

新たな時代に対応した 評価技術の検討

～BMD 法の更なる活用に向けて～

2018 年 7 月

食品安全委員会評価技術企画ワーキンググループ

目次

○食品安全委員会委員名簿.....	2
○食品安全委員会評価技術企画ワーキンググループ専門委員名簿.....	2
○審議の経緯.....	3
I. はじめに.....	4
1. 評価技術企画ワーキンググループの設置.....	4
2. 食品安全委員会が行う毒性評価における BMD 法の活用に向けた検討.....	4
II. BMD 法の開発・活用の経緯.....	6
1. BMD 法を用いた毒性評価.....	6
2. ソフトウェア等の開発状況.....	7
3. BMD 法の活用に関する指針（ガイダンス）の整備状況.....	8
4. 食品分野における毒性評価での活用事例.....	8
III. 食品安全委員会が行う毒性評価における BMD 法の活用に向けて.....	10
IV. 指針策定に向けて整理した論点.....	11
1. BMD 法を適用する試験・研究の選択.....	11
(1) 前提となる考え方.....	11
(2) 動物試験.....	12
(3) 疫学研究.....	12
(4) 試験・研究データの統合と原データの確認.....	14
2. BMR の設定.....	15
(1) 二値データと連続値データ.....	15
(2) 疫学研究データ.....	17
3. 数理モデルの選択.....	17
(1) 前提となる考え方.....	17
(2) モデル選択の基準及び手順.....	18
4. HBGV 算出に使用する POD.....	21
V. おわりに.....	21
別紙 1：略称一覧.....	22
別紙 2：用語説明.....	23
参照.....	25

＜食品安全委員会委員名簿＞

(2018年6月30日まで)

佐藤 洋 (委員長)
山添 康 (委員長代理)
吉田 緑
山本茂貴
石井克枝
堀口逸子
村田容常

(2018年7月1日から)

佐藤 洋 (委員長)
山本茂貴 (委員長代理)
川西 徹
吉田 緑
香西みどり
堀口逸子
吉田 充

＜食品安全委員会評価技術企画ワーキンググループ（第7回）専門委員名簿＞

川村 孝 (座長)	小坂 健
広瀬明彦 (座長代理)	小関成樹
赤堀有美	山田隆志
岡田 孝	

＜食品安全委員会評価技術企画ワーキンググループ（第8回）専門委員名簿＞

川村 孝 (座長)	岡田 孝
広瀬明彦 (座長代理)	小関成樹
赤堀有美	山田隆志

(専門参考人)

小坂 健	西浦 博
------	------

＜食品安全委員会評価技術企画ワーキンググループ（第9回）専門委員名簿＞

川村 孝 (座長)	岡田 孝
広瀬明彦 (座長代理)	小関成樹
赤堀有美	山田隆志

(専門参考人)

井上 薫	西浦 博
小坂 健	

＜食品安全委員会評価技術企画ワーキンググループ（第10回）専門委員名簿＞

川村 孝 (座長)	小関成樹
広瀬明彦 (座長代理)	西浦 博
赤堀有美	山田隆志
小坂 健	

(専門参考人)

井上 薫	岡田 孝
------	------

＜審議の経緯＞

2017年 8月31日	第 7回評価技術企画ワーキンググループ
2017年10月20日	第 8回評価技術企画ワーキンググループ
2018年 3月30日	第 9回評価技術企画ワーキンググループ
2018年 6月 7日	第 10回評価技術企画ワーキンググループ
2018年 7月10日	第704回食品安全委員会（報告）

I. はじめに

1. 評価技術企画ワーキンググループの設置

食品安全委員会は、2003年の設立以来、食品に含まれる化学物質、食中毒原因微生物等について、国際的に合意されたリスクアナリシスの考え方にに基づき、人の健康に与えるリスクを科学的に評価してきた。

そのうち、化学物質のリスク評価に当たっては、評価対象となる物質の特性、入手可能な毒性試験データの質等に応じて、最も適切と考えられる評価方法を随時活用してきた。

その一方で、測定技術の進展等に伴って評価対象となる化学物質が多様化し、毒性試験をめぐる社会的情勢が変化している中で、より科学的に妥当性の高い食品健康影響評価を行うためには、これまでに活用してきた評価方法に加えて、新しい評価方法についても活用していく必要性が生じている。さらに、情報技術の発展とレギュラトリーサイエンスの発展とが相まって、新しいアプローチによる様々な評価方法が開発されてきており、今後、これらを活用した安全性評価が想定されるようになっている。

食品安全委員会は、科学技術の発展に応じ、より科学的に妥当性の高い食品健康影響評価の実施を常に目指す必要がある。このため、国内外の動向を踏まえ、今後の積極的な活用が見込まれる評価方法について現状及び課題を整理し、今後の取組の方向性について提言する目的で、2016年4月に評価技術企画ワーキンググループ（以下「WG」という。）を設置した（参照1）。

WGは、これまでに、溶出物質、不純物、代謝物等の微量な化学物質のうち、毒性試験データが乏しい物質について、より科学的に妥当性の高い評価を行うことを目指し、コンピュータ上での化学物質の毒性評価方法について検討を行い、2017年7月、議論の経過を「新たな時代に対応した評価技術の検討～化学物質の毒性評価のための(Q)SAR及びRead acrossの利用～」として取りまとめ、食品安全委員会に報告している（参照2）。

2. 食品安全委員会が行う毒性評価における BMD 法の活用に向けた検討

ベンチマークドーズ（Benchmark Dose: BMD）法は、動物試験や疫学研究で得られた化学物質のばく露量と毒性発生の頻度又は量との関係（用量反応関係）に、数理モデルを当てはめて得られた用量反応曲線から、バックグラウンドに比して一定の反応量の変化（Benchmark Response: BMR）をもたらす用量であるベンチマークドーズ（BMD）及びその信頼区間の下限值である Benchmark Dose Lower Confidence Limit: BMDL を算出する手法である。

BMD 法は、主に汚染物質の毒性評価に利用されており、米国環境保護庁（EPA）や欧州食品安全機関（EFSA）等が、BMD 法を活用するための手順等を整理した

指針（ガイダンス）を公表している。

食品安全委員会においても、これまでに、メチル水銀、ヒ素、加熱時に生じるアクリルアミド等、主に汚染物質の評価で BMD 法を用いてきた。BMD 法の活用に当たっては、その都度、用量反応関係が成立し、かつ、試験設計や研究デザインが妥当な動物試験や疫学研究のデータを用いて、毒性学的、臨床的及び社会的に有意と判断された反応が生じる用量やばく露量を算出してきた。

BMD 法は、個々の動物試験又は疫学研究において得られた全ての用量（ばく露量）－反応データを用いて最もフィットする数理モデルを設定することにより、低用量域における反応レベルを推定でき、また、試験動物数・研究対象者数の多寡、測定値のばらつき等、試験・研究のデータの質の問題及び統計学的な不確実性をある程度克服できる等の特徴を有している。

