cadmium; exposure assessment

镉及其代谢产物在人体蓄积会产生毒作用^[1],因此,制定合理的生物接触限值有利于保障镉暴露者身体健康。基准剂量(benchmark dose, BMD)^[2]是基于剂量-反应模型,估计某种毒性反应增加到某一特定水平(一般为1%~10%)时所对应的剂量,特定的反应水平称为基准反应。BMD及其95%置信区间下限(lower confidence limit of benchmark dose, BMDL)已被国内外学者广泛应用于镉暴露的风险评估^[3-7]。关于镉暴露的BMD估计多采用最优模型法,即在多个不同函数关系的剂量-反应备选模型中,选择赤池信息准则(Akaike information criterion, AIC)最小的模型作为最优模型来计算BMD和BMDL^[8],如Yu等^[5]、Li等^[3]的研究基于不同数据选择了不同剂量-反应模型,模型选择的不确定性显而易见。事实上,单一模型难以表征生物体内复杂的反应,选择某个单一模型可能忽略了模型本身的不确定性,舍弃了其他模型提供的有用信息,从而降低BMD和BMDL估计的可信度^[9]。

贝叶斯模型平均(Bayesian model averaging, BMA)可以考虑所有备选模型信息及模型选择的不确定性,以模型的后验概率为权重,加权组合得到最终的模型估计^[10]。有国外学者研究报告基于BMA的BMD的估计^[11-13],而国内目前主要使用单一最优模型估计镉暴露BMD^[6, 7, 14-15]。因此,本研究将BMA应用于我国镉污染的暴露评估,通过模拟研究,探讨与传统BMD估计法相比,BMA法的优劣及其适用性,并应用于李浩等^[16]公开发表的汇总数据,估计当地尿镉的BMD,以期为有害物质的健康暴露评估提供科学有效的方法学支持。

1 资料与方法

1.1 资料来源

1.1.1 模拟数据 模拟数据的参数设置见表1,其中,剂量-反应关系曲线的参数来源于李浩等^[16]公开发表

的汇总数据的拟合结果,以此为参数的各模型称为“真模型”(为模拟数据实际使用的已知参数的模型);剂量设置参考中国5个镉污染区的尿镉范围(0.05~57.27 $\mu\text{g}\cdot\text{g}^{-1}$,以尿肌酐校正,余同)^[17]。数据模拟过程如下:(1)模拟尿镉浓度,范围为0~57.27 $\mu\text{g}\cdot\text{g}^{-1}$,按照剂量组数设置在每个剂量组内均匀抽取尿镉值($<0.025 \mu\text{g}\cdot\text{g}^{-1}$ 的剂量赋值为 $0.025 \mu\text{g}\cdot\text{g}^{-1}$);(2)根据剂量-反应关系计算各尿镉浓度下的反应率 p_i ;(3)根据反应率 p_i 从 $P = p_i$ 的伯努利分布中抽样获得反应数据(0/1);(4)模拟数据检查。对每组数据模拟1000次进行分析。模拟研究将反应数据为“1”的定义为 β_2 -微球蛋白异常,以 β_2 -微球蛋白异常率(反应数据为“1”的人数占该剂量组总人数的比例)为效应指标。

表1 模拟数据参数设置
Table 1 Parameter settings of simulated data

模拟数据参数 (Parameter of simulated data)	取值(Value)
真模型(True model)	
Gamma	$\gamma_0=0.0615, \beta_1=0.0428, \beta_2=0.0010$
Log-logistic	$\gamma_0=0.0884, \beta_1=-4.0105, \beta_2=1.5779$
Log-probit	$\gamma_0=0.0745, \beta_1=0.0313, \beta_2=1.2205$
Two-stage	$\gamma_0=0.0996, \beta_1=-2.4862, \beta_2=0.9711$
Weibull	$\gamma_0=0.0800, \beta_1=0.0833, \beta_2=1.3734$
等权重混合模型 (Equal weight mixed model)	各备选模型的等权重加权平均 (Equal weighted for all alternative models)
剂量组数 (Number of dosage groups)	
5	0.025~, 2.0~, 6.0~, 12.0~, 20.0~57.27
8	0.025~, 2.0~, 4.0~, 6.0~, 8.0~, 10.0~, 12.0~, 20.0~57.27
单剂量组样本量 (Sample size of single dosage)	50, 100, 200

[注] γ_0 、 β_1 、 β_2 为根据文献数据使用不同模型(Gamma、Log-logistic、Log-probit、Two-stage和Weibull模型)的拟合参数。

[Note] γ_0 , β_1 , and β_2 are fitting parameters for different models (Gamma, Log-logistic, Log-probit, Two-stage, and Weibull models) from literature data.

