

## 基准剂量估计 R 软件包 bmd 的改进与应用\*

袁 苗<sup>1</sup> 彭雯洁<sup>1</sup> 王青青<sup>1</sup> 崔 靖<sup>1</sup> 张 婷<sup>2</sup> 周小林<sup>2,3,Δ</sup> 余红梅<sup>1,4,Δ</sup>

**【摘要】** 目的 本研究旨在构建一个更加便捷的 R 扩展包简化基准剂量估计(benchmark dose, BMD)的计算和操作过程。方法 使用 R 4.1.3 对原有 R 软件包 bmd 进行修改和整合,构建一个 R 扩展包。采用研究中封装完成的 R 扩展包将微囊藻毒素 LR(microcystin-LR, MCLR)的数据与 18 种剂量-反应模型拟合,根据 Akaike 信息标准(Akaike information criterion, AIC)最小原则,得到最佳拟合模型下 MCLR 的 BMD 和基准剂量下限(benchmark dose lower confidence limit, BMDL)。结果 封装后的 R 扩展包将原有 bmd 包的运算过程进行简化。MCLR 的基准剂量估计中 2 个参数的 Weibull type 2 模型 AIC 值最小,在该模型下, MCLR 的 BMD 和 BMDL 分别为 1.39 mg/kg 和 0.49 mg/kg。结论 封装后的 R 扩展包简化了 BMD 的计算和操作过程,可以同时得到 18 种剂量-反应模型的 AIC、BMD 和 BMDL,直观选择 AIC 最小的模型相对应的 BMD 及 BMDL。

**【关键词】** 基准剂量估计 剂量-反应 微囊藻毒素

**【中图分类号】** R195

**【文献标识码】** A

**DOI** 10.11783/j.issn.1002-3674.2024.05.029

健康风险评估包括危害鉴定、暴露评估、剂量-反应关系和风险度表征四个步骤。剂量-反应关系是其关键一步,首先要确定一个安全剂量的起始值(point of departure, POD),再由该起始值外推到与人类健康的相关性。最初通常使用未观察到有害作用剂量(no observed adverse effect level, NOAEL)或观察到最低水平有害作用剂量(lowest observed adverse effect level, LOAEL)作为 POD<sup>[1]</sup>,但其存在许多缺点<sup>[2]</sup>,比如 NOAEL 忽略了剂量-反应关系曲线的斜率,并且高度依赖剂量的选择和样本量等。为此,1984 年 Crump<sup>[3]</sup>提出将数学模型与所收集到的数据拟合,得到相应的剂量-反应曲线,并根据剂量-反应曲线计算基准剂量(benchmark dose, BMD),将其称为基准剂量法,该方法在一定程度上弥补了 NOAEL 的缺点,对剂量依赖性小,充分利用剂量-反应信息<sup>[4]</sup>,基准剂量估计法得出的 BMD 统计学可信区间下限(benchmark dose lower confidence limit, BMDL)可作为 POD 的一般表示方法。1995 年,美国环境保护署(Environmental Protection Agency, EPA)正式提出用 BMD 法代替 NOAEL<sup>[5]</sup>。对于试验数据,基准剂量风险评估方法可以对其毒性效应进行量化,并将其作为推导健康风险评估的基础<sup>[6]</sup>。

开源软件 R 已经发展为一个强大的统计软件,其中包含的各种 R 包有不同的功能,2020 年 Jensen<sup>[7]</sup>等

人提出可以利用 R 软件中的 bmd 包和 drc 包进行基准剂量估计。BMD 的计算过程较复杂,虽然 R 软件为其提供了一个环境,且有相应的 bmd 包进行模型拟合计算,但一次 bmd 包的应用只能计算一次拟合,若进行多个模型拟合,操作繁琐且耗时,占计算机运行内存。本研究的目的是在原有 bmd 包的基础上进行修改和整合,构建一个 R 扩展包简化 BMD 的计算和操作过程,并结合实例说明该 R 扩展包在基准剂量估计中的应用,以期为今后计算 BMD 和 BMDL 提供参考。

