

食品健康影響評価における
ベンチマークドーズ法の活用に関する指針

令和元年（2019年）10月
（令和6年（2024年）4月改正）

食品安全委員会

第 1	目的	3
第 2	用語の説明	3
第 3	食品安全委員会が行う食品健康影響評価における BMD 法の活用	5
	I. 動物試験で得られた用量反応データへの適用	5
	1. BMD 法を適用する用量反応データの収集、選択	5
	(1) 動物試験結果の収集	5
	(2) データセットの選択	5
	(3) 原データの入手	6
	(4) データセットの統合	6
	2. BMR の設定	6
	(1) 二値データ	6
	(2) 連続値データ	7
	3. 用量反応モデリングと結果の評価及び POD の決定	8
	(1) 用量反応モデリング	8
	(2) 用量反応モデリング結果の評価	8
	(3) POD の決定	9
	4. 結果の文書化	10
	II. 疫学研究で得られた用量反応データへの適用	11
	1. BMD 法を適用する用量反応データの収集、選択	11
	(1) 疫学研究結果の収集	11
	(2) データセットの選択	11
	(3) 原データの入手	12
	(4) データセットの統合	12
	2. BMR の設定	12
	(1) 二値データ	12
	(2) 連続値データ	12
	3. 用量反応モデリングと結果の評価及び POD の決定	13
	(1) 用量反応モデリング	13
	(2) 用量反応モデリング結果の評価	14
	(3) POD の決定	15
	4. 結果の文書化	15
第 4	指針の見直し	16
別添	ベイズ推定を活用した BMD 法を使用する際の考え方について	17

改正履歴

発行時期（版）	主な改正概要
2019年10月（初版）	「動物試験で得られた用量反応データへの適用」として指針を作成
2022年6月（第2版）	ベンチマークドーズ法に係る研究事業の成果を踏まえて修正。
2023年9月（第3版）	「ベイズ推定を活用したBMD法を使用する際の考え方について」を追加。
2024年4月（第4版）	「疫学研究で得られた用量反応データへの適用」の項を追加

第1 目的

本指針は、化学物質に係る食品健康影響評価において、ベンチマークドーズ法を活用して適切な POD¹を求める場合の基本的な考え方、手順等を整理することにより、食品安全委員会が、同法を活用するに当たってのさらなる一貫性及び透明性の確保に資することを目的とする。

第2 用語の説明

本指針における用語の説明は以下のとおりとする。

1. BMR (Benchmark Response)

バックグラウンド反応からの、リスク評価上意義のある反応レベル²。化学物質のばく露量と、当該物質によりもたらされる影響の発生の頻度又は量との関係（用量反応関係）全体に数学的関数（数理モデル）を当てはめて得られた関数（用量反応曲線）において、BMD 関連指標を求めるために使用する。

2. BMD (Benchmark Dose)

一定の BMR をもたらす化学物質のばく露量。

3. ベンチマークドーズ法 (BMD 法)

化学物質のばく露量と、当該物質によりもたらされる影響の発生の頻度又は量との関係（用量反応関係）は、特定の数学的関数（数理モデル）に従うという仮定の下、BMD 及びその信頼区間を算出する方法。

4. BMDL (Benchmark Dose Lower Confidence Limit)

BMD の信頼区間の下限値（信頼下限値）。通常、BMD の 90%信頼区間（片側信頼区間とすれば 95%信頼区間）の下限値が BMDL として用いられ

¹ Point of Departure。各種の動物試験や疫学研究から得られた用量反応評価の結果から得られる値で、通常、無毒性量 (No-Observed-Adverse-Effect Level : NOAEL) や BMDL を指す。健康影響に基づく指標値 (Health-Based Guidance Value : HBGV) を設定する際や、ばく露マージン (Margin of Exposure : MOE) を算出する際等に用いられる。国際的には、Reference Point ということもある。出典：食品の安全性に関する用語集 http://www.fsc.go.jp/yougoshu/kensaku_hyouka.html (食品安全委員会)

² 「バックグラウンド反応からの、リスク評価上意義のある反応レベル」の考え方は、エンドポイントの特性あるいは研究のデザインに応じて、以下のような例がある。

- ① 「見込まれる最大変化量」に対する「あるばく露量における変化量」
- ② 「無ばく露又は最小ばく露の状態におけるエンドポイントの値（ベースライン）」に対する「あるばく露量における変化量」