1.1.2 实例数据 本研究实例数据来源于已公开发表文献的汇总数据^[16], 研究对象为甘肃白银某镉污染区的 1685 名居民, 其中女性 915 人, 男性 770 人, 平均年龄 55.72 岁。按照《环境重金属污染健康监测技术指南(试行)》要求采集 150~200 mL 的中段晨尿, 采集后及时进行肌酐或比重测定, 弃去肌酐质量浓度 < 0.3 g·L⁻¹ 或 > 3 g·L⁻¹、比重 < 1.010 或 > 1.030 的尿样并重新采样, 当天完成尿肌酐测定。检测尿镉、尿肌酐和尿 β₂-微球蛋白, 各指标均用肌酐进行校正, 不同剂量组尿镉水平及 β₂-微球蛋白异常(> 17 μg·g⁻¹) 人数见表 2。效应指标为 β₂-微球蛋白异常率(β₂-微球蛋白异常人数占该剂量组总人数的比例)。

表 2 不同剂量组尿镉水平及 β₂-微球蛋白异常率
Table 2 Cadmium levels and abnormal rates of β₂-microglobulin in different dose groups

剂量组(Dose group)/(μg·g ⁻¹)	n	尿镉(Urinary cadmium)/(μg·g ⁻¹), 几何均值(Geometric mean)	尿β ₂ -微球蛋白异常(Abnormal urinary β ₂ -microglobulin)	
			n	%
0~	305	0.94	32	10.5
2~	576	3.35	106	18.4
5~	548	7.08	193	35.2
10~	256	15.41	156	60.9

1.2 BMA 法

BMA 法^[10]以备选模型的贝叶斯后验概率为权重, 得到模型的加权估计。本研究所用的剂量-反应模型包括 Gamma、Log-logistic、Log-probit、Two-stage 和 Weibull 模型共 5 个二分类结局模型, 各模型表达式及详细参数说明详见相关文献^[18]。

假设有 k 个备选模型 M₁, ..., M_k, Δ 为感兴趣的量(如 BMD), 则在给定数据 D 时 BMD 的后验分布为:

$$Pr(\Delta|D) = \sum_{i=1}^k Pr(\Delta|M_i, D) Pr(M_i|D)$$

(1)

其中 Pr(M_i|D) 是模型 M_i 的后验概率, 通过式(2) 计算:

$$Pr(M_i|D) = \frac{Pr(D|M_i) Pr(M_i)}{\sum_{t=1}^k Pr(D|M_t) Pr(M_t)}$$

(2)

式中 M_i 是备选模型中的第 i 个模型, Pr(D|M_i) 是模型 M_i 的边际似然概率。目前, 常用的估计贝叶斯后验概率的方法是基于马尔科夫蒙特卡洛^[19]模拟的估计, 通过 Gibbs 抽样实现此算法。

1.3 统计学分析

以尿镉为镉暴露标志物, β₂-微球蛋白异常率为效应指标, 分别拟合 Gamma、Log-logistic、Log-probit、Two-stage 和 Weibull 模型, 基准反应设定为 10%, 使用最优模型法和 BMA 法分别估计 BMD。BMDL 估计采用 bootstrap 法重复 1000 次获得; BMA 的 Gibbs 抽样通过 JAGS 4.2.0 实现, 设置 3 条平行链, 设置不同参数的初始值, 前 10000 次迭代结果作为退火, 保留 10000 次迭代结果用于后续分析。根据迭代轨迹图^[20]、Geweke 统计量^[21]、收缩因子以及蒙特卡洛标准误^[22]诊断迭代是否收敛。对于模拟数据, 分别估计正确模型(即直接采用模拟数据时使用的模型)、最优模型(基于模型数据选择的模型)和 BMA 下的 BMD, 并计算其与“真模型”下 BMD 的相对偏差(1000 个估计的 BMD 与“真模型”BMD 之差占“真模型”BMD 比例的均数)。对于实例数据, 比较不同方法的拟合结果, 所有分析均使用 R 3.3.2 进行。