### 原理与方法

#### 1. BMD 概述

BMD 是指在反应中能够引起预定变化的剂量,该过程中的反应变化称为基准反应(benchmark response, BMR),BMD 的统计学可信区间下限值是 BMDL<sup>[8]</sup>,同时可得到其统计学可信区间上限值(benchmark dose upper confidence limit, BMDU)。BMDL 被用于代替 NOAEL 作为健康危险度评价的起始点,从而推导人类安全暴露水平的基准剂量。BMD 的计算需要以下四个步骤:

- (1) 确定进行剂量-反应数据的类型
- (2) 确定 BMR

BMR 的选择直接关系到 BMD 的计算结果,小的 BMR 会增加 BMDL 对剂量-反应模型的依赖性,不同的数据类型, BMR 的选择不同,对于计数资料的数据建议使用 10% 的 BMR<sup>[9]</sup>。而在连续数据中,通常使用 5% 或 10% 的 BMR。对于生殖、发育毒性的数据,常用 5% 的 BMR,但是对于流行病学研究和人体相关研究,由于其反应灵敏度高,通常使用 1% 的 BMR<sup>[10]</sup>。

- (3) 拟合并选择最佳的剂量-反应模型

模型的选择依赖于数据类型,在一定程度上也会

\* 基金项目:生态环境部重点地区环境与健康调查、监测和风险评估(21111011101EHSM);国家环境保护辐射环境与健康重点实验室平台开放基金(YP21020203)

1.山西医科大学公共卫生学院卫生统计教研室(030001)

2.中国辐射防护研究院

3.国家环境保护辐射环境与健康重点实验室

4.重大疾病风险评估山西省重点实验室

Δ通信作者:余红梅, E-mail: yu@sxmu.edu.cn;周小林, E-mail: xiaolin0824@sohu.com

影响 BMD 和 BMDL 的数值, BMD 法需要将数据和多个模型进行拟合, 为了能选择到合适的模型, 可以根据 Akaike 信息标准 (Akaike Information Criterion, AIC)

选择合适的模型, 即选择 AIC 最小的模型, 本研究共纳入了 18 种剂量-反应模型, 具体见表 1。

(4) 计算 BMD 和 BMDL

表 1 18 种剂量-反应模型

模型	简写	参数个数	模型方程
log-logistic	LL.2	2	$f(x, (b, e)) = (1 + (x/e)^b)^{-1}$
	LL.3	3	$f(x, (b, d, e)) = d (1 + (x/e)^b)^{-1}$
	LL.3u	3	
	LL.4	4	$f(x, (b, c, d, e)) = c + (d - c) (1 + (x/e)^b)^{-1}$
	LL.5	5	$f(x, (b, c, d, e, f)) = c + (d - c) (1 + (x/e)^b)^{-f}$
Weibull type 1	W1.2	2	$f(x, (b, e)) = \exp(-(x/e)^b)$
	W1.3	3	$f(x, (b, d, e)) = d \exp(-(x/e)^b)$
	W1.4	4	$f(x, (b, c, d, e)) = c + (d - c) \exp(-(x/e)^b)$
Weibull type 2	W2.2	2	$f(x, (b, e)) = 1 - \exp(-(x/e)^b)$
	W2.3	3	$f(x, (b, d, e)) = d (1 - \exp(-(x/e)^b))$
	W2.4	4	$f(x, (b, c, d, e)) = c + (d - c) (1 - \exp(-(x/e)^b))$
Asymptotic regression	AR.3	3	$f(x, (c, d, e)) = c + (d - c) (1 - \exp(-x/e))$
Michaelis-Menten	MM.2	2	$f(x, (d, e)) = dx (x + e)^{-1}$
log-normal	LN.2	2	$f(x, (b, e)) = \varphi(b(\log(x) - \log(e)))$
	LN.3	3	$f(x, (b, d, e)) = d\varphi(b(\log(x) - \log(e)))$
	LN.4	4	$f(x, (b, c, d, e)) = c + (d - c)\varphi(b(\log(x) - \log(e)))$
Exponential decay	EXD.2	2	$f(x, (d, e)) = d \exp(-x/e)$
	EXD.3	3	$f(x, (c, d, e)) = c + (d - c) \exp(-x/e)$

2. 统计方法

使用 R 4.1.3 对原有 R 软件包 bmd 进行修改和整合, 为了推广 BMD 函数包, 本次研究封装成独立的扩展包可以将计量资料或计数资料与 18 种剂量-反应模型拟合。R 程序的实现包括以下几个步骤:

#1. 加载包

```
Install.package("drc")
```

```
Install.package("bmd")
```

#2. 加载数据文件

```
data <- read.csv("Path location of the data set.csv")
```

#3. 通过定义 BMD 函数读取模型、设置数据结果框, 并拟合模型, 最后输出结果。

```
result <- BMD(data, bmr = 0.05, type_data = 0, type_distributed = 0)
```

```
BMD <- function(data, bmr, type_data, type_distributed)
```

BMD 扩展包的流程图见图 1, 扩展包流程中数据集的固定形式为: 计数资料 (Response ~ Total ~ Dose), 计量资料 (Response ~ Dose)。18 种模型依次命名为: LL.2、LL.3、LL.3u、LL.4、LL.5、W1.2、W1.3、W1.4、W2.2、W2.3、W2.4、AR.3、MM.2、LN.2、LN.3、LN.4、EXD.2、EXD.3。

附加条件包括: 风险反应 (bmr)、数据类型 (type\_data)、分布情况 (type\_distributed)。附加条件中的 bmr 一般取值为 0.05, 用于动物数据向人群数据的外推。计数资料的 type\_distributed 为 binomial, 计量资料的 type\_distributed 包括 Normal 和 Poisson 两种。

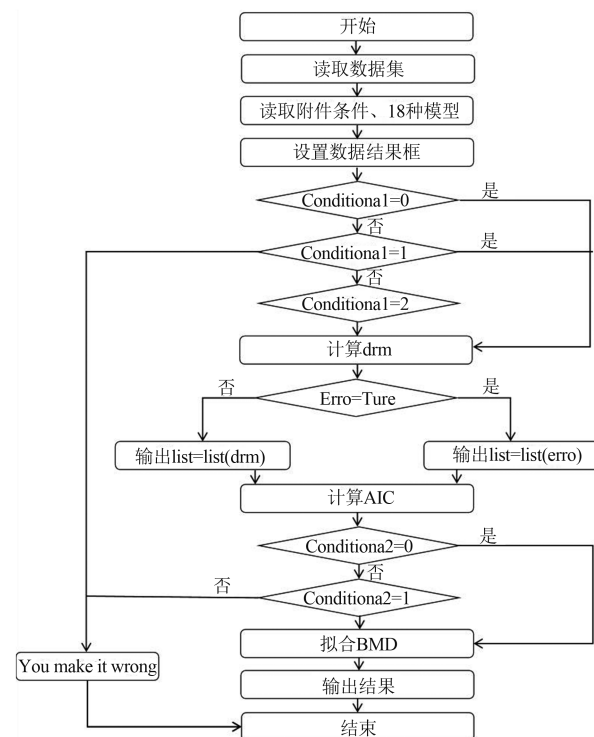


图 1 BMD 扩展包流程图

流程中依次循环 18 个模型, 使用 suppressWarnings() 函数, 消除不符合模型的警告提示, 赋值 Nan, 使用 tryCatch() 函数对拟合错误的模型赋值 Erro。结果报告每个模型的 AIC、BMD 的相关值。

实例分析

1. 数据来源

研究数据来源于美国风险信息系统 (Integrated

Risk Information System, IRIS) (<https://www.epa.gov/iris>) 的公共数据库,其数据内容为毒理学审查报告中的实际毒理学数据。该数据集的研究对象为 BALB/c 老鼠,染毒物质为不同剂量下的 10 种微囊藻毒素(microcystin,MC),观察结果为老鼠的肝脏损害,根据 Chernoff<sup>[11-12]</sup> 等人的基础研究,微囊藻毒素 LR (microcystin-LR, MCLR) 主要毒性在肝脏,因此本研究选用不同剂量的 MCLR 作为研究指标,效应指标为老鼠肝脏是否有损害,为二分类数据。具体数据见表 2。

表 2 MCLR 的染毒剂量及肝脏损害情况

剂量(mg/kg)	N	反应数
3	12	5
5	12	11
7	12	12
9	16	16
11	13	13

### 2.MCLR 的基准剂量估计

利用封装好的 R 扩展包可得到 18 种模型的 AIC、BMDU、BMD 和 BMDL,由表 3 可知 18 种模型均可计算出 AIC 值,其中 2 个参数的 Weibull type 2 的模型 AIC 最小,值为 8.89。说明在这 18 种模型中,该模型的拟合效果最好,在该模型下,MCLR 的 BMD、BMDU 和 BMDL 分别为 1.39 mg/kg、2.30 mg/kg 和 0.49 mg/kg,MCLR 剂量-反应关系图见图 2。