このため、適切な条件が整えば、化学物質の毒性評価において BMD 法を用いることで、高い妥当性を持った毒性に関する評価値が得られる。しかしその一方で、BMR の設定方法や数理モデルの選択等についての考え方が国際的に統一される段階には至っていないなど、同法の活用に当たっては技術的な課題が複数存在する。

このような状況を踏まえ、WG は、国内外における BMD 法の活用状況を把握するとともに、食品安全委員会が、引き続き一貫性及び透明性をもって同法を活用するための課題について検討した。

II. BMD 法の開発・活用の経緯

1. BMD 法を用いた毒性評価

BMD 法は、用量反応関係全体に数理モデルを当てはめ、関数化した用量反応曲線としてとらえることにより、低用量域におけるリスクを考察する試みである。

毒性評価においては、1954 年に Lehman と Fitzhugh により無毒性量 (No-Observed-Adverse-Effect Level: NOAEL) の概念が提唱されて以降、耐容摂取量等の健康影響に基づく指標値 (Health-Based Guidance Value: HBGV) を算出する際の出発点 (Point of Departure: POD) として、NOAEL が広く利用されてきた。

一方、BMD 法と同様に、用量反応関係に数理モデルを当てはめて、低用量域での用量反応評価を行う手法も提案されてきた。1961 年に Mantel と Bryan は、低用量域での発がんリスク評価の手順として、数理モデルを用いて、最小試験用量又は腫瘍の 1%過剰発生率をもたらす推定用量 (現在の BMD に相当) での腫瘍過剰発生率の信頼上限値を算出する方法を提示した (参照 3)。また、Gaylor and Kodell (1980)、Van Ryzin (1980) 及び Farmer ら (1982) は、試験での最小用量又は有害影響の 1%過剰発生率をもたらす推定用量 (現在の BMD に相当) での信頼上限値から、過剰リスクがゼロとなる点へ直線外挿する方法を提示した (参照 4)。

1984 年に Crump は、健康影響に関する反応の程度をバックグラウンドに比べて一定量押し上げる用量の信頼下限値を “benchmark dose” (現在の BMDL に相当) と定義し、この “benchmark dose” を POD として HBGV を得る方法が、NOAEL を POD として HBGV を得る方法に置き換わりうるものであると主張した (参照 5)。

EPA は、化学物質のリスク評価のうち発がん影響の評価において、用量反応データに適合する数理モデルを用いて低用量域に外挿することで、低用量域における生涯発がんリスクを算出し、実質安全量 (Virtually Safe Dose: VSD)¹を推定する方法を採用してきた。2005 年には、発がん物質のリスク評価ガイドラインにおいて、生物学的に適正なモデルが得られない場合の方法として、10%過剰発がんリスクの信頼下限値である BMDL を POD として、原点への直線外挿により VSD を算出する方法を提案した (参照 6)。その一方で、非発がん影響の評価においては、主に NOAEL を POD として HBGV を算出する方法を採用していたが、1995 年に、EPA の Risk Assessment Forum において、非発がん影響のリスク評

¹ 閾値が存在しない遺伝毒性発がん物質等の毒性に対し、生涯にわたり摂取した場合のリスクが、許容できるレベルとなるようなばく露量のこと。発がん性の場合、10 万分の 1 あるいは 100 万分の 1 というような低い確率でがんを発生させるようなばく露量となり、通常の生活で遭遇する稀なリスクと同程度の非常に低い確率となるようなばく露量 (実質的に安全な量) と解釈される。

価における BMD 法の利用方法が報告書として取りまとめられ（参照 7）、同年、EPA は、化学物質の非発がん影響（メチル水銀の発達神経毒性）について、BMD 法を用いた毒性評価を初めて行うとともに（参照 8）、コンピュータ上で BMD 及び BMDL を算出するためのソフトウェアの開発に着手している。

2. ソフトウェア等の開発状況

リスク評価における BMD 法の導入支援を目的として、コンピュータ上で BMD 及び BMDL を算出するためのソフトウェアやオンラインツールが開発されてきた。

EPA は、1995 年にソフトウェアの開発に着手し、外部評価や性能確認試験を経て、2000 年に Benchmark Dose Software: BMDS を公開した。BMDS はその後も改良が進められ、現在、国際的に最も知られているソフトウェアとなっている（参照 9）。

また、欧州では、オランダ国立公衆衛生環境研究所（RIVM）が、2009 年にソフトウェア PROAST を開発した。同ソフトウェアは、BMD 等の算出に統計解析ソフトウェアを要するものの、共変量を含む解析が可能等の特徴を有し、EFSA での毒性評価を中心に使用されている（参照 10）。

我が国では、村田らが開発した統計ソフトウェア SPBS に、疫学研究のデータから BMD 及び BMDL を算出するプログラムが組み込まれ、公開されている（参照 11）。

また、2007 年、Wheeler と Bailer が、数理モデルの予測の良さを示す AIC（Akaike Information Criterion: 赤池情報量規準）等の統計量を重みとして、当てはめ可能な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数理モデルを導き、BMD 及び BMDL を算出する方法（モデル平均化）を提唱し（参照 12）、2008 年には、モデル平均化を行うためのソフトウェア Model Averaging for Dichotomous Response Benchmark Dose: MADr-BMD を開発している（参照 13）。

2017 年には、EFSA が、BMD 及び BMDL をオンライン上で算出可能なウェブサイトを公開した。同サイトでは、BMD 等の算出には PROAST が、モデル平均化には MADr-BMD が使用されている（参照 14）。

BMD 及び BMDL 算出に利用するソフトウェアでは、用量反応データについて非連続データ（二値データ）である場合及び連続データである場合のそれぞれにおいて、用量反応曲線の作成に複数の数理モデルが利用可能であり、利用可能な数理モデルを当てはめて得られる BMD 及び BMDL とともに、それぞれの数理モデルについて、用量反応データへの適合度合いを示す統計量も併せて算出される。なお、現在、BMD 法では、logistic モデル、probit モデル等既存の確率分布に基

づく数理モデル、Weibull モデル、multistage モデル等有害影響の発現機序に関する理論に基づく数理モデル等が利用されており、BMDS (ver. 2.7) には 30 種類のモデルが収載されている。

3. BMD 法の活用に関する指針（ガイダンス）の整備状況

BMD 法を活用するための手順等を整理した指針（ガイダンス）が、海外のリスク評価機関等からいくつか公表されている。

EPA は、1986 年に、化学物質の発がん影響の評価において、用量反応データに数理モデルを当てはめ、低用量域での発がんリスクを算出する方法を整理したガイダンスを公開した。その後、非発がん影響も含め、リスク評価における BMD 法の利用に向けた検討を進め、2000 年に、用量反応データへの数理モデルの当てはめ及び選択、BMR の設定等 BMD 法の活用における各プロセスの手順を取りまとめたテクニカルガイダンスの案を公表し、その後、運用経験等に基づき、2012 年に同ガイダンスを正式に策定、公表した（参照 4）。