2 结果

2.1 模拟研究

采用正确模型、最优模型以及 BMA 法得到的 BMD 中位数(第 5 百分位数和第 90 百分位数)以及相对偏差结果见表 3 和图 1。随着剂量组数或单剂量组样本含量的增加, 第 5 百分位数至第 90 百分位数之间的区间均呈现变窄的趋势。

表 3 三种方法得到的模拟数据 BMD 估计
Table 3 BMD estimation of simulated data by three methods

真模型 (True model)	剂量组数 (Number of dosage groups)	BMD估计方法 (Method for estimation of BMD)	样本含量(Sample size), M(P ₅ , P ₉₀)		
			50	100	200
			单位 (Unit): μg·g ⁻¹		
Two-stage	5	正确模型(Correct model)	2.25(1.56, 3.66)	2.22(1.64, 2.99)	2.19(1.74, 2.70)
		最优模型(Optimal model)	3.07(1.43, 5.25)	2.80(1.66, 4.60)	2.62(1.80, 4.11)
		BMA	3.30(1.79, 5.18)	3.24(1.99, 4.54)	3.18(2.28, 4.02)
	8	正确模型(Correct model)	2.31(1.69, 3.37)	2.27(1.75, 2.95)	2.27(1.85, 2.66)
		最优模型(Optimal model)	3.15(1.65, 4.95)	2.99(1.79, 4.48)	2.80(1.95, 4.14)
		BMA	3.37(1.96, 4.81)	3.33(2.22, 4.32)	3.29(2.47, 3.98)

续表 3

真模型 (True model)	剂量组数 (Number of dosage groups)	BMD估计方法 (Method for estimation of BMD)	样本含量(Sample size), $M(P_5, P_{90})$		
			50	100	200
Log-logistic	5	正确模型(Correct model)	3.12(1.49, 4.98)	3.09(1.79, 4.26)	3.05(2.12, 3.89)
		最优模型(Optimal model)	2.86(1.16, 4.74)	2.87(1.22, 4.20)	2.96(1.45, 3.89)
		BMA	2.52(1.41, 4.18)	2.46(1.54, 3.49)	2.41(1.70, 3.16)
	8	正确模型(Correct model)	3.23(1.75, 4.57)	3.21(2.11, 4.17)	3.15(2.35, 3.75)
		最优模型(Optimal model)	3.03(1.30, 4.47)	3.11(1.57, 4.18)	3.20(1.76, 3.89)
		BMA	2.57(1.56, 3.76)	2.57(1.76, 3.43)	2.57(1.93, 3.09)
Weibull	5	正确模型(Correct model)	2.61(1.18, 4.22)	2.55(1.38, 3.60)	2.52(1.65, 3.27)
		最优模型(Optimal model)	3.11(1.35, 5.18)	2.90(1.54, 4.62)	2.80(1.78, 4.27)
		BMA	3.14(1.67, 4.99)	3.07(1.87, 4.21)	3.01(2.16, 3.81)
	8	正确模型(Correct model)	2.71(1.46, 3.94)	2.65(1.68, 3.59)	2.64(1.92, 3.22)
		最优模型(Optimal model)	3.17(1.65, 4.87)	3.14(1.78, 4.47)	2.98(1.90, 4.30)
		BMA	3.20(1.91, 4.55)	3.16(2.14, 4.14)	3.13(2.36, 3.78)
Log-probit	5	正确模型(Correct model)	3.43(1.72, 5.10)	3.41(2.09, 4.47)	3.36(2.46, 4.12)
		最优模型(Optimal model)	2.91(1.13, 4.75)	2.91(1.22, 4.22)	3.04(1.48, 3.94)
		BMA	2.57(1.38, 4.14)	2.53(1.57, 3.51)	2.47(1.76, 3.20)
	8	正确模型(Correct model)	3.49(2.06, 4.84)	3.47(2.42, 4.42)	3.47(2.69, 4.03)
		最优模型(Optimal model)	3.05(1.34, 4.48)	3.14(1.66, 4.21)	3.28(1.78, 3.92)
		BMA	2.60(1.57, 3.78)	2.58(1.77, 3.49)	2.61(1.96, 3.12)
Gamma	5	正确模型(Correct model)	2.78(1.02, 4.66)	2.73(1.27, 3.95)	2.66(1.58, 3.59)
		最优模型(Optimal model)	3.11(1.34, 5.11)	2.93(1.50, 4.53)	2.90(1.75, 4.25)
		BMA	3.05(1.66, 4.81)	3.03(1.84, 4.09)	2.94(2.09, 3.78)
	8	正确模型(Correct model)	2.86(1.34, 4.35)	2.84(1.61, 3.97)	2.81(1.89, 3.51)
		最优模型(Optimal model)	3.20(1.61, 4.92)	3.22(1.77, 4.48)	3.10(1.90, 4.29)
		BMA	3.12(1.89, 4.44)	3.08(2.10, 4.05)	3.08(2.32, 3.67)
等权重混合模型 (Equal weight mixed model)	5	正确模型(Correct model)	2.89(1.57, 4.56)	2.82(1.75, 3.84)	2.78(1.99, 3.57)
		最优模型(Optimal model)	3.05(1.26, 4.99)	2.92(1.42, 4.43)	3.03(1.65, 4.21)
		BMA	2.91(1.57, 4.56)	2.82(1.73, 3.90)	2.78(1.99, 3.60)
	8	正确模型(Correct model)	2.92(1.78, 4.24)	2.89(1.97, 3.86)	2.88(2.17, 3.45)
		最优模型(Optimal model)	3.16(1.41, 4.77)	3.22(1.66, 4.36)	3.27(1.83, 4.18)
		BMA	2.93(1.76, 4.23)	2.90(1.96, 3.88)	2.91(2.18, 3.51)