表 3 MCLR 的基准剂量估计结果

模型	AIC	BMDU (mg/kg)	BMD (mg/kg)	BMDL (mg/kg)
LL.2	9.24	2.63	1.96	1.30
LL.3	10.87	NaN	1.60	NaN
LL.3u	10.87	12.20	4.08	-
LL.4	12.87	NaN	3.20	NaN
LL.5	14.87	NaN	3.19	-
W1.2	9.34	2.87	2.39	1.91
W1.3	10.88	2.67	2.24	1.82
W1.4	12.87	455.09	3.45	-
W2.2	8.89	2.30	1.39	0.49
W2.3	13.90	NaN	2.42	NaN
W2.4	12.87	90.14	3.02	-
AR.3	10.86	7.52	0.17	-
MM.2	9.70	1.86	0.27	-
LN.2	9.02	2.66	1.96	1.26
LN.3	10.87	2.35	1.90	1.44
LN.4	12.87	204.33	2.90	-
EXD.2	35.04	84.07	30.41	-
EXD.3	10.87	NaN	13.56	NaN

注:NaN 代表模型未给出估计值,-代表值为负数。

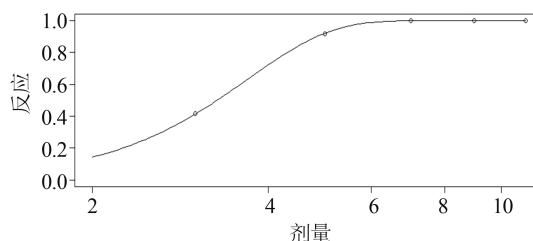


图 2 MCLR 的剂量-反应关系图

## 讨论

Chernoff<sup>[11-12]</sup> 等人使用该数据进行 10 种微囊藻毒素同系物对小鼠毒性及其剂量-反应关系的研究,结果表明 MCLR 的主要损害在肝脏,将 3 mg/kg 定为 MCLR 的 NOAEL,而在本研究中,2 个参数的 Weibull type 2 模型的拟合效果最好,在该模型下,MCLR 的 BMD 和 BMDL 分别为 1.39 mg/kg 和 0.49 mg/kg,经过基准剂量估计得到 MCLR 的 BMDL 值远远小于其 NOAEL,相比于 NOAEL, BMD 法可以得到更加安全的参考剂量,发现更多的风险人群。

BMD 法同样可以用于人群数据,直接计算人群接触某些有害物质的 BMD 及 BMDL,Zou<sup>[13]</sup> 等人对焦炉排放物进行基准剂量估计,将暴露于焦炉排放物的工人作为研究对象,所导致的氧化损伤作为效应指标,分别计算暴露于焦炉排放物的男性和女性的 BMD 及 BMDL。但由于各种因素,一些研究会考虑采用动物接触数据外推到人群数据的方法。有研究表明氧化亚氮对生殖健康有不良影响,由于人群可能会暴露在其他影响生殖健康的因素下,存在数据不充分的问题,Menon<sup>[14]</sup> 等人在关注氧化亚氮职业接触推荐限值问题时使用动物作为研究对象,计算氧化亚氮的 BMD 及 BMDL,并由此计算出一个职业接触的限值。在动物数据向人类数据外推的过程中,需要考虑可变性和不确定性的问题,其中重要的是种群差异和种内差异。

BMD 的计算过程较复杂,尤其在模型拟合过程中,拟合 18 种剂量-反应模型所花费的时间较长,速度较慢,因此需要一个简便的方法计算 BMD 及 BMDL。对原有的 bmd 包进行修改和整合,构建一个 R 函数扩展包,简化 BMD 的计算和操作过程。本研究所采用的 R 函数包涵盖了计量资料和计数资料,应用简便,只需将数据导入,即可得到 18 种模型的 AIC、BMD 及 BMDL,可以直观选择 AIC 最小的模型相对应的 BMD 及 BMDL,大大简化了应用 R 软件 bmd 包的计算和操作过程。

(下转第 774 页)