る。

5. AIC (Akaike Information Criterion)

以下の式で定義され、異なる数理モデル間で、モデルの複雑さと測定データとの適合度とのバランスを比較するための指標。

$$AIC = -2 \log(L) + 2k$$

$\log(L)$: モデルの最大対数尤度

k : モデルに含まれる推定パラメータ³数

AIC の値が小さくなるモデル、すなわち少ないパラメータ数で適切な当てはめを実現するモデルが好ましいとされている。

6. 二値データ

化学物質へのばく露により各個体で観察された影響を、死亡又は生存、腫瘍発生の有無等、起こりうる帰結が特定の 2 つのいずれかであるとして分類した非連続データ (分類データ) のこと。

7. 順序カテゴリーデータ

化学物質へのばく露により各個体で観察された影響をカテゴリーに分類した非連続データのうち、尿糖検査値 (−、±、+、++他)、がんのステージ等、影響の程度等の順序はあるが量的概念はないカテゴリーに分類したデータのこと。

8. 連続値データ

化学物質へのばく露による影響として、臓器重量、酵素濃度等の連続量について測定したデータのこと。なお、発生した腫瘍の数等の離散値データについても連続値データに準じて扱う場合もある。

9. Restriction

パラメータ最適化のプロセスの中でパラメータが取る値の範囲に制限を設けること。BMD 法においては、数理モデルから得られる用量反応曲線が、生物学的に説明できない用量反応曲線とならないように、数理モデルに含まれるパラメータが取る値の範囲に制限を設けることとなる。

10. ベイズ推定

ベイズ統計学に基づくパラメータ推定のこと。

³ $y=f(x, a)$ が、 a の種々の値に対して応答変数 y と説明変数 x の関係を表すと考えるとき、 a をパラメータという (変数の数が多い場合も同様)。

第3 食品安全委員会が行う食品健康影響評価におけるBMD法の活用

化学物質の食品健康影響評価において BMD 法を活用する際は、本指針に基づき BMD 及びその信頼区間の算出を行う。動物試験と疫学研究とはデータの特性が異なるため、BMD 法を適用する際には、それぞれのデータの特性に基づいて、データの収集、選択、用量反応モデリング等を検討する。(I.及びII.参照)

また、BMD 等の算出に必要な生物学的及び統計学的な判断は、各分野の専門家の意見に従うものとし、BMD 等の算出に当たってベイズ推定を利用する場合には、別添の「ベイズ推定を活用した BMD 法を使用する際の考え方について」を参照する。

なお、本指針に拠らない考え方又は手法を採用する場合は、BMD 等の算出結果の報告時にその内容及び採用した理由を併記する。

I. 動物試験で得られた用量反応データへの適用

1. BMD 法を適用する用量反応データの収集、選択

(1) 動物試験結果の収集

評価対象とする化学物質（以下「評価対象物質」という。）に関して、その用量と特定の反応との関係を観察した動物試験結果のうち、利用可能なものを全て収集する。

その際、利用可能な用量と反応との関係を観察したデータ（以下「用量反応データ」という。）の種類及び数によっては、専門家の意見に基づき、対象とする毒性試験を先に選択した上でエンドポイントを選択する、又は対象とするエンドポイントを先に選択した上で毒性試験を選択するという手順を取り得る。

(2) データセットの選択

毒性試験とエンドポイントの組合せからなる各データセットのうち、用量反応関係が成立しているものについて、以下の条件に該当するか否かを専門家の関与の下で確認する。

- ① 動物の種及び数、投与方法、投与量等の試験設計が適切であること。
- ② エンドポイントが毒性学的に意義のある影響であること。なお、疫学研究の結果も踏まえ、ヒトへの外挿性についても考慮する。
- ③ データセットが、BMD 法を適用する上で必要となる以下の情報を含む

こと。なお、用量反応モデリング⁴の際に共変量⁵となりうる変数（生殖・発生毒性における同腹児数等）が想定されるデータセットについては、それらの情報を含むことも確認する。

- a 二値データ：各用量群における試験動物数及び反応を示した動物数。
- b 順序カテゴリーデータ：各用量群における試験動物数、反応を示した動物数及びその反応のグレード⁶。
- c 連続値データ：各個体データ。これが利用可能でない場合は、各用量群における試験動物数、反応量の代表値（平均値又は中央値）及び反応量のばらつきに関する指標（標準偏差又は四分位範囲）

（3）原データの入手

原データ（個体データ）を活用することで、BMRに対応する用量域における毒性について、より詳細な解析が可能となる。そのため、可能な限り、それぞれ個体データまでさかのぼり、活用することが望ましい。

（4）データセットの統合

同質性が期待できるデータセットが複数存在する場合、標本数を増やし、より精緻な検討を行うため、専門家の関与の下、当該データセットを統合して解析することが可能である。この場合、用量反応関係が不明確になること等を防ぐため、各データセットにおける用量の範囲、動物種、標本サイズ、データ測定方法等の試験設計の違いを把握し、事前に生物学的及び統計学的な観点からその妥当性を判断する必要がある。