[注] BMD: 基准剂量; BMA: 贝叶斯模型平均。

[Note] BMD: Benchmark dose; BMA: Bayesian model averaging.

当正确模型为单一的 Two-stage、Log-probit、Log-logistic 和 Weibull 模型时, BMA 法估计的相对偏差高于其他两种方法; 3 种方法估计的尿镉 BMD 的相对偏差大小关系均为正确模型 < 最优模型 < BMA。

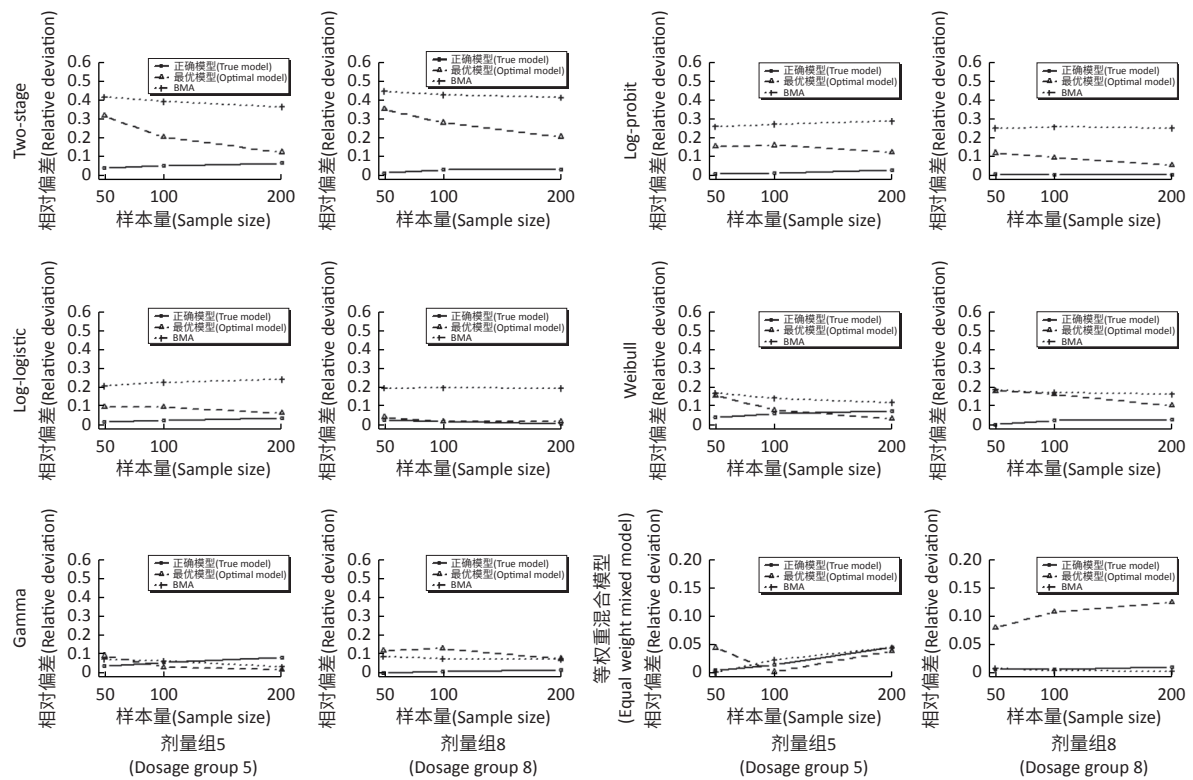
当正确模型为 Gamma 模型时, 不同剂量组数对结果有影响; 当剂量组数为 5 时, 3 种方法相对偏差无法比较; 当剂量组数为 8 时, BMA 法估计的尿镉 BMD 的相对偏差小于最优模型, 大于正确模型。