EFSA は、2009 年に、BMR の設定に係る課題、数理モデルの選択方法等を整理したガイダンスを公表した（参照 15）。EFSA は、数理モデルの不確実性と標本誤差（サンプリングエラー）による不確実性を同時に説明する最も良い方法としてモデル平均化を採用し、それまでの活用経験も踏まえ 2016 年にガイダンスの内容を改訂した（参照 16）。

また、国際化学物質安全性計画（International Programme on Chemical Safety: IPCS）は、2009 年に、用量反応評価のために用量反応データに数理モデルを当てはめる際の考え方や手順を解説したガイダンスを策定し、その中で数理モデルの選択、BMD 及び BMR の設定等 BMD 法の手順について整理している（参照 17）。

我が国では、食品安全委員会の食品健康影響評価技術研究「用量反応性評価におけるベンチマークドース法の適用に関する研究」（平成 22 年度～平成 24 年度）において、BMD 法を活用する際に、BMR の設定や数理モデルの選定等で推奨される手法を整理した「ベンチマークドース法の適用に関するガイダンス」が研究成果として取りまとめられている（参照 18）。

4. 食品分野における毒性評価での活用事例

1995 年、EPA は、メチル水銀の評価において、ヒトでの食品を通じた中毒事例のデータに BMD 法を初めて適用し、その発達神経毒性を評価した。その後、2001 年に、疫学研究で得られた臍帯血中水銀濃度と子の神経発達に係る影響指標のデータに BMD 法を適用し、メチル水銀の発達神経毒性を再評価している（参照 8）。

食品中の汚染物質に係る国際的なリスク評価機関である FAO/WHO 合同食品添加物専門家会議 (Joint FAO/WHO Expert Committee on Food Additives: JECFA) は、2005 年、第 64 回会合でのアクリルアミド、エチルカーバメート及び多環芳香族炭化水素類の評価において、動物試験で得られた発がん影響に関する用量反応データに BMD 法を適用し、得られた BMDL と食品を通じた推計摂取量の比 (Margin of Exposure: MOE) の値から、各化合物がヒトの健康に悪影響を及ぼす懸念の程度について評価した (参照 19)。

食品安全委員会は、2004 年に魚介類等に含まれるメチル水銀の評価において、疫学研究で得られた母親毛髪中の水銀濃度と子の言語能力に関する影響指標のデータに BMD 法を適用し、得られた BMDL 等を POD として耐容週間摂取量を設定した (参照 20)。

また、2013 年には、食品中のヒ素の評価において、疫学調査で得られた飲料水中のヒ素濃度の区分ごとに発がん影響及び非発がん影響の発症率を整理したデータに BMD 法を適用した。各エンドポイントにおいて設定した BMR の幅だけ発症率を増加させる飲料水中の無機ヒ素濃度 (Benchmark Concentration: BMC) 及びその信頼下限値 (Benchmark Concentration Lower Confidence Limit: BMCL) を算出するとともに、BMCL から各エンドポイントにおける無機ヒ素経口摂取量の BMDL を算出し、日本人の無機ヒ素推計摂取量との比較及び評価を行った (参照 21)。

さらに 2016 年には、加熱時に生じるアクリルアミドの評価において、動物試験で得られた発がん影響及び非発がん影響に関する用量反応データに BMD 法を適用し、得られた BMDL と食品を通じた日本人の推計摂取量から算出した MOE を基に、各エンドポイントについて、アクリルアミドが日本人の健康に悪影響を及ぼす懸念の程度を評価した (参照 22)。

Ⅲ. 食品安全委員会が行う毒性評価における BMD 法の活用に向けて

食品健康影響評価を実施するに当たり、毒性評価において試験・研究の結果から耐容摂取量等の HBGV を算出するためには、低用量域における化学物質の用量反応関係を把握することが重要である。

前述のとおり、BMD 法は、個々の試験・研究において得られた全ての用量（ばく露量）－反応データを用いて最もフィットする数理モデルを設定することにより、低用量域における反応レベルを推定できる、また、試験動物数・研究対象者数の多寡、測定値のばらつき等、試験・研究のデータの質の問題及び統計学的不確実性のある程度克服できる等の特徴を有している。このため、適切な条件を整えば、化学物質の毒性評価において、妥当性の高い POD が得られることから、可能な範囲で BMD 法を用いることが望まれる。

一方、BMD 法の活用にあたっては技術的な課題が複数存在する。例えば、BMR については、どの値に指定するのが合理的か、あるいは、毒性学的、臨床的に意味のある値となるかという課題があり、国際的にも議論が成熟したとは言えない状況である。

また、用量反応データに当てはめた複数の数理モデルからどのモデルを選択するかについては、海外のリスク評価機関等の間でも考え方が統一されていないため、同一の試験・研究データに BMD 法を適用した際に異なるモデルが選択された結果、機関によって異なる POD が算出された事例がある。

さらに、疫学研究においては、そのデータに BMD 法を適用するにあたって、交絡の調整等動物試験のデータとは異なる特有の課題が存在する。

WG は、今後、食品安全委員会がより一貫性及び透明性を確保して化学物質の毒性評価で BMD 法を活用するために、BMD 法の活用にあたっての技術的な課題や食品安全委員会での使用経験を踏まえ、一定の考え方を整理する必要があると考え、その第一段階として、技術的な課題について整理し、その議論の経過を以下のとおり取りまとめた。

今回取りまとめた検討結果に基づき、更に研究による検証や知見の収集等を行った上で、WG として食品安全委員会における BMD 法の活用に関する指針を作成していくこととする。なお、これまでの議論では、BMD 法を用いて妥当な POD を算出するためには、統計学の見地からの専門家の判断が必要であり、BMD 法の活用を進めていく上では専門性を有する人材の有無も重要な観点であるとの意見もあり、人材の確保・育成についても別途検討を進めていく必要がある。

