ベンチマークドーズ法に関するガイダンス及び
導入支援ソフトウェア等の現状調査
成果報告書

平成 **31** 年 **3** 月

一般財団法人 化学物質評価研究機構

目次

調査の概要 1

1. 調査の目的及び実施内容 2

1.1　調査の背景及び目的 2

1.2　調査方法 3

1.2.1　有識者の選定 3

2. BMD 法に関連するガイダンスの収集、整理 5

2.1　Principles for Modelling Dose-Response for the Risk Assessment of Chemicals (EHC 239) (WHO, 2009) 6

2.1.1　概要 6

2.1.2　用量反応モデリングの原則 7

2.1.3　BMD 及び BMR の選択 19

2.1.4　BMD 法の適用方針 22

2.2　Principles and Methods for the Risk Assessment of Chemicals in Food (EHC 240) chapter 5 Dose-Response Assessment and Derivation of Health-Based Guidance Values (FAO/WHO, 2009) 23

2.2.1　概要 23

2.2.2　HBGV 導出における BMD 法の手順 23

2.2.3　BMD 法の適用方針 26

2.3　Safety evaluation of certain contaminants in food (Food Additive Series: FAS No. 55) ANNEX 4 / ANNEX 5 (FAO/WHO/JECFA, 2006) 28

2.3.1　概要 28

2.3.2　BMD 法に関する手順 28

2.3.3　BMD 法の適用方針 30

2.4　Benchmark Dose Technical Guidance (EPA, 2012) 32

2.4.1 概要 32

2.4.2 BMD 法の手順 33

2.4.3 BMD 法の適用方針 49

2.5 Use of the benchmark dose approach in risk assessment Guidance of the Scientific Committee (EFSA, 2009) 51

2.5.1 概要 51

2.5.2 BMD 法に関する手順 52

2.5.3 その他 57

2.5.4 BMD 法の適用方針 58

2.6 Update: use of the benchmark dose approach in risk assessment (EFSA, 2017) 59

2.6.1 概要 59

2.6.2 BMD 法に関する手順 59

2.6.3 BMD 法の適用方針 71

2.7 The use of advanced risk assessment methods in answering various types of risk management questions (RIVM, 2006) 73

2.7.1 概要 73

2.7.2 BMD の適用方針 73

2.8 Valeurs toxicologiques de reference Guide d'élaboration de l'Anses (英訳: Toxicological Reference Values ANSES Development Guide) (ANSES, 2017) 76

2.8.1 概要 76

2.8.2 BMD 法の適用方針 76

2.9 Updating of the PNNS guidelines: revision of the food-based dietary guidelines (ANSES, 2016) 77

2.9.1 概要 77

2.9.2 汚染物質のリスク評価における BMDL の利用 77

2.10 Risk Analysis in Food Regulation (FSANZ, 2013) 78

2.10.1　概要 78
2.10.2　化学物質のハザード評価及びリスク評価における BMD 法の利用 78
2.11　BMD 法に関するガイダンスにおける BMD 適用方針の比較 79
3. BMD 法に関するソフトウェア等に関連する情報等の収集整理 81
3.1　BMDS (EPA) 82
3.1.1　名称及びバージョン 82
3.1.2　BMDS の概要 82
3.2　PROAST (RIVM) 113
3.2.1　名称及びバージョン 113
3.2.2　PROAST の概要 113
3.2.3　利用方法に関する情報 118
3.3　Benchmark Dose Modelling (Efsa) 151
3.3.1　名称及びバージョン 151
3.3.2　Benchmark Dose Modelling の概要 151
3.3.3　利用方法に関する情報 153
3.4　SPBS (秋田大学) 178
3.4.1　名称及びバージョン 178
3.4.2　SPBS の概要 178
3.1.1. 3.4.3　利用方法に関する情報 180
3.5　MADr-BMD 193
3.5.1　名称及びバージョン 193
3.5.2　MADr-BMD の概要 193
3.5.3　利用方法に関する情報 194
3.6　BMDExpress 207

3.6.1　名称及びバージョン 207

3.6.2　BMDExpress の概要 207

3.6.3　利用方法に関する情報 213

3.7　ソフトウェアの主要な機能の比較 223

4. BMD 法に関連する論文等の収集 225

5. 資料の和訳 226

調査の概要

本事業では、ベンチマークドーズ法 (以下、BMD 法) に関連するガイダンス等の情報収集を行うと共に、BMD 算出を支援するソフトウェアについて概要をまとめ、食品健康影響評価における BMD 法活用のための有益な基礎情報を得ることを目的に、(1) BMD 法に関連するガイダンス等の収集、整理、(2) BMD 法に関するソフトウェア等に関連する情報等の収集、整理、(3) BMD 法に関する文献の収集、(4) 資料の和訳を実施した。なお、調査の実施にあたっては、有識者から意見の聞き取りを行い、調査方針、内容、結果等を確認した。

調査対象とするガイダンスについては、BMD 法の活用を積極的に進めている国際機関や国で作成されたものを対象とし、ソフトウェアについては、上記の国際機関や国で多く利用されているものを中心に調査した。収集する論文については、最新の動向としてモデル平均化手法を扱ったものや、ベイズ法についての論文を重点的に収集した。

本調査により、BMD 法の利用に関する国際機関や国の方針を整理し、それらを比較整理すると共に、BMD ソフトウェアに収載されたモデルやパラメータ推定法、信頼区間の設定方法等の機能について整理することができた。これらの成果により、食品安全委員会が、今後 BMD 法を活用し、更に一貫性及び透明性の確保された評価を実施するための有用な情報を整理することができた。

1. 調査の目的及び実施内容

1.1 調査の背景及び目的

食品安全委員会における取組み

食品安全委員会では、評価技術企画ワーキンググループにおいて取りまとめた「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～」と題した報告書を 2018 年 7 月に公表し、BMD 法を活用する際の技術的な課題について議論の経過をとりまとめた。同報告書では、海外のリスク評価機関において、BMD 法を利用する際の考え方やモデル選択が不統一である現状を挙げ、わが国の食品健康影響評価において一貫性及び透明性を確保して BMD 法を活用するために一定の考え方を整理する必要があるとしている。また、2018 年 12 月に開催された第 12 回食品安全委員会評価技術企画ワーキンググループでは、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 (素案) として、動物試験で得られた用量反応データへの適用について審議が行われ、継続審議となっている。

本事業の目的

本事業では、BMD 法に関連する情報収集を行うと共に、BMD 算出を支援するソフトウェアについて概要をまとめ、食品健康影響評価における BMD 法活用のための有益な基礎情報を得ることを目的とした。

具体的な実施事項を、以下に示す。

(1) BMD 法に関連するガイダンス等の収集、整理

(2) BMD 法に関するソフトウェア等に関連する情報等の収集、整理

(3) BMD 法に関する文献の収集

(4) 資料の和訳

1.2　調査方法

1.2.1　有識者の選定

調査を実施するに当たり、毒性学及び化学物質のリスク評価方法、統計学及び BMD 法に関する有識者として、2 名の有識者を招聘した（表 2.2-1）。

表 1.2-1　検討会の有識者

専門分野	氏名	所属
統計学、数理モデルを利用した分析	西浦　博	北海道大学大学院医学研究員教授
毒性学、化学物質のﾘｽｸ評価方法	広瀬明彦	国立医薬品食品衛生研究所 安全性生物試験研究センター 安全性予測評価　部長

表 1.2-2　聞き取り調査での主な確認事項

検討会	主な協議事項
第 1 回 (2018 年 12 月 14 日面談)	● 本調査事業での調査方針について ● 調査対象とするガイダンス及び調査内容について ● 調査対象とするソフトウェア及び調査内容について ● 文献調査の対象とするデータベースについて ● 和訳対象とする文書について
第 2 回 (2019 年 2 月 7 日 mail での聞き取り)	● 文献調査の結果、収集する文献について
第 3 回 (2019 年　月 mail での聞き取り)	● ガイダンスの調査結果について ● ソフトウェアの調査結果について

検討会における主な指摘事項やコメント等を以下に示す。

・調査対象とするガイダンスについて

① 主要なガイダンスは網羅されており、問題ない。

② カナダや南アフリカ等も独自のガイダンスがあると聞いているが、文献等で引用されるものは網羅されており、問題ない。

③ モデル選定の基準や、信頼区間算出等における計算方法を整理した情報は有益なものとなる

・調査対象とするソフトウェアについて

④ MADr‐BMD、BMD Express も網羅されており、問題ない。

⑤ 信頼区間算出等におけるアルゴリズムの情報は有益である。

・文献調査対象とするデータベースについて

⑥ BMD 法の方法論自体を整理した論文の収集、整理を優先すべきである。

⑦ EPA 等で検討されている Bayesian Benchmark Dose Estimation について検索対象とする。

・和訳対象とする文書について

⑧ Benchmark dose software (BMDS) user manual version 2.6 には、収載モデル毎に、パラメータの意味や BMDL 等の計算方法といった詳細が記載されており、本資料全体を和訳する意義は大きい。

・文献調査の結果、収集する文献について

⑨ 最近 5 年間の大きな動き（Bayesian method やモデル平均化）に関してキーとなる文献のうち、重要なものは抜け落ちていない。

2. BMD 法に関連するガイダンスの収集、整理

下表に調査対象としたガイダンスを示す。

表 2-0 調査対象のガイダンス

作成機関	発行年	名称
世界保健機関 (WHO)	2009	Principles for Modelling Dose-Response for the Risk Assessment of Chemicals[1] (環境保健クライテリア (EHC) 239)
国際連合食糧農業機関 (FAO)/ WHO	2009	Principles and Methods for the Risk Assessment of Chemicals in Food (EHC240)[2] chapter 5 Dose-Response Assessment and Derivation of Health-Based Guidance Values
FAO/WHO 合同食品添加物専門家委員会 (JECFA)	2006	Safety evaluation of certain contaminants in food (Food Additive Series: FAS No. 55) ANNEX4 The Formulation of Advice on Compounds that are both Genotoxic and Carcinogenic / ANNEX5 Approach to dose-response modelling[3]
米国環境保護庁 (EPA)	2012	Benchmark Dose Technical Guidance[4]
欧州食品安全機関 (EFSA)	2009	Use of the benchmark dose approach in risk assessment Guidance of the Scientific Committee[5]
	2017	Update: use of the benchmark dose approach in risk assessment[6]
オランダ国立公衆衛生研究所 (RIVM)	2006	RIVM report 320016001/2006 The use of advanced risk assessment methods in answering various types of risk management questions
フランス食品環境労働衛生安全庁 (ANSES)	2017	Valeurs toxicologiques de reference Guide d'élaboration de l'Anses[7]
	2016	Updating of the PNNS guidelines: revision of the food-based dietary guidelines[8]
オーストラリア・ニュージーランド食品基準機関 (FSANZ)	2013	Risk Analysis in Food Regulation[9]

[1] http://www.inchem.org/documents/ehc/ehc/ehc239.pdf
[2] https://www.who.int/foodsafety/publications/chemical-food/en/
[3] http://whqlibdoc.who.int/publications/2006/9241660554_eng.pdf?ua=1
[4] https://www.epa.gov/sites/production/files/2015-01/documents/benchmark_dose_guidance.pdf
[5] http://www.efsa.europa.eu/en/efsajournal/pub/1150
[6] https://www.zora.uzh.ch/id/eprint/144774/1/Hardy_et_al-2017-EFSA_Journal.pdf
[7] https://www.anses.fr/fr/system/files/SUBSTANCES2017SA0016Ra.pdf
[8] https://www.anses.fr/fr/system/files/NUT2012SA0103Ra-1EN.pdf
[9] http://www.foodstandards.gov.au/publications/riskanalysisfoodregulation/Documents/risk-analysis-food-regulation-full-pdf.pdf

2.1 Principles for Modelling Dose-Response for the Risk Assessment of Chemicals (EHC 239) (WHO, 2009)

2.1.1 概要

本ガイダンスは、リスク評価者が、危害要因判定 (hazard characterization) の際に、用量反応モデリングを行うときの手引きとして世界保健機関 (WHO) が作成した環境保健クライテリア (EHC) である。用量反応モデリングに関して、その基本概念、原理・原則、利用の背景と利用場面 (NOAEL 法と BMD 法の比較を含む) 及びその結果をどのように活用し、並びにリスクコミュニケーションに活用するかが示されている。本ガイダンスの中で、BMD の算出は、用量反応モデリングにおける特に重要な活用の 1 つとされており、BMD 法に基づく ADI の算出の概要、利点、適用上の課題等が示されている。

用量反応モデリングに用いるデータは、作業負荷を軽減するため、利用可能なデータの中から、より鋭敏なエンドポイントのデータを選択する必要がある。使用するモデルの選択は、用いるデータの種類 (連続値データ、順序カテゴリーデータ、非連続データ又はカウントデータ (count data)) に依存するとともに、それらのモデルには、用量反応関係を表すモデルと、反応データのばらつきを表すモデルが含まれるとしている。そして、データセットにモデルをフィッティングさせた後、適合度検定の結果から、各モデルが用量反応データをどの程度説明できているかを評価した上で、更に AIC 等の指標を用いて、用量反応データを説明できている程度を相互に比較するとしている。

また、このようなモデルを用いた推測が有する不確実性には、大きく分けて、試験した標本内における反応量のばらつきに起因する統計学的な不確実性、試験条件等の違いに起因する試験間におけるばらつき、真の用量反応関係について異なるモデルを用いて様々な近似を行うことに起因する不確実性の 3 種類があるとした上で、用量反応解析では、可能な限り常に、これら 3 つの不確実性の要因全てについて対処する必要があるとしている。

そして、適切な用量反応データが利用可能な場合は、ADI 又は TDI 等の指標値を計算する際には、NOAEL 又は LOAEL よりも BMD を用いる方がより良い方法であるとしている。なお、外挿が必要な場合には、予測に伴うあらゆる不確実性を明示する必要があるとした上で、この不確実性の中にモデルに起因する不確実性を含めることは特に重要な場合が多いとしている。

以下に、本書に記載された用量反応モデリングの基本手順を整理するとともに、BMD 及び BMR の選択に関する記載を示す。なお、BMD 法の手順については、同じ EHC の「240 Principles and methods for the risk assessment of chemicals in food」第 5 章 (本報告書の 2.2 項参照) の記載と重複することから、詳細は 2.2 項で整理することとし、本項では、用量反応モデリング全般に関する基本手順を中心に整理した。

2.1.2　用量反応モデリングの原則

表 2.1-1 に用量反応モデリングの基本手順を示す。

表 2.1-1　用量反応モデリングの基本手順

ステップ	説明	選択肢
1. データの選択	モデリングする反応を決定し、適切な用量反応データを選択する。	エンドポイント、データの質、サンプルサイズ、データの実用性 (utility)、データの入手可能性
2. モデルの選択	データにフィッティングさせるモデルの種類を選択する。	エンドポイント、データの入手可能性、目的
3. 統計学的関連付け (statistical linkage)	用量反応データを説明する統計学的な確率分布を仮定する。	エンドポイント、データの種類、モデル選択、ソフトウェアの入手可能性
4. パラメータの推定	適切なコンピュータプログラムを用いて、最初の 3 ステップを統合・実施し、モデルのパラメータを推定する。	リンク関数 (linkage function)、ソフトウェアの入手可能性、分散
5. 解析の実行	推定したモデルのパラメータとモデル式を用いて反応量/用量を予測する。	結果の出力、解析対象の選択、モデル予測、BMD、直接外挿
6. 解析結果の評価	解析で用いた仮定に対する予測結果の感度を検証する。	モデルの比較、不確実性

以下に、各ステップの概要を示す。

(1)　ステップ 1：データの選択

(1)-1　データの選択

入手可能な毒性試験データで観察された各エンドポイントそれぞれに対して用量反応解析を行うことは効果的ではなく、最初のステップとして、他の試験と比べて明らかに大きい NOAEL を示す試験は除外することを検討する。例えば、最も感受性の高い種に対する、特定の毒性反応 (例: 慢性毒性、発生毒性) の種類を選択する。また、明確な用量反応を示さないエンドポイントは除外することが可能である。

続いて、反応量の見かけ上の大きさと毒性学的な影響度合いに基づき、モデリング対象の候補となるエンドポントを選択する。提出された試験データに、各エンドポイントにおける観察データポイントの (表に加えて) プロット、できればプロットにフィッテ

ィングさせた用量反応曲線があれば、エンドポイントの選択プロセスにおいて非常に有用である。

解析対象の候補となるエンドポイントを選択した後、それぞれの用量反応データセットにおいて用量反応解析が可能か否かを判断する。一般的に、少なくとも 3 又は 4 の異なる用量群 (対照群を含む) を有することが望ましい。さらに、影響のレベルが、各用量群で互いに異なっている必要があり、少なくとも 3 つの異なる反応レベルを有することが望ましい。

(1)-2　データの種類

反応に関する主な分類として、非連続データと連続値データの別がある。非連続データは、個別の試験個体 (試験動物又はヒト) で、反応が観察されたか、又は観察されなかったかに関するデータである。従って、各用量群における試験個体数の中から反応を示した個体数が報告される。連続値データは、個々の試験個体に関する定量的な測定結果である。両者の中間タイプのデータとして、(順序化された) 重症度のカテゴリーを反映した順序カテゴリーデータがあり、定性的データではあるが階層順になっている (例: 病理組織学的データ)。カテゴリーに順序がない場合は、カテゴリカルデータと呼ばれるが、用量反応データとしては稀である。カウントデータ (count data) は、最終的には別の種類のデータ (すなわち、離散データ) となるが、実際には連続値データとして扱われることが多い。

データの種類は統計学的な理由からも重要であるが、非連続データと連続値データとの区別は、解析結果の解釈及びリスク評価での利用の観点からも重要である。非連続用量反応データの場合、用量変化に伴う発現率の変化の情報は、ある特定の度合いの影響に対して得られる。例えば、口蓋裂の発現率は、用量の増加により上昇する可能性があるが、「非口蓋裂」と「口蓋裂」というカテゴリーには、影響の程度に関する情報はない。これに対して、順序カテゴリーデータと連続値データでは、影響の程度と発現率の両方が用量の関数として得られる。例えば、口蓋裂は、「裂け目なし」、「軽度の裂け目」、「中度の裂け目」、「重度の裂け目」のレベルを用いた順序カテゴリーからなる変数に分類することや、例えば閉鎖部分の割合といった連続値からなる変数を用いて定量的に扱うことも可能であろう。平均反応量と用量との関係から、ばく露により影響の程度がどれほど変化するかの情報が得られる。例えば、(平均) 赤血球数のプロットは、用量に対する関数として、平均赤血球数の減少 (すなわち、影響の程度) を示す場合がある。個々のデータポイントも検討することで、発現率の情報も同様に得られる。例えば、ある臨界値未満の赤血球数を有する個体の割合を推計することができる。

ヒトにおける反応を表すモデルとして、試験動物のデータを使用する場合、観察された用量反応の情報は、ヒトでの用量反応にある程度類似していると仮定される。この

仮定が、ある影響の発現率よりも影響の程度に対してより尤もらしいかどうかは、議論があるかもしれない。動物で観察される用量反応関係は、使用する動物のばらつきが大きく反映されるが、研究室で行う試験では、このばらつきが非常によくコントロールされていることから、動物試験で観察されるばらつきは、ヒトでのばらつきとは異なる可能性がある。

(2)　ステップ 2：モデルの選択

(2)-1　連続値データからなる用量反応データのモデル

個々の試験個体における用量と連続値スケールからなる反応量の大きさとの関係を説明する際に使用される数理モデルの一部を表 2.2-2 に示す。これらの数式は、統計学的な確率分布 (例えば、正規分布や対数正規分布) と組み合わせることで、反応量の推定中央値に対する連続値モデルとして、ある集団における用量と連続反応量との関係を説明することにも使用できる。

用量反応データは、個別の試験個体における観察値から、対照群の (平均) 値を差し引くことにより調整されることが多い。しかし、この手順は、対照群における反応レベルにおいて、試験群における反応レベルと同様にサンプリングエラーが生じるという事実を検討していない。より良いアプローチは、モデルのバックグラウンド時の反応を、データから推定されるパラメータを用いて説明する方法である。このための方法のうち、最も簡単な 3 つのモデルを以下に示す。

1. $y = a + f(D)$
2. $y = a \times f(D)$
3. $y = f(a + D)$

ここで、D は用量、a はバックグラウンドパラメータ、f は用量反応関数である。毒性メカニズムに関する情報から、適切な式を判断できる場合がある。例えば、上記のモデル 1. は、独立作用を生じる反応のモデリングに適しており、モデル 2. はバックグラウンド反応の一部分として反応を標準化する考えに対応しており、モデル 3. は、同一のメカニズムによって作用する他の物質の寄与を反映している。

表 2.2-2　連続値データからなる用量反応データを説明するモデル

モデル名	モデルの説明	反応式	パラメータの説明
Michaelis-Menten law of mass action	酵素又は受容体活性に関し、その作用速度が、結合速度 (ka) と解離速度 (kd) の関数であるとする理論的説明。	$=R_{Max}\frac{[S]}{K_M+[S]}$	R_{Max}: 最大反応速度 [S]: 基質濃度 K_M: Michaelis-Menten 係数 (この係数は ka/kd に等しい)
Hill equation log-logistic (Hill の式 対数ロジスティック)	作用の発現には、複数の部位又は受容体の占有が必要であるとする、Michaelis-Menten 式の変形式。	$=R_{Max}\frac{[D]^n}{K_D^n+[D]^n}$	R_{Max}: 最大反応速度 D: 用量 K_D: 薬物－受容体相互作用 n: (理論的) 結合部位数
First-order exponential (一次指数関数)	化学物質と標的部位間の相互作用が不可逆な場合、反応速度は結合速度 (ka) のみによって決定される。	$= R_{Max}(1-e^{-rD})$	R_{Max}: 最大反応速度 D: 用量 r: 指数関数的速度定数
Power	単純な指数関数モデル。	$=\beta D\alpha$	D: 用量 α: 形状パラメータ β: スケールパラメータ
Linear	通常、本モデルが提唱される生物学的理論はないが、線形モデルはその簡潔さから正当化されることが多い。	＝Dm	D: 用量 m: 傾き

(2)-2　非連続データからなる用量反応データのモデル

非連続データからなる用量反応の関数は、1 つの集団における用量と特定のアウトカムの発現頻度との関係を説明する (表 2.2-3)。均一又は同一に近い個体群では、用量と頻度の関係は、与えられたある用量において、全ての個体が反応するか、反応しないかというステップ関数で説明できる。しかしながら、生物間にはばらつきが遍在しているので、非連続データからなる用量反応データは、通常、用量の増加とともに発現率が

徐々に増加する。これに対する 1 つの解釈は、個体間で、評価物質への耐性が異なることであり、これが、統計学的な許容限界の分布 (statistical tolerance distribution) によって示される。従って、いずれの累積分布関数も、非連続データからなる用量反応の関数として使用できるだろう。他のモデルは、gamma multi-hit model のように、物質が生体内でどのように影響を生じさせるかについての統計学的な仮定から導かれてきた。