2. BMRの設定

（1）二値データ

- ① 通常の動物試験における試験動物数から得られる統計学的精度に加え、化学物質間又はエンドポイント間における比較及び NOAEL 法を用いた

⁴ 化学物質のばく露量と当該物質によりもたらされる影響との間の用量反応関係に、数理モデルを当てはめることで用量反応曲線を得ること。

⁵ 主要な関心の対象でないが、応答変数に影響を与えると考えられ、解析のために付け加えられる変数。

⁶ 順序づけされた観察された影響（病理所見等）の程度のこと。なお、グレード別の発生頻度が記録されている順序カテゴリーデータを、二値データとして取り扱うことは、慎重に検討するべきである。具体的には、対照群における病理所見の発生頻度を考慮しつつ、病理所見の種類や、そのグレードの付け方及び発生頻度に基づき、専門家の関与の下で毒性学的意義があると判断されたグレード以上の病理所見の有無を二値データとして取り扱うことを検討する。

毒性評価との連続性⁷も考慮し、BMR として過剰リスク⁸10%を用いた BMD 関連指標は基本的に算出する。

- ② 当該過剰リスク 10%は、BMR の値として、あらゆる種類のデータセットに最適な値になるとは限らない。各 BMR の値は、生物学的な意義付けのほか、用いる用量反応データの特性を考慮しつつ、専門家の関与の下で設定する。例えば、過剰リスク 10%が、最小用量で観察された反応よりもかなり小さく、最小用量よりも大幅に低用量側の推計を伴う場合は、より大きい値が妥当な場合もある。また、生殖発生毒性のうち次世代への影響に関する用量反応データを用いる場合は、過剰リスク 10%より小さい値が妥当な場合もある。

(2) 連続値データ

- ① 生物学的に意味のある反応量の変化を、科学的根拠に基づいて設定できる場合は、その変化量を BMR として設定する。
- ② 連続値の用量反応データは多様な分布を示すことが想定されるため、生物学的に意味のある反応量の変化が不明な場合は、用量反応データごとに専門家の関与の下で BMR を設定する。
- ③ 科学的に合理的と思われるカットオフ値を設定できる場合は、反応量の分布を仮定した上で、カットオフ値を超える反応量を示す個体割合が、対照群に比べて、あらかじめ定めた増加分 (BMR) だけ増える投与量を BMD として定義した上で、BMD 関連指標を算出するハイブリッド法⁹等を取り得る。
- ④ 科学的に合理的と思われるカットオフ値を設定でき、かつ、個体データが利用可能である場合は、情報量の減少を伴う点に留意する必要があるが、カットオフ値に基づいて連続値データを二値データに変換した上で BMR を設定し BMD 関連指標を算出することが有用である場合もある。

⁷ S. Sand *et al.*, A Signal-to-Noise Crossover Dose as the Point of Departure for Health Risk Assessment. *Environmental Health Perspectives*. 2011, 119: 1766-1774.を参照。

⁸ 過剰リスクは、記録された反応の発生率のバックグラウンド時からの増加分である追加リスク： $P(d) - P(0)$ を、バックグラウンド時の反応の非発生率で除した値であり、過剰リスク = $\{P(d) - P(0)\} / \{1 - P(0)\}$ で定義される。

P(d) : 用量 d における反応の発生率

P(0) : バックグラウンド時 (対照群) における反応の発生率

⁹ ハイブリッド法の手順の詳細については、同法が提案された K. S. Crump, Calculation of Benchmark Doses from Continuous Data. *Risk Analysis*, 1995. 15: 79-89 や、U.S. EPA, Benchmark Dose Technical Guidance. 2012.を参照。

3. 用量反応モデリングと結果の評価及び POD の決定

(1) 用量反応モデリング

- ① 鋭敏なエンドポイントの排除を防止する等の観点から、1 (2) の条件を満たしたデータセットについては、全て用量反応モデリングを行う。
- ② 生物学的な根拠がある場合を除き、データセットに含まれる全ての用量反応データを用いて用量反応モデリングを行う（最高用量群等の特定のデータを用量反応モデリングの際に除外することは行わない。）。
- ③ 特定の化学物質とエンドポイントの組合せに関する用量反応関係は、理想的には、普遍的なトキシコキネティクス及びトキシコダイナミクスを説明する、生物学的見地に基づく単一の数理モデルで説明されることとなる。