IV. 指針策定に向けて整理した論点

WGは、化学物質の毒性評価に当たってBMD法を活用する際の主要な作業について、留意点や更なる情報収集が必要な点を整理した。

1. BMD 法を適用する試験・研究の選択

(1) 前提となる考え方

海外のリスク評価機関等が策定したBMD法の指針（ガイダンス）の内容は以下のとおりであり、BMD法を適用する試験・研究の選択について国際的に考え方が統一された状況にはなっていない。

a. IPCS（参照17）

- ・ 利用可能な試験・研究で観察された全てのエンドポイントについて、用量反応データをモデル化することは効果的ではない場合がある。最も感受性が高い種における特定の毒性反応の選定等により、他に比べて明らかに高用量での反応が見られる試験・研究を除外することが考えられる。
- ・ ある試験・研究が多数のエンドポイントを観察している場合、明確な用量反応関係を示していないエンドポイントの目視による除外や、毒性学的影響及び見かけ上の反応量の大きさに基づくエンドポイントの選択を行うことができる。しかし、目視によりエンドポイントを選択することは容易でない例があり、この場合、全てのエンドポイントをモデル化する必要がある。
- ・ モデル化に当たっては、エンドポイントの選択後、各用量反応データについて、実際に用量反応の分析が可能か否かを判断する。

b. EPA（参照4）

- ・ 毒性試験データのレビュー後、BMD法を適用する試験・研究を選択する。試験・研究の選択プロセスは、BMDの算出が可能となるように、モデル化が可能な試験・研究を特定することを意図しており、関連する全ての試験・研究について検討する。
- ・ モデル化の可能性について試験・研究を評価した後、通常は、リスク評価者が評価対象物質のばく露に関連すると判断した試験・研究における全てのエンドポイントについて、モデル化を検討する。多数の試験・研究又は多数のエンドポイントを観察した試験・研究を対象とした場合、標的臓器又は当該試験・研究を代表する反応をエンドポイントのサブセットとして選択できる場合もある。
- ・ PODの算出に使用される可能性があるエンドポイント、特に異なる不確実係数が適用される可能性がある試験及びエンドポイントについては、全てモデル化すべきである。

c. EFSA（参照16）

- ・ 評価対象物質において、用量反応関係のデータセットが多数存在する場合、他のデータセットに比べてより低用量での反応が見られるデータセットを目視で予備的に選択することも考えられるが、目視による選択のみではサンプリングエラーを考慮できないことから、理想的には全てのデータ

セットをモデル化した結果に基づいて、重要な単一又は複数のデータセットを選択する。

WGは、海外のリスク評価機関等の考え方も踏まえて検討した結果、評価対象物質において用量反応関係が成立している全ての試験・研究データを対象として、BMD法の適用を検討すべきであることを確認した。

また、WGでは、試験・研究の選択と対象とするエンドポイントの決定の手順に関して以下の意見があった。

- ・ 対象とするエンドポイントを先に決めるべきである。
- ・ 試験・研究を先に選択してからエンドポイントを決定すべきである。
- ・ 利用可能なデータの種類によって、実際のリスク評価の過程で決めていくことが重要である。

さらに、WGは、対象とするエンドポイントを検討する場合に、各エンドポイントの発現機序や毒性学的意義、臨床的意義、カットオフ²値の妥当性等を考察し、健康影響の事象を正しくとらえているか検討する必要があること、特に検査値等の代替指標を用いる場合は、観察された変化が、ヒトへの健康影響において実質的な意味や公衆衛生学上の意味を持つのかどうか、十分に考察する必要があることを確認した。

（２）動物試験

WGは、動物試験で得られた用量反応データにBMD法を適用する際、特に文献報告の場合には、その試験設計（動物種、匹数、投与方法、投与量の設定等）が適切であるか確認をした上で、評価に用いるデータセットとして採用するか否かを判断すべきであることを確認した。

なお、WGでは、病理組織学的所見を観察した用量反応データにBMD法を適用する際には、試験によっては、発生頻度に加えて病変の程度が記載されているケースが想定されるため、二値データとして取り扱うことの妥当性を慎重に検討すべきという意見があった。

（３）疫学研究

WGでは、疫学研究がヒトを対象とした研究であることを重視し、BMD法の適用を検討する際に疫学研究のデータが利用可能な場合には、その利用を必ず検討するという認識を関係者間で共有すべきという意見があった。

一方、疫学研究は、動物試験とは異なる特有の課題を有することから、WGは、BMD法の適用を検討するに当たって、事前に研究内容の妥当性を確認することが重要との意見で一致し、確認すべき事項について以下のとおり整理した。

² 臨床における意思決定の閾値として使用するために設定された値。確定診断、効果判定など目的に応じて任意に設定される。

① 対象集団

疫学研究においては、対象集団に関する情報を十分に吟味することが重要である。具体的には、集団の属性（性別、年齢、人種、小児、妊婦等のハイリスクグループが含まれているか等）が妥当であるか、対象地域（集団の居住する環境が一般環境（非汚染地域）か又は汚染地域か等）が適切に選択されているか等を確認する必要がある。もし、幅広い範囲のばく露量が認められる対象集団であれば、より正確に健康影響を評価できる可能性がある。

また、疫学研究は、数万人を対象とした大規模コホート研究から数十人規模の研究まで、対象集団の規模が研究により大きく異なるため、評価対象物質による健康影響を評価する上で十分な対象者数が確保されているかどうかを確認する必要がある。

さらに、対象集団において、母集団に対する代表性が保証されていることを確認することも重要である。

② ばく露状況の把握

疫学研究は、化学物質へのばく露が管理された条件下にないことから、評価対象物質のばく露状況を正確に測定することは困難であり、対象集団が実際にばく露された量を推定する必要がある。

食品中に含まれる化学物質のリスク評価を行う際には、食品由来のばく露量を把握することから、経口ばく露量の推定方法について吟味すべきである。経口ばく露量の推定に当たり、血液中、尿中、毛髪中の濃度等のバイオモニタリングデータを活用する場合には、その値を経口ばく露量に換算する方法の妥当性を確認する必要がある。

また、食品又は水の汚染実態データを活用する場合には、摂取する濃度、量、期間等のばく露シナリオについて、エンドポイントの発現機序等を参考に、ばく露期間が十分に長期であるか等、ばく露量把握のための研究の設定が妥当であるかを確認する必要がある。特に、様々な食品由来のばく露を想定する場合、各食品の摂取量と摂取頻度の分布も加味して考慮する必要がある。近年、食物摂取頻度調査票（Food Frequency Questionnaire: FFQ）に基づくばく露量推定技術が改善されている状況にあるものの、栄養評価を主目的とする調査票は、食品中に微量に混入、又は生成した物質のばく露量把握には向いていないことに留意する必要がある。

さらに、疫学研究で用量反応関係を解析する際には、ばく露量を一定の範囲で区分し、対象者を集団に分け各集団の反応を解析する方法がとられる場合が多い。この際のばく露量の区分は、ばく露量をそのまま又は対数変換して等間隔に区分する方法や、対象者数を等分する方法等の算術的に説明が容易な方法で行われることもあれば、研究報告者が任意に区分することもある。特に、研究報告者による任意の区分については、統計学的有意差を生じさせるためにばく露量を恣意的に区分する可能性が完全には否定しきれないことに留意し、区分の考え方に合理性があるか、各区分に十分な対象者数が含ま

れているか等を慎重に確認する必要がある。

③ 交絡

疫学研究で認められた関連から因果関係を明らかにするためには、研究デザインに応じた交絡の調整が必要であり、BMD法の適用にあつては、その調整内容について確認する必要がある。