バックグラウンド時の反応率は、連続値データの場合のように、用量反応モデルに追加のパラメータを組み入れることで説明される。この場合の最も簡潔なものは以下の形である。

1. y=a+(1–a)f(x)
2. y=f(x+a)

ここで、f(x)は、任意の用量反応関数 (0 から 1) である。連続値データのように、用量反応解析に先立ちバックグラウンド時における反応のデータを収集することは、統計学的な根拠は希薄である。バックグラウンド時の反応レベルは、用量反応モデルによって同時に推定されるべきであって、他の用量群で観察された反応と同様に取り扱われるべきである。

表 2.2-3　非連続データからなる用量反応データを説明するモデル

モデル名	基本理論	頻度 (F) に対する式	パラメータの説明
Step function (ステップ関数)	変数なし。	D<T の場合、F=0 D≥T の場合、F=1	D: 用量 T: 閾値パラメータ
One-hit (single-hit)	ヒット理論モデルは、要因 (例: 分子) と標的 (例: ヒト集団) 間の相互作用を説明するために rate (速度) を採用する。	$=1-e^{-(\alpha+\beta D)}$	D: 用量 e: Euler 定数 α: 位置パラメータ β: 傾きパラメータ
Gamma multi-hit	One-hit model の拡張型であり、特定の影響が生じるためには、複数のヒット又は事象が必要との考えに基づく。	=Γ(gamma*D, k)	Γ ( ): 不完全 gamma 累積分布関数 D: 用量 gamma: rate (速度) パラメータ k: 特定の影響が生じるのに必要なヒット数
Probit normal	正規分布又はガウス	=Φ (α+D*β)	Φ ( ): 正規累積分布関

モデル名	基本理論	頻度 (F) に対する式	パラメータの説明
	分布に基づく記述モデル。		数 D: 用量 α: 位置パラメータ β: 傾きパラメータ
Logistic	統計学的なロジスティックモデルも、理論的な基礎を有さない記述ツールである。	$= \frac{1}{1+e^{-\alpha-D\times\beta}}$	D: 用量 α: 位置パラメータ β: 傾きパラメータ
Weibull	本来は、人口統計学において生存データを示すために開発されたフレシキブルな記述モデル。	$=e^{-(\alpha+(\beta\times D)\gamma)}$	D: 用量 α: バックグラウンドパラメータ β: 傾きパラメータ γ: 指数

(2)-3　閾値

「閾値」という用語は、3 つの異なる意味で使用される。1 つ目は、影響が起こらないばく露レベルを示す科学的な意味合いがある (例: 物理的刺激はあるが反応はない)。2 つ目として、閾値は、影響が起こるか起こらないかのレベルであるが、観察するにはあまりにも小さ過ぎる (例: NOAEL) レベルであると考えられる。この場合は、実際の試験の対象というよりも、観察者又は解析者の知覚的な限界を示すものである。3 つ目である「実際の閾値」は、影響が些細であり更なる検討は不要と判断される反応である。用量反応モデルには、1 つ目の意味の閾値を組み込むことができるであろう。閾値パラメータを用い、閾値用量における用量反応関係を以下のように切り捨てる。

・閾値未満では、影響を及ぼす用量はゼロである。

・閾値を超えると、影響を及ぼす用量は、当該用量から閾値を引いた値である。

一般的に、閾値パラメータを正確に推定することは難しく、大きな信頼限界を有する。

(2)-4　重症度 (影響の程度)

用量反応モデリングにおいて、毒性反応の重症度が定性的な手法 (例: 腫瘍形成 vs 受胎能の低下) 以外で使用されることは稀である。しかし、単一エンドポイントのレベルでは、反応の重症度や程度を定量的に捉えることもある。上述のように、連続値で表されるエンドポイントの用量反応、例えばヘマトクリット値の減少率や体重の変化率は、用量に対応した影響の程度の変化として直接解釈される場合がある。ここで、BMD を

算出するために、特定の影響の程度 (例えば体重の 5%低下) を選択すると、BMDL は、当該エンドポイントにおける影響の程度に関して特定 (例: 5%) の変化をもたらす用量として定義付けされる。

連続値で表される反応データを基に、影響の程度の観点から BMR を設定する場合の重要な利点は、有害でないと考えられる BMR の値が観察範囲内にある、又は観察範囲に近接することである。そのため、連続値で表されるエンドポイントについて検討した場合、低用量側への外挿は必要ないか、あってもその外挿の程度は僅かであると考えられる。

順序カテゴリーデータが得られる病理組織学的なエンドポイントの場合、用量反応関数は、カテゴリカル回帰によってフィッティングされ、特定の影響の程度 (例: 最小又は軽度) に対応した BMDL が推計される。

カテゴリカル回帰をより高いレベル、つまり複数試験の統合解析に適用することもある。カテゴリカル回帰をこのように適用する場合、重症度のカテゴリーが異なるエンドポイントをカバーするように設定する。これらの適用例のほとんどが、観察用量レベルにおいて、与えられた重症度のカテゴリーが発生する尤度の推定に焦点を当てている。

(2)-5　共変量を伴うモデリング

状況によっては、ばく露量に係る変数に加えて、追加の変数を用量反応モデルに含めることが望ましい。例として、疫学研究では、ばく露だけでなく、年齢、性別、社会経済的状況、喫煙状況及び病状に関連する可能性のあるその他の測定結果の観点から見た疾病リスクをモデリングすることが一般的である。これらのばく露以外の要因は、サンプリング方法に起因してばく露の状況と相互関係を示す可能性がある。その際、適切な共変量がばく露と健康に関するエンドポイントとの関係を表すモデルに含まれていない場合は、ばく露の影響は正しく推定されない。試験動物を各投与群にランダムに割り振る動物試験では、こういった交絡は原則的に生じないが、測定に関するばらつきの一部を説明するために、性別等の共変量を含めることが有用な場合がある。

(2)-6　生物学に基づいた用量反応モデル

生物学に基づいた用量反応モデルは、統計学的なモデルと比較してはるかに複雑であり、毒性物質の初回ばく露から最終的な病理学的アウトカムに至る生物学的な詳細をモデリングするために、明確にデザインされている。通常、このようなモデルは、親化合物と毒性を有する代謝物の分布と代謝やその他のメカニズムを説明する生理学的トキシコキイネティクス (毒物動態) モデル及び標的組織中濃度と最終的な反応を関連付けるトキシコダイナミクス (毒物力学的) モデルを含んでいる。モデルの毒物力学的な部分は、比較的単純な場合 (例: クロルピリホスのモデルにおいて、アウトカムがアセチル

コリンエステラーゼ阻害)、又は発がんに関する完全かつ精緻な確率論的モデルのように複雑な場合がある。このようなモデルは、実際に、一連の生物学的仮説を定量的に表現したものであって、重要な試験が厳密に実施された場合に、試験結果を、コントロールされた試験下では再現が困難又はコストを要するばく露領域に外挿する際に信頼のおけるツールとなる。このようなモデルは、投入する資源や時間の両面から構築にかなり費用がかかるため、懸念の程度が最も大きいばく露と毒性に関してのみ、十分な開発が期待される。

(3) ステップ 3：統計学的分布

(3)-1 連続型分布

正規分布、ガウス分布は対称な形をとり、マイナス無限大からプラス無限大まで定義される。両分布とも、平均と標準偏差の 2 つのパラメータを有し、それぞれ、分布の配置とスケールを定めている。多数の小さな影響からなる集合は、概ね正規分布となる傾向があることから、ばらつきと測定誤差の変動を表現する際に正規分布が用いられることが多い。

対数正規分布は、幾何平均と幾何標準偏差の 2 つのパラメータを有する。観察又は予測される値の対数が正規分布となることが仮定された、正規分布の派生物と見なされる。対数正規分布は、対数変換前のスケールでは非対称の形状の分布 (skewed distribution) を形成する。対数正規分布を用いることでもたらされるもう一つの重要な効果は、負の値が生じないことであり、このため対数正規分布は、正の値のみからなるデータセットの表現により適している一方、負の値を有するデータセットには適さない。多くの分布は非対称の形状であって正の値のみ含んでいるため、対数正規分布によってうまく説明できることが多い。さらに、多数の小さな影響の積は、ほぼ対数正規分布する傾向にある。生物学的な測定値で表される影響は、相加的よりも、むしろ相乗的(比例的) となる傾向があるため、一般的に、対数正規分布は生物学的な測定値 (の表現) に適している。

Weibull 分布は、状況によっては、物理的システム/製品や生物学的システムの生存率又は「寿命」の分布を表すのに最も一般的に使用されている。多くの適用事例で、Weibull 分布が適切である、又は使用されるべきとする明確な理論的根拠はないが、極値分布を表す分布クラスにおいて何らかの理論的な裏付けを有する。曲線適合性の観点から、分布の関数式は、単純に指数関数モデルのべき乗変換であり、そのモデルのデータ説明に柔軟性を与える。Multi-hit model は Weibull model の特殊な事例である。

(3)-2 離散型分布

離散型分布は、有限又は無限のスケールで、好ましくはカウントデータ (count data) で反応を説明する。特別な例として 0 又は 1 の二値の非連続的アウトカムとなる反応が

ある。

ベルヌーイ分布のアウトカムは 1 又は 0 であり、これは、無限に連続する変数からなる試行において頻度 f で事象が発生する、又は発生しないという結果に対応するものである。このためベルヌーイ分布では、単純に頻度 f で「1」、頻度 1-f で「0」となる。ベルヌーイ試行は二項分布の基礎であり、二項分布の定義にはベルヌーイ試行が含まれる。

二項分布は、指定された数のベルヌーイ試行の合計の分布として定義され、特定の事象が発生する又は発生しない場合をそれぞれ 1 又は 0 として示す。毒性学に適用する場合、試行数は試験デザインによって決まり、特定の事象が発生する対象集団の割合を推定する。このため、二項分布は、非連続データからなる用量反応を示すモデルのパラメータを推定するときに通常使用される分布である。

ポアソン分布は、正の離散値からなる反応についての 1 パラメータ分布である。反応変数の領域は任意の正の整数である。この分布は元々、稀な事象の分布として導き出された。具体的には、試行数が多く、1 試行当たりの事象の確率 (P) が低い場合における一連のベルヌーイ試行で発生する事象数 (n) である。従って、ポアソン分布は、n が大きく、P が小さい場合、二項分布の近似として使用することができる。ポアソン分布は、ある期間にわたって対象コホートを前向きに追跡調査するような試験デザインであり、有害事象の発現率がコホートの規模と比較して小さいと予想される場合、疫学データの解析に一般的に用いられる。

(4) ステップ 4：モデルのフィッティングとパラメータの推定

モデルのフィッティングには基本的に 2 つの方法がある。1 つの方法は、目的関数が最小又は最大になるようにパラメータを選択する従来法である。もう 1 つの方法はベイズ推定であり、ベイズ推定ではデータセットの情報をモデルパラメータに関する事前情報と組み合わせ、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。歴史的、計算上の理由から、用量反応解析と非線形モデリングを実行するために設計された「ユーザーフレンドリーな」ソフトウェアは、従来法を用いたものに限られていた。一方、ベイズ推定は、より広範なプログラミングと統計学の詳細をより十分に理解する必要があるソフトウェアパッケージに実装されている。そのようなソフトウェアでは、ベイズ推定をうまく使用するために十分な統計学の理解が必要であり、本文書の範囲外である。従来法を用いる場合であっても、ソフトウェアを適用した結果を適切に解釈するためには、基本的原理を理解する必要がある。いくつかの一般的な注釈を以下に示す。

(4)-1　評価関数 (Criterion function)

モデルフィッティングの一般的な方法は、データに最適にフィットするモデルのパラメータ値を見いだすことである。そのために、モデルのフィッティングを反映する評

価関数 (criterion function) を定義する。目標は、評価基準値を最適化するパラメータ値を見いだすことである。通常使用される多くのモデルでは、このようなパラメータ値は、「試行錯誤」を繰り返しながらでないと見いだせない。

多くのアプリケーションでは、尤度関数の対数が評価基準として使用される。尤度は、データ内の分散に対して仮定された分布から直接得られる。非連続データでは、通常、二項尤度が使用される。連続値データでは、観察される反応量自体又は対数変換した反応量に対して、正規尤度が使用されることが多い。正規分布をとると仮定されるデータの尤度関数を最大にすることは、実際には二乗和を最小化することに匹敵する。

(4)-2　検索アルゴリズム

コンピュータソフトウェアは、モデルのデータへのフィッティングを最適化するパラメータ値を見いだすアルゴリズムを用いているので、ユーザーは計算の実際の状況を懸念する必要はない。しかし、結果を解釈するためは、検索プロセスについて、ある程度の基本的理解が必要となる。

反復検索アルゴリズムは、試行錯誤の過程でパラメータ値を変更することによりフィッティングが改善されるかどうかを評価するプロセスを通して、「より良い」パラメータ値を見いだそうと試みる。より高度なアルゴリズムは、複数のパラメータ値の変更に対してフィッティングが改善されるときの尤度の勾配を評価することによって機能する (基本的には、尤度が最大となる尤度関数上の極大値を見いだす)。アルゴリズムは、パラメータの開始値 (starting value) がある場合のみ、検索を開始できる。ソフトウェアは多くの場合、開始値 (starting value) を合理的に推定するが、ユーザーが開始値 (starting value) を変更しなければならない場合もあるであろう。最終的な推定結果が選択した開始値 (starting value) に依存することは珍しいことではなく (特に、データ内の情報が目的とするパラメータを推定するのに全く十分ではない場合)、ユーザーはこのことを認識する必要がある。

アルゴリズムは、中断基準が満たされるまでパラメータ値を変更し続ける。アルゴリズムが検索プロセスを中断する理由は、主に以下の 2 つである。

1. アルゴリズムが収束した (例: 対数尤度関数において明らかな最大値を見いだした) 場合。この場合、関連するパラメータ値は、尤度が最大化された場合、「最適な」推定値 (例: 最尤推定値) と見なせる。しかし、対数尤度関数は 1 つではなく複数の (局所的な) 極大値を持つ場合がある。このため、開始値 (starting value) を変更してアルゴリズムを再度実行すると、推定結果が異なる可能性がある。
2. アルゴリズムが収束しなかった場合 (すなわち、尤度関数において明らかな最適値を見いだすことはできなかったが、反復 (試行) 回数が最大値を超えたため中断する場合) 。関連するモデルが、データからかなり乖離しているなど、開始値 (starting value) が適切に選択されなかった場合や、データ内の情報が推定対象のパ

ラメータ数に比べて少ない場合に起こる可能性がある。例えば、5 つの未知パラメータを含む用量反応モデルは、4 つの用量群で実施された試験結果からは推定できない。別の例として、用量群内の観察値間のばらつきが用量反応の全体の変化に比べて大きい可能性が挙げられる。このような場合、尤度関数は非常に平坦となり、アルゴリズムは、関数の増加と減少の間で変化する点を見いだすことができない。実際にはモデル推定結果が変わらない場合であっても、パラメータ推定値間の高い相関関係からこうした状況を認識できる場合がある (つまり、1 つのパラメータ値の変化が別のパラメータの変化により補正されることがある)。

(5)　ステップ 5 : 解析の実行

推定したモデルパラメータとモデル式を用いて反応/用量の予測を実行する。

(6)　ステップ 6 : 解析結果の評価

(6)-1　モデルの比較

モデルを評価する基本的な基準は、選択したモデルが、特に推定が必要とされる用量反応領域において、データをよく説明していることである。ほとんどのフィッティング方法で、全体的な適合度の尺度が得られ、通常は p 値が得られる。これらの尺度は、モデルによる予測結果がどの程度データと一致しているかを定量化している。p 値が小さい場合は、データへのフィッティングが良くないことを示す。データが適切に説明されることが特に重要であるため、適合度の臨界値の算出において $p = 0.05$ 又は 0.01 という一般的な値ではなく、$p = 0.1$ を使用することを推奨する。

用量反応データからのずれを見いだすもう一つの方法として、グラフの目視がある。プロットにより常に適合度検定を補足しなければならない。連続値データのプロットでは、各データポイントのばらつきの尺度も含めてプロットすることが非常に役立つ。データが単調用量反応でない場合や、ある低用量ではバックグラウンド反応のみしか示さないものの、次の用量で反応が急激に増加する場合のような観察データには、用量反応モデリングに用いられる典型的なモデルは適合しない。このような場合、データの調整 (例: 用量の変換) 又はモデルの調整 (例: ばく露と無関係な死亡に対する調整) が有用であり得る。

多くの異なるモデルを同じデータにフィッティングさせる場合、全てのモデルで同じフィッティングが得られるわけではなく、モデルを選択する際には注意が必要である。この課題に統計理論を当てはめた場合、以下の 4 つの状況のうち 1 つが起こり得る。

1. 同一分布族でネストした一連のモデル:

フルモデル (full model) が存在するという観点で、フルモデル (full model) について、より多くのパラメータに固定値を設定する、又は逆に、より多くのパラメータ

をモデル内に組み込むことにより、フルモデル (full model) から「制限された」モデルを導き出す。尤度比検定により、追加のパラメータを推定することによってフィッティングが改善するかを評価することができる。尤度比検定では、$2\times(LL_{full}-LL_{restricted})$ (ここで、LL は対数尤度) を算出し、自由度 $P_{full}-P_{restricted}$ (Px はモデル x の推定パラメータ数) のカイ二乗分布の臨界値と比較する。

2. 同一分布族のモデルであるが、ネストしたモデルではない:

モデルの比較にいくつかの統計量、特に赤池情報量規準 (AIC)[10] を使用できる。この場合、最小の AIC 値を有するモデルを選択するが、AIC 値が近い (差がおよそ 4 以下の) モデルはおそらく同等である。

3. 同一分布族のモデルではないが、確率分布の仮定が同じである (例: 全て対数正規分布を用いる、又は全て正規分布を用いる):

この場合も、最良モデルを識別するために AIC を利用できるとの意見もあるが、これらのモデルは非常に似た構造を有し、類似したパラメータ数を含むことから、AIC に基づいて一般的に使用される非連続データの用量反応モデルを差別化することは困難との意見もあり、いまだに統計学的研究上の課題となっている。現時点では、本ガイドラインが変更される可能性を認識した上で、AIC を用いて前述の事例のようにモデルを選択することでおそらく十分だろう。

4. モデルの確率分布の仮定が異なる:

この場合、統計学的な指針はほとんどなく、個体データの分布からデータの分布の仮定が尤もらしいかを検討する必要がある。しかし、連続値データは平均及び標準偏差に要約されて報告されることが多く、この場合は、分布の仮定を検討できない。このような場合は、モデリング対象のエンドポイントにおける過去の経験に基づき、妥当な確率分布を選択することが最善策となる。

(6)-2　不確実性の説明

与えられたモデルから推定されるパラメータや予測結果は推定値に過ぎず、多かれ少なかれ不確実である。この不確実性は少なくとも 3 つの原因から生じる。

1. サンプリングエラー: 単一の試験からより大きな母集団を推測することで生じる誤差。
2. 試験誤差: 試験デザイン、プロトコールの違い、又は管理されていない状況により生じる誤差。
3. モデルの誤差:「真の」モデルが未知であることによる、各用量間の内挿に伴う不確実性。観察結果のある用量範囲外に外挿する場合は、さらに不確実性が増す。

[10] AIC=-2LL+2P で定義され、LL はパラメータに対する最尤推定値における対数尤度、P はモデルの自由度

モデルの不確実性は、特に低用量外挿において問題となる。リスク評価においてモデルの不確実性を要約して表す方法の 1 つとして、確率ツリー (probability tree) を利用する方法がある。例えば、3 つのモデルについて同等に十分な裏付けがある場合、各モデルの確率は 0.33 となる。1 つのモデルの重みが他のモデルより 6 倍高い場合、そのモデルの確率は 0.75 となり、他のモデルの確率は 0.125 となる。モデルの確率は、根拠による裏付けの強さのみに拠るのでなく、他にどのようなモデルが検討されているかにも拠ることに注意する必要がある。裏付けの少ないモデルでも、他の全てのモデルの裏付けがより少ない場合、確率が高くなる可能性がある。モデルの優先度の定量的指標を組み合わせて、全体的なランクを作成する、又は、証拠の重み (weight of evidence) の正式な指標とすることができる。

モデルに確率を割り当てる定量的方法は全て、検証できない仮定又は判断要素にある程度依存する。このため、モデルに確率を割り当てる最も簡単でわかりやすい方法は、単純に全てのモデルに同じ重みを与えることである。種々のモデルの予測結果が単純に列挙されている場合、この方法がとられる。もう一つの比較的簡単な方法は、専門家判断により尤もらしい理論を特定し、それらの理論に確率を割り当てる方法である。これらの割り当てられた確率は、ベイズ推定によりデータへ追加情報を組み込むことで更新することができる。しかしながら、モデルに重み又は確率を割り当てる方法は数多くある。モデルの選択基準と同じ基準を用い、検討対象とする各代替モデルに重み付けや確率の割り当てを行うといった略式の方法を利用することもできる。ブートストラップ法を使用することにより、モデルの不確実性とサンプリングエラーを統合することもできる。ブートストラップ法では、データセットから繰り返し無作為抽出を行い、各データセットを一連のモデルに再度フィッティングさせる。各ブートストラップから得られる最適モデルがパラメータの不確実性とモデルの不確実性の両方を表す確率ツリー (probability tree) に残される。

あるいは、データに十分に適合すると考えられるモデルのサブセットを選択することで、不確実性に対処する場合がある。データに十分に適合すると考えられるモデルであり十分な適合度を有するモデルを、範囲及び分散と共にまとめる。最終的な BMD 値を選択する際、これらの値の平均又は幾何平均をとって総計し、中心点推定を行うか、又は専門家判断により 1 つの値を選択する。

2.1.3 BMD 及び BMR の選択

用量反応モデリングの重要な活用方法の 1 つが BMD の算出である。BMD は、ある特定の、あらかじめ指定されたレベルの反応が起こると推定される用量である。この方法論

は、ADI などの量を決定するための用量反応評価において、NOAEL 及び LOAEL の使用に代わるものとして導入された。NOAEL 及び LOAEL の代わりに BMD を利用する主な利点は、BMD 法では用量反応データをより完全に使用している点と、BMD の不確実性は、統計学的方法論を用いて定量化できるという点である。BMD の不確実性は、信頼区間として表すことができ、片側 95%信頼区間の下限値を BMDL と呼ぶ。また、BMD の不確実性は、ベイズ推定に基づく事後確率分布として表すこともできる。

BMR は、対応する BMD が計算される反応レベルである。本文書では、BMR の選択における技術的課題、BMR をどこに設定すべきかという方針を決定することの重要性について説明するが、特定の値の選択については直接的には取り上げない。

(1)　二値エンドポイントの BMR

BMR の表現方法は、モデル化される反応変数の種類によって決まる。二つの状態を有する (影響を受けた／受けなかった) エンドポイントの場合、BMR は通常、バックグラウンドについて補正する方法で表される。以下の 2 つの方程式が一般的である。