值得注意的是, 当正确模型为等权重混合模型时, 不同剂量组数对结果有影响; 当剂量组数为 5 时, 3 种方法

难分伯仲; 但当剂量组数增加时, BMA 法估计的尿镉 BMD 相对偏差小于其他两种方法; 3 种方法估计的尿镉 BMD 的相对偏差大小关系为 BMA < 正确模型 < 最优模型。

2.2 实例研究

研究结果显示, Gamma、Log-logistic、Log-probit、Two-stage 和 Weibull 模型拟合优度检验的 P 值均大于 0.10。其中, AIC 最小的模型为 Log-probit 模型, 估计的 BMD(BMDL)为 $3.46(2.68) \mu\text{g}\cdot\text{g}^{-1}$; 其次为 Log-logistic 模型, 估计的 BMD(BMDL)为 $3.16(2.33) \mu\text{g}\cdot\text{g}^{-1}$; BMA 估计的 BMD(BMDL)为 $2.92(2.07) \mu\text{g}\cdot\text{g}^{-1}$ (见表 4)。



[注] BMD: 基准剂量; BMA: 贝叶斯模型平均。
[Note] BMD: Benchmark dose; BMA: Bayesian model averaging.

图 1 模拟研究中不同方法估计的 BMD 相对偏差
Figure 1 The relative deviations of BMD estimated by different methods in simulation studies

表 4 不同模型计算尿镉 BMD 结果
Table 4 BMD of urinary cadmium calculated by different models

模型(Model)	尿镉(Urinary cadmium)/(μg·g ⁻¹)		AIC	P
	BMD	BMDL		
Gamma	2.87	1.84	1814.96	0.47
Log-logistic	3.16	2.33	1814.57	0.72
Log-probit	3.46	2.68	1814.46	0.93
Two-stage	2.34	1.84	1815.46	0.31
Weibull	2.71	1.80	1815.10	0.42
BMA	2.92	2.07	-	-

[注] BMD: 基准剂量; BMDL: 基准剂量下侧限值; AIC: 赤池信息准则; BMA: 贝叶斯模型平均。
[Note] BMD: Benchmark dose; BMDL: Lower confidence limit of benchmark dose; AIC: Akaike information criterion; BMA: Bayesian model averaging.

3 讨论

BMA 能够考虑模型选择的不确定性,在基于剂量-反应关系模型估计 BMD 时,可以整合多个备选模型,符合正确模型普遍未知的情形,从而提供更为稳定的结果。本研究将 BMA 应用于我国镉污染的暴露评估,通过模拟研究探讨 BMA 在不同正确模型、不同剂量组数以及不同样本含量下的效果,并应用于甘肃省白银市某镉污染区的实际人群数据,对比不同 BMD 估

计策略对结果的影响。
本研究结果显示,当正确模型为单一模型时,BMA 法没有优势;当正确模型为等权重混合模型时,BMA 法估计 BMD 的相对偏差小于最优模型法;说明在 BMD 和 BMDL 估计时,正确模型为单一确定模型时,使用最优模型法依然是值得推荐的方法。但现实条件下,不同人群来源、不同暴露方式或不同剂量组数的数据可能具有不同的数据特征,正确模型未知或者有上述多种成分存在时,BMA 法估计的结果优于最优模型法。
有研究指出,模型平均基于模型贝叶斯后验概率的加权测量最终得到的预测可能优于任何单一模型^[23]。当污染物潜在的生物学机制尚不清楚时,为解决模型选择不确定的问题,采用 BMA 法估计 BMD 是欧盟食品安全管理局科学委员会推荐的首选方法^[24]。美国环保署于 2018 年在其 BMD 估计的官方软件 BMDS v3.0 中纳入二分类结局 BMA 估计方法^[25],2020 年纳入连续结局 BMD 的贝叶斯估计,但构建连续结局的模型尚未经过美国环保署的正式审查和批准^[26]。本研究探讨了暴露评估中 BMD 估计的模型平均法的性能,可为有害物质暴露评估的方法学选择提供依据。