なお、WGは、交絡の調整に関して、一般の人口集団を対象とした疫学研究では以下の課題があることを考慮し、交絡の影響を排除できる可能性がある操作変数（instrumental variable）法³等の新しい方法論を用いた因果推論も念頭に置きながら、より豊富なデータを用いて検討した結果を基に議論を深める必要があることを確認した。

- ・ 調整すべき要因が多岐にわたるため、その影響を全て特定して調整することは困難であること。
- ・ 未測定の影響が多く存在し、仮に比較的強い関連が得られたとしても因果の推論は容易でないこと。
- ・ 食品の摂取は、生活習慣の一部であり社会環境的因子等も複雑に絡むことから、通常のコホート研究で測定される因子を用いて完全に調整することは困難であること。

（４）試験・研究データの統合と原データの確認

① 試験・研究データの統合

プールドアナリシス等のメタアナリシスを実施して、複数の研究の結果を数量的に統合し、再解析することがある。WGは、この背景にある考え方として、ばく露量の範囲や標本数を可能な限り増やして真の用量反応関係をとらえる目的があること、また、標本数を増やすことにより標本の母集団に対する偶然の偏りを減らして信頼区間の幅を狭くすることで、より精緻な検討を可能とする目的があることを確認した。

また、WGでは、複数の試験・研究データを統合する場合の留意点として以下の意見が出された。

- ・ データ統合の妥当性を判断するためには、ばく露量の範囲、人種・動物種、標本サイズ、データ測定方法等の違いを把握した上で、生活環境等の交絡因子の調整の可否についても検討する必要がある。
- ・ 標本数を増やせば統計学的有意差も得られやすくなることから、有意な反応の差が得られるばく露量が算出された場合であっても、その差が毒性的、臨床的に意義のある変化かどうか十分に検討する必要がある。
- ・ 評価対象物質によっては、動物愛護の観点から複数の動物試験が実施されず、データ統合の実現可能性が少ないと思われる。
- ・ 一方、一部の汚染物質の様に広く動物試験が実施されている物質もある。

³ 要因には影響するが転帰には直接関与しない因子である操作変数を介在させることによって擬似的にランダム化を行う方法。操作変数が遺伝子多型であるものをメンデルのランダム化（Mendelian randomization）と呼ぶ。

- ・ データの統合が困難である場合に、複数の研究結果の不確実性を全て取り込んだ上でモデル化するデータ統合技術「Bayesian evidence synthesis」の活用により対応できる可能性も示唆されており、更なる知見の収集が望まれる。

② 原データの確認

前述のとおり、疫学研究では、研究報告者が恣意的に区分した、離散的な区分間の統計学的有意差をもって影響の有無を主張する可能性も完全には否定しきれない。このため、WGは、低用量域の用量反応関係について十分な議論を尽くすためには、原データまで遡って再解析をすることが重要であることを確認した。

また、WGでは、原データの確認について以下の意見があった。

- ・ 複数の研究データを統合する場合にも、各研究の原データまで遡って検討することが望ましい。
- ・ 理想的には、試験・研究を開始する時点で原データの公開を促すような仕組みがあることが望ましい。
- ・ 近年は、陰性所見（ネガティブデータ）も収載する雑誌も刊行され、有意な結果が得られなかった試験・研究のデータについても入手可能な環境が整えられつつあり、従来起こりがちであった出版バイアス⁴が避けられつつある。将来的にはこれらの原データも活用し、低用量域における毒性をより詳細に検討できるようになることが期待される。

2. BMR の設定

WGは、BMR の設定について、リスクがその幅で上昇することによって国民全体にどのような影響を与えるのかということを十分説明できるように議論を深める必要があることを確認した。

（1）二値データと連続値データ

WGは、BMRについて、BMD法を適用する用量反応データが二値データの場合及び連続値データの場合のそれぞれについて議論した結果、いずれの場合においても、毒性学的、臨床的な観点から意義のあるBMRを設定すべきであることを確認した。

① 二値データ

二値データを利用する場合におけるBMRの設定について、海外のリスク評価機関等における考え方は以下のとおりである。

a. IPCS（参照17）

- ・ 有害でないと見なしうる反応レベル（BMR）は明確でないことが多く、その選択にあっては毒性学者及び臨床医間での議論が求められる。

⁴ 有意な結果が出た研究は文献として刊行されやすく、そうでない研究は世に出にくいことに起因する情報バイアス。

- ・ BMRの選択には技術的及び政策的な観点の両方が関係してくる。BMRに特定の値を選択することは取り上げない。

b. EPA (参照4)

- ・ BMRの値として、エンドポイント間の比較を目的とする場合は、発がんに関するバイオアッセイの大部分及び非発がんに関するバイオアッセイの一部における反応量の検出限界を基に、過剰リスク10%が推奨されるが、PODの決定等を目的とする場合、過剰リスク10%はBMRのデフォルトでない。
- ・ PODの決定を目的とする場合に、BMRとして5%若しくは5%より低い値、又は10%より高い値を用いる際は、統計学的及び生物学的な検討を必要とする。

c. EFSA (参照16)

- ・ 様々な研究が、NOAELにおける過剰リスクの上限の中央値が10%に近いことを推計している。動物試験から得られた二値データを利用する場合におけるBMRのデフォルト値として、過剰リスク10%を提案する。
- ・ デフォルト値は、統計学的又は生物学的検討結果に基づき変更される可能性があり、デフォルト値と異なる値を用いる場合は、その生物学的妥当性を議論するとともに、根拠を文書化し説明する必要がある。

WGでは、二値データを利用する場合におけるBMRのデフォルト値を設定すること及び設定する場合の具体的な値については、NOAELとの比較だけでなく、毒性学的、臨床的な意義付けについても検討すべきであるとの意見があり、WGは、本論点について科学的情報に基づく更なる検討が必要と判断した。

② 連続値データ

連続値データを利用する場合におけるBMRの設定について、海外のリスク評価機関等における考え方は以下のとおりである。

a. IPCS (参照17)

- ・ 有害でないと見なしうる反応レベル (BMR) は明確でないことが多く、その選択に当たっては毒性学者及び臨床医間での議論が求められる。
- ・ BMRの選択には技術的及び政策的な観点の両方が関係してくる。BMRに特定の値を選択することは取り上げない。

b. EPA (参照4)

- ・ エンドポイントにおける変化の最小レベルが、通常、生物学的に有意な影響とみなされる場合は、その変化量をBMRとして設定するアプローチが望ましい。
- ・ 個別のデータが利用可能であり、合理的に有害とみなされる反応レベルを決定できる場合は、当該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することが可能であり、BMRを二値データの場合と同様に設定することができる。

- ・ 有害とみなされる反応レベルが分からない場合には、**BMR**として、対照群の標準偏差に相当する平均値からの変化レベルを用いるべきである。

c. EFSA (参照16)