そのような「発現機序の本質を捉えた」数理モデルがある場合は、当該モデルから **BMD** 等を算出することが優先される。

- ④ 「発現機序の本質を捉えた」数理モデルが利用できない場合は、**BMD** 等の算出を行うソフトウェアやオンラインツール（以下「ソフトウェア」という。）を使用し、それらに収載されている数理モデルを用いて用量反応モデリングを行う。
- ⑤ ソフトウェアは、使用実績及び **BMD** 等の計算方法に関する理論的背景も踏まえつつ、専門家の関与の下で **BMD** 等の算出結果が妥当であると判断したソフトウェアを用いる。
- ⑥ 使用するソフトウェアにおいて、モデル平均化¹⁰が可能である場合は、その機能を用いた用量反応モデリングを併せて行う。
- ⑦ 個別モデルを用いた用量反応モデリングを行う際にモデルのパラメータの値を制限する **Restriction** が可能な場合は、同一モデルであっても、**Restriction** を行った場合 (**ON**) と行わなかった場合 (**OFF**) は基本的には別個のモデルであるとの前提に立ち、**ON** と **OFF** 両方のモデルを用いた用量反応モデリングを行う。

(2) 用量反応モデリング結果の評価

データセットごとに、モデル平均化の結果及び各数理モデルを用いた用量反応モデリング結果を、以下の場合分けに基づき評価する。

① モデル平均化を行っている場合

モデル平均化に供した個別数理モデルの用量反応モデリング結果につ

¹⁰ AIC 等の指標を重みとして、収束した数理モデルに基づく推定反応量を用量別に加重平均した結果から新たな数理モデルを導き、**BMD** 関連指標を算出する方法。

いて、用量反応データへの適合度等を評価した結果も勘案しつつ、モデル平均化の結果が妥当であることを専門家の関与の下で確認したときは、当該結果を採用する。

なお、その際、モデル平均化に供した数理モデルのうち、視覚的に又は適合度検定¹¹等によって用量反応曲線が用量反応データを明らかに説明できていないと判断されたものを除外した上で、モデル平均化を行い、その結果を評価するという手順を取り得る。

② モデル平均化により妥当性若しくは信頼性が十分に担保された結果は得られ難いと判断された場合、又はモデル平均化を行っていない場合

個別の数理モデルを用いた用量反応モデリング結果を以下の手順により評価する。

a 用量反応モデリングの結果得られた用量反応曲線について、専門家の関与の下、用量反応データへの適合度を視覚的に又は適合度検定等によって評価した結果、用量反応データを明らかに説明できていないと判断されたものは、あらかじめ評価対象から除外する。

なお、適合度検定を行う場合の有意水準は、基本的に 0.1 を用いることとし、用量反応モデリングに当たって特定のモデルを使用する理由がある場合は、状況に応じて異なる有意水準を用いる。

b 専門家の関与の下、各用量反応モデリング結果について、以下の項目全てについて評価する。

(a) 用量反応曲線の形状に生物学的な矛盾がないこと

(b) AIC の値が、同値の最も小さい用量反応モデリング結果と比べて、+3 以下の範囲にあること

(c) 得られた BMD の信頼区間の幅が、他の用量反応モデリング結果と比べて明らかに大きくないこと

(d) 得られた BMD 及びその信頼区間の下限値が、最小用量と比べて著しく低用量側でないこと

c 評価の結果、全ての評価項目を満たす用量反応モデリング結果を選択する。なお、該当する用量反応モデリング結果が存在しない場合は、BMD 法を適用する是非について専門家の関与の下で判断する。

(3) POD の決定

(2) の結果、データセットごとに単一又は複数の独立した用量反応モデリング結果が得られることとなる。

¹¹ 用量反応モデリングの結果得られた用量反応曲線と用量反応データの適合度（あてはまりの度合い）を評価する検定。

それらの中から、専門家の関与の下、ヒトへの外挿性等を踏まえた各エンドポイントの意義、観察用量範囲等の用量反応データの特性、BMD 及びその信頼下限値である BMDL の値等を基に用量反応モデリング結果を選択し、当該モデリング結果において算出された BMD の 90%信頼区間の下限値 BMDL を、POD とする。