1 つ目は、追加リスク (added risk (AR)) を表す式である。

$BMR_{AR} = f(BMD) - f(0)$

ここで、f(x)は用量 x で評価される用量反応関数である。

もう 1 つは、おそらく最も広く利用されている過剰リスク (extra risk (ER)) を表す式であり、追加リスクを、ばく露されていない集団における影響を受けていない個体の割合で除したものである。

$$BMR_{ER} = \frac{f(BMD) - f(0)}{1 - f(0)}$$

また、3 つ目として、疫学研究で広く用いられ、動物試験にも適用できる相対リスク (relative risk (RR)) がある。

$$BMR_{RR} = f(BMD) / f(0)$$

(2)　連続値エンドポイントの BMR

連続値エンドポイントの BMR は、平均反応レベルの変化量をもって直接的に、又はある臨界値を超える (又は下回る) 試験動物の割合をもって間接的に表すことができる。例えば、成人の平均体重で考えると、平均体重が対照群の体重の 90%未満に低下した際の用量を、又は脳中アセチルコリンエステラーゼ活性が対照レベルに対して 10%抑制される用量 (多くの場合、この用量が臨界影響サイズと呼ばれる) を BMD として選択できる。ま

た、例えば、神経伝導速度の平均値が一定の速度を下回る、又はばく露されていない集団における速度と比較して一定の差が生じる用量を選び、一定の値又は平均値における一定の減少値を指定してもよい。酵素誘導のようにS字状の反応を示すエンドポイントについては、過剰リスクと同様の式を用いることが提案されている。これらのエンドポイントでは、BMDは、反応のダイナミックレンジ (例: バックグラウンドと推定される最大誘導レベルとの差) 全体に対して特定の比率に該当する反応をもたらす用量として最もよく特徴付けられる。反応の定義において、ある特定の設定 (つまり、1%の変化) を検討し、得られたBMDの不確実性と「ハイブリッド」法において特定の条件で推定されたBMDの不確実性を比較した事例があるが、得られた結論は一般的な条件 (例: ダイナミックレンジ全体に対して反応の5%又は10%の変化を検討する) では成り立たない可能性がある 。

間接法 (「ハイブリッド」法) では、直接的アプローチと同様の方法で、連続値からなる変数 (連続変数) の平均と用量との関係をモデリングする。次に、有害と考えられる変数の臨界値を決定し、BMDの計算に用いる過剰 (又は追加) リスクで表現されるBMRを選択する。臨界値は生物学的考察に基づくことが望ましいが、そうでなければ、対照群における変数の分布の末端部の値とすることもできる。平均反応が増加するにつれ、事前に決定した臨界値を超える個体の割合も増加する。BMDは、あらかじめ設定した臨界値を超える個体の割合が、非連続データと同様に定義したBMR (例: BMR_{ER}) に達する用量である。

連続変数のBMRを、動物試験で使用される動物のばらつきに基づいて特定する方法もある。対照群における「小さな」(例: 0.1～2%) リスク及び10%程度の過剰リスクに対応する臨界値でのBMDを近似により求めることができ、このBMDは、おおよそ、反応変数の平均値を対照群の標準偏差に等しい量だけ対照群の平均値からずらす用量に対応する。

上述の、ばらつきに基づく2つのハイブリッド法は、ばらつきが真の個体間のばらつきであり、大きな測定誤差によるものではないことが前提となる。これらの方法は、ばく露されていない集団末端にある個体は、ばく露集団で同分位点と同じ値を示した個体と同様の影響を受けるとみなされるという仮定に大きく依存している。

用量反応モデルにおいて、用量が唯一の独立変数ではない場合もある。例えば、疫学調査では、ばく露の影響を推定する際のバイアスを軽減するため、多くの共変量を解析に含めることが多い。発生毒性試験では、モデリングにおいて母動物又は同腹児の全体としての特徴 (例: 着床数) を共変量として使用することで、このような試験でみられることが多い腹毎のさらなるばらつきを説明するのに役立つ可能性がある。成獣のみのげっ歯類の動物試験でも、通常は雄と雌で分けて考える。このような場合、評価者は通常、どの共変量

に基づき BMD 値を計算するかを決定しなければならない。共変量が離散値であってとりうる値がほとんどない場合は、各データセットについて、別々の BMD (例: 雄と雌のサブグループそれぞれの BMD) を算出することも合理的である。動物試験において、共変量が連続値 (又は着床数のように、共変量を連続値として取り扱う) の場合、通常は、対照群において典型的な値を選ぶ。しかし、BMD が連続変数の値と共に変化する場合、依存性について詳細に解析を行う必要がある (例: 当該共変量の関数として BMD をモデリングする)。当該変数がヒトへの外挿において意味がある場合、いくつかの共変量値について BMD を計算することは、ヒトにおける共変量値の範囲における BMD の検出感度を評価する上で有益であろう。

2.1.4 BMD 法の適用方針

以上を踏まえ、EHC 239 (WHO, 2009) における BMD 法の適用方針のポイントを以下に整理する。

- BMD 法の適用手順として、データの選択、モデルの選択、統計学的関連付け (statistical linkage)、パラメータの推定、解析の実行、解析結果の評価の 6 つのステップが示されている。
- データの種類は、連続値データ、非連続データ、順序カテゴリーデータ、カテゴリカルデータ、カウントデータ (count data) に整理されている。
- BMD 法におけるデータの選択では、少なくとも 3、可能であれば 4 以上の異なる用量 (対照群を含む) が設定され、明らかな用量反応関係を示すデータが望ましい。さらに、各用量群で影響レベルが異なっている必要があり、少なくとも 3 つの異なる反応レベルを有することが望ましい。
- BMR の具体的な値については言及されていない。
- モデルの適合度の基準として、適合度検定の $p > 0.1$ を用いるとともに、目視によるプロットの確認を行うことが推奨されている。
- モデルの比較では、比較を行うモデルの種類や分布の仮定に応じて、尤度比検定や AIC を用いることが推奨されている。
- モデルのフィッティングとパラメータ推定の中でベイズ推定について触れられているが、本ガイダンスの対象外とされている。

2.2 Principles and Methods for the Risk Assessment of Chemicals in Food (EHC 240) chapter 5 Dose-Response Assessment and Derivation of Health-Based Guidance Values (FAO/WHO, 2009)

2.2.1 概要

EHC240 は、FAO/WHO 合同食品添加物専門家会議 (JECFA) 及び FAO/WHO 合同残留農薬専門家委員会 (JMPR) が実施する食品中の化学物質の評価におけるリスク評価手法に関するモノグラフとして策定された食品添加物及び食品中の汚染物質の安全性評価の原則に関する EHC (EHC 70) 及び食品中の残留農薬の毒性評価の原則に関する EHC (EHC 104) について、近年の分析、毒性評価、リスク評価手法の進歩を反映させ改訂したものである。EHC240 は、化学物質の同定、危害要因特定 (hazard identification) 、危害要因判定 (hazard characterization)、用量反応評価、健康影響に基づく指標値 (health-based guidance value (HBGV)) の導出等の幅広い項目について記載している。BMD については「第 5 章 用量反応評価及び HBGV」の中で NOAEL 法に代わる手法として紹介されており、基本的に前項 2.1.で整理した ECH 239 の内容と同等のものとなっている。以下に、EHC 240 の 5.2.5 章に記載された HBGV 導出のための BMD 法の手順を示す。

2.2.2 HBGV 導出における BMD 法の手順

NOAEL 法を代替する方法として、BMD の概念が紹介されてきた。NOAEL 法とは対照的に、BMD 法では、リスク評価の出発点 (Point of Departure (POD)) として、影響サイズ又は反応レベルがゼロではないばく露レベルを定義する。BMD 法には、統計解析において全ての用量反応データを用いることによりデータ内の不確実性の定量化が可能となるなど、いくつかの利点がある。データの不確実性が高い場合、例えば、試験群のサイズが小さい場合や群内のばらつきが大きい場合などは、HBGV が低くなる。

非連続データを例に、Weibull model を用いた BMD 法に基づく HBGV の導出方法と NOAEL 法に基づく HBGV の算出方法との比較を表 2.2-1 に示す。

表 2.2-1　NOAEL 法と BMD 法に基づく HBGV 導出方法の比較

ステップ	NOAEL 法に基づく ADI 導出	BMD 法に基づく ADI 導出
1. データの選択	十分なサンプルサイズ (統計学的に有意な影響がみられない用量群が少なくとも 1 つ必要)。適切な生物種の適切なエンドポイントであることが重要。	異なる反応レベルが得られている十分な数の用量設定数と十分な数の個体数
2. モデルの選択	なし	用量反応モデル (例: Weibull model) のフィッティング
3. 統計学的関連付け (statistical linkage)	各用量群と対照群のペアワイズ比較	推定される画分 (predicted fractions) と実際に観察された画分 (observed fractions) との関連付けを行い、フィッティングの評価関数 (criterion function) (例: 仮定した分布に基づく尤度関数) を最適化することによりそれらの“距離 (distance)”を最小化する。
4. パラメータ推定	パラメータは使用しない。 NOAEL は試験における 1 用量である。	試験で観察された反応レベルの範囲から適切な反応レベル (p) を選ぶ。BMDLp (BMDp の片側 95%信頼区間の下限値) を推定する。 $\frac{R(BMDp) - R(0)}{1 - R(0)} = p$
5. 実行	ADI = NOAEL / UFs (UF: 不確実係数)	ADI = BMDLp / UFs (UF: 不確実係数)
6. 評価	試験が評価対象とする影響を検出するのに十分な感度を有するか確認するため、統計学的検出力を解析する。	種々のモデルをフィッティングさせることにより、モデルの選択に応じた BMD の感度を確認する。

以下に、各ステップにおける BMD 法の手順を示す。

(3) ステップ 2　データの選択

BMD 法でのモデリングにおけるデータ選択では、NOAEL 法と同様の基本的検討を行う。BMD 法において、POD は有害影響が検出されないばく露レベルに基づくものではな

いため、各用量群の大きさはそれほど重要ではない。有意な用量依存傾向を有し段階的な単調応答を示す試験が、モデリングを行う上では最良の試験データといえる。

BMD 法の主な課題は、一定の反応レベルである BMR の選択が必要となることである。一般に、選択するレベルは、当該研究の検出限界に近いレベル又は一般に無視できる程度の健康影響を表すと考えられるレベルである。BMR として 1%、5%、10%などの種々の反応レベルが選択される。選択した BMR の違いにより、種々の規制機関間の HBGV に相違が生じる可能性がある。

(4) ステップ 2　モデルの選択

利用可能なデータの種類とモデリングを行う反応の特性に応じてモデルを選択する。単純なモデルに比べ、複雑なモデルは多くの用量群を必要とし、データの種類に応じていくつかのモデルが提案されている。米国環境保護庁 (EPA) の Benchmark Dose Software (BMDS) プログラムでは、日常的に使用される複数のモデルを収載している。複数のモデルを同じデータに適用した際に大きく異なる推定値が得られた場合、BMDL を計算するために特定のモデルを選択する必要があり、適合度の指標 (例: AIC) に基づくモデルの選択、モデル平均化又は最小 BMDL が得られるモデルの選択が提案されている。

(5) ステップ 3　統計学的関連付け (statistical linkage)

データとモデル間の統計学的関連付け (statistical linkage) においては、多くの異なる形が仮定され得る。非連続データについては、データが各用量群において二項分布をとると仮定するのが適切である。

(6) ステップ 4　POD の選択

BMD 法では、モデリング結果として、選択した BMR に対応する BMD 及び BMDL の値が出力されるだけであるため、同法における POD の選択は、実際には BMR の選択である。どの程度の反応レベル (BMR) が無視できる程度の健康影響を表していると見なすことができるかは、明確でない場合が多い。BMR の選択には、毒物学者と臨床医の間での議論が必要である。BMR の明確な説明は、NOAEL に対応する可能性がある一般的に未知の反応レベルとの比較の観点から発展してきたものの、BMR の選択は、合意形成を必要する上、依然として、本質的に数学的手法を用いる主観的な「専門家」判断のままである。BMR の選択における別の方法として、試験の検出限界に近く、それ以下は試験データが十分に支持しない過剰反応 (多くの場合 10%) を選択することが挙げられるが、当該 BMR で生じる反応レベルが健康への影響を及ぼす可能性についての問題は未解決のままである。

(7) ステップ 5　HBGV の算出

HBGV は、以下の式により算出する。

$$\mathrm{HBGV} = \frac{\mathrm{BMDL}}{\mathrm{UFs}}$$

この計算では、不確実係数の値は NOAEL 法で使用した値と同一か、又は BMDL は NOAEL と僅かに異なると解釈して補正した値を用いる場合もあり得る。NOAEL 法とは異なり、全ての用量レベルが (対照群と比較して統計学的に) 有意なレベルの有害影響をもたらす場合に、追加の不確実係数は必要ない (実際、このようなデータは BMD モデリングにより適している)。実証的な検討結果からは、多数かつ代表的な化合物群において BMDL の値は NOAEL の値と類似し、互いに置き換えても同様の HBGV が得られるであろうことが示されている。

(8) ステップ 6　評価

NOAEL 法とは異なり、BMD 法では、所定の用量での反応、所定の反応をもたらす用量及びそれらの信頼限界を決定する。生物学的に観察可能な用量範囲を下回る領域に用量反応モデルを外挿すると、特定された (より低い) 用量レベルでの反応と、特定の反応レベルをもたらす用量を推定できる。しかし、観察範囲のデータにフィットする単一のモデルからの外挿は、同様にデータにフィットする他のモデルが、同じ低用量でのリスクに対してかなり異なる推定値を与える可能性もあるため、正当化できないことに注意すべきである。簡便な低用量領域への外挿方法として、BMD_{10} からの直線外挿が用いられることが多いが、JECFA の第 64 回会合では、「POD からの直線外挿は、ばく露マージン (MOE) 法を超える利点はなく、また、BMD 法では数値が推定されることから実際のリスクが定量化されたものと見なされる可能性があり、結果は誤解されやすい」と結論している。

2.2.3　BMD 法の適用方針

以上を踏まえ、EHC 240 (WHO, 2009) における BMD 法の適用方針のポイントを以下に整理する。

- BMD 法の適用手順は EHC 239 と同じく、データの選択、モデルの選択、統計学的関連付け関連付け (statistical linkage)、パラメータの推定、解析の実行、解析結果の評価の 6 ステップである。
- BMD 法におけるデータの選択では、各用量群の大きさは重要ではなく、有意な用量依存傾向を有し段階的な単調応答を示す試験データが適切である。
- BMD 法の主な課題は BMR の選択であり、専門家により合意形成が必要である。試験の検出限界に近く、それ以下では試験データ十分に支持されない過

剰反応 (多くの場合 10%) の選択が例として挙げられるが、BMR の選択の課題は残されたままである。

- 複数のモデルを同じデータに適用した際に大きく異なる推定値が得られた場合、BMDL を計算するために特定のモデルを選択する必要があり、適合度の指標 (例: AIC) に基づくモデルの選択、モデル平均化又は最小 BMDL が得られるモデルの選択が提案されてきた。
- 統計学的分布として、多くの異なる形が仮定され得る。非連続データについては、データが各用量群において二項分布をとると仮定するのが適切である。
- NOAEL 法とは異なり、BMD 法では、所定の用量での反応、所定の反応をもたらす用量、及びそれらの信頼限界を決定することができる。一方で、低用量領域への外挿については、単一モデルからの外挿だけでは十分でない。また、POD からの直線外挿については、MOE 法を越える利点はない。

2.3 Safety evaluation of certain contaminants in food (Food Additive Series: FAS No. 55) ANNEX 4 / ANNEX 5 (FAO/WHO/JECFA, 2006)

2.3.1 概要

➢ 背景

本ガイダンスのうち ANNEX 4 は、JECFA が、食品中の汚染物質のうち、遺伝毒性及び発がん性の両方の性質を持つ化合物に関するリスク評価を行い、リスク管理者にヒトで起こりうる健康影響の程度を勧告する際の手法や考え方を整理した文書である。ANNEX 5 は、JECFA が 2005 年の第 64 回会合で BMD 法を用いた用量反応評価を行った際の手法を整理した文書である。

➢ **ANNEX 4 の概要**

危害要因判定 (用量反応評価) は、発がんに関する入手可能な用量反応データに基づいており、多くの場合、これらのデータはヒトの推定摂取量よりも数桁高い用量を投与したげっ歯類の試験から得られる。WHO のワークショップでは、動物の発がん性試験データが用量反応モデリングに適している場合、危害要因判定の開始点 (starting point) として BMD の信頼区間の下限値 (BMDL) の使用を推奨している。

BMD は、多くの場合で、BMDL に比べてモデル間でのばらつきが少なく、異なる化合物間で有害影響の可能性を順位付けする場合に適した指標となる可能性がある一方、BMDL は用いた用量反応データの質を反映する指標であるため、リスク判定に適している。ただし、BMDL の導出と解釈には、かなりの統計学的及び生物学的専門知識が必要である。

遺伝毒性及び発がん性の両方の性質を有する化合物は、推定ばく露量と BMDL との比較によりリスク判定、具体的には、ばく露マージン (MOE) に基づくリスク判定を行うことを推奨している。

2.3.2 BMD 法に関する手順

用量反応モデリングでは、データの選択、モデルの選択、統計学的分布、パラメータ推定、モデリングの実施及び評価について検討する必要がある。このうちデータの選択、モデリングの実施及び評価 (リスク判定) に関して、以下の記述がある。

(9) データの選択について

発がん性については、モデリングを行う用量反応データの選択においては、特定の部位、特に感受性が最も高い部位における腫瘍の部位特異的な発現率と、明確な臓器特異性を示さない化合物における組合せデータ (例えば、腫瘍を有する動物数) の両方を検討する必要がある。腫瘍を有する動物数に基づく分析は、遺伝毒性と発がん性の両方を有する

化合物の複雑な混合物の評価にも適用できる可能性がある。

用量反応関係の判定 (dose-response characterization) は、観察された実験的な用量反応関係の下限を確実に定義する最小レベルの反応量 (BMR) での、ヒトの健康に関連する発がんの BMDL を定義することを目的としている。10%未満の発現率では、異なるモデル間のばらつきがより大きくなるため、発がん性試験データのモデリングにおいては、10%の発現率を BMR とするのが最も適していると考えられる。同じ値の BMR (10%) を用いることで、遺伝毒性発がん性を有する種々の化合物のリスクの比較が促進されるだろう。遺伝毒性発がん性を有する種々の化合物の非発がん影響を同様のアプローチで解析することができ、得られた BMDL の値とそれらの傾きの比較結果は、当該化合物のリスク評価において、重要な意味を持つ有害影響の特定に役立つ可能性がある。

(10) モデル選択及び BMD 等の算出 (ANNEX 5)

通常、飽和モデル (full model と同義、データポイント 1 つにつき、1 つのパラメータを有するモデル) に比べ、フィッティングが有意に悪いとされなかったモデルは許容されると見なした。具体的には、飽和モデルとの対数尤度の差が、統計学的に有意なフィッティングの改善を示す値 (臨界値 0.05) を下回ることを確認、判断した。

許容されると見なした各モデルのフィッティング結果から、BMR として 10%の過剰リスクを用いた際の BMD 及び BMDL をそれぞれ算出した。

(11) 用量反応モデリング結果の評価 (リスク判定) について

遺伝毒性と発がん性の両方を有する化合物のばく露評価方法は他の汚染物質の方法と同じである。リスク判定は、推定ばく露量と特定されたBMDLとの比較を含む様々な方法で実施され、主要なものとして次の3種類の方法がある。

1) MOE法 (ばく露マージンの計算)

MOE は、ヒトでの推定摂取量に対するBMDLの比である。MOE法は様々な汚染物質について、 (リスク管理の) 優先順位を付与するのに用いられる。MOEの許容度はその桁数と、究極的にはリスク管理上の決定に拠る。リスク評価者は、この決定を支援するため、リスク評価結果の通知に当たり、毒性データとばく露量データの両方について、含まれる不確実性の質と大きさの情報を提供すべきである。

2) 観察された用量範囲外における用量反応解析

定量的な用量反応解析により、ヒトの推定ばく露量でのがん発現率又は特定の発現率 (例: 100万分の1) に関連するばく露量を理論的に計算できる。数学的なモデリングにおいては、発がん性試験で用いられる高用量データに対して、ヒトの摂取量がはるかに低い場合の用量反応関係を検討する必要がある。このような情報は動物試験研究で入手できるデータから導き出すことはできない。将来的には発がんに関連する重要な生物学的活性 (代

謝活性、解毒プロセス、DNA結合、DNA修復、細胞増殖、アポトーシス等) に関する用量反応や濃度反応に関係するデータを取り入れることができると考えられる。現時点では、がん発現率の推定は経験的な数式に基づいており、生物学的な複雑性を反映していない可能性がある。低用量領域への外挿のために多くの数理モデルが提案されており、リスク推定値は使用する数理モデルに依存する。用量が低くなるにつれて逸脱が大きくなり、非常に低い発現率においては、種々の数式に基づく推計結果は大きく異なる可能性がある。

3) 出発点 (POD) からの直線外挿

低用量領域で推定されたリスクはモデル依存的であるため、一部の機関では、種々の理論的がん発現率に対応するばく露レベルの計算において、保守的で容易に適用できる手法として、BMDLからの直線外挿を政策方針として使用してきた。使用される発現率はがんの生涯リスクの上限推定値と見なされ、実際のリスクは計算された上限推定値とゼロの間にある。

直線外挿を用いて10%の発現率でBMDLから100万分の1の発がんを引き起こす摂取量を計算することは、単純にBMDLを100,000で除することと等しいため、この方法は、MOEの計算以上の情報はもたらさない。

上記の3つの方法のうち、MOE法とPODからの直線外挿が現時点で最も実用的な方法である。PODからの直線外挿はMOE法を超えるメリットはなく、推定された数値が実際のリスクを定量化したものと見なされる可能性があることから、結果は誤解されやすい。

結論として、遺伝毒性と発がん性の両方の性質を持つ化合物に関する (リスク評価上の) 勧告は、推定したMOEに基づくべきである。

2.3.3 BMD 法の適用方針

以上を踏まえ、Safety evaluation of certain contaminants ANNEX 4 / ANNEX 5 (FAO/WHO/JECFA, 2006) における BMD 法の適用方針のポイントを以下に整理する。

- BMD 法の適用手順は EHC 239 と同じく、データの選択、モデルの選択、統計学的関連付け (statistical linkage)、パラメータの推定、解析の実行、解析結果の評価の 6 ステップである。
- 発がん性データの BMD モデリングにおけるデータの選択では、感受性が最も高い部位における腫瘍の部位特異的な発現率と、明確な臓器特異性を示さない化合物における組合せデータ (例えば、腫瘍を有する動物数) の両方を検討する必要がある。発がん影響における BMR としては 10%の過剰リスクが適切である。
- 飽和モデルに比べ、フィッティングが有意に悪いとされないモデルは許容されると見なし、許容されると見なした各モデルのフィッティング結果から、BMR と