- ・ 動物試験から得られた連続値データを利用する場合における**BMR**のデフォルト値として、5%（平均反応量における変化）を提案する。
- ・ デフォルト値は、統計学的又は生物学的検討結果に基づき変更される可能性があり、デフォルト値と異なる値を用いる場合は、その生物学的妥当性を議論するとともに、根拠を文書化し説明する必要がある。

WGでは、連続値データを利用した場合における**BMR**について、統一的なデフォルト値を設定することに対し慎重な意見があったほか、以下の意見があり、WGは、本論点について科学的情報に基づく更なる検討が必要であると判断した。

- ・ データが多様な分布を示すことが想定されることから、統一的なデフォルト値を示すことは困難である。
- ・ 数値が非正規分布を示す場合、標準偏差で処理することは不適切である。
- ・ 連続値データは二値データに変換するという考え方を優先したほうが良い可能性がある。
- ・ 連続値データを二値データに変換すると情報量を減らすことになり、データの扱いとして必ずしも正しいわけではない。また、二値データに変換するにしても、毒性の有無を分けるカットオフ値の設定には、エンドポイントごとに専門家が関与して慎重に行うべきである。

（2）疫学研究データ

疫学研究で得られたデータに**BMD**法を適用する場合、ヒトのデータであり、一般的には動物試験に比べて標本数が多いことに留意して**BMR**を検討する必要がある。

食品安全委員会は、2013年に取りまとめた「食品中のヒ素」の食品健康影響評価において、**BMR**は「がんのような重大疾病の発症割合では厳しく（例えば1%又は5%）、非致死的な疾患ではそれほど厳しくなく（例えば5%又は10%）、**IQ**などの検査値は代替指標（surrogate marker）であることから、疾患あるいは健康度とのつながりが明確なものに限定し、臨床的に意味のある最小の差（minimal clinical important difference）に従うべき」としている（参照21）。

WGは、個人における数値の変化と集団における数値の変化では、公衆衛生上の意味合いが異なる点も十分に考慮して**BMR**を議論すべきであるとの認識で一致した。

3. 数理モデルの選択

（1）前提となる考え方

WGは、化学物質とエンドポイントとの組合せにより、その用量反応関係は、普遍的に単一の数理モデルを取るはずであり、そのような「発現機序の本質を捉えた」数理モデルが存在していれば、その数理モデルを優先することが望ましいことを確認する一方、「発現機序の本質を捉えた」数理モデルが存在するエンドポイントはごく限られているため、現状としてはBMDs等のソフトウェアに収載されている数理モデルの中から選択せざるを得ないことも確認した。

また、WGでは、「発現機序の本質を捉えた」数理モデルが不明である場合は、観察された用量反応関係を少数のパラメータで説明し得る複数の数理モデルについて、その適合度等を比較することでモデルの不確実性に対処すべきであるとの意見があった。

さらに、モデル選択に当たっては、エンドポイントの発現機序に関するレビュー結果を考慮すべきであるが、発現機序に関する情報がどの程度入手できれば「発現機序の本質を捉えた」数理モデルを導くことが可能となるのかについては、更なる検討が必要であるとの意見があった。

このほか、WGでは、試験・研究の結果が母集団から抽出された標本のデータであることから、母集団に対する標本の代表性の確保について以下の意見があった。

- ・ 別標本の試験・研究も勘案して母集団を推定することに努めるべきである。
- ・ 一つの評価対象物質に関して、別標本の試験・研究結果が入手できない場合は、当該物質の試験結果のうち類似の発現機序が想定されるエンドポイントの試験結果又は類似物質の試験結果のうち同じエンドポイントの試験結果を参考にすることが望ましい。
- ・ 「類似」の範囲の定義づけが難しく、単一の試験でBMD等を判断せざるを得ないケースが多い。

（２）モデル選択の基準及び手順

WGは、試験・研究データのサンプリングエラーを考慮して、明らかに妥当でない数理モデル以外は検討対象とすべきとの認識で一致した。

また、生物学的妥当性の観点から、用量反応データに当てはめた数理モデルに基づく用量反応曲線の原点付近における傾きが無限大になることを防ぐこと等を目的として、EPA（2012）は、モデルのパラメータの値を制限するRestrictionを必要に応じて行うこととしている（参照4）。この点について、WGは、同一数理モデルであっても、Restrictionを行った場合（ON）と行わなかった場合（OFF）は基本的には違うモデルであるとの前提に立ち、ONとOFF両方のモデルを検討対象とすべきであることを確認した。

用量反応データに当てはめた各モデルの適合度合いを評価し、モデルを選択する基準及び手順について、海外のリスク評価機関等における考え方は以下のとおりである。

a. IPCS (参照17)

- ・ モデルに基づく予測結果と用量反応データとの適合度合いを、適合度検定により臨界値0.1を用いて評価する。
- ・ 用量反応データからのずれを検出する他の方法としてグラフの目視があり、モデルをプロットすることで常に適合度検定が補完される。
- ・ 統計学的理論に基づく単一又は複数モデルの選択方法として考え得る方法
 - ー 同一の分布族 (family) に属し、ネストしているモデルに対しては、尤度比検定を用いて、パラメータの追加による適合度の改善が妥当か否かを判断する。
 - ー 同一の分布族に属しているが、ネストしていないモデルに対しては、AIC値が最小のモデルを選択する⁵。
 - ー 同一の分布族に属していないが、基礎となる確率分布に同一の仮定を用いているモデルに対しては、現時点でAIC値を用いたモデル選択が恐らくは妥当であるが、これには異論もあり、今後、ガイダンス内容を変更する可能性がある。
 - ー 同一の確率分布を基礎としていないモデルにおいては、統計学的なモデル選択のガイダンスはほとんど得られず、個別データの分布を確認することで、仮定したデータ分布の妥当性を検査する必要がある。同様の検査ができない場合は、同一のエンドポイントをモデル化した過去の経験に依拠して妥当な確率分布を選択することが最善の方法である。

b. EPA (参照4)

- ・ 各モデルの用量反応データへの適合度合いを、適合度検定により臨界値0.1 (特定のモデルを使用する理由がある場合は0.05又は0.01) を用いて評価し、選択する。
- ・ 残差、グラフ及びデータの観察結果から、用量反応関係のうち関連する低用量部分を明らかに説明できていないモデルを更に除外する。
- ・ 残ったモデルは、理論的にはBMDLの決定に全て利用可能であるため、これらのモデルの中からBMDLの算出に用いるモデルを選定するための以下の基準は、必然的にやや恣意的なものであって初期設定としての提案である。
 - ー 各モデルから算出されるBMDLの値が十分に近いものであり、かつ、各モデルに特有の影響を反映していない場合、AIC値が最小のモデルをBMDLの算出に利用することができる。AIC値が最も小さいモデルが複数ある場合には、各BMDLの平均値を利用することができる。
 - ー 各モデルから算出されるBMDLの値が十分に近くない場合、それらの結果を信頼する上で不確実性が大きすぎないか専門家による統計学的判断が有用であり、その範囲が妥当なものであると判断された際には、妥当かつ保守的な評価結果として最小のBMDLを選択することができる。

c. EFSA (参照16)