4. 結果の文書化

解析の結果は、以下の情報を含む形で文書化し、各専門調査会等における評価対象物質の食品健康影響評価の審議で利用する。

(1) 使用したソフトウェアの名称及びそのバージョン

(2) BMD 法を適用した各データセットの情報

- ① 試験設計概要（試験動物種、エンドポイント、投与方法、投与期間等）
- ② 公開データの場合は、出典（著者名、雑誌名、年号等）
- ③ 試験結果概要
 - a 二値データの場合は、各用量群における試験動物数及び反応を示した動物数
 - b 順序カテゴリーデータの場合は、各用量群における試験動物数、反応を示した動物数及び各個体の反応のグレード並びに当該データを二値データ化した場合はその手順及び根拠
 - c 連続値データの場合は、各用量群における試験動物数並びに反応量の代表値（平均値、中央値）及びばらつき（標準偏差、四分位範囲等）
- ④ 共変量として用いた変数がある場合は、その変数
- ⑤ データセットを統合した場合は、統合後のデータセットに関する③及び④の内容、測定方法等統合前のデータセット間で異なる試験設計及びデータ統合の妥当性に関する考察

(3) BMR の値及びその値を用いた根拠

(4) BMD 法を適用した各データセットにおける用量反応モデリング結果

- ① 各用量反応モデリング結果のプロット（モデル平均化並びに **Restriction ON** 及び **OFF** の両方により得られた用量反応曲線を含む。）
- ② 各用量反応モデリング結果における指標（AIC 等）並びに BMD 及びその 90%信頼区間
- ③ BMD 等の信頼区間においては、その推定に用いた統計学的手法（プロファイル尤度法、ブートストラップ法等）
- ④ モデル平均化を行った場合は、重み付けに用いた指標及び各モデルの重み

(5) 各データセットにおける用量反応モデリング結果の評価の手順等

- (6) POD の値並びに当該 POD を導いたエンドポイント及びデータセットを採用した理由

II. 疫学研究で得られた用量反応データへの適用

1. BMD 法を適用する用量反応データの収集、選択

(1) 疫学研究結果の収集

ヒトへの有害な作用が確認された評価対象物質の、ヒトへのばく露量とその結果生じる健康影響又はその指標（以下「健康影響等」という。）との関係が観察、記録されている利用可能な全ての疫学研究結果を系統的に収集する。その際、専門家の関与の下、動物実験の結果やメカニズム等に基づき、因果関係が確認又は示唆されている健康影響等を、総合的に判断し選定する。

(2) データセットの選択

(1) で収集した全ての疫学研究結果について、以下の項目を整理し、BMD 法の適用を検討した上で、専門家の関与の下、信頼性が高く、POD を求めることが妥当と判断し得る適切なデータセットを選択する。

なお、用量反応関係の統計的な有意性は考慮しない。

- ① 研究デザイン
 - a 分類（観察研究、介入研究等）
 - b 研究実施時期
 - c 追跡、遡及期間等（該当する場合）
- ② 研究対象集団に関する情報
 - a 研究対象集団の属性（性別、年齢、人種等）
 - b 研究対象者数
 - c 研究対象地域
- ③ 評価対象物質のばく露に関する情報
 - a ばく露理由（日常生活、職業性、事故等）
 - b ばく露経路
 - c ばく露期間
 - d ばく露量の推定方法（ばく露指標とその測定方法）
 - e ばく露量の範囲
 - f ばく露量を区分している場合はその区分方法（等間隔、対数等間隔、分位等）及び各区分の対象者数等
- ④ 研究対象とした健康影響等及びその判定基準（臨床的意義等）
- ⑤ 研究内で想定した交絡要因とその調整の有無及び方法

⑥ バイアス及びそれに対する対策又は考察

(3) 原データの入手

原データ（個別の研究対象者のデータ）を活用することで、**BMR**に対応するばく露量での健康影響等について、より詳細な解析が可能となる。そのため、可能な限り、原データ（個別の研究対象者のデータ）又はその他必要なデータを入手し、活用することが望ましい。

(4) データセットの統合

データセットが複数存在する場合、ばく露量の範囲を広げて研究対象者数を増やし、より精緻な検討を行うため、専門家の関与の下、用量反応関係を検討した上でデータセットを統合して解析することが可能である。この場合、用量反応関係が不明確になること等を防ぐため、各データセットにおけるばく露量の範囲、研究対象集団の属性、標本サイズ、データ測定方法等の同質性を把握し、事前に臨床医学的及び統計学的な観点からデータセットの統合に関する妥当性を判断する必要がある。

2. BMR の設定

疫学研究で得られた用量反応データに **BMD** 法を適用する場合、毒性学的又は臨床医学的若しくは公衆衛生的観点から、専門家の関与の下で **BMR** を設定する。その際、対象とした健康影響等の大きさ（重症度、病悩期間、罹患率等）とその精度（ばらつき）を考慮する。なお、ヒトの疫学研究であっても **BMR** によって導出された **BMD** 関連指標がそのまま健康影響に基づく指標値（**HBGV**）とはならないことに留意する。