目前国内有研究以尿 β_2 -微球蛋白异常率作为效应指标得到尿镉 BMD, 如张洁莹^[15]、Wang 等^[4]采用 Log-logistic 模型, 钱海雷等^[27]、Li 等^[3]、Jin 等^[28]采用 Logistic 模型, 罗磊等^[14]和 Yu 等^[5]采用线性模型, Hong 等^[29]采用 Probit 模型, Ke 等^[17]采用 Log-logistic 模型或者 Gamma 模型, Wu 等^[30]采用 Hill 模型或者指数模型, Lv 等^[31]采用 Log-probit 模型, 而本次实例研究结果显示 Log-probit 模型为最优模型, 表明不同来源的数据选择的剂量-反应模型不同。同时, 上述研究所得 BMD 和 BMDL 变化范围分别为 $1.16\sim 25.39\ \mu\text{g}\cdot\text{g}^{-1}$ 和 $1.00\sim 14.42\ \mu\text{g}\cdot\text{g}^{-1}$, 变化范围较大, 显示了 BMD 估计的不确定性, 也从实际角度说明了模型平均的必要性。我国职业卫生接触限值规定镉暴露量小于 $5\ \mu\text{g}\cdot\text{g}^{-1}$, 对于大多数人来讲, 低于职业人群的镉暴露水平可能不会对机体造成损害。2009 年欧洲食品安全局污染物专家小组规定镉每周允许摄入量为 $2.5\ \mu\text{g}\cdot\text{kg}^{-1}$ (以体重计)^[32]。2010 年, 联合国粮农组织和世界卫生组织下的食品添加剂专家委员会提出暂定镉的每月允许摄入量为 $25\ \mu\text{g}\cdot\text{kg}^{-1}$ (以体重计), 即每周允许摄入量为 $5.8\ \mu\text{g}\cdot\text{kg}^{-1}$ (以体重计)^[32]。本研究实例数据分析时采用 BMA 法估计的 BMD 和 BMDL 更低, 源于其考虑了模型的不确定性, 提供了更保守的参考值, 这与目前镉暴露风险研究保护大多数人的研究理念一致。

当正确模型为单一的 Two-stage、Log-probit、Log-logistic 和 Weibull 模型时, 剂量分组对正确模型、最优模型以及 BMA 法估计 BMD 的优劣次序没有影响; 仅在 Gamma 模型和等权重混合模型时, 可能受剂量分组影响, 但由于本研究仅考虑了两种可能的剂量组数, 剂量组数对结果的影响并不直观, 更多的剂量组数是本研究可以进一步探讨的因素。

本研究仍存在以下局限性: 第一, 备选模型只考虑了剂量-反应模型软件中常用的 5 种模型, 实际应用时应结合毒理学知识选择备选模型并讨论备选模型的确定策略; 第二, 本研究未考虑一些重要的协变量, 如年龄、性别、吸烟状况等, 可能会提高模型拟合的不确定性; 第三, 本研究所用实例数据来自镉污染区, 数据可能代表性不足, 未来扩大研究人群的进一步研究对健康指导值的建立是必要的。

综上所述, 在正确模型已知时使用最优模型法依然是值得推荐的方法; 但在正确模型未知或者有多种成分存在时, BMA 法不失为一种可选择的方法。目前 BMA 法没有得到广泛应用, 本研究结果可为模型平均法在 BMD 估计方面的应用推广提供有价值的帮助。