して 10%の過剰リスクを用いた際の BMD 及び BMDL をそれぞれ算出する。

- リスク判定手法については、MOE 法 (BMDL のヒトの推定摂取量に対する比の算出) を推奨する。

2.4 Benchmark Dose Technical Guidance (EPA, 2012)

2.4.1 概要

➢ ガイダンス作成の背景

本ガイダンスは、EPA 及び外部のコミュニティに対して、様々な用途における BMD 法の適用に関する指針を提供する目的で作成された。本ガイダンスの草案は、BMD に関するワークショップやシンポジウムの内容を反映して 2000 年に作成され、当該草案や外部のピアレビュー委員会のコメント、EPA のリスク評価における手法の適用によって得られた経験を反映して最終的なガイダンス及び推奨オプションが作成された。なお、本ガイダンスの記載内容は文書作成時点における知識や理解、経験に基づいており、新たな手法が利用可能となった場合や、手法の代替や追加等が生じた場合には改訂されることになっている。

➢ **BMD 法の利用状況**

BMD は、化学物質のリスク評価における用量反応情報を得る目的の他、化学物質間の相対効力や部分集団間の相対感度の比較、生態リスク評価等の分野にも適用可能であるが、本ガイダンスは、動物試験の毒性データを用いた用量反応モデリングに焦点を当てている。なお、ヒトのデータについては、モデリングの機会が限られていること、ヒトの研究は動物試験よりも標準化が難しいこと、及び共変量の補正等の追加的な検討事項があることから、ヒトのデータについてモデリングを実施する場合は、より事例毎に特異的な手法が適用される。

➢ 今後の課題

本ガイダンスでは、現在までに得られていない、又はガイダンスに記載できるレベルにない事項として以下が示されている。

・BMR の選定における判断の指針
・モデル選定における事前の絞り込みの指針
・多変量解析、カテゴリカル回帰、時間-反応解析、分布解析、ブートストラップ法、モデル平均化、ベイズ推定に基づく手法等のより専門的な事項[11]

今後の BMD 法の経験の蓄積により、これらについても指針が示されることが期待される。

[11] 最新の BMDS 3.1 にはモデル平均化機能が収載されており、二値モデルを用いる解析において、ベイズ推定に基づくモデリングの結果得られた各モデルの BMD 及び BMDL 推定について、ベイズ推定に基づくモデル平均の値が得られる。

2.4.2　BMD 法の手順

本ガイダンスで示されている BMD 及び BMDL の算出手順の概要を以下に示す。

データの評価	試験デザインやデータの報告内容等に基づき、モデリングを行う試験及びエンドポイントを選定する。
BMR の選定	データの種類、試験デザイン、評価エンドポイントにおける変化レベルの有害性に関する判断等に基づいて選定する。
モデリング	データタイプに特有のモデル構造及び試験デザインを用いて、用量反応データをモデリングする。
モデルの適合度の評価	p 値や残差、モデリング結果のプロットから、モデルの適合度を検討する。
BMD 及び BMDL の算出	データに適切にフィットするモデルを使用して、候補 BMD の 95% 信頼区間の下限値 (BMDL) を計算する。
BMDL の選定	データに適切にフィットするモデルの中から BMDL を選定する。
BMD 算出の文書化	BMD 算出過程における判断の根拠等を明確に文書に残す。

以下に、各ステップにおける具体的な手順を示す。

(12)　データの評価

本ガイダンスで言及されているデータの種類を表 2.4.-1 に示す。

表 2.4.-1　データの種類

種類	説明
連続値データ	間隔尺度、比尺度とも呼ばれる、連続した尺度のデータ (体重、酵素活性等) 。
非連続データ	名義尺度、順序尺度とも呼ばれる、カテゴリー分類尺度のデータ。毒性試験では一般に、任意のカテゴリーに含まれる個体の割合として表される。
カテゴリカルデータ (順序データ)	非連続データのうち、複数の定義されたカテゴリー (無影響/軽度/中程度/重度等) を含むデータ。
二値データ	カテゴリカルデータのうち、2 つのカテゴリーのいずれかに分類されるデータ (影響の有無等) 。

(1)-1　データの評価と BMD 解析のフィージビリティの検討

データの評価として、入手可能な毒性データについて、試験デザイン、データの報告内容等の評価を行った上で、モデリングを行う試験及びエンドポイントを選定し、選定したデータセットについて BMD 解析のフィージビリティの検討を行う。以下に、データ評価の項目と各評価項目における評価のポイントを示す。

1)　試験デザイン

用量反応モデリングに適用する試験は、用量群がより多く、特に BMR 付近でより多くの用量が設定された試験が望ましい。また、段階的な単調反応を示す試験データが望ましい。単用量の試験やバックグラウンド時の反応レベル又は最大反応レベルと同じレベルの反応のみしか示さない試験は、BMD の解析に適しているとは考えられない。

2)　データの報告内容

データの報告内容については、可能であれば原著を確認し、用量群や試験対象動物等の設定等の意図や根拠を理解することが望ましい。各個体の情報が得られていることが望ましいが、文献情報は多くの場合、要約データであることから、要約データが BMD 解析を適用するのに十分な情報かどうかを確認する必要がある。

二値データの BMD モデリングにおいては、反応を示した個体数と群全体の個体数の情報が必要となる。連続値データについて各個体のデータが得られない場合は、各用量群の個体数、反応の平均値、ばらつきの尺度 (例: 標準偏差 (SD)、標準誤差 (SE)、分散) の情報が必要である。

3)　モデリングを行う試験の選定

ヒトのばく露状況、試験の質、報告内容の十分性及びエンドポイントの関連性に基づき、BMD 解析に用いる試験を選定する。妥当性のある全ての試験についてモデリングを行うことを検討すべきである。

4)　モデリングを行うエンドポイントの選定

モデリングを行う試験について検討した後、用量反応関係に焦点を当ててモデリングを行うエンドポインの選定を行う。基本的にはばく露に関連すると判断される全てのエンドポイントについてモデリングを行うことを検討する。

5)　BMD モデリングのための最小データセット基準

重要なエンドポイントを選定した後、データセットについて BMD 解析のフィージビリティを検討する。BMD モデリングのための最小データセット基準を以下に示す。

・選択したエンドポイントにおいて、少なくとも統計学的又は生物学的に有意な用量依存傾向がある。

・データセットにおいて、対照群における反応レベルと観察された最大反応レベルの間で用量反応関係がある (理想的には BMR に近いデータポイントを有する)。

(1)-2　BMD 算出のためのデータの統合について

BMD 算出のためのデータの統合について、統計学的及び生物学的に互換性のあるデータセットは、用量反応モデリングの前に統合でき、これにより、BMD 算出において統計学的及び生物学的信頼性が増す。データセットを統合する最も簡単な方法は、データ全てが同時に集計されたように取り扱うことであり、複数のデータセットで用量反応が同質である (例: 複数のデータセットに共通する用量での反応が本質的に同一であり、統計学的に区別できない) 場合には、この方法は妥当である。一方で、データセット間にばらつきが存在する方が一般的であり、この場合、データセットの統合に際して、統計学的に適切な方法と生物学的な判断に基づきその正当性を説明する必要があるが、具体的なガイダンスを提供できる状況にはない。

(1)-3　用量の調整について

試験動物への投与量をヒト等価用量 (HEDs) に換算するために、PBPK モデル等に基づく用量調整が行われる場合が多い。こうした調整については本ガイダンスの対象外であるが、用量反応モデリングを行う前に用量調整について検討を行うことが重要となる状況について言及されている。

調整が用量に比例する場合 (例: 継続的なばく露に対する一定の調整)、モデリングの

前に用量を調整するか、モデリングの後に用量を調整するかは、最終的な結果に影響しないが、調整が用量に比例しない場合、用量反応モデリングの前に用量を調整する方が適切である可能性がある。

(13)　BMR の選定

BMR の選定には、データセットの統計学的・生物学的な特徴や、BMD 及び BMDL の使用目的に基づいた判断が必要となる。しかし、現時点で EPA にはこの判断をサポートする指針はない (発がんリスク推定のための BMR 選定については EPA の発がん物質リスク評価ガイドライン[12]が対応している)。本ガイダンスでは、BMR の選択における一般的なアプローチの概要を特定の事例に沿って示すとともに、化合物又はエンドポイント間の比較を行う目的での標準的な報告要件としての BMR について示す。

以下に、データの種類別に BMR の選定の一般的なアプローチを示す。

(2)-1　非連続 (二値) データ

BMR の選択は BMR 及び BMDL の利用目的に応じて個別に検討する必要があるが、一般的なアプローチは以下の通り要約される。なお、化合物又はエンドポイント間での比較を行う目的から、常に 10%過剰リスクに対応する BMD 及び BMDL を報告することを推奨する。

- 化合物又はエンドポイント間の比較を行う目的では、標準的な BMR として 10%過剰リスクを推奨する。
- 生物学的検討の結果に応じて、リスク評価のための参照値 (reference value) の出発点 (POD) の根拠として、重篤な影響 (frank effects) 等のある種の影響の場合は 5%又はより低い BMR、あるいは前駆効果のような影響の場合 10%を上回る BMR を用いるのが妥当である可能性がある。
- 生物学的検討に基づき設定した 10%を下回る BMR が、観察可能な範囲に含まれる場合がある。統計学的な観点から、ネストした[13]生殖発生毒性試験では BMR として 5%が推奨され、疫学研究のデータでは BMR として 1%が用いられる場合が多い。BMR が観察可能な範囲を下回る場合、推定の不確実性が増すことに留意する必要がある。このような場合には、BMD と BMDL が大きく乖離していないかを確認する。また、測定データの範囲を下回るとモデルの不確実性が増すことに留意する。

[12] U.S. EPA. (2005a) Guidelines for carcinogen risk assessment. Federal Register 70(66)177650-18717. Available online at: http://www.epa.gov/raf/pubalpha.htm.

[13] データがネストした構造になること。生殖発生毒性試験の場合、児のデータは親のデータに紐付いている (ネストしている) と見なすことが出来る。

(2)-2　連続値データ

連続値データにおける BMR の選定方法として様々な方法があるが、一般的な手順は以下に要約される。なお、非連続 (二値) データにおいて 10%過剰リスク[14]に対応する BMD 及び BMDL を報告することを推奨するのと同様に、比較を行う目的から、対照群における平均反応量の標準偏差 (SD) 1 つ分の変化に等しい変化を平均反応レベルにもたらす BMD 及び BMDL を常に報告することを推奨する。

- 望ましい方法：評価対象エンドポイントにおいて、生物学的に有意であると考えられる最小の変化レベルが分かる場合、当該変化レベルを BMR とする。
- 個体のデータが得られており、有害影響と考えられる個体レベルを決定できる場合、当該データをハイブリッド法[15]により暗に (implicitly)、又はカットオフ値を用いて明確に (explicitly) 二値化し、二値データの場合と同様に BMR を設定することができる。この場合、情報が失われる (カットオフ値に基づく) 明確な二値化よりも、ハイブリッド法による二値化が望ましい。
- 有害と考えられる反応レベルが不明な場合、対照群の平均反応量の標準偏差 (SD) 1 つ分に等しい平均反応量の変化レベルを BMR とする。重篤な影響の場合、より低い BMR (0.5SD 等) が妥当な場合もある。

(3)　モデリング

BMD 算出のためのモデル選定は十分に文書化されたモデリングソフトウェアで行うことが望ましく、EPA は、本ガイダンスに記載された複数のモデルとパラメータ制限等のデフォルトの設定を含む BMDS ソフトウェア[16]を開発している。

異なる対象物質間において候補となっているエンドポイントや用量反応関係の形状は様々であるため、現在のところ、条件に応じて BMD の算出に用いる少数のモデルを事前に指定することは不可能だが、今後の BMD 法の経験の蓄積によって、特定のシナリオに対するモデルをある程度絞り込めるようになる可能性がある。

(14)-1　フィッティングを行うモデルの選定

フィッティングを行うモデルの選定では、エンドポイントの種類、試験デザインに応じてモデルを選択し、パラメータ数、パラメータの制限、共変量の設定等を行う。

以下に、フィッティングを行うモデルの選定における検討事項を示す。

---

[14] EPA (2012) Benchmark Dose Technical Guidance で使用される「過剰発現率」「過剰リスク」という用語は、発現率の増加や、コントロール又はバックグラウンド時の反応を超えるリスクの増加を広く意味する。

[15]各個体が、有害とみなされる反応レベルの範囲に含まれる確率を、用量の関数として直接推定する手法((3)-1 1) を参照)。

[16] http://epa.gov/NCEA/bmds/

1)　エンドポイントの種類

数理モデルの選択においては、評価対象エンドポイントの測定変数の種類を検討することが重要となる。データの種類別に推奨モデル分布族を表 2.4.-2 に示す。

表 2.4.-2　データの種類別の推奨モデル分布族

データの種類		推奨モデル分布族
非連続 (二値) データ		Logistic、Probit、Weibull などの確率密度モデル
連続値データ	BMR が連続値データの一定の変化レベルとして定義される場合	Polynomial、Power、Hill などの連続値モデル
	データが二値化され、BMR が特定の値を超えた個体の割合として定義される場合	二値データのモデル (Logistic、Probit、Weibull などの確率密度モデル)

連続値データについては、BMR の選定において BMR がどのように定義されたかによってモデリング法が決定される。BMR が連続値データの一定の変化レベルとして定義された場合、連続値データに対応するモデル (連続値モデル) を適用する。データが二値化され、特定の値を超えた個体の割合として BMR が定義された場合、二値データとしてのモデルを適用することができるが、この方法は連続値データの情報が失われる結果となるため推奨しない。これらに代わる連続値データのモデリング手法として、ハイブリッド法がある。ハイブリッド法は、各個体が、有害とみなされる反応レベルの範囲に含まれる確率を、用量の関数として直接推定する手法である。ハイブリッド法では、対照群における各個体の測定値の確率分布を特定し、有害とみなされるパーセンタイル値 (例: 測定値が低い場合に有害とみなされる場合、1 パーセンタイル等) を設定した上で、連続値モデルをデータにフィッティングさせて各用量での平均反応量とその分散を推定することにより、有害とみなされる反応レベルの範囲に含まれる画分の割合を推定し、特定した BMR に対する BMD を計算する。

2)　試験デザイン

通常、用量反応曲線のパラメータ数は用量群の総数を超えられないため、用量群の数が推定可能なパラメータの数に影響する。

また、クラスター分析が行われた試験では、個々の測定値を互いに独立していると見なすことができないため、特殊な近似手法を使用する必要がある (本ガイドラインの対象外) 。

3)　制限と共変量

i)　パラメータの制限

推定結果の生物学的妥当性を補強するため、パラメータがとり得る値に制限が必要な場合がある。パラメータ制限の例を表 2.4.-3 に示す。

表 2.4.-3　パラメータ制限の例

対象	制限の内容
予測値が正となることが想定される生物学的測定データ	予測値が正となるようにパラメータを制限する
用量の増加に伴い影響が単調増加する毒性試験データ	単調増加を示すようにパラメータを制限する (例: Multistage model ではパラメータが負にならないように制限)
全投与群で反応レベルが BMR を超え、投与群間の差が小さいデータ (Weibull モデル等の、推定される用量パラメータがべき乗となるモデルにおいて用量反応曲線の傾きが低用量で急峻になるデータ[17])	べき乗パラメータを 1 以上に制限する
バックグラウンドパラメータにより、ばく露がない場合にも影響が生じる確率を 0 より大きい値として算出する非連続データに対応するモデル (非連続モデル)	(ばく露がない場合には影響が生じないことが明らかな場合のみ) バックグラウンドパラメータの値を 0 に設定する

ii)　共変量

用量反応モデルのフィッティングにおいて、個体データに共変量を含めることが望ましい場合もある。このような例として、発生毒性試験で得られた動物試験データをモデリングする際の同腹児数、疫学データのモデリングの際の共変量 (年齢、出産回数等) が挙げられる。(因果関係はないものの) 用量に相関し、結果に影響を与えるいずれの変数も共変量として取り扱うことを検討すべきである。なお、共変量は試験集団のサブセットを定義することから、各レベルの共変量に対してそれぞれ BMR が必要となる可能性がある。

(3)-2　モデルのフィッティング

モデルフィッティングの目標は、全てのモデルパラメータについて、フィッティング後のモデルが用量反応データを可能な限り説明するものとなるような値を見いだす

[17] このようなデータでは、見かけの信頼区間の適用範囲は (漸近的に) 正確ではなく、また、真 (未知) のべき乗パラメータが 1 未満の場合には、これを 1 以上に制限することで、意図した適用範囲よりもはるかに狭くなる可能性がある。そのため、どのようなデータであったか報告する必要がある。

「パラメータ推定」を行うことである。そのための一つの方法は、関数の全体的な最小値 (又は最大値) に対応するパラメータ値が、求められるモデル予測をもたらすという特性を持った全パラメータ及び全データの関数 (目的関数) を特定することである。本ガイダンスでは、一般的なデータの種類別に、分布の仮定及び目的関数の構築・最適化手法について紹介されている。

1) データの種類別の分布の仮定

一般的なデータの種類と分布の仮定を表 2.4.-4 に示す。

表 2.4.-4 一般的なデータの種類との分布の仮定

データの種類	分布の仮定
二値データ	反応個体数は二項分布に従う。
連続値データ	通常、正規分布又は対数正規分布する (特に複数の観測値の平均値の場合)。
発生毒性試験データ	同腹児に関するデータは独立していないため、有害事象を伴う動物数の分布は β-二項分布であると見なすことが多い[18]。

2) 目的関数の構築・最適化手法

目的関数の構築・最適化の手法とその概要を表 2.4.-5 に示す。

表 2.4.-5 目的関数の構築・最適化手法

手法	概要
最尤法	尤度関数が最大となる (観察データを得る確率が最も高い) パラメータを推定する方法。データの分布について合理的な仮説を立てられる場合に目的関数を導出するための一般的な手法。
非線形最小二乗法	用量群の平均周辺の個々の観測値の分散は用量間で一定であると仮定し、観測値と推定値の差 (残差) の二乗の和を目的関数とし、目的関数が最小となるパラメータを推定する方法。観測値を独立と見なすことができる連続変数に対する一般的な手法。
一般化非線形最小二乗法	非線形最小二乗法の変法。観測値を独立と見なすことができるが、用量群の平均周辺の個々の観測値の分散が用量間で一定であると仮定できない場合に用いる。フィッティングさせるデータがほぼ正規分布していると推定できる場合に重要な手法。

[18] ただし、この方法では、同腹児への影響を誤ってモデリングした場合に頑健性が欠如する。この場合、準尤度や一般化された推定式に基づく分析法で代替できる。また、単純なデータ変換を使用するアプローチが一般化推定方程式 (GEE) や最尤法と同程度に効率的とされる。

手法	概要
擬似尤度法及び一般化推定方程式 (GEE)	データの平均、分散及び相関構造のみを指定するパラメータ推定手法。一般に最も広い範囲のデータタイプに適用可能で、臨床研究や反復神経行動試験のような多数の反復測定デザインにも有効と考えられる。

3) パラメータの決定

目的関数を特定した後、パラメータの決定を行う。通常は、ソフトウェアを用いてパラメータ値を繰り返し推計し、収束する一連の推定値を生成する。多くの場合この方法で正しい推定値に収束するが、モデルやデータセットによっては初期値について複数の推定が必要な場合もある。異なる初期推定によって複数の推定値に収束する場合は、専門家の判断により、これらのうち 1 セットのみを採用する。

(15) モデルのフィッティング結果の評価

(4)-1 適合度の指標による評価

フィッティングさせたモデルの重要な選択基準の 1 つは、特に BMR 領域においてモデルがデータを適切に説明しているか否かである。ほとんどのフィッティング方法では、全体の適合度の指標 (通常、p 値) が得られる。p 値はモデルによって予測される各用量群における反応レベルの平均値が実際の値とどの程度差があるかを表し、p 値が小さいほどモデルはデータへの適合度が低いことを表す。

BMD モデリングは、通常、モデル一式のカーブフィッティングの繰り返しであり、データが BMD 算出の目的で適切にモデリングされていることが重要であることから、臨界値[19]として用いる p 値として、通常用いられる有意水準 (α) = 0.05 又は 0.01 の代わりに α = 0.1 の使用を推奨する[20]。特定のモデルを選択する経験的な根拠がある場合は、例外的に α = 0.05 又は 0.01 の使用を検討する。p 値は試験結果と一致するモデルの特定のみに用いられ、モデル間での比較はできない。なお、モデル内に共変量がある場合は計算がより複雑になる。

なお、モデルはデータポイントと大きく離れていない (つまり、適合度検定の p 値はそこまで小さくない) が、常に投与群の平均値の片側にある場合や、反応が広範囲に存在し高用量反応領域はよく説明されているが低用量反応領域を適切に説明していない

[19] χ^2 適合度検定の場合、臨界値は適切な自由度における χ^2 分布の 1-α パーセンタイルである。この目的のために選択された第 1 種の過誤 (偽陽性) の限界確率である α 未満の p 値に相当する、大きな値の χ^2 を棄却する。

[20] 一部のケースでは、適合度 p 値のみに注目した場合、p 値が 0.1 未満に基づくとほとんどの使用可能なモデルが適切ではないようにみえる可能性がある。しかし、専門家の判断のもとで他の基準 (エンドポイントのばらつきの質、目視によるフィッティングの確認、データ範囲の中で最も関連性のある領域における残差等) を検討すると、いくつかのモデルで満足のいく適合度が得られることがある。

場合などは、適合度検定のp値に有意な差はないが、フィッティングを慎重に評価すべきである。こうした状況を検出する手法は2通りある。1つは、残差 (モデルにより推定した反応と実際のデータの偏差) の表を用いた整理やプロットでの表示による検出である。残差が推定値のばらつき (SE) によって補正される場合において、その補正された又は標準化された残差の絶対値が2を超える場合は、モデルのフィッティングをさらに検討する必要がある。もう1つは、グラフに表す方法である。プロットは常に適合度検定を補足するので、データポイントに加え、信頼限界などの分散の尺度も含めてプロットすることは非常に有用である。

(4)-2 モデルのフィッティングの改善

利用可能なモデルのいずれにおいてもデータへの適切なフィットが得られない場合は、以下 1)〜4) を順に検討する。

1) データの質や試験上の問題の検証

データの質や試験設計に関して、最初のデータ評価で見過ごしがなかったか (再)検証を行う。この場合、データの補正 (用量の対数変換や無関係の死亡に対する補正等) を検討する。データのプロットがプラトーに達したり、非単調反応を示したりする場合は、より広範な反応の定義 (他の反応への進行や、他の反応によるマスキング) を検討する。また、より複雑なモデル (応答時間を説明するモデル等) の適用や、薬物動態データ等によって、他のモデルに適合するような適切な用量反応関係が得られる可能性もある。