(1) 二値データ

本指針の「I. 動物試験で得られた用量反応データへの適用」では、二値データにおける **BMR** の設定について、「**BMR** として過剰リスク 10%を用いた **BMD** 関連指標は基本的に算出する。」としている。一方、疫学研究は、研究対象や研究デザイン等が多様であるため、動物試験の場合のように基本的に算出する **BMR** を設定することは困難である。このため、疫学研究で得られた二値データに **BMD** 法を適用する場合には、データセット及び健康影響等ごとに専門家の関与の下で **BMR** を設定する。

(2) 連続値データ

- ① 公衆衛生学的に集団レベルで意味のある検査値等の変化を、科学的根拠に基づいて設定できる場合は、その変化量を **BMR** として設定する。
- ② 公衆衛生学的に意味のある検査値等の変化が設定できない場合、真の転帰と検査値等との関連並びにバックグラウンドの反応率（反応量）及び

そのばらつき等を踏まえ、専門家の関与の下で **BMR** を設定する。

- ③ ①②の他、科学的に合理的と思われるカットオフ値を設定できる場合は、反応量の分布を仮定した上で、カットオフ値を超える反応量を示す個人の割合が、バックグラウンドに比べて、あらかじめ定めた増加分(**BMR**)だけ増える投与量を **BMD** として定義した上で、**BMD** 関連指標を算出するハイブリッド法等を取り得る。
- ④ 科学的に合理的と思われるカットオフ値を設定でき、かつ、個別の研究対象者のデータが利用可能である場合は、情報量の減少を伴う点に留意する必要はあるが、カットオフ値に基づいて連続値データを二値データに変換した上で **BMR** を設定し **BMD** 関連指標を算出することが有用である場合もある。

3. 用量反応モデリングと結果の評価及び POD の決定

(1) 用量反応モデリング

- ① 鋭敏なエンドポイントの排除を防止する等の観点から、1 (2) で選択したデータセットについては、全て用量反応モデリングを行う。
- ② 生物学的な根拠がある場合を除き、データセットに含まれる全ての用量反応データを用いて用量反応モデリングを行う (最高用量群等の特定のデータを用量反応モデリングの際に除外することは行わない)。
- ③ **BMD** 等の算出に有用と考えられる共変量が利用できる場合は、利用の是非について専門家の関与の下で判断する。
- ④ 特定の化学物質と健康影響等の組合せに関する用量反応関係は、理想的には、普遍的なトキシコキネティクス及びトキシコダイナミクスを説明する、生物学的見地に基づく単一の数理モデルで説明されることとなる。

そのような「発現機序の本質を捉えた」数理モデルがある場合は、当該モデルから **BMD** 等を算出することが優先される。

- ⑤ 「発現機序の本質を捉えた」数理モデルが利用できない場合は、**BMD** 等の算出を行うソフトウェアやオンラインツール (以下「ソフトウェア」という。) を使用し、それらに収載されている数理モデルを用いて用量反応モデリングを行う。
- ⑥ ソフトウェアは、使用実績及び **BMD** 等の計算方法に関する理論的背景も踏まえつつ、専門家の関与の下で **BMD** 等の算出結果が妥当であると判断したソフトウェアを用いる。
- ⑦ 使用するソフトウェアにおいて、モデル平均化が可能である場合は、そ

の機能を用いた用量反応モデリングを併せて行う。

- ⑧ 個別モデルを用いた用量反応モデリングを行う際にモデルのパラメータの値を制限する **Restriction** が可能な場合は、同一モデルであっても、**Restriction** を行った場合 (**ON**) と行わなかった場合 (**OFF**) は基本的には別個のモデルであるとの前提に立ち、**ON** と **OFF** 両方のモデルを用いた用量反応モデリングを行う。

(2) 用量反応モデリング結果の評価

データセットごとに、モデル平均化の結果及び各数理モデルを用いた用量反応モデリング結果を、以下の場合分けに基づき評価する。

① モデル平均化を行っている場合

モデル平均化に供した個別数理モデルの用量反応モデリング結果について、用量反応データへの適合度等を評価した結果も勘案しつつ、モデル平均化の結果が妥当であることを専門家の関与の下で確認したときは、当該結果を採用する。

なお、その際、モデル平均化に供した数理モデルのうち、視覚的に又は適合度検定等によって用量反応曲線が用量反応データを明らかに説明できていないと判断されたものを除外した上で、モデル平均化を行い、その結果を評価するという手順を取り得る。

② モデル平均化により妥当性若しくは信頼性が十分に担保された結果は得られ難いと判断された場合、又はモデル平均化を行っていない場合

個別の数理モデルを用いた用量反応モデリング結果を以下の手順により評価する。

- a 用量反応モデリングの結果得られた用量反応曲線について、専門家の関与の下、用量反応データへの適合度を視覚的に又は適合度検定等によって評価した結果、用量反応データを明らかに説明できていないと判断されたものは、あらかじめ評価対象から除外する。