参考文献

- [1] 苗亚琼, 林清. 广西土壤重金属镉污染及对人体健康的危害[J]. 环境与可持续发展, 2016, 41(5): 171-173.
MIAO YQ, LIN Q. Pollution caused by heavy metal cadmium to the soil in Guangxi and its harm to human health[J]. Environ Sustain Dev, 2016, 41(5): 171-173.
- [2] CRUMP KS. A new method for determining allowable daily intakes[J]. Fundam Appl Toxicol, 1984, 4(5): 854-871.
- [3] LI Y, WANG H, YU J, et al. An assessment of sensitivity biomarkers for urinary cadmium burden[J]. BMC Nephrol, 2020, 21(1): 385.
- [4] WANG X, WANG Y, FENG L, et al. Application of the benchmark dose (BMD) method to identify thresholds of cadmium-induced renal effects in non-polluted areas in China[J]. PLoS One, 2016, 11(8): e0161240.
- [5] YU T, HOU W, HOU Q, et al. Safe utilization and zoning on natural selenium-rich land resources: a case study of the typical area in Enshi County, China[J]. Environ Geochem Health, 2020, 42(9): 2803-2818.
- [6] 林立丰, 潘尚霞, 王凤岩, 等. 基准剂量法在突发性镉污染饮用水应急阈值研究中的应用[J]. 环境与健康杂志, 2018, 35(4): 295-298.
LIN LF, PAN SX, WANG FY, et al. Application of benchmark dose method in establishment of emergency health advisory value for accident of drinking water contamination by cadmium[J]. J Environ Health, 2018, 35(4): 295-298.
- [7] 王艳莉, 李成国, 祝白春, 等. 2014—2018年南京市居民蔬菜和食用菌中铅和镉膳食暴露评估[J]. 卫生研究, 2022, 51(1): 118-123, 127.
WANG YL, LI CG, ZHU BC, et al. Risk assessment of lead and cadmium exposure of Nanjing residents through vegetables and edible fungus from 2014 to 2018[J]. J Hyg Res, 2022, 51(1): 118-123, 127.
- [8] AKAIKE H. A new look at the statistical model identification[J]. IEEE Trans Autom Control, 1974, 19(6): 716-723.
- [9] DRAPER D. Assessment and propagation of model uncertainty[J]. J Roy Stat Soc, 1995, 57(1): 45-70.
- [10] HOETING JA, MADIGAN D, RAFTERY AE, et al. Bayesian model averaging: a tutorial[J]. Stat Sci, 1999, 14(4): 382-401.
- [11] FANG Q, PIEGORSCH WW, SIMMONS SJ, et al. Bayesian model-averaged benchmark dose analysis via reparameterized quantal-response models[J]. Biometrics, 2015, 71(4): 1168-1175.
- [12] UCHE UI, NAIDENKO OV. Development of health-based exposure limits for radiofrequency radiation from wireless devices using a benchmark dose approach[J]. Environ Health, 2021, 20(1): 84.
- [13] WANG YH, WU CF, LIU CC, et al. A probabilistic approach for benchmark dose of melamine exposure for a marker of early renal dysfunction in patients with calcium urolithiasis[J]. Ecotoxicol Environ Saf, 2020, 200: 110741.
- [14] 罗磊, 李娜, 罗振华, 等. 职业性镉作业工人尿中酸性鞘磷脂酶活力和 β_2 微球蛋白含量的比较研究[J]. 工业卫生与职业病, 2021, 47(2): 89-92, 99.
LUO L, LI N, LUO ZH, et al. Comparative study on acid sphingomyelinase activity and β_2 microglobulin content in urine of the cadmium exposure workers[J]. Ind Health Occup Dis, 2021, 47(2): 89-92, 99.
- [15] 张洁莹. 广东省某地区镉污染损害人体健康调查研究[D]. 北京: 北京交通大学, 2016.
ZHANG JY. Effects of cadmium pollution on human health in a region of Guangdong province[D]. Beijing: Beijing Jiaotong University, 2016.
- [16] 李浩, 柯岫, 贾闻婧, 等. 白银市某镉污染区人群尿镉基准剂量研究[J]. 公共卫生与预防医学, 2014, 25(6): 13-16.
LI H, KE S, JIA WJ, et al. Calculation of biological exposure limit of cadmium