- ・ 用量反応データに当てはめたモデルのうち、そのAIC値が、ヌルモデル

⁵ AIC 値の差が 4 以内にあるモデルは高い確率で同等であるとしている。

(null model) ⁶のAIC値に比べて2以上小さいものを全て選択する。このとき、該当するモデルがない場合は、用量反応関係がないと判断し、BMD法の適用を中止する。

- ・ ヌルモデル及びフルモデル (full model) ⁷を除く選択した全てのモデル (ネストしている⁸モデルを使用している場合は、分布族ごとにAIC値が最小のモデルをあらかじめ選択しておく。) を対象としてモデル平均化を行い、BMD値及びその信頼区間を得る⁹。
- ・ 単一のデータセットに数理モデルを適合させる方法に加えて、性別、種、ばく露期間等の特定の事項以外は類似し、かつエンドポイントが同一であるデータセットの組合せに、既存のモデルを適合させることも可能である。

WGでは、これらの考え方に対して以下の意見があった。

- ・ EFSAが2009年に策定したガイドライン (参照15) で提唱していた最も低いBMDLが得られるモデルを選択する方法は、より保守的なPODを得るためにBMDの信頼下限を重視する方法であるが、信頼性が低い (すなわち信頼区間が広い) 研究を積極的に採択するおそれがある。
- ・ AICを用いてモデルを選択する方法は、試験データへの当てはまりが良いモデルを選択する方法であるが、高用量域における当てはまりがAIC値に反映され、低用量域でのモデル予測値の数値的妥当性がAIC値のみからは保証できないことが多い場合があり、そのことにAICの限界を認めやすい点は留意する必要がある。
- ・ AICの絶対値が2～3程度の範囲内にあるモデルは同等にフィットしている可能性が常にあるため、この範囲内のAICを示すモデルは全て検討対象とするべきである。

WGでは、検討対象となった複数の数理モデルの中から、PODの算出に用いるモデルを選択するために適用する基準及び手順について以下の意見があり、WGは、本論点について、更なる検討が必要であると判断した。

- ・ 評価対象物質において、同一のエンドポイントを検討した複数の試験・研究が利用可能である場合、全ての試験・研究を総覧して、グラフの目視、適合度検定p値及びAIC値を用いて普遍的に最も適合性の良いモデルを選択する。選択しきれない場合はAIC値が近いモデルを対象にモデル平均化を行う。決定したモデルを複数の試験・研究データに当てはめてBMD関連指標を算出する。
- ・ 高用量域における当てはまりがAIC値に反映され、低用量域でのモデル予測値の数値的妥当性が保証できないというAICの限界から、BMD関連指標もモデルを選択する際の基準の一つとして使用する場合がある。

⁶ パラメータが1つ (切片) のみの数理モデル。

⁷ バックグラウンドを含む用量区分数と同数のパラメータを使用した数理モデル。

⁸ ある数理モデルが、他のモデルに1以上のパラメータを追加することにより拡張されたモデルであるとみなせること。

⁹ 選択した全モデルのAIC値が、フルモデルのAIC値より2以上大きい場合は、モデルの修正等を検討するため、BMD法の専門家への相談を推奨している。

- ・ 単一又は少数の試験・研究における複数のエンドポイントについて検討した場合、エンドポイントそれぞれに対して、用量反応データに当てはめた複数の数理モデルの中から、グラフの目視、適合度検定p値、AIC値、さらにはBMD関連指標も基準として、最も適合性の良いモデルを選択する。

4. HBGV 算出に使用する POD

WGでは、HBGV算出に使用するPODに関して以下の意見があり、WGは、本事項について更なる検討が必要であると判断した。

- ・ 最尤値であるBMDが最小となった研究又はエンドポイントを選択し、選択された研究又はエンドポイントから算出されるBMDLをPODとする。
- ・ 信頼区間の下限值であるBMDLが最小となった研究又はエンドポイントを選択し、選択された研究又はエンドポイントから算出されるBMDLをPODとする。
- ・ BMDとBMDLのどちらを優先すべきかという議論は、論点の一つにすぎず、モデルのフィッティングや意味論、低用量領域が良く検討されているか、高用量領域も見ているかなど使用する用量反応データの特性も含め、あらゆる要素を検討しなければならない。用量反応データの特性によってはNOAEL等を参考にする場合もある。その他、毒性発現の機序等を基にヒトへの外挿性の観点からも各エンドポイントの意義を確認するとともに、さらには、参照用量や安全係数/不確実係数についても議論を尽くさなければならない。
- ・ 食品分野でのリスク評価においては、国民全体を考慮する必要があることといった同評価に特有の性質を考慮する必要がある。
- ・ HBGV算出に使用するPODを決定する際の基準及び手順を機械的に決めることは困難であり、学問上の原理・原則と実際のリスク評価の過程で検討することを分けて考えた方がよい。

V. おわりに

現在、食品安全委員会では、WGにおけるこれまでの議論で明らかになった検討課題について、食品健康影響評価技術研究の実施等を通じた科学的知見の集積を進めているところである。WGは、本報告の内容を基に、国内外における活用実績を踏まえつつ、指針の骨子を作成するとともに、同研究の成果等を基に集積された科学的知見を組み込むことにより指針を取りまとめることとする。

別紙 1 : 略称一覧

AIC	Akaike Information Criterion : 赤池情報量規準
BMC	Benchmark Concentration : ベンチマーク濃度
BMCL	Benchmark Concentration Lower Confidence Limit : ベンチマーク濃度の信頼区間の下限值
BMD	Benchmark Dose : ベンチマークドーズ
BMDL	Benchmark Dose Lower Confidence Limit : ベンチマークドーズの信頼区間の下限值
BMR	Benchmark Response : ベンチマークレスポンス
EFSA	European Food Safety Authority : 欧州食品安全機関
EPA	Environmental Protection Agency : 米国環境保護庁
FAO	Food and Agriculture Organization of the United Nations : 国際連合食糧農業機関
HBGV	Health-Based Guidance Value
IPCS	International Programme on Chemical Safety : 国際化学物質安全性計画
JECFA	Joint FAO/WHO Expert Committee on Food Additives : FAO/WHO 合同食品添加物専門家会議
MOE	Margin of Exposure : ばく露マージン
POD	Point of Departure
(Q)SAR	(Quantitative) Structure-Activity Relationship : (定量的)構造活性相関
RIVM	Rijksinstituut voor Volksgezondheid en Milieu (英名: Netherlands' National Institute for Public Health and the Environment) : オランダ国立公衆衛生環境研究所
VSD	Virtually Safe Dose : 実質安全量
WHO	World Health Organization : 世界保健機関