2) モデルのフィッティングプロセスの検証

モデルのフィッティングプロセスの検証として、例えば非線形モデルへのフィッティングが最適な推定結果をもたらしているか、又は不均一分散による影響を適切に考慮しているか等を検証する。非線形モデルへのフィッティングにおいては、パラメータの初期値を変えた試行を推奨する。また、連続値モデルのフィッティングにおいては不均一分散が悪影響を及ぼす可能性がある。この場合、分散を平均反応のべき乗に比例するものとしてモデリングする手法がある。この手法は、変動係数が一定で、分散が平均の 2 乗に比例し、かつ変動係数が比例定数の平方根である状況にも対応できる。また、加重最小二乗法及び一般化非線形最小二乗法などの手法の利用も考えられる。

3) より柔軟な経験的モデルの検討

なおも適切なフィットが得られない場合は、用量反応関係を適切に説明可能な、より柔軟な経験的モデルを検討する。このアプローチは、全てのデータを解析に組み込

むことができる点が長所である一方、特定の用量範囲のデータに当てはめようとすると別の用量範囲、より直接的に関心がある範囲の用量反応曲線が歪む可能性があるが、多くの場合、BMD は試験における最低用量に近いため、BMD 領域におけるモデルの適合度の評価は可能である。

4)　その他の手法

用量反応パターンに対する生物学的に尤もらしい解釈が可能な場合であっても、生物学的メカニズムに対応するモデルは利用できない場合が多く、そのようなモデルに適切にフィットさせるために十分な量の追加データが得られない場合もある。有害物質に対する生物学的反応のメカニズムが不明の場合、BMR よりもはるかに大きい反応を示すばく露データからは BMR 領域の反応の形状に関する十分な情報が得られない可能性がある。しかしその一方で、そのようなばく露データが、フィッティングさせたモデルの BMD 領域の形状に非常に強い影響を与える可能性もある。

BMD 法を用いて低用量領域への外挿を行う場合において、適合度検定 $p < 0.10$ などの客観的基準に基づきいずれの利用可能なモデルも適切なフィッティングが得られない場合は、最大用量のデータを除外した残りのデータに各モデルを再度フィッティングさせることもできる[21]。この手法では、適切なフィッティングが得られるまで最大用量のデータを除外するプロセスを繰り返してよい。ただし、毒性データでは、用量群の数が少ないため、この手法の適用に限界が生じる可能性がある。また、この手法の適用に付随し、自由度をはじめとする様々な損失がもたらされる。特に、連続値データにモデルを当てはめる際に高 (及び中) 用量を除外すると、分散をモデリングするための情報が失われる可能性がある[22]。そのため、用量群の除外は慎重に行い、結果はモデリングされたデータ範囲に限定されたものであることに留意する。モデリング結果の報告においては、除外した群を明確にして手法の詳細や正当性を示し、除外した用量を含めた場合の結果も提示した上で、データを除外した場合の影響について考察すべきである。また、この手法は、既に適切なフィッティングが得られたモデルのフィッティングをさらに改善する目的では適用しない。

(16)　モデルの比較選定

与えられたデータセットに対して複数のモデルが適当なフィッティング結果をもたらす場合が多々ある。ベイズ推定又は他の方法を用い、各モデルから算出される推計値の重み付けを通じた統合を行うモデル平均化法も検討されている。こうした手法により、リスクの推定におけるモデルの不確実性の影響を説明できる可能性があるが、その適用は容易で

[21] 漸近項を推定するモデル (Hill モデルなど) を使用する場合は、高用量データの削除には注意が必要。
[22] 分散分析モデルがデータを適切に説明できない場合は、より柔軟なモデルを検討する必要がある。また、この状態は外れ値の存在を示している可能性がある。

はなく、異なる多様な結果が得られる可能性もあることから、明確な指針が求められている。このため、現時点では、より良いフィッティングを示しかつ妥当モデルを選択することを推奨する。

適合度の統計値では異なるモデルの比較を行うことはできず、あるモデルの適合度検定の p 値が高いからといって、p 値がより低い他のモデルよりも適合度が高いとは限らない。以下に、ガイダンスに記載されているモデル比較手法や選定方法を示す。

(5)-1　同一分布族に属するモデルの比較

尤度比検定により、追加パラメータの推定によりフィッティングが適切に改善されるかどうかを評価する。これは、一般的に用量反応モデルでは追加パラメータの導入によってフィッティングが向上することを利用している。

(5)-2　異なる分布族に属するモデルの比較

同一のフィッティング方法 (最小二乗法、二項分布のパラメータに対する最尤推定法等) を使用したモデルを比較する場合は、AIC などの統計学的指標を用いる。このような方法は正確ではないが、モデル選択においては有用な指針をもたらす。

(5)-3　類似エンドポイントをもつ複数のデータセットが存在する場合の比較

あるモデルが全てのデータセットを説明できると仮定した上で、同じ種類のモデルを全てのデータにフィットさせ、試験間で BMD 算出結果を比較することによりモデルを比較できる場合がある。また、特定のデータをフィッティングさせるための既存の手法が存在する場合、それを用いることもできる。ただし、いずれの方法も、フィッティングが良くないモデルの使用を正当化するものではない。また、可能な限り、モデルのパラメータは少ないほうが良いと見なされる場合が多い。

(17)　BMDL の算出

適切なフィッティングを行った後、BMDL (BMD の信頼区間の下限値) の算出を行う。BMDL は以下の手順で算出する。

1) エンドポイントの選択
2) BMR (あらかじめ設定する対照群と比べた反応量の変化レベル) の特定
3) 適切な推定手法を用いた、データに適切にフィットするモデルの構築
4) 用途に応じた、信頼区間 (片側又は両側) 及び信頼水準 (例: 95%) の選択
5) BMD の推定に用いたモデル及び推定手法を用いた、あらかじめ選択した BMR における信頼区間の算出

BMDL の算出には、統計学的問題を正確に組み立て特別なプログラムを作成し、かつ、

作成したプログラムが正確な結果をもたらすものであるか検証する必要がある。このため、EPA の BMDS のように方法論がしっかり文書化されたものを使用することが望まれるとともに、特別に構築されたプログラムがしっかりと文書化されたものであることが望まれる。

信頼区間 (confidence interval (CI)) は、サンプリングエラー及び/又は測定誤差に起因したパラメータ推定における不確実性を表すものである。用途に応じて信頼水準 (ガイダンスの推奨値は 95%) を設定する。信頼区間 (CI) の両端が信頼限界と呼ばれる。なお、算出された信頼区間 (CI) は近似値であり、CI やその上下端である信頼限界は、モデリングされた動物データと検討対象となるヒト母集団との間の対応関係を説明又は仮定するものではないことに留意する必要がある。

BMR を超える反応を示さないことが高い確率で (例: 95%) 保証される用量を、BMD の信頼区間の下限値 (BMDL)を算出することで得る。設計や手順がより優れた試験のデータを用いることで、この信頼区間 (CI) はより精密になり (狭まり)、より高い値の BMDL が得られる。BMDL 又はベンチマーク濃度の信頼区間の下限値 (benchmark concentration lower bound (BMCL)) の計算のためのいくつかの手法が Gaylor et al. (1998)[23] により示されている。

信頼限界を算出する手法は、概してモデルから BMD を推定する手法と関連する。BMD の推定手法やデータタイプ別の BMDL 算出法を以下に示す。

- 最尤推定法によりパラメータ推定を行った場合、CI は尤度比の漸近分布 (プロファイル尤度) 又は最尤推定値 (MLE) の漸近分布に基づく。カバー領域がより広い傾向があることから、通常、最尤推定法により推定したパラメータについては、CI は尤度比の漸近分布に基づくことが望ましい (Crump and Howe (1985)[24])。尤度比の漸近分布に基づくパラメータの片側 100×(1-α) %の CI を計算するためには、当該モデル中の全パラメータの MLE を計算した上で、CI の計算を行う対象パラメータ (μ とする) を他のパラメータと分離する。そして、その他のパラメータの尤度が最大となるように調整したときに、対数尤度が MLE での値から正確に $\chi^2_{(1,1-2\alpha)/2}$ だけ減少するような μ の値を求める (ここで $\chi^2_{(1,1-2\alpha)/2}$ は自由度 1、上側確率 2 における χ^2 分布の分位値) (例として Crump and Howe (1985)[23]; Venzon and Moolgavkar (1988)[25]を参照)。
- 非連続データの場合、まず、標準的な統計理論 (特に、デルタ法 (Gart et al (1986)[26])

[23] Gaylor, DW; Ryan, L; Krewski, D; Zhu, Y. (1998) Procedures for calculating benchmark doses for health risk assessment. Regul Toxicol Pharmacol 28:150-164.
[24] Crump, KS; Howe, R. (1985). A review of methods for calculating statistical confidence limits in low-dose extrapolation. In: Clayson, DB; Krewski, D; Munro, I; eds. Toxicological Risk Assessment. Boca Raton, FL: CRC Press, Inc.
[25] Venzon, DJ; Moolgavkar, SH. (1988) A method for computing profile-likelihood-based confidence intervals. Appl Stat 37:87-94.
[26] Gart, JJ; Krewski, D; Lee, PN; Tarone, RE; Wahrendorf, J. (1986) Statistical methods in cancer research. Vol. 3.

により BMD の分散を推定し、この推定値を用いて BMD の信頼区間の下限値を求める。正値の BMDL を確実に得るためには、対数化した用量を使用する。

- クラスター化されたデータ (生殖発生影響データ等) の場合、モデリングに用いた手法で CI を直接算出できる。生殖発生毒性が生殖過程の多くの段階で生じる場合は、用量反応モデルの開発及び信頼限界の算出に関して多くの手法が利用可能である (Chen et al. (1991)[27]; Ryan et al. (1991)[28]; Catalano and Ryan (1992)[29]; Ryan (1992b)[30]; Catalano et al. (1993)[31]; Zhu et al. (1994)[32]; Krewski and Zhu (1994)[33]; Krewski and Zhu (1995)[34])。

(18) POD 算出に使用するモデルの選択

POD とする BMDL の算出に使用するモデルの選択は、下記の手順で行う。

1) 臨界値の決定に α= 0.1 (特定のモデルを使用する理由がある場合は α= 0.05 又は α= 0.01) を用いて決定した臨界値を用いて適合度検定を行う。
2) 残差及び目視によるプロットの確認により、低用量領域における用量反応関係を適切に説明していないモデルをさらに除外する。
3) 理論上、1)及び 2)で除外されなかったモデルは、いずれも BMDL 算出に利用できる。このため、以下の選定基準は、必然的にいくらか恣意的であって、デフォルトとしての提案である。
4) 各モデルの BMDL 推定値が十分近い値をとる場合は、推定値に対する個別モデル特有の影響は見られないと考えられるため、最小 AIC を有するモデルを POD とする BMDL の算出に用いても良い。最小 AIC を有するモデルが複数存在する場合は、最小 AIC を有するモデルにより算出された BMDL の算術平均又は幾何平均を使用してもよい。ただし、これは「モデル平均化」ではない。また、算出された平均値は個々の推計値の統計学的性質を失った単なる特定の BMDL の平均値であり、平均し

The Design and Analysis of Long-Term Animal Experiments. Lyon, France: International Agency for Research on Cancer.

[27] Chen, C; Farland, W. (1991) Incorporating cell proliferation in quantitative cancer risk assessment: approaches, issues, and uncertainties. In: Butterworth, B; Slaga, T; Farland, W; et al., eds. Chemical induced cell proliferation: implications for risk assessment. New York, NY: Wiley-Liss, pp. 481‑499.

[28] Ryan, LM; Catalano, PJ; Kimmel, C; Kimmel, G. (1991) Relationship between fetal weight and malformation in developmental toxicity studies. Teratology 44:215-223.

[29] Catalano, PJ; Ryan, LM. (1992) Bivariate latent variable models for clustered discrete and continuous outcomes. J Am Stat Assoc 87:651-658.

[30] Ryan, L. (1992b) Quantitative risk assessment for developmental toxicity. Biometrics 48:163-174.

[31] Catalano, PJ; Scharfstein, DO; Ryan, LM; Kimmel, CA; Kimmel, GL. (1993) Statistical model for fetal death fetal weight and malformation in developmental toxicity studies. Teratology 47:281-290.

[32] Zhu, Y; Krewski, D; Ross, WH. (1994) Dose-response models for correlated multinomial data from developmental toxicity studies. Appl Stat 43:583-598.

[33] Krewski, D; Zhu, Y. (1994). Applications of multinomial dose-response models in developmental toxicity risk assessment. Risk Anal 14:613-627.

[34] Krewski, D; Zhu, Y. (1995) A simple data transformation for estimating benchmark doses in developmental toxicity experiments. Risk Anal 15:29-39.

た BMD の 95%信頼下限値ではないことに留意する。

5) 各モデルの BMDL 推定値が十分近い値をとらない場合は、推定値がいくらかのモデル依存性を有していることが示唆されるため、この時点で、結果の一部又は全部を信頼する上でモデルの不確実性が大きすぎるか否かを専門家が統計学的に判断する。推定値の範囲が合理的であると判断された場合、それらのモデルの中から特定のモデルを選択する明確な生物学的又は統計学的根拠はなく、合理的で保守的な推定値として、最小 BMDL を選択してもよい。さらなる解析と考察事項としては、追加モデルの検討、使用したモデルのパラメータ値の検討、BMD においても BMDL と同様の傾向が存在するかどうかを判断するための評価等が挙げられる。そして、これらの決定過程における検討内容を明記しておく。

6) モデリングを試みても有用な結果が得られないケースもある。試験におけるエンドポイントは適切と考えられるものの、データがモデリングに適さないと考えられる試験に由来する場合、NOAEL (又は LOAEL) を POD として用いる。

(19) 推奨される報告事項

BMD 及び BMDL 算出時における報告事項として、次の事項を明記することを推奨する。

1) BMD 算出のために選択した試験

a) 試験選択の論理的根拠

b) エンドポイント選択の論理的根拠

c) 使用した用量反応データのリスト

2) 各試験について選択した用量反応モデル

a) 論理的根拠

b) モデルパラメータの推定手順 (最尤法、最小二乗法、一般化推定方程式等)

c) モデルパラメータの推定値

d) 適合度 (χ^2 検定結果等)、対数尤度及び AIC

e) 標準化残差 ((観測反応値－予測反応値) / SE)

3) 各試験についての BMR の選択

a) 論理的根拠

b) 連続値データの場合、使用した手法

4) 各試験についての BMD の計算

5) 各試験についての BMD の信頼区間の下限値 (BMDL) の算出

a) 信頼区間の算出手順 (プロファイル尤度法、デルタ法、ブートストラップ等)

b) BMDL 値

6) 図

a) 各データポイントとエラー (SD) バーを含む用量反応曲線のプロット

b) 用量反応曲線の信頼区間のプロット (任意。これを含む場合は、算出に使用した方法についても説明。)

c) 同定した BMD と BMDL

7) 標準 BMR に対応する BMD 及び BMDL (比較用)

a) 二値データの場合、過剰リスク 0.10 に対応する BMD と BMDL

b) 連続値データの場合、対照群の 1SD に等しい平均反応値の変化に対応する BMD と BMDL

8) 推定の適用と実行可能性に応じて、BMD の信頼区間の上限値 (BMDU)

2.4.3 BMD 法の適用方針

以上を踏まえ、Benchmark Dose Technical Guidance (EPA, 2012) における BMD 法の適用方針のポイントを以下に整理する。

- データの種類は、連続値データ、非連続データ (二値データ、カテゴリカルデータ) に整理されている。
- モデリングを行う試験及びエンドポイントの選定においては、ヒトのばく露状況、試験の質、報告内容の十分性及びエンドポイントの関連性に基づき妥当性のある全ての試験についてモデリングを行うことを検討し、ばく露に関連すると判断される全てのエンドポイントについてモデリングを行うことを検討することが推奨されている。
- BMD 法に適用できるデータの基準として、以下が挙げられている。

・選択したエンドポイントにおいて、統計学的又は生物学的に有意な用量依存傾向が見られること。

・対照群における反応レベルと観察された最大反応レベルの間で用量反応関係がある (BMR に近いデータポイントを有することがより望ましい)。

- BMD に用いるデータについて、統計学的及び生物学的に互換性のあるデータセットについて用量反応のモデリング前に統合できると言及されている。
- BMR の選択についての明確な指針はないとしつつ、非連続データ、連続値データについて以下の一般的なアプローチが示されている。

・非連続データ：

- 比較を行う目的から、標準的な BMR として、常に 10%過剰リスクに対応する BMD 及び BMDL を報告することを推奨する。
- 生物学的検討の結果に応じて、リスク評価のための参照値の POD の根拠として、重篤な影響 (frank effects) 等のある種の影響については 5%又はより低い BMR、あるいは前駆効果のような影響の場合 10%を上回る BMR を用いるのが妥当である可能性がある。
- 統計学的な観点から、ネストした生殖発生毒性試験では BMR として 5%が推奨され、疫学研究のデータでは BMR として 1%が用いられる場合が多い。

・連続値データ

- 生物学的に有意であると考えられる最小の変化レベルが分かる場合、当該変化レベルを BMR とする。
- 個体のデータが得られており、有害影響と考えられる個体レベルを決定できる場合、当該データをハイブリッド法又はカットオフ値を用いて二値化し、二値データの場合と同様に BMR を設定することができる。この場合、二値化よりも、ハイブリッド法による二値化が望ましい。
- 比較を行う目的から、対象群の平均の標準偏差 (SD) 1 つ分の変化に等しい平

均反応レベルに対応する BMD 及び BMDL を常に報告することが推奨される。

- フィッティングを行うモデルの選定では、エンドポイントの種類、試験デザインに応じてモデルを選択し、パラメータ数、パラメータの制限、共変量の設定等を行う。
- 用量に相関し、結果に影響を与える変数を共変量としてモデルに含めて解析することが推奨されている。
- モデルの適合度の基準として適合度検定における p 値 ＞0.1 が示されている。ただし、p 値がそこまで小さくないがデータとモデルの乖離がある場合や、低用量域での反応を適切に説明していない場合があるため、残差及び目視によるプロットの確認を行うことが推奨されている。
- BMDL の算出に使用するモデルの選択においては、適合度検定の結果に加え、残差目視によるプロットの確認を行う。これらの基準に適合するモデルはいずれも BMDL 算出に利用できるが、これらの BMDL が十分に近い場合は AIC が最も低いモデルを選択して BMDL を採用してもよいとされている。これらの BMDL が十分に近いといえない場合は、モデルの不確実性について検討した上で、保守的アプローチとして最小 BMDL を選択してもよいとされている。

2.5 Use of the benchmark dose approach in risk assessment Guidance of the Scientific Committee (EFSA, 2009)

2.5.1 概要

本ガイダンスは、EFSA 科学委員会 (Scientific Committee) 及び科学パネル (Scientific Panel) がリスク評価を行う際に、透明性があり、かつ科学的妥当性のある解析手法を使用することが必要であるとの考えに基づき、NOAEL 法の代替として BMD 法に関し以下の検討を行った結果をまとめたものである。

- EFSA は BMD 法を使用すべきか
- どのような状況において BMD 法を使用するのが適切か
- (動物) 試験研究に由来する用量反応データを解析する際に BMD 法をどのように用いるべきか
- 観察的疫学研究に由来するデータを BMD 法に適用する方法
- 参照点 (reference point (RP)) の導出に BMD 法を用いる場合に不確実係数の選択が必要か

科学委員会は、リスク評価における RP の導出方法として BMD 法と NOAEL 法の長所、短所を比較した結果、BMD 法が利用可能な用量反応データの拡大利用が可能であり、かつ用量反応データに含まれる不確実性を定量化できることから、NOAEL 法よりも科学的により進歩した方法であると結論した。BMD 法を用いることで、特定の BMR の帰結として、より一貫性ある RP が導出される。

BMD 法で得られる健康影響に基づく Health-based guidance value (HBGV) は、NOAEL 法で得られたものと同様に、ヒトの健康を保護するものであることが期待される。そのため、現在用いられている不確実係数のデフォルト値は BMDL 法においても適切な値であって、追加の不確実係数は不要である。

科学委員会は、BMD 法と NOAEL 法から得られる結果は概して同等であることから、過去に実施された全ての評価を BMD 法で再評価する必要はなく、ヒトのばく露量が ADI 値に近接している等、過去のリスク評価内容を再評価する必要がある 場合には、BMD 法を適用することが有用であるとしている。BMD 法は食品中の全ての化学物質に対して、そのカテゴリーや起源によらず適用可能であり、特に以下の場合に有益である。

- NOAEL の同定に不確実性が伴う場合
- 遺伝毒性や発がん性を有する物質等で MOE を算出するための RP を示したい場合
- 観察的疫学データを用いて用量反応評価を行いたい場合

長期的には、HBGV や MOE を導出する際の RP を決定する方法として、BMD 法が利用されることが期待される。EFSA における BMD 法の導入と広範な利用には実務上考慮すべ

き点があり、また、その適用に当たっては一定レベルの専門家の判断と用量反応モデリングの専門知識が必要であることから、科学委員会は科学パネルと EFSA のユニットの専門家が用量反応モデリング及びソフトウェアの習熟のための訓練を受けることを提案している。

なお、本ガイダンスは、その後の BMD 法の適用事例のレビュー結果を踏まえ、2017 年に改訂されている。改訂ガイダンスについては本報告書の 2.6 項に示す。

2.5.2 BMD 法に関する手順

BMD 法の適用手順として、以下の 5 つのステップが示されている。

1) 用量反応データの特定
2) BMR の特定
3) 用量反応モデル候補の選択
4) フィッティングと許容モデルの特定
5) BMD の推定、RP としての BMDL の設定

以下に各ステップの概要を示す。

(20) 用量反応データの特定

用量反応関係のないエンドポイントについては、通常、BMD は算出しない。対象外とするエンドポイントの決定は、目視による確認によって行える場合があり、正式には BMD 解析を試行して、用量反応関係が統計学的に有意でないことを確認することで行う。

反応データには様々なタイプがあり、その中には連続値 (continuous)、非連続 (quantal)、順序カテゴリー (ordinal) データが含まれる。順序カテゴリー (ordinal) データは中間的なデータタイプと解釈することができ、各個体に病理組織学的検査等における重症度分類 (極めて軽度、軽度、中等度等) を割り当てたデータである。順序カテゴリー (ordinal) データは非連続データに変換することもできるが、これにより情報が失われるこため、推奨されない。連続値データでは、理想的には個体別の観察結果が BMD 解析の入力データとなる。各個体のデータが得られず要約データのみの場合、各用量群における平均値、標準偏差 (又は平均値の標準誤差) 、個体数の組み合わせデータに基づき BMD 解析を行う。要約データを用いると、個体別のデータを用いた場合と結果が多少異なる場合がある。非連続データについては、各用量群の個体数と影響を受けた個体数のデータが必要となる。

(2) BMR の特定

BMD 法を用いて RP を算出するためには、BMR として特定の値を選択する必要がある。本ガイダンスで提示する BMR のデフォルト値は“平均すると BMDL が NOAEL と一致する”レベルである。