- by using benchmark dose method[J]. J Public Health Prev Med, 2014, 25(6): 13-16.
- [17] KE S, CHENG XY, ZHANG JY, et al. Estimation of the benchmark dose of urinary cadmium as the reference level for renal dysfunction: a large sample study in five cadmium polluted areas in China[J]. BMC Public Health, 2015, 15: 656.
- [18] FILIPSSON AF, SAND S, NILSSON J, et al. The benchmark dose method—Review of available models, and recommendations for application in health risk assessment[J]. Crit Rev Toxicol, 2003, 33(5): 505-542.
- [19] CARLIN BP, CHIB S. Bayesian model choice via Markov chain Monte Carlo methods[J]. J Roy Stat Soc, 1995, 57(3): 473-484.
- [20] 茆诗松, 汤银才. 贝叶斯统计[M]. 2版. 北京: 中国统计出版社, 2012.
- MAO S S, TANG Y C. Bayesian statistics[M]. 2nd ed. Beijing: China Statistics Press, 2012.
- [21] 刘文忠, 张沅. 贝叶斯方差组分估计中Gibbs链的收敛诊断[C]//中国农学会. 中国青年农业科学学术年报. 北京: 中国农业出版社, 2002.
- LIU W Z, ZHANG Y. Diagnosis of convergence of Gibbs chains in Bayesian variance component estimation[C]//Chinese Agricultural Society. Annual Report of China Youth Agricultural Science. Beijing: China Agricultural Press, 2002.
- [22] KASS RE, CARLIN BP, GELMAN A, et al. Markov chain Monte Carlo in practice: a roundtable discussion[J]. Am Stat, 1998, 52(2): 93-100.
- [23] MADIGAN D, RAFTERY AE. Model selection and accounting for model uncertainty in graphical models using Occam's window[J]. J Am Stat Assoc, 1994, 89(428): 1535-1546.
- [24] EFSA Scientific Committee, HARDY A, BENFORD D, et al. Update: use of the benchmark dose approach in risk assessment[J]. EFSA J, 2017, 15(1): e04658.
- [25] U. S. EPA. Bayesian models and model averaging software (external review draft, 2018)[EB/OL]. [2022-01-05]. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NCEA&dirEntryId=343986.
- [26] U. S. EPA. About benchmark dose software (BMDs)[EB/OL]. [2022-01-05]. <https://www.epa.gov/bmds/about-benchmark-dose-software-bmds>.
- [27] 钱海雷, 金泰虞, 黄波, 等. 环境镉接触人群尿金属硫蛋白排泄与镉致骨损伤效应[J]. 卫生研究, 2007, 36(1): 1-5.
- QIAN H L, JIN T Y, HUANG B, et al. Excretion of urinary metallothionein and osteal damage induced by cadmium in an environmentally cadmium exposed population[J]. J Hyg Res, 2007, 36(1): 1-5.
- [28] JIN T, WU X, TANG Y, et al. Environmental epidemiological study and estimation of benchmark dose for renal dysfunction in a cadmium-polluted area in China[J]. Biometals, 2004, 17(5): 525-530.
- [29] HONG F, JIN T, ZHANG A. Risk assessment on renal dysfunction caused by co-exposure to arsenic and cadmium using benchmark dose calculation in a Chinese population[J]. Biometals, 2004, 17(5): 573-580.
- [30] WU X, WEI S, WEI Y, et al. The reference dose for subchronic exposure of pigs to cadmium leading to early renal damage by benchmark dose method[J]. Toxicol Sci, 2012, 128(2): 524-531.
- [31] LV Y, WANG P, HUANG R, et al. Cadmium exposure and osteoporosis: a population-based study and benchmark dose estimation in southern China[J]. J Bone Miner Res, 2017, 32(10): 1990-2000.
- [32] EFSA. Scientific opinion of the panel on contaminants in the food chain on a request from the European commission on cadmium in food[J]. EFSA J, 2009, 980: 1-139.

(英文编辑: 汪源; 责任编辑: 汪源)

· 告知栏 ·

关于假冒《环境与职业医学》编辑和诈骗邮件的声明

近期发现有个人或机构通过电子邮件和微信冒用《环境与职业医学》编辑身份, 以组稿或确认稿件信息名义, 请求添加作者的微信; 以稿件转让为由, 广发邮件进行诈骗, 涉及学术不端行为。此类行为严重侵犯本刊的学术声誉, 我们对此表示深恶痛绝。本刊严正声明, 《环境与职业医学》杂志是一本优秀的学术期刊, 绝无此类买卖稿件等有损学术声誉的行为。希望广大作者和专家不要上当受骗。

另外, 我们也特此郑重声明: 本刊编辑不会以私人微信添加作者帐号, 所有编辑均使用以 scdc.sh.cn 为后缀的邮箱, 请作者务必注意甄别。

本刊网址: www.jeom.org; 微信公众号: 环境与职业医学; E-mail: jeom@scdc.sh.cn; 电话: 021-62084529。

提醒: 不要轻易点开网址链接, 不要扫码或添加来路不明的微信号, 不要向假冒个人或机构转账或发送稿件。如有疑问, 请及时与编辑部沟通确认。

《环境与职业医学》编辑部

2022 年 9 月 25 日