別紙2：用語説明

1. ベンチマークドーズ法 (Benchmark Dose approach)

化学物質や要因のばく露量と当該物質等によりもたらされる有害影響の発生の頻度又は量との関係（用量反応関係）に、数理モデルを当てはめて得られた用量反応曲線から、有害影響の発現率等の反応量に関してバックグラウンドに比して一定の変化（Benchmark Response: BMR）をもたらす用量（Benchmark Dose: BMD）及びその信頼区間の下限値であるBenchmark Dose Lower confidence limit: BMDLを算出し、それをリスク評価におけるPODとして役立てる方法。

2. ベンチマークドーズ (Benchmark Dose: BMD)

ある有害影響の発現率（発生頻度）又はある生物学的な影響に関する測定値について、バックグラウンド反応に比して一定の反応量の変化（BMR）をもたらす化学物質等のばく露量。用量反応関係に数理モデルを当てはめて得られた用量反応曲線を基に算出される。

3. BMR (Benchmark Response)

ある有害影響の発現率（発生頻度）又はある生物学的な影響に関する測定値について、バックグラウンド反応に比して一定の反応量の変化（増加）のこと。反応量には病変ないし得点化された反応の発症率等が用いられる。

4. BMDL (Benchmark Dose Lower Confidence Limit)

BMDの信頼区間の下限値。通常、BMDの90%信頼区間（片側信頼区間としては95%信頼区間）の下限値がBMDLとして用いられる。なお、同信頼区間の上限値はBenchmark Dose Upper Confidence Limit: BMDUと呼ばれる。

5. 用量反応評価

リスク評価における危害要因判定（Hazard Characterization）の段階において、化学的、生物学的又は物理学的な物質・要因へのばく露の大きさ（用量）と、それにより生物や集団にもたらされる健康への悪影響（反応）の程度及び/又は頻度との関係を決定すること。

6. AIC (Akaike Information Criterion)

赤池情報量規準。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。以下の式で定義される。

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

$\log(L)$ ：モデルの最大対数尤度

p ：モデルに含まれる推定パラメータ数

同じデータを説明するために使用する数理モデルの中で、AIC値が最も小さい数理モデルが、最良の予測が得られるモデルであるとされる。

7. 交絡

ある2変数間の関係が他の変数によってゆがめられてしまうことで、目的とする要因を持つばく露集団において、ばく露が生じないと仮定した場合に、その集団での結果が非ばく露集団の結果と一致しないとき、交絡が生じていると定義される。例えば、注目している物質とは別の要因が転帰の原因となっているのに、両者が密接に関連しているため、注目物質が転帰の原因になっているように見えてしまうことが起こり得る。BMD法の適用場面で言えば、用量を変えた群間で比較可能性が失われてしまう。交絡因子が調査結果に影響を与えないように、データ補正（交絡の調整）を行う必要がある。

8. サンプルングエラー

母集団の値（真値）と特定の標本から得られた値の差であり、母集団から標本を抽出したときに、どの標本（個体の組合せ）が選ばれたかという偶然の変動にともなう誤差であり、標本誤差とも呼ばれる。サンプルングエラーの大きさは標本の大きさに確率的に依存し、標本が大きければ大きいほどサンプルングエラーは小さくなることが期待される。

9. 二値データ

化学物質又は生物学的若しくは物理学的な要因へのばく露により各個体で観察された影響を、死亡と生存、特定の腫瘍発生の有無等、起こりうる帰結が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。

10. 連続値データ

化学物質又は生物学的若しくは物理学的な要因へのばく露による影響として、臓器重量や酵素濃度等の連続量について測定したデータのこと。

11. Restriction

パラメータ最適化のプロセスの中でその範囲に制約を設けることを指す。BMD法では、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とる値の範囲に制限を設けること。

参照

1. 食品安全委員会, 評価技術企画ワーキンググループの設置について (平成 28 年 3 月 29 日 食品安全委員会決定) . 2016.
2. 食品安全委員会評価技術企画ワーキンググループ, 新たな時代に対応した評価技術の検討～化学物質の毒性評価のための(Q)SAR 及び Read across の利用～. 2017.
3. N. Mantel and W.R. Bryan, "Safety" testing of carcinogenic agents. J Natl Cancer Inst, 1961. **27**:455-70.
4. U.S. EPA, Benchmark Dose Technical Guidance. 2012.
5. K.S. Crump, A new method for determining allowable daily intakes. Fundam Appl Toxicol, 1984. **4**(5):854-71.
6. U.S. EPA, Guidelines for Carcinogen Risk Assessment. 2005.
7. U.S. EPA, The use of the benchmark dose approach in health risk assessment. 1995.
8. U.S.EPA/IRIS, Methylmercury (MeHg); CASRN 22967-92-6.
9. U.S. EPA ホームページ Benchmark Dose Tools.
<https://www.epa.gov/bmds>
10. RIVM ホームページ PROAST.
https://www.rivm.nl/en/Documents_and_publications/Scientific/Models/PROAST
11. 秋田大学医学部ホームページ 教育用統計ソフトウェア SPBS.
<http://www.med.akita-u.ac.jp/~eisei/link.html>
12. M.W. Wheeler and A.J. Bailer, Properties of model-averaged BMDLs: a study of model averaging in dichotomous response risk estimation. Risk Anal, 2007. **27**(3):659-70.
13. M.W. Wheeler and A.J. Bailer, Model Averaging Software for Dichotomous Dose Response Risk Estimation. J Stat Softw, 2008. **26**(5):1-15.
14. EFSA ホームページ Benchmark dose modelling approach updated.
<https://www.efsa.europa.eu/en/press/news/170124>
15. EFSA, Guidance of the Scientific Committee on a request from EFSA on the use of benchmark dose approach in risk assessment. EFSA Journal, 2009. **1150**:1-72.
16. EFSA, Update: use of the benchmark dose approach in risk assessment. EFSA Journal, 2017. **15**(1):4658.
17. WHO/IPCS, Principles for modelling dose-response for the risk assessment of chemicals. 2009.
18. 国立医薬品食品衛生研究所ホームページ BMDL シミュレーション結果データベース ベンチマークドース法の適用に関するガイダンス.
<http://dra4.nihs.go.jp/bmd>
19. JECFA (Joint FAO/WHO Expert committee on food Additives): Sixty-fourth meeting, Safety evaluation of certain contaminants in food. WHO Food Additives Series 55. 2006.

20. 食品安全委員会, 魚介類等に含まれるメチル水銀について. 2005.
21. 食品安全委員会, 化学物質・汚染物質評価書 食品中のヒ素. 2013.
22. 食品安全委員会, 評価書 加熱時に生じるアクリルアミド. 2016.