なお、適合度検定を行う場合の有意水準は、基本的に 0.1 を用いることとし、用量反応モデリングに当たって特定のモデルを使用する理由がある場合は、状況に応じて異なる有意水準を用いる。

- b 専門家の関与の下、各用量反応モデリング結果について、以下の項目全てについて評価する。

(a) 用量反応曲線の形状に生物学的な矛盾がないこと

(b) AIC の値が、同値の最も小さい用量反応モデリング結果と比べて、+3 以下の範囲にあること

(c) 得られた **BMD** の信頼区間の幅が、他の用量反応モデリング結果と

比べて明らかに大きくないこと

- c 評価の結果、全ての評価項目を満たす用量反応モデリング結果を選択する。なお、該当する用量反応モデリング結果が存在しない場合は、BMD 法を適用する是非について専門家の関与の下で判断する。

(3) POD の決定

(2) の結果、データセットごとに単一又は複数の独立した用量反応モデリング結果が得られることとなる。

それらの中から、専門家の関与の下、各健康影響等の意義、観察されたばく露量の範囲等の用量反応データの特性、BMD 及びその信頼下限値である BMDL の値等を基に用量反応モデリング結果を選択し、当該モデリング結果において算出された BMD の 90%信頼区間の下限値 BMDL を、POD とする。

4. 結果の文書化

解析の結果は、以下の情報を含む形で文書化し、各専門調査会等における評価対象物質の食品健康影響評価の審議で利用する。

(1) 使用したソフトウェアの名称及びそのバージョン

(2) BMD 法を適用した各データセットの情報

- ① 情報源（著者名、雑誌名、年号、原データ管理機関等）
- ② 研究設計概要（研究対象集団、健康影響等、ばく露経路、ばく露期間等）
- ③ 研究結果概要
 - a 評価対象物質のばく露に関する情報
 - b 各ばく露量における研究対象者数
 - c 各ばく露量における健康影響等の反応量に関する情報
 - d cに関するばらつき
- ④ データセットに含まれる健康影響等の情報を変換（二値化等）した場合は、その手順及び根拠
- ⑤ 共変量として用いた変数がある場合は、その変数
- ⑥ データセットを統合した場合は、統合後のデータセットに関する③及び④の内容、測定方法等統合前のデータセット間で異なる試験設計及びデータ統合の妥当性に関する考察

(3) BMR の値及びその値を用いた根拠

(4) BMD 法を適用した各データセットにおける用量反応モデリング結果

- ① 各用量反応モデリング結果のプロット（モデル平均化並びに Restriction ON 及び OFF の両方により得られた用量反応曲線を含む。）
- ② 各用量反応モデリング結果における指標（AIC 等）並びに BMD 及び

その 90%信頼区間

- ③ BMD 等の信頼区間においては、その推定に用いた統計学的手法（プロファイル尤度法、ブートストラップ法等）
 - ④ モデル平均化を行った場合は、重み付けに用いた指標及び各モデルの重み
- (5) 各データセットにおける用量反応モデリング結果の評価の手順等
- (6) POD の値並びに当該 POD を導いた健康影響等及びデータセットを採用した理由

第4 指針の見直し

BMD 法に関する国際的な動向や、食品健康影響評価における活用実績等を踏まえ、必要に応じて本指針の見直しを行う。

ベイズ推定を活用したBMD法を使用する際の考え方について

1. はじめに

近年、計算科学の進歩やソフトウェアの開発により、BMD法においてベイズ推定を活用した手法の導入の検討が急速に進んでいる。過去の研究結果等を事前分布としてBMD関連指標の算出に反映させることにより、その推定精度を高めることができる等の利点があることから、令和2年(2020年)の Principles and Methods for the Risk Assessment of Chemicals in Food (Environmental Health Criteria 240, WHO) chapter 5の改正において、「Bayesian methods are generally preferred in this guidance」と記載され、令和4年(2022年)に改正されたGuidance on the use of the benchmark dose approach in risk assessment (EFSA)においても、「Bayesian model averaging is recommended as the preferred approach, as it brings the following main advantages compared to the frequentist model averaging approach recommended in the previous version of this guidance」とベイズ推定を推奨することが明記されるなど、国際的にベイズ推定を活用した手法がBMD法の主流に変わりつつある。このような事情を踏まえると、我が国においても今後、ベイズ推定を活用したBMD法の導入を図っていく必要があると考えられるため、その使用に当たって考慮すべき点等を整理することとした。