非連続データの BMR は、バックグラウンド時の反応と比較したときの病変/スコア化された反応の発現率の増加として定義される。一般的には、追加リスク又は過剰リスクが用いられる。本ガイダンスでは、非連続データにおける BMR のデフォルト値として 10%過剰リスクを推奨する。

連続値データの BMR は様々な方法で定義されるが、本ガイダンスでは、バックグラウンド時の反応レベルと比較した反応変数の平均値におけるパーセント変化量として定義することを推奨する。そして、連続値データにおける BMR のデフォルト値として 5% (例: 赤血球数の 5%減少) を推奨する。一部の連続値データ (例: 認知スコア) では、異常な反応を起こした個体数の 5%増加といったように、非連続データと同様の方法で BMR を定義する方が適切な場合がある。このような BMR の設定方法は、ハイブリッド法と呼ばれる。

なお、科学委員会は、BMR の初期設定値について統計学的あるいは毒性学的考察に基づいて修正する可能性がある。

(3) 用量反応モデル候補の選択

(3)-1 推奨モデル

主な BMD ソフトウェアパッケージ (BMDS や PROAST) は、様々なタイプのデータセットの解析に適切に対応可能なモデルセットを収載しており、他のソフトを用いる場合も同様のモデルセットを用いることを推奨する（表 2.5-1 参照)。いくつかモデルはネストした一連のモデルの一部として収載されている。ネストした一連のモデルを使用する利点は、どのモデルが適切なパラメータ数を有するかを決められることである。具体的には、はじめにパラメータがほとんどない単純なモデルをフィッティングさせ、その後、モデルにパラメータを追加し、フィッティングが改善するかを確認しながら統計学的に最適なモデルを見いだすことができる。フィッティングが改善するか否かの判断は、尤度比検定による有意差検定に基づき実施する。

表 2.5-1　BMD 法での使用が推奨されるモデル [4)]

モデル	パラメータ数	モデル式 用量 (x) の関数である反応 (y)	制限
連続値データ			
指数分布族 (Exponential family)			
Model 1 [1)]	1	$y = a$	$a>0$
Model 2	2	$y = a \exp(bx)$	$a>0$
Model 3	3	$y = a \exp(bx^d)$	$a>0, d>1$
Model 4	3	$y = a [c – (c-1)\exp(- bx)]$	$a>0, b>0, c>0$
Model 5	4	$y = a [c – (c-1)\exp(- bx^d)]$	$a>0, b>0, c>0, d>1$
Hill 分布族 (Hill family)			
Model 2	2	$y = a [1 - x/(b+x)]$	$a>0$
Model 3	3	$y = a [1 - x^d/(b^d+x^d)]$	$a>0, d>1$
Model 4	3	$y = a [1 + (c-1)x/(b+x)]$	$a>0, b>0, c>0$
Model 5	4	$y = a [1 + (c-1)x^d/(b^d+x^d)]$	$a>0, b>0, c>0, d>1$
非連続データ [2)]			
Logistic	2	$y = 1 / (1 + \exp(-a – bx))$	$b>0$
Probit	2	$y = \mathrm{CumNorm}(a + bx)$	$b>0$
Log-logistic	3	$y = a + (1-a) / (1 + \exp(-\log(x/b) / c))$	$0\leq a \leq 1, b>0, c>1$
Log-probit	3	$y = a + (1-a)\ \mathrm{CumNorm}(\log(x/b) / c)$	$0\leq a \leq 1, b>0, c>0$
Weibull	3	$y = a + (1-a) \exp((x/b)^c)$	$0\leq a \leq 1, b>0, c>1$
Gamma	3	$y = a + (1-a)\ \mathrm{CumGam}(bx^c)$	$0\leq a \leq 1, b>0, c>1$
多段階線形分布族 (Linearized multistage (LMS) family) [3)]			
One-stage	2	$y = a + (1-a) \exp(– bx)$	$a>0, b>0$
Two-stage	3	$y = a + (1-a) \exp(– bx – cx^2)$	$a>0, b>0, c>0$
Three-stage	4	$y = a + (1-a) \exp(– bx – cx^2– dx^3)$	$a>0, b>0, c>0, d>0$

a, b, c, d: モデルをデータにフィッティングさせて推定される未知のパラメータ
CumNorm: 累積 (標準) 正規分布関数
CumGam: 累積 Gamma 分布関数
1) Model 1 はいずれの用量反応モデルに対してもネストしたモデルと見なされる (= 水平線)。
2) ここで示す制限により、モデルは用量反応曲線を昇る結果となる。
3) One-stage は BMDS に導入された非連続 Linear モデルと同じである。BMDS ではこのモデルは“multistage”と呼ばれ、段階の数は多項式モデルの次数により定義される (例: two-stage model では 2)。
4) 疫学研究では、追加のモデル (例: y=a+bx) も利用される。

＜連続値データの場合＞

連続値データでは、以下の特徴から、指数分布族 (Exponential family) と Hill 分布族 (Hill family) のモデルを推奨する。

- 単調な用量反応 (常に増加又は減少) を示す。
- 最大反応量で横ばいになる、あるいは閾値が存在するような曲線を示す。
- 実際に遭遇するほぼ全ての用量反応データを表現するのに十分な柔軟性を示してきた。
- 毒性学的に意味のある方法で、共変量を組み入れることができる。

用量反応データが負の値を含む (例えば高用量で体重増加量が正から負の値に減少する) ような場合には、追加のバックグラウンドのパラメータを有するモデルが必要となる場合がある。

＜非連続データの場合＞

表 2.5-1 に示す 7 つのモデルが推奨されており、このうち多段階線形 (Linearized multistage (LMS)) 分布族のみがネストしたモデルとなっている。よって、この分布族からは one-multi stage モデルのみを選択し、BMD 解析に用いるべきである。

(3)-2　パラメータの制限

望ましくない特性を有するモデルを避けるため、モデルのパラメータにある一定の制限が課される。例えば、連続反応量は通常、負の値とはならないことから、連続値データからなる用量反応を示すモデル (連続値モデル) においてはバックグラウンド時の反応を表すパラメータが正となるような制限が加えられる。非連続データからなる用量反応を示すモデル (非連続モデル) では、バックグラウンド時の反応を表すパラメータが 0 と 1 の間 (反応としては 0%と 100%の間) になるように制限される。

さらに用量 0 で曲線の傾きの大きさが無限大となることを避けるために、一部のモデルにある”shape (形状) ”パラメータは以下のように制限される場合が多い。

(連続値データ) 指数分布族と Hill 分布族の M3 と M5 モデル：　　$d > 1$

(非連続データ) Log-logistic、Weibull、Gamma モデル：　　$c > 1$

(4) フィッティングと許容モデルの特定

フィッティングとは、用量反応曲線がデータに限りなく近づくようにモデル中の未知パラメータの値を見いだすことである。この状態は最適フィット (best fit) と呼ばれ、対数尤度を最大化することで達成される。

BMD 法は、統計学的に最適な単一の推定値を見いだすことを目指しているのではなく、むしろ実データと符合する尤もらしい値を全て見いだすことを目指している。このため、最適フィットをもたらす単一のモデルを見いだすことではなく、許容できるフィットが得られるモデルを見いだすことがゴールである。

フィッティングしたモデルが許容できるか否かは、2 つの原則に基づいて判断する。1 つ目の原則は、同一モデル族の種々のモデルの対数尤度を尤度比検定で比較した結果に基づき、ネストしたモデルの中からは 1 つのモデルだけが選定されることである。パラメータがより少ないモデルで有意にフィッティングが悪くなければ、パラメータがより少ないモデルの方が望ましい。2 つ目の原則は、適合度検定で P 値 > 0.05 であれば、モデルのフィッティングは当該用量反応データを合理的に説明するというものである。適合度検定の方法は複数あり、尤度比検定は選択肢の一つである。尤度比検定では、フィッティングしたモデルに関連する対数尤度を、いわゆる“フルモデル (full model)”の対数尤度と比較し、検定する。検定の結果、当該モデルのフィッティングがフルモデル (full model) より有意に劣っていなければ、当該モデルは許容される。

(5) BMD の推定、RP とする BMDL の設定

試験・研究の中で観察された重要と考えられるエンドポイントについて、BMD 解析を行う。

・重要と考えられるエンドポイントについて一連のモデルを適用し、許容モデルごとに BMD とその信頼区間を算出する。この信頼区間の下限値が BMDL である。信頼水準のデフォルトは 95% (片側) である。

・得られた BMDL の中で最小の BMDL (つまり、各エンドポイントにおいて最も保守的な値) を、重要と考えられる各エンドポイントにおける BMDL とする。

・全てのエンドポイントにおいて同様に算出された BMDL の中から最小の BMDL を決定する。この値が当該試験・研究全体を代表する BMDL となる。

同一エンドポイントの中で最小の BMDL を選択することが保守的であると考えられるが、この手法が推奨されるのは、“モデル平均化”のような、より進歩的な方法が完全に開発され正当化されるまでとする。

このように、BMDL は様々なエンドポイントについて決定されうる。BMD 解析の前に重要なエンドポイントを特定できない場合は、NOAEL 法で取られる方法と同様に、BMD 解析結果に基づき、ヒトの健康影響に関連する影響であることが確認されたものの中から、最小の BMDL がもたらされるエンドポイントを選択してもよい。

その他、BMDL の設定に用いる用量反応データの妥当性と BMDL の設定に関する見解を以下に示す。

・連続反応レベルの間に大きなギャップがある場合など、用量反応関係を定義できないデータにフィットさせた種々のモデルは、異なる形状となることが予想され、得られる BMDL 値の幅は大きくなる。このような BMDL は HBGV の設定のための確実な基準とはならないであろう。

・得られた BMDL の範囲に基づき用量反応データの妥当性を判断する基準は、まだ確立されていない。

・一般的なルールとして、用量反応データについて、許容される種々のモデルから推定される BMDL の範囲が一桁を超えるべきではない。この値を超えるときはケースバイケースで、例えば BMR を増加させる、モデルセットを再評価する、モデル平均化を行う等の対応を検討すべきである。

2.5.3　その他

(1) 共変量について

統計解析における共変量の使用は、ヒトのデータ解析では一般的な手法であるが、動物試験の解析では今のところ限られている。

ヒトでの研究における反応変数は、観察目的とする用量反応に干渉する可能性のある交絡因子の影響を受ける。この交絡因子を考慮しないことは、BMD の過少/過大推計につながる可能性があり、共変量としてモデルに組み込むことで対処する。

例えば、高齢者になると傷病にかかりやすくなるが、このような影響の変容 (effect modification) によって付加的な問題が生じる場合は、最も脆弱な特定集団における反応を反映した BMDL を採用すべきである。このときに関連する共変量を含めることで、影響の変容 (effect modification) をモデリングすることができる。

このような交絡因子や影響の変容 (effect modification) に関する補正は BMDS ソフトでは実行できないが、PROAST ソフトでは一部可能である。

(2)　ヒトの用量反応データを使う場合の課題

観察的疫学研究に由来する用量反応データは、いくつかの点で典型的な動物の毒性データと異なる。

観察研究は非ばく露対照群を含まないことが多く、例えば大気汚染物質や食品汚染物質などの場合は、全ての対象者がある程度ばく露されている。このような場合でも、用量反応曲線をフィッティングさせるのに必ずしもばく露量がゼロの場合の観察値は必要としないため、BMD法を適用することができる。しかし、ばく露量がゼロでの反応量を低用量外挿によって推定する必要があるため、疫学研究データから得られたBMDはモデル依存性が高いといえる。

また、ヒトのばく露量の推定値は不正確であることが多く、この点に考慮を払わないと用量反応関係の評価にバイアスが生じる可能性がある。この課題に対処し、かつバイアスの小さい用量反応関係及びBMD値を推定するための統計学的手法が開発されている。

(3)　BMD 解析結果の報告について

BMD 解析結果は、他者がその解析法を追従可能な方法で報告するべきである。ある特定の試験データについて BMD を解析した結果の報告に当たっては、必ずしも解析を行った全てのエンドポイントの情報を提供する必要はなく、重要なエンドポイントについてのみ報告すればよい。報告すべき情報は以下のとおりである。

・BMD の解析結果のうち報告対象としたエンドポイントに関する用量反応データを要約した表
・選択した BMR の値とその根拠
・使用ソフトウェアとそのバージョン

・モデルのフィッティング手順における設定条件及び仮定条件
・使用モデルと対数尤度の表、フィッティングさせたモデル及び許容したモデルに関する情報、許容モデルから算出された BMD と BMDL
・重要なエンドポイントに関する用量反応データにフィッティングさせたモデルのうち、少なくとも一つのプロット (BMDL を含む)
・当該試験・研究から導出された BMDL に関する結論

2.5.4 BMD 法の適用方針

以上を踏まえ、Use of the benchmark dose approach in risk assessment1 Guidance of the Scientific Committee (EFSA, 2009) における BMD 法の適用方針のポイントを以下に整理する。

- 用量反応関係のないエンドポイントは BMD 算出に適用できない。エンドポイントの決定は、目視による確認によって行える場合があり正式には、BMD 解析を試行して、統計学的に有意でないことを確認することにより行う。
- 非連続データの BMR のデフォルト値として 10% (過剰リスク)、連続値データの BMR のデフォルト値として 5% (反応量の変化レベル) が推奨されている。
- モデルの選択においては、パラメータ数が少ないモデルから検討を始め、パラメータを追加し、フィッティングが改善するかを確認しながら統計学的に最適なモデルを見いだすことを推奨している。フィッティングが改善するか否かの判断は、尤度比検定による有意差検定に基づき実施する。
- 望ましくない特性を有するモデルを避けるため、モデルのパラメータに制限を課すことが言及されている。
- 適合度の基準として適合度検定の P 値 ＞0.05 が示されている。
- 信頼区間設定における信頼水準のデフォルトは 95% (片側) とされている。
- モデル平均化法等が開発されるまで、保守的なアプローチとして最小 BMDL の選択が推奨される。
- BMD 法は、データに合う全ての尤もらしい (plausible) 値を見いだすことを目的としており、許容できるフィットが得られる全てのモデルを見いだすべきとしている。
- 許容される種々のモデルから推定される BMDL の範囲が一桁を超える場合は、BMR を増加させる、モデルセットを再評価する、モデル平均化を行う等の対応を検討すべきであるとしている。

2.6 Update: use of the benchmark dose approach in risk assessment (EFSA, 2017)

2.6.1 概要

➢ 背景

2009 年、EFSA 科学委員会 (Scientific Committee) が、リスク評価における BMD 法の利用に関するガイダンスを公表した (本報告書の 2.5 項参照)。

その後、2015 年、同科学委員会が、BMD 法の適用によって得られた経験及びリスク評価分野における近年の方法論的な発展内容を基に、EFSA における BMD 法の導入・実装についてレビューした結果、2009 年に公表したガイダンスの改訂が必要であると結論し、その改訂が行われた。

➢ 概要

本ガイダンスは、2009 年に作成されたガイダンスのうち、主に実際の BMD 法の適用方法に関するセクションについて改訂したものである。その中で、モデル平均化について、個別のツールは開発途上であることを認識した上で、同法が BMD の信頼区間の計算に望ましい方法であるとして推奨している。また、BMD 解析に用いるデフォルトモデルが見直され、用量反応データセットに対する種々の数理モデルの相対的な適合度を明らかにするために、対数尤度に代えて AIC が導入された。さらに、BMD 法を適用する場合には、BMD 値ではなく、BMD の信頼区間を報告することが推奨されている。BMD の信頼区間の下限値である BMDL は潜在的な RP として、BMD の信頼区間の上限である BMDU は BMDU/BMDL 比を求めるためにそれぞれ必要であり、これらによって BMD の推定における不確実性が表されるとしている。

科学委員会は本ガイダンスの中で、BMD 法、特にモデル平均化手法を RP の導出に用いるべきであるとしている。また、科学委員会は、科学パネルや EFSA ユニットの専門家が用量反応モデリングや BMD ソフトウェアの使用に関する訓練を受けることを推奨している。さらに、BMD の解析に関して、EFSA の専門家や職員が必要に応じて相談する常設のワーキンググループを設置すべきとしている。

そのほか、科学委員会は、ヒトのデータの解析における BMD 法の使用に関する特定のガイダンスの必要性について確認するとともに、今後 BMD 法が幅広く適用されると予想されることから、現行の毒性試験ガイドラインの再検討が必要であるとしている。

2.6.2 BMD 法に関する手順

BMD 法の適用手順として以下の 5 つのステップが示されている。

1) 用量反応データの特定
2) BMR の特定
3) 用量反応モデル候補の選択

4) 用量反応モデル候補のフィッティングと BMD の信頼区間の算出
5) 種々のモデリング結果の単一の BMD の信頼区間への統合、RP としての BMDL の設定

これらの適用手順は、基本的には 2009 年のガイダンスと同様であるが、主に以下に点が改訂されている。

- 連続値データの BMR のデフォルト値を 10%とする根拠が追記された。
- 連続値データの BMR の特定において、EPA の Benchmark Dose Technical Guidance に記載された標準偏差 (SD) 1 つ分の変化に対応する反応レベルを BMR として利用することについて問題があるとの見解が加えられた。
- 用量反応モデルについて、2009 年のガイダンスに記載された推奨モデルから一部のモデルの削除等が行われた。
- パラメータの制限に関する記述が追加された。バックグラウンド時の反応パラメータは連続値モデルにおいて正の値 (> 0)、非連続モデルでは0から1の値になるように制限するとともに、曲線の急峻さを示す形状パラメータ (急峻度パラメータ) については、用量0で曲線の傾きの大きさが無限大となることを回避するためにとられることがある1より大きい値 (> 1) への制限は適切でなく、人為的に高い値のBMDLにつながる可能性があるため、適用すべきではないとしている。
- 用量反応モデル候補のフィッティング結果の評価指標として、対数尤度に代えて AIC が導入された。
- モデル平均化法に関する記載が追記された。
- 特定のエンドポイントにおける用量反応データセットからBMDの信頼区間及びBMDLを算出するフローが追加された。BMDの信頼区間及びBMDLの算出において、モデル平均化機能を有するソフトウェアが利用可能な場合は、モデル平均化により算出されるBMDの信頼区間及びBMDLを採用する方針が示された。
- モデル平均化機能を有するソフトウェアが利用できない場合は、AIC値が最小のモデリング結果及びAIC値が最小値の2単位以内の範囲にあるモデリング結果から、それぞれBMDの信頼区間を算出し、得られたBMDLの最小値及びBMDUの最大値を、当該データセットから算出されるBMDの信頼区間として採用する方針が示された。

以下に、これらの主な改訂点を中心に、ガイダンスの記載内容を示す。

(1) BMR の特定

(1)-1 非連続データ

非連続データの BMR は、バックグラウンド時の発現率と比較したときの病変/スコ

ア化された反応の発現率の増加として定義される。毒性学においては、そのような増加を反映する一般的な単位として、追加リスク (additional risk) (特定の用量での発現率から対照群での発現率を差し引いたもの) と過剰リスク (extra risk) (追加リスクを、対照群において影響を受けない集団の発現率で割ったもの) がある。

試験動物で観察された非連続反応データにおける BMR の値として、当初は追加又は過剰リスク 1%、5%又は 10%が提案されていたが、多くの研究で NOAEL における過剰リスクの上限の中央値が 10%に近いと推定されており、$BMDL_{10}$ が適切なデフォルト値であることが示唆されている。また、より低い BMR では BMDL が用量反応モデルの選択に大きく依存する可能性があることからも、非連続データにおける BMR の値としては 10%が好ましいと考えられる。

(1)-2 連続値データ

連続値データの場合、BMR は様々な方法で定義できる。1 つの方法は、BMR を標準偏差 (SD) として測定される対照群の変動と比較した平均反応量の変化率として定義する方法である。EPA の Benchmark Dose Technical Guidance (本報告書の 2.4 項参照) では、1 SD に相当する平均値の差として定義された BMR に対応する BMD を報告することを推奨している。しかし、この BMR の定義方法における一つの問題点として、対応する (真の) BMD が、試験に固有の要因 (測定誤差、投与誤差、試験条件の不均一性) が原因で、試験に依存してしまう点が挙げられる。また、1SD を用いる方法のもう 1 つの問題点として、この BMR に対応する BMD の推定値は、ヒトを含む、より大きな群内変動を有する集団における等価用量に変換できない点が挙げられる。BMR を定義するためのもう 1 つの方法は、BMR を平均反応量の変化率として定義することである。この方法で定義される BMR に対応した BMD は、群内変動に依存せず、同じ用量反応関係を調べた異なる試験間及び異なる集団間においてより安定している。このため、科学委員会は、BMR をバックグラウンド時と比較した平均反応量の変化率として定義することを推奨する。

多くの NTP 試験を再解析した結果、平均すると、$BMDL_{05}$ が同じデータセットから得られた NOAEL に近いことが示され、ほとんどのデータセットでその差は一桁以内であった (Bokkers and Slob (2007)[35])。胎児体重データの研究においても同様の結果が報告されている (Kavlock et al. (1995)[36])。連続値データでは BMR のデフォルト値として 5%が推奨されるが、毒性学的又は統計的検討に基づき変更される可能性もある。

結論として、科学委員会は、動物試験のデータについて、非連続データでは 10% (過

[35] Bokkers BGH and Slob W, 2007. Deriving a data-based interspecies assessment factor using the NOAEL and the Benchmark dose approach. Critical Review in Toxicology Journal, 37, 353-377.

[36] Kavlock RJ, Allen BC, Faustman EM and Kimmel CA, 1995. Dose-response assessments for developmental toxicity IV. Benchmark doses for fetal weight changes. Fundamental and Applied Toxicology, 26, 211-222.

剰リスク)、 連続値データでは 5% (平均反応量の変化率) を BMR のデフォルト値として用いることを提案する。

なお、BMR のデフォルト値は統計学的又は生物学的考察に基づいて修正できる。例えば、BMR が最低用量で観察された反応よりもかなり小さく、観察範囲を大きく外れた外挿が必要となるような場合、より大きな BMR を選択することができる。このような場合、デフォルトの BMR から逸脱する論理的根拠を文書化して説明する必要がある。

(2) 推奨される用量反応モデル

(2)-1 推奨モデル

表 2.6-1 に本ガイダンスで推奨するモデルを示す。これらのモデルは一般に毒性データセットを分析するのに適していると考えられている。連続値データと非連続データで、使用するモデルは異なるが、双方に関連する特別なモデルが「フルモデル (full model)」と「ヌルモデル (null model)」である。

フルモデル (full model) は、特定の用量反応を想定せず、試験された用量で観察された (平均) 反応量により用量反応関係を単純に説明する。しかし、フルモデル (full model) はモデルの分布部分を含むことから、あらゆる用量反応モデルの適合度の評価に用いることができる。ヌルモデル (null model) は、用量反応の傾向がなく、すなわち反応反応関係が水平線であることから、用量反応の傾向の存在を統計学的に評価するために利用できる。

表 2.6-1　BMD 法での使用が推奨されるモデル

モデル	パラメータ数	モデル式 用量 (x) の関数である平均反応 (y)	制限
フルモデル(i) (full model)	対照を含む用量群の数	各用量で観察された平均値又は発現率の組み合わせ	
ヌルモデル(ii) (null model)	1	$y = a$	連続値データ: $a>0$ 非連続データ: $0<a<1$
連続値データ			
指数分布族 (Exponential family)			
3-parameter Model (iii)	3	$y = a \exp(bx^d)$	$a>0, d>1$
4-parameter Model (iv)	4	$y = a [c – (c-1)\exp(– bx^d)]$	$a>0, b>0, c>0, d>1$
Hill 分布族 (Hill family)			
3-parameter Model (iii)	3	$y = a [1 - x^d/(b^d+x^d)]$	$a>0, d>1$
4-parameter Model (iv)	4	$y = a [1 + (c-1)x^d/(b^d+x^d)]$	$a>0, b>0, c>0, d>1$
非連続データ			
Logistic	2	$y = 1 / (1 + \exp(-a – bx))$	$b>0$
Probit	2	$y = \mathrm{CumNorm}(a + bx)$	$b>0$
Log-logistic	3	$y = a + (1-a) / (1 + \exp(-\log(x/b) / c))$	$0\leq a \leq 1, b>0, c>0$
Log-probit	3	$y = a + (1-a)\ \mathrm{CumNorm}(\log(x/b) / c)$	$0\leq a \leq 1, b>0, c>0$
Weibull	3	$y = a + (1-a) \exp((x/b)^c)$	$0\leq a \leq 1, b>0, c>0$
Gamma	3	$y = a + (1-a)\ \mathrm{CumGam}(bx^c)$	$0\leq a \leq 1, b>0, c>0$
Linearized multistage (LMS) (two-stage) model	3	$y = a + (1-a)(1-\exp(-bx-cx^2))$	$a>0, b>0, c>0$
上記の連続反応モデルに基づく Latent Variable models (LVMs) (v)	ベースとなる連続反応モデルに依る	連続反応を仮定し、データから推定される (潜在的な) カットオフ値に基づき反応有り／反応なしに二値化	連続反応モデルを参照

a, b, c, d: モデルをデータにフィッティングさせて推定される未知のパラメータ

CumNorm: 累積 (標準) 正規分布関数

CumGam: 累積 Gamma 分布関数

(i) フルモデル (full model) は、検討対象のデータセットにおいて、 (ある統計学的仮定条件における) 対数尤度の最大値が得られる。

(ii) ヌルモデル (null model) は、いずれの用量反応モデルに対してもネストしたモデルと見なせ、用量反応関係がない状態 (=水平線) を示す。

(iii) PROAST では model 3 と呼ばれ、BMDS の指数分布モデルと類似。

(iv) PROAST では model 5 と呼ばれ、BMDS の指数分布モデルと類似。

(v) LVMs は PROAST に導入されている。

＜連続値データのモデル＞

指数分布族 (Exponential family) と Hill 分布族 (Hill family) のモデルを推奨する。これらのモデルには以下の特徴がある。

- 常に正の値を予測する。例えば、臓器重量は 0 以下とはならない。
- 単調な用量反応 (常に増加又は減少) を示す。
- 最大反応量で横ばいになるデータに適している。
- 幅広いエンドポイントの用量反応データセットを適切に説明する。
- 毒性学的に意味のある方法で共変量を組み入れることができる。
- 最大 4 つのパラメータ (a, b, c, d) を含み、これらは両方のモデル族で共通である。ここで、a は「用量 0」での反応、b は化学物質の潜在的可能性を示すパラメータ (集団における感受性)、c はバックグラウンド時の反応量と比較した反応の最大比、d は傾きの急峻度を示すパラメータである。

指数分布族と Hill 分布族の両方のモデルについて、それぞれ 3 パラメータモデル、4 パラメータモデルが推奨される。2009 年のガイダンスには他のモデルも含まれたが、パラメータ d が実際には 1 でない場合に、BMD の信頼区間の範囲が狭くなる傾向があることから、本ガイダンスでは推奨しない。

EPA の BMDS には、連続値データのモデルとしてこの他にいくつかのモデル (power model 及び (線形を含む) 多項式モデル) が含まれるが、これらのモデルは、バックグラウンド時の反応量に関して加法的であり、負の値を予測する近似曲線が得られる可能性がある。よって科学委員会はこれらのモデルの使用を推奨しない。

＜非連続データのモデル＞

Two-stage model は、linearized multistage (LMS) model 族にネストしたモデルである。Two-stage model は、他の多くの非連続反応モデルと同様に、尺度パラメータ (a 及び b) に加えて形状パラメータ (c) を有するため、LMS モデル族を使用することを推奨する。2009 年のガイダンスでは経験的に推奨されていた three-stage model は、データにフィットすることがほとんどないことから、推奨モデルから削除した。

潜在変数モデル (Latent Variable models (LVMs)) は PROAST に実装されており、一般に非連続データを適切に説明することが知られている。特にモデル平均化を適用する場合、BMD 解析に利用される。

(2)-2　パラメータの制限

望ましくない特性を有するモデルを避けるため、モデルのパラメータに一定の制限が

課される。例えば、連続反応量は通常、正の値であるため、バックグラウンド時の反応パラメータ (a) は連続値モデルにおいて正となるように制限される。非連続モデルでは、バックグラウンド時の反応パラメータが 0 から 1 の間 (すなわち、0%から 100%の応答) になるように制限される。

表 2.6-1 に示す制限に加えて、曲線の急峻さを示す形状パラメータ (急峻度パラメータ) に関する追加のパラメータ制限 (すなわち、非連続モデルのパラメータ c (c>1) 及び連続値モデル (指数分布及び Hill) のパラメータ d (d>1)) が課される場合が多い。この制限は、用量反応曲線が、用量 0 で曲線の傾きの大きさが無限大となることを回避するためにある。ほとんどのモデルでは、これは急峻度パラメータを (> 0 ではなく) > 1 に制限することで実現できる。一見したところ、これは生物学的観点からの合理的な制限のように見えるが、この制限は誤った議論に基づいており、実際の用量反応データと矛盾する (Slob and Setzer, (2014)[37])。急峻度パラメータを ＞1 とする制限は不適切であり、人為的に高い値の BMDL につながる可能性があるため、適用すべきではない。

(3) モデルのフィッティング

現時点で入手可能なEPAやRIVMのBMDソフトウェアでは、モデルのフィッティングを重要視している。モデルのフィッティングとは、用量反応曲線をデータに限りなく近づけるようなモデル内の未知のパラメータの値を見いだすことである。これは当該モデルにおける最適フィット (best fit) と呼ばれ、各モデルの対数尤度を最大にすることによって達成される。

(3)-1 収束

現在利用可能なBMDソフトウェアでは、推奨モデルに数値アルゴリズムを適用することによってフィッティングを行う。対数尤度が改善しなくなるまで、種々のパラメータ値を用いてモデルのフィッティング結果を繰り返し再評価する。アルゴリズムが最大尤度を見いだすことができる場合、ソフトウェアはアルゴリズムが「収束」したと報告する。一方、アルゴリズムが収束しなかった (非収束) と報告するケースもある。これには様々な理由が考えられるが、通常は、データがモデル内の全てのパラメータを適切に推定するために十分な情報を提供していないことを示している。信頼性のあるBMDの信頼区間を得る上で、収束することは重要ではない可能性があり、非収束のメッセージは、必ずしもモデルを棄却すべきであることを意味するわけではないが、警告と見なす必要がある。

(3)-2 AIC

[37] Slob W and Setzer RW, 2014. Shape and steepness of toxicological dose-response relationships of continuous endpoints. Critical Reviews in Toxicology, 44, 270-297.

AICは、対数尤度とモデルのパラメータ数を1つの値として直接統合したものであるため、異なるモデルのフィッティングを比較する上で便利な基準である。AICはlog (L)をモデルの対数尤度、pをパラメータ数とするとき、-2 log (L) + 2pとして計算される。モデルをデータに近似させると、はじめは、-2 log (L) は減少する。一方、AICには2pという項が含まれ、パラメータ数が増えるとAICの値が増える。よって、AICが比較的低いモデルは、それほど多くのパラメータを使用せずに良好なフィッティングをもたらすと見なすことができる。

AICの差が2単位以内の異なるモデルは、データを同等にうまく説明していると見なすことができる。さらに、AICはフルモデル (full model) で最小、ヌルモデルで (null model) で最大となる傾向がある (多くの用量群が存在する場合はこの限りではない)。AICは、フィッティングさせたモデルが用量反応関係を示すことの統計学的根拠として利用することができ、科学委員会では、その基準として$AIC < AIC_{null} - 2$ (得られたAIC値が、null modelのAIC値から2を減じた値より小さい) を提案する。

AICは、任意のモデルのフィッティングをフルモデル (full model) のフィッティングと比較する目的でも使用できる。フィッティングさせたモデルのAICは、理論的には、フルモデル (full model) のAICよりも2単位以上大きくならないようにする必要がある。AICが最小のモデルがフルモデル (full model) のAICよりも2単位以上大きい場合 ($AIC_{min} > AIC_{full} + 2$)、その理由としては、不適切な用量反応モデルの使用 (例: パラメータ数の不足)、モデルの分布部分の説明が不正確 (例: 同腹効果の無視)、又はデータ内の非ランダムエラーの存在が可能性として挙げられる。

(3)-3 共変量

用量反応モデルを単一のデータセットにフィッティングさせることに加えて、性別、種、ばく露期間など、特定の側面は異なるが、それ以外は類似するデータセットの組み合わせに対して、特定のモデルをフィッティングさせることができる。ただし、反応パラメータ (エンドポイント) は同じでなければならない。いわゆる共変量と呼ばれる特定の因子を解析に含め、組み合わせたデータセットに用量反応モデルをフィッティングさせることにより、AICのような統計学的原則に基づき、部分集団の用量反応にどのような差があるかを検討することができる。

一般に、解析結果として以下の3つの可能性が考えられる。

- 部分集団は類似の用量反応を示し、単一の曲線で全ての部分集団を統合して説明できる可能性がある。
- 部分集団間の用量反応は異なるが、その違いは部分的である。例えば、バックグラウンド (用量ゼロ) 反応量は異なるものの、化学物質に対しては等しい感受性を示す場合、あるいは、感受性が異なる場合でも、用量反応は同じ形状をしている場合がある。後者の場合、解析によって、各部分集団特有のBMDの信頼区間が

得られる。

・部分集団は、モデル内の全てのパラメータが異なると考えられる。この場合、用量反応解析を組み合わせた結果は、部分集団を別々に解析した場合と同じになる。

PROASTのような適切なソフトウェアを用いて、組み合わせ解析を実行することができ、組み合わせたデータセットがどのように最適に説明され得るかが示される。用量反応解析において、データセットを統合し共変量を含めて解析することにより、部分集団間の用量反応の潜在的な差を定量化することができる。また、データセットを統合し、共変量を含めた解析を行うことにより、推定BMDの精度を向上させる、すなわち、より小さいBMDの信頼区間を得ることができる。

(4) モデル平均化

BMD法は、統計学的に最適な単一のBMD推定値を見いだすことを目指しているのではなく、むしろ実際のデータと符合する尤もらしい値を全て見いだすことを目指している。このため、最適なフィットをもたらす単一のモデルを見いだすことではなく、フィッティングを行った全てのモデルから得られる結果を考慮することがゴールである。このための方法として“モデル平均化法”が推奨される。

モデルの不確実性を検討し、同時にデータのサンプリングエラーに関連する不確実性を検討するには、モデル平均化を使用したマルチモデル推定及び推論が最善の方法である。モデル平均化では、個々のモデルの結果は重み付けされながら統合され、データによりよくフィットするモデルでは重みが大きくなる。これらの重み付けはAICの観点から定義されることが多い。

簡潔に述べると、モデル平均化は主に2つのステップから成る。最初のステップで、フィッティングを行った用量反応モデルの加重平均をとることによって、多数の用量について平均反応量を計算する。これにより、平均化したモデルにおけるBMDが計算される。次のステップで、平均化したモデルに基づいて多数のデータセットを人為的に作成し、各データセットに対して最初のステップを繰り返す。これにより、多数のBMDが得られ、上下5パーセンタイルに該当するBMDの範囲がBMDの90%信頼区間として定義される。

(5) BMDの信頼区間の設定

(5)-1 BMDの信頼区間

図2.6-1に、特定のエンドポイントの用量反応データセットに対するBMDの信頼区間及びBMDLを導出するフローチャートを示す。

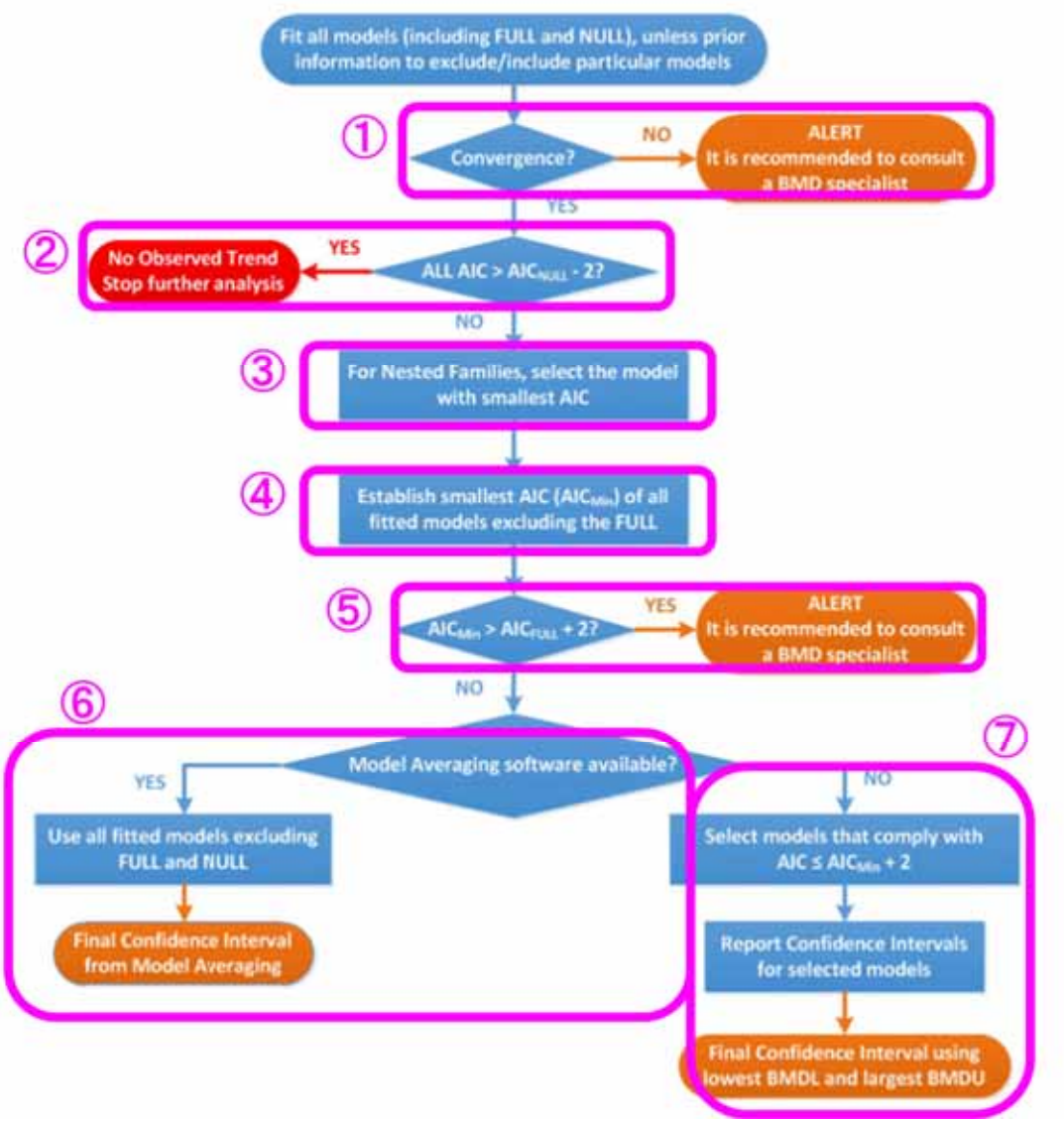


図2.6-1　特定のエンドポイントの用量反応データセットに対する BMDの信頼区間及びBMDLを導出するフローチャート

① ソフトウェアが複数のモデルに対して「非収束」と報告した場合、これは警告 (明らかに、データの有用性が低い、又はモデルのパラメータが過剰である) と見なされる可能性がある。収束しない場合は、BMD解析の進め方について、用量反応モデルの専門家に相談する。

② 少なくとも1つのモデルが用量反応関係を明らかに示すかどうか、すなわち各モデルのAIC値がヌルモデル (null model) のAIC値から2を減じた値より小さいかどうかを確認する。これが満たされない場合、反応量に用量相関の傾向はなく、BMD解析は中止する。

③ ネストしたモデルが使用されている場合 (すなわち連続値データの場合)、モデル族ごとに最小AICが得られたモデルを選択する。

④ フィッティングを行った全てのモデルの中から最小のAIC値を決定する。

⑤ ④の最小AICが、フルモデル (full model) のAIC値に2を加えた値より大きい ($AIC_{min} > AIC_{full} + 2$) 場合、警告と見なされる場合がある。この場合には、まず、モデルの分布部分について調整が必要か否か検討する。次に、データに他の問題 (例: 用量群間において、試験上の要素が意図せず異なってしまったことで生じた系統的な誤差等) がないかを考察する。$AIC_{min} > AIC_{full} + 2$ となるもう1つの理由として、当該データセットにフィットするモデルがなかった可能性が挙げられるが、比較的良好なデータセットにおいて、そのような事態は稀である。

⑥ フィッティングを行ったモデルの結果を統合して、望ましくはモデル平均化により、最終的なBMDの信頼区間を設定する。ここでは、フィッティングが低いモデルも含めて検討する。モデル平均化では、AICによる重み付けを行うため、フィッティングが悪いモデルは最終的なBMDの信頼区間の設定にはほとんど影響しない。

⑦ モデル平均化ソフトウェアを利用できない場合は、フィッティングが比較的良好なモデルとフィッティングが悪いモデルを区別する。フィッティングが比較的良好なモデルのセットは、AICが最小 (AIC_{min}) のモデル及びAICの最小値から2単位以内のAICを有する全てのモデルとする。これらのモデルから算出される最小のBMDLと最大のBMDUを用いてBMDの信頼区間を定義する。この場合に得られるBMDの信頼区間は、一般に、個々のモデルから得られるBMDの90%信頼区間より大きくなる。従って、この場合に得られるBMDLは、通常、モデル平均化により得られるBMDLよりも小さくなる。

(5)-2　BMDの信頼区間の幅の判断

試験データから用量反応関係に関する十分な情報が得られている場合、各モデルは類似のBMDの信頼区間をもたらし、HBGVの設定又はMOEの計算におけるRPを定義するための適切なベースが得られる。

一方、用量反応関係がデータによって明確に示されていないケースもあり、その場合は、あてはめたモデル間の一部又はその全てでBMDの信頼区間が大きく異なる場合や、非常に幅広い (数桁) 場合があり得る。統合したBMDの信頼区間の幅が数桁に及ぶ場合、BMDLは、より良いデータが利用可能な場合に導出される真のBMDに比べて桁違いに低くなる可能性がある。従って、結果として得られるRPや最終的にそれらから導出されるHBGV、MOEは、それぞれはるかに大きくなることがある。そのような場合に、より良いデータを要求できる可能性もあり得るが、多くの場合は不可能である。その代替として、ヒストリカルデータから入手可能であれば、例えば形状パラメータを制限したり、ベイズ推定で事前分布を適用したりすることによって、形状パラメータの典

型的な値に関する事前情報を検討して、データを再解析する方法がある。いかなる選択肢を適用する場合も、明確に文書化すべきである。このオプションは、以下のような理由から、統合した信頼区間幅が広い場合に検討すべきである。

(i) 試験における動物 (又はその他の実験単位) の総数が少ない
(ii) 用量の増加に伴う連続 (平均) 反応量のばらつきが著しい
(iii) 試験デザインの用量設定数が少ない、又は異なる反応が観察される用量数が少ない
(iv) 最高投与量における反応量が比較的小さい
(v) 最低投与量における反応量が比較的大きい

(5)-3　RPの決定

フローチャートに従い、特定のエンドポイントに関連する特定の用量反応データセットについて最終的なBMDの信頼区間が得られる。BMDの信頼区間は、関連性があると考えられる (RPにつながる可能性がある) 全てのデータセットについて導出すべきであり、その結果、検討したエンドポイントに関して、真のBMD周辺の不確実性の範囲を示す信頼区間のセットが得られる。このBMDの信頼区間のセットは、利用可能なデータによって得られる情報を簡潔に反映しており、リスク評価者がRPを導出するための出発点となる。一つの進め方は、最も低いBMDLを有するエンドポイントを選択し、その値をRPとして使用することである。ただし、この手順は全ての場合に最適というわけではないため、リスク評価者は、(単なるBMDLではなく) BMDの信頼区間、関係するエンドポイントの生物学的な意味、導出されるHBGV又はMOEへの影響など、関連するあらゆる側面から検討する。このプロセスはケースバイケースで異なり、リスク評価者の責任において、どのBMDLをRPとして使用するかを決定する。

例えば、最も小さいBMDLを含むBMDの信頼区間が数桁の幅を持つ場合がある。これは、真のBMDがBMDLよりはるかに高い可能性があることを意味し、そのBMDLが適切なRPであるかどうかという疑問が生じる。その疑問に答えるために、以下の事項を検討する：

・特定のBMDLに基づいて設定されたHBGVが、推定ばく露量よりはるかに高い、又はMOEが10,000よりはるかに大きい場合、信頼区間が広くRPの不確実性は高いが、危害要因判定 (Hazard Characterization) には影響がない。ただし、ばく露推定値は固定値ではなく (将来変わる可能性があり)、不確実である。
・場合によっては、選択したRPが最も低いBMDLとならないことがある。例えば、最も低いBMDLが他の評価項目によっても示される影響 (肝臓壊死と血清酵素の組み合わせなど) に関連する場合、信頼区間はより狭くなるが、BMDLの値は高くなる。それらの類似したエンドポイントに対する真のBMDはおそらく類似しているが、(大きな測定誤差などにより) それらの1つがかなり広い信頼区間の算出につながった可能性がある。

(6) BMD解析結果の報告情報

ガイダンスの改訂内容を反映し、BMD解析結果の報告において提供すべき情報についても以下のように改訂されている。

A) BMD 解析で報告されるエンドポイントに関するデータの要約表

B) 選択した BMR の値、デフォルト値と異なる場合はその根拠

C) 使用ソフトウェアとそのバージョン

D) 推奨されているデフォルトの手順から逸脱した場合のモデルのフィッティング手順の設定内容と統計学的仮定及びそのような手順を採用した論理的根拠

E) 使用したモデルを整理した表 (ヌルモデル (null model)、フルモデル (full model) 及びそれらの AIC、BMD の信頼区間、BMDL 及び BMDU の値)