一方、その導入に当たって今後さらに検討を要する課題もあり、速やかに既存の手法から変更することは困難と考えられることから、現時点では本指針の別添として本文書を取りまとめることとした。

2. ベイズ推定を活用したBMD法とは

頻度論では、パラメータは固定された定数とみなして推論を行う。このため、頻度論を用いたBMD法では、最尤法¹によって数理モデルのパラメータを推定して、得られた数理モデルからBMDを算出する。

一方、ベイズ統計学では、パラメータを確率変数とみなして推論を行う。具体的には、最初に推論の出発点となるパラメータの確率分布として事前分布を想定し、これを観察されたデータで更新することで事後分布を求めて、パラメータの値に関する推論を行う。したがって、ベイズ推定を活用したBMD法

¹ 観察されたデータが得られる確率が最大となるパラメータの値を推定値とする方法。

では、得られた用量反応データで数理モデルのパラメータに関する事前分布を更新し、これらのパラメータの不確実性の程度を示す事後分布を求めることで、BMDを確率分布として得る。

令和2年度～令和3年度食品健康影響評価技術研究「ベイズ推定を活用したベンチマークドーズ法の評価手法検討と国際動向の研究（課題番号JPCAFSC20202001）」において、ベイズ推定を活用したBMD法の利点として、少数データのもとでも推定結果が安定しやすいこと、過去の研究結果等を事前分布として考慮することでBMD関連指標の推定精度を高めることができると等が挙げられている。

3. 事前分布の設定

ベイズ統計学では、事前分布を観察されたデータで更新して事後分布を得るため、使用するデータが増えると、事後分布に与える事前分布の影響が相対的に小さくなる。しかし、一般的に評価に使用される用量反応データは、サンプルサイズが限られていることが多いため、PODとして利用されるBMD関連指標は、事前分布の影響を受けやすい。このため、食品健康影響評価におけるベイズ推定を活用したBMD法の使用に当たっては、事前分布の設定が極めて重要である。

事前分布には、無情報事前分布²と有情報事前分布³がある。事前情報がほとんどない又は事前情報があっても使用することが適切でない場合には、無情報事前分布を使用する。有情報事前分布を使用する場合には、BMD関連指標の推定精度を高めることができる一方、偏った結果を導く可能性もある。

事前分布を使用する際には、次の点を考慮すべきである。

- ・評価毎に専門家の関与の下で、事前分布を設定する。
- ・無情報事前分布を含む複数の事前分布から得られた結果等を比較して、事前分布の影響等を検討する。

4. モデル平均化と個別数理モデルの選択

² 事前情報がないものとして設定される事前分布。無情報事前分布の一例は、一様分布（ある範囲の確率（密度）が全て等しい確率分布）。ただし、狭い範囲で一様分布を設定した場合には、強い情報を有し得るため、一様分布が必ずしも無情報事前分布になるとは限らない。無情報事前分布を用いる場合、恣意性が低く、頻度論と同様の結果が得られやすい等の特徴がある。

³ 過去の研究結果、文献、専門家の知見等に基づき、設定される事前分布。わずかな情報に基づく有情報事前分布は、弱情報事前分布（weakly informative prior）と呼ばれる。

モデルの不確実性に対処するため、基本的にはモデル平均化の結果を優先する。

モデル平均化で得られた信用区間が広すぎる場合等、モデル平均化で得られた結果が適切でないと思われるケースでは、**Bayesian Information Criterion (BIC)**⁴等の情報量規準等を判断基準として、専門家の関与の下で個別数理モデルを選択することも選択肢の一つである。

5. 結果の評価における留意点

有情報事前分布を使用する際には、3. で記載したように、**BMD**関連指標の推定精度を高めることができる一方、偏った結果を導く可能性もあるため、無情報事前分布を含む複数の事前分布から得られた結果等を比較して、事前分布の影響等を検討し、結果を評価すること。

6. 結果の文書化

用いた事前分布及びその根拠について文書化すること。なお、事前情報があっても無情報事前分布を使用した場合は、有情報事前分布を使用しなかったその理由を記載すること。有情報事前分布を使用した場合は、恣意的な選択ではないことを示すために、当該有情報事前分布を使用した根拠等を記載すること。

7. 今後の取組

ベイズ推定を活用した**BMD**法の今後の展開や新たな考え方について、国際的な動向や国内外の科学的知見等を引き続き注視し、必要に応じて本別添の見直しを行う。

⁴ AICと同様に、異なる数理モデル間で、モデルの複雑さと測定データとの適合度とのバランスを比較するための指標の一つ。BICの値が小さくなるモデルが好ましいとされている。