F) フィッティングを行った平均化モデルのプロット (モデル平均化を行わない場合、重要なエンドポイントについてフィッティングを行った全てのモデルのプロット)
ネスト化したモデル族の場合、各モデル族について選択した 1 モデルのプロット

G) 結論

2.6.3 BMD 法の適用方針

以上を踏まえ、Update: use of the benchmark dose approach in risk assessment (EFSA, 2017) における BMD 法の適用方針のポイントを以下に整理する。

- BMR のデフォルト値として、試験動物を用いた試験では、連続値データについては 5% (平均応答の変化)、非連続データについては 10% (過剰リスク) を用いることを提案している。
- モデルの制限に関して、急峻度パラメータを 1 より大きくするという制限は不適切であり、人為的に高い BMDL につながる可能性があるため適用すべきではないとしている。
- モデルのフィッティングの評価には AIC を用い、フィッティングを行ったモデルの AIC 値がヌルモデル (null model) の AIC 値から 2 を減じた値より小さいものを全て選択する。ネストしたモデルが使用されている場合、モデル族ごとに最小 AIC が得られたモデルを選択した上で、フィッティングを行った全てのモデルの中から最小の AIC 値を決定する。
- エンドポイントが同一であり、性別、種、ばく露期間など、特定の側面は異なるが、それ以外は類似するデータセットは統合が可能であるとしている。
- モデル平均化により最終的な BMD の信頼区間を確立するのが望ましいとしている。モデル平均化機能を有するソフトウェアが利用できない場合は、AIC 値が最小のモデリング結果及び AIC 値が最小値から 2 単位以内の AIC を有するモデリン

グ結果から、それぞれ BMD の信頼区間を算出し、得られた BMDL の最小値及び BMDU の最大値を、当該データセットから算出される BMD の信頼区間として採用することが推奨されている。

2.7 The use of advanced risk assessment methods in answering various types of risk management questions (RIVM, 2006)

2.7.1 概要

本ガイダンスは、オランダ国立公衆衛生環境研究所 (RIVM) が 2006 年に発表した報告書である。本ガイダンスでは、化学物質のリスク管理上の質問に答える形で、従来の古典的な方法を用いたアプローチと、新規のより高度な方法である BMD 法及び確率論的リスク評価を用いたアプローチを整理、比較し、どのような状況下であれば、後者の使用が特に有益となるのか、また、使用する際のコストも含めて解説している。

近年、BMD 法や確率論的なリスク評価等のより高度なリスク評価法に対する関心が高まっている。これらの方法は、関連する不確実性について、より現実的かつ一貫した方法で考慮することを目的としている。本報告書では、5 つのリスク管理上の課題に対して、古典的なリスク評価法及びより高度な方法それぞれで対応する場合の概要、問題点等を総説している。また、リスク管理上の課題ごとに、用量反応モデリング、BMD 法、確率論的ばく露評価 (probEA)、確率論的ハザード評価 (probHC) といったより高度な方法の使用を検討し、また PBPK モデリングについても考察している。さらに、どのような状況下であれば、これら高度な方法の使用が特に有益となるのか、また使用する際に必要なコスト (コンピューター演算に要する時間の観点から) について示している。

一般的には、ヒトでの推定ばく露量が ADI などのばく露限界よりも何桁も低い場合には古典的な方法で十分であると考えられ、推定ばく露量が ADI などのばく露限界より高い場合に、ヒト健康リスクの可能性を推定するためにより高度な方法が必要となる。古典的な方法で得られたばく露限界がヒトの推定ばく露量と何桁も高くない場合においても、高度な方法が推奨される。

2.7.2 BMD の適用方針

BMD は NOAEL (つまり、有害影響の閾値) の代替として導入されてきた。ゼロ影響は測定できないため、小さな大きさの影響をゼロ影響の代用として選択している。非連続値データで表されるエンドポイント (非連続エンドポイント) の場合、この小さな大きさの影響をベンチマークレスポンス (BMR) と呼び、連続値データで表されるエンドポイント (連続値エンドポイント) では臨界影響サイズ (CES) と呼ぶ。用量反応モデルをフィッティングさせると、BMR 又は CES に対応する用量 (非連続値のエンドポイントでは BMD、連続値のエンドポイントでは CED) が推定される。通常、この推定用量の信頼区間の下限値が出発点 (POD) と見なされる。閾値のない影響の場合、BMDL を、現在使用している POD や T25 の代替として使用してもよい。

(21)　BMD 法が適用できるデータ

BMD 法を適用する用量反応データは、少なくとも、異なる影響レベル (反応量) を示す 3 以上の用量が設定されていることが必要である。

(22)　BMR の選択

BMD 法では、適切な BMR の選択が重要となる。非連続エンドポイント (例: 肝病変の発現率) の場合、BMR は通常、過剰リスクの観点から定義される。連続値エンドポイント (例: 肝臓重量) の場合、BMR (又は CES) は通常、対照群の平均値と比較した変化率として定義される。非連続データにおける BMR の選択はリスク管理上の選択を伴うと考えられる一方、連続値エンドポイントにおける CES の選択は、科学的 (毒性学的) な問題である。望ましい (許容可能) と考えられる BMR 又は CES が観察できるとは限らない。例えば、がん又は発生奇形の許容リスクレベルは、標準的な毒性試験で観察できるリスクレベルよりもはるかに低い可能性がある。このような場合、観察可能な状況に基づいて BMR を選択してもよい。観察できるリスクレベルが、望ましいリスクレベルよりもはるかに高い状況下では、論争のさなかにある低用量外挿の問題に直面する。連続値エンドポイントの場合、低用量外挿の問題はあまり顕著ではない。連続値データからなる用量反応データでは、観察可能で、かつ (おそらくは) 有害ではない CES に基づいて重要な臨界影響用量 (critical effect dose) の評価が行われる場合が多い。

(23)　BMD 法を用いた HBEL 導出手順

(3)-1 非連続エンドポイント (影響の発現率)

1. 用量反応データを確認し、BMR＝10%で BMD(L) を算出する (用量反応データから、さらに低い BMR の設定が可能であれば当該 BMR を選択してもよいが、そのようなケースは稀である)。
2. 影響の種類の重要性を基に、ヒトの集団で 10%の過剰リスクが許容できるか否か決定する。許容できない場合は 3 へ、他の場合は 4 へ進む。
3. 低用量外挿係数を POD に適用する。(注: 直線外挿を保守的な方法と見なしてもよい。保守的ではない方法は利用できない)。4 へ進む。
4. (群内の) 分布に関する (デフォルトの) 評価係数を適用する。(注: ラボ内の試験動物と比較してヒトの集団ではばらつきが大きくなる可能性を考慮するためには、種間係数以外に種内係数も必要となる)。

(3)-2 連続値エンドポイント (影響の程度)

1. 用量反応データを確認し、CES (BMR) ＝5%で BMD(L)を算出する (評価対象とする影響の種類において、生物学的根拠に基づきより高い CES の設定が可能である場合はその CES を選択してもよい)。

2. (群内の) 分布に関する (デフォルトの) 評価係数を適用する。

- **BMD 法と NOAEL 法の比較**

科学的な観点から、POD として BMD を使用する方が、NOAEL の使用よりも常に望ましい。NOAEL は、用量反応データが用量反応モデリングに適さない (例: 異なる影響を示す用量群の数が少なすぎる) 場合のみ使用する。用量反応データを利用可能な場合、完全な BMD 解析の方が NOAEL 解析よりもより精緻なものとなる。しかしながら、試験を計画する際に BMD 法を考慮に入れておけば、使用した試験動物数が同一でより多くの情報を得られる (又はより少ない動物数から同一量の情報が得られる) という意味で、高い効率を期待できる。NOAEL 法に基づき既に HBEL (Health-based exposure limits) が算出されている状況で、ヒトでの推計ばく露量が現行の HBEL に近い場合には、BMD 法による再評価の実施を検討することもあり得る。

BMD 法の適用が特に有用な状況は、最も鋭敏な試験又はエンドポイントの用量反応データが「LOAEL のみ」を示す場合である。多くの事例で、そのようなデータにおける BMD(L)の算出が可能である。

これまでの経験上、BMD 法は、平均的に見れば、高い値にも低い値にもならない。一部の事例で、BMD が NOAEL より低くなる場合や高くなる場合があるが、BMD は NOAEL よりも正確な推定値をもたらす。

2.8 Valeurs toxicologiques de reference Guide d’élaboration de l’Anses (英訳: Toxicological Reference Values ANSES Development Guide) (ANSES, 2017)

2.8.1 概要

フランス食品環境労働衛生安全庁 (ANSES) は、2004 年に毒性学的基準値 (TRVs) の設定と検証のための国家プロジェクトを開始した。本ガイダンスは ANSES による TRVs 設定の作業を理解し、分かりやすく再現性のある方法で TRVs を提案するために、情報の共通の基盤を提供することを目的として作成された。本ガイダンスは 2015 年 9 月に作成された旧版の改訂版であり、既に公開されている生殖毒性物質と発がん性物質の TRVs 設定に関する 2 つのガイダンスに新たな科学的知見を追加し、TRVs 設定に必要な情報を統合したのもである。本ガイダンスにより、政策方針を決定する立場の人々や公衆衛生分野の従事者に専門家委員会 (CES) による科学的選択を説明することが可能になる。また、本ガイダンスは、データが不足しているが安全性に疑念が持たれる化学物質に関して、Indicative Toxicological Value (TIV) と呼ばれる毒性学的基準値を設定することを提案している。

2.8.2 BMD 法の適用方針

リスク評価に用いられる毒性学的データのほとんどは動物を用いた試験に由来している。動物試験は、NOAEL(C)、LOAEL(C)、又は BMD(C) 値が得られるようにデザインしなければならない。

BMD/濃度 (BMD / C) 法は、現在用いられている手法であり、観察から得られたデータをモデリングした後に、特定の反応レベルに対応した用量レベル (BMR) を決定する。専門家委員会 (CES) は、得られるデータを用いて可能な場合には、常に BMD 法を適用することを選択している。

BMD とは、対照群に対して一定のレベルの反応に対応した測定可能な効果を生じる用量である。通常、95%又は 90%信頼区間の下限値 (BMDL95% 又は BMDL90%) が用いられる。この手法は、用量反応曲線全体を検討した試験データのモデリングに基づいている。投与群の反応の解析に関しては、影響 (組織の変化、病理学的所見など) がみられた動物の数 (影響がみられた個体の比率) を問題とする場合には、二値又は非連続反応と見なされる。また、生理学的あるいは生物学的パラメータ (例えば体重増加、赤血球数、肝酵素の血中濃度など) を問題とする場合には連続反応と見なされる。

近年、EPA による BMDS、RIVM による PROAST などの BMD 値導出に特化したソフトウェアが開発されている。それらの使用に関しては、用量反応関係の理論的モデル、調整法、及び反応レベルに関する選択が必要である。

専門家委員会 (CES) では、BMR として、二値の用量反応データに関して動物試験では

10%、ヒトのデータでは検出感度や影響の重症度及び対照群での発現頻度の基底値に応じて 1～10%を適用している。連続値からなる反応量の場合には、BMR として 5%を適用している。モデルの選択に関しては、試験データに最もよくフィットする BMD 又は BMDL 値を与えるモデルを優先的に選択する。

2.9 Updating of the PNNS guidelines: revision of the food-based dietary guidelines (ANSES, 2016)

2.9.1 概要

本ガイダンスは 2012 年 4 月の Director General for Health (DGS) からの国家栄養健康計画 (National Health and Nutrition Programme: PNNS) の食生活指針 (food-based dietary guidelines) 改訂の要請に対応して、ANSES が作成した意見書である。新たなガイドラインの枠組み作成のためにワーキンググループと専門家委員会が調査と評価の作業を行い、2016 年に意見書が最終的に確定された。本ガイダンスでは、本調査の対象である BMD 法の手順に関する記載はないが、BMD 法に関する記載として、「3.2.3　汚染物質のばく露の制限」の中で、リスク評価に用いる参照用量として ADI、TDI 等と並んで BMDL が挙げられ、「3.2.3.3　環境汚染物質」の中で参照用量として BMDL が選択された場合のリスク判定の考え方が示されている。以下にその内容を示す。

2.9.2 汚染物質のリスク評価における BMDL の利用

食生活指針 (food-based dietary guidelines) の改訂作業では、食品中の汚染物質についてリスク評価が行われた。この中で、閾値のない汚染物質 (遺伝毒性物質の場合) や BMDL が毒性学的参照用量として選択されている物質のデフォルトのばく露量は、最大値として、第 2 回トータルダイエットスタディ (TDS2) で推定された集団のばく露量の中央値が選択された。これらの汚染物質については、ばく露量の閾値を選択できないため、最大ばく露量が現在の集団のばく露量を超えることを防ぐことが決定された。この場合、リスク判定では、遺伝毒性発がん物質の MOE 又は一定の閾値から影響が現れる非遺伝毒性発がん物質の安全マージン (MOS) の算出が行われた。これらの MOE 又は MOS は、臨界ばく露量 (例えば BMDL) と最大ばく露量の比である。集団に対するリスクについて結論付けるために、国内機関又は国際機関によって BMDL が確立されている場合には、定義された臨界マージンとこれらの MOE 又は MOS の比較によりリスク判定が行われた。

2.10 Risk Analysis in Food Regulation (FSANZ, 2013)

2.10.1 概要

本ガイダンスはオーストラリア・ニュージーランド食品基準局 (FSANZ) が行っている食品におけるリスク解析を解説したものである。本調査の対象である BMD 法の手順に関する記載はないが、化学物質のハザード評価及びリスク評価における BMD 法の利用に関する記載がある。以下にこれらの内容を示す。

2.10.2 化学物質のハザード評価及びリスク評価における BMD 法の利用

➢ **化学物質のハザード評価における BMD 法の利用**

多くの化学物質のばく露では、それ以下の量では健康への有害影響がみられない閾値が存在する。ハザード評価では、ある一定の期間 (例えば生涯、又は 24 時間) 摂取しても健康リスクが見られない量の設定が重要であり、この用量を健康影響に基づく指標値 (HBGV) と呼ぶ。HBGV は多くの化学物質の場合、動物試験で有害影響がみられない最大用量として得られる NOAEL と安全 (又は「不確実」) 係数を用いて算出される。しかしながら、遺伝毒性発がん物質のように、閾値が同定されない化学物質も存在する。そのような化学物質については、NOAEL に基づく方法の代替法として BMD 法が用いられる。BMD とは反応レベルは低いが測定可能な有害反応、すなわち、ある特定の有害影響に関して、あらかじめ決定された増加 (通常 5 又は 10%) を生じる用量である。BMD は、信頼区間の下限値である BMDL としても表記される。

➢ **化学物質のリスク評価における BMD 法の利用**

リスク評価においては、閾値が明らかでない場合、相対的なリスクの評価法として MOE 法が用いられる。MOE 法では、ある化学物質の BMD (又は BMD が求められない場合には LOAEL) とその物質の食物からのばく露量を比較する。MOE が高い場合 (例えば 10,000 超) はリスクが低いことを示すと考えられるが、MOE はリスクの定量的評価ではなく、導出方法と MOE を用いる評価の限界について説明を併記する必要がある。

2.11　BMD 法に関するガイダンスにおける BMD 適用方針の比較

	WHO (EHC)	WHO (JECFA)	EPA	EFSA	RIVM	ANSES
BMD 法を適用できる試験・研究	・一般的に、少なくとも 3 又は 4 の異なる用量群 (対照群を含む) を有するデータが望ましい。 ・各用量群で影響レベルが異なり、少なくとも 3 つの異なる反応レベルを有するデータが望ましい。 ・有意な用量依存傾向を有する段階的な単調応答を示すデータが適切である。	発がん性データのデータの選択では、感受性が最も高い部位での腫瘍の部位特異的発生率と、明確な臓器特異性が示されない化合物の組合せデータ (例えば、腫瘍を有する動物数) の両方を検討する必要がある。	BMD モデリングのための最小データセット基準 ・選択したエンドポイントにおいて、少なくとも統計学的又は生物学的に有意な用量反応関係がある。 ・対照群における反応レベルと観察された最大反応レベルの間で用量反応関係がある (BMR に近いデータポイントを有することがより望ましい)。 モデリングが可能なエンドポイントについては全てモデリングを行う。	・用量反応関係のないエンドポイントは BMD 算出に適用できない。 ・エンドポイントの決定は、目視による確認で行える場合があり、正式には BMD 解析を試行し統計学的に有意でない傾向を確認することで行う。 ・理想的には全てのデータセットのモデリングを行う。	少なくとも、異なる影響レベルを示す 3 以上の用量を備えた用量反応データ。	言及なし。
試験・研究データの統合	具体的な言及はないが、カテゴリカル回帰が複数試験の統合解析に適用されることがあるとの記載がある。	言及なし。	統計学的及び生物学的に互換性のあるデータセットについて用量反応のモデリング前に統合できる。	言及なし。	言及なし。	言及なし。
BMR の設定	具体的な値について言及されていない。	発がん反応の BMR としては 10%が適切である	明確な指針はないとしつつ、以下の一般的なアプローチとして以下が記載されている。 非連続データ： ・比較を行う目的から、標準的な BMR として 10%過剰リスクに対応する BMD 及び BMDL を報告する。 ・生物学的検討の結果に応じて、リスク評価のための参照値の POD の根拠として、重篤な影響等については 5%又はより低い BMR、あるいは前駆効果のような影響の場合 10%を上回る BMR を用いるのが妥当である可能性がある。 ・統計学的な観点から、ネストした生殖発生毒性試験では BMR として 5%が推奨され、疫学研究のデータでは BMR として 1%が用いられる場合が多い。 連続値データ ・生物学的に有意であると考えられる最小の変化レベルが分かる場合、当該変化レベルを BMR とする。 ・個体のデータが得られており、有害影響と考えられる個体レベルを決定できる場合、当該データをハイブリッド法又はカットオフ値を用いて二値化し、二値データの場合と同様に BMR を設定することができる (二値	デフォルト BMR として、実験動物を用いた試験では以下を提案。 ・連続値データ: BMR5% (平均応答の変化) ・非連続データ: BMR10% (過剰リスク)	・非連続エンドポイント: BMR＝10%で BMD(L) を導出する (用量反応データからさらに低い BMR の設定が可能である場合はその BMR を選択してもよい) ・連続値エンドポイント: CES＝5%で BMD(L) を導出する (生物学的根拠に基づきより高い CES の設定が可能である場合はその CES を選択してもよい)	二値の用量反応データ: 動物実験では 10%、ヒトのデータでは検出感度や影響の重症度及び対照群での発生頻度の基底値に応じて 1～10%の BMR を適用。 ・連続値からなる反応量: 5%の BMR を適用。

	WHO (EHC)	WHO (JECFA)	EPA	EFSA	RIVM	ANSES
			化よりも、ハイブリッド法による二値化が望ましい)。 ・比較を行う目的から、対象群の平均の標準偏差 (SD) 1 つ分の変化に等しい平均反応レベルに対応する BMD 及び BMDL を報告する。			
モデルの適合度の評価基準	適合度検定の p > 0.1 を用いるとともに、目視によるプロットの確認を行う。	飽和モデルに比べフィッティングが有意に悪いとされないモデルは許容される。	p 値 ＞0.1 p 値がそこまで小さくないがデータとモデルの乖離が大きくない場合や、低用量での反応をうまく記述していない場合があるため、残差及び目視によるプロットの確認を行う。	AIC。 フィッティングを行なったモデルの AIC がヌルモデルの AIC - 2 単位より低いかどうかで評価。	言及なし。	言及なし。
モデルの比較に用いる基準	比較を行うモデルの種類や分布の仮定に応じて、尤度比や AIC を用いる。	言及なし。	適合度の基準を満たすモデルについて BMDL を算出し、BMDL が十分に近い場合は AIC が最も低いモデルを選択して BMDL を採用する。BMDL が十分に近いといえない場合は、モデルの不確実性について検討した上で、保守的アプローチとして最小 BMDL を選択してもよい。	AIC 値が最小のモデリング結果及び AIC 値が最小値の 2 単位以内の AIC を有する全てのモデリング結果を BMD の信頼区間の設定に用いる。	言及なし。	実験データに最もよくフィットする BMD 又は BMDL 値を与えるモデルを優先的に選択。
モデル平均化に関する記載	・具体的な記載はない (モデルのフィッティングとパラメータ推定の中でベイズ推定について触れられているが、ガイダンスの対象外としている)。	言及なし。	複数のモデルからの推定値の統合を可能とするモデル平均化法が検討されているとの記載があるが、ガイダンスに記載できるレベルにない事項とされている。	モデル平均化により最終的な BMD 信頼区間を確立するのが望ましい。フィッティングを行なったモデルの結果を統合して最終的な BMD の信頼区間を設定する。	言及なし。	言及なし。
低用量領域への外挿について	・単一モデルからの外挿だけでは十分でない。 ・POD からの直線外挿については、MOE 法を越える利点はない。	直線外挿を用いて 10% の発生率で BMDL から 100 万分の 1 の発がんを引き起こす摂取量を計算することは、単純に BMDL を 100,000 で除することと等しいことから、MOE の計算以上の情報にはならない	低用量での反応を適切に記述していない場合があり、残差及び目視によるプロットの確認を行う。	BMR が最低用量で観察された反応よりもかなり小さく、観察範囲を大きく外れた外挿が必要となるような場合、より大きな BMR を選択することができるとの記載あり。	連続値エンドポイントの場合、低用量外挿の問題はあまり顕著ではない。	言及なし。
リスク評価への BMD の利用	LOAEL が NOAEL よりはるかに高い場合は、BMD 法を適用する方がよい。	BMD 法は MOE 法を超える利点はなく、MOE 方を推奨する。	BMD は、化学物質のリスク評価における用量反応情報を得る目的の他、化学物質間の相対効力や部分集団間の相対感度の比較、生態リスク評価等の分野にも適用可能。	リスク評価における基準点 (RP) の導出方法として BMD 法が NOAEL 法よりも科学的により進歩した方法であると結論。	科学的な観点から、BMD を POD として使用する方が、NOAEL を使用するよりも常に良好。	得られるデータを用いて可能な場合には、常に BMD 法を適用。

3. BMD 法に関するソフトウェア等に関連する情報等の収集整理

本調査の調査対象ソフトウェアを表 3-0 に示す。各ソフトウェアの概要を整理し、利用方法についてまとめた。

表 3-0 調査対象ソフトウェア

ソフトウェア名	開発者	バージョン等
BMDS[38]	米国環境保護庁 (EPA)	BMDS Version 3.1[1)]BMDS Version 2.7
PROAST[39]	オランダ国立公衆衛生環境研究所 (RIVM)	PROAST 65.5[1)]
Benchmark Dose Modelling[40]	欧州食品安全機関 (EFSA)、RIVM	PROAST (version 66.24)[2)]
SPBS[41]	秋田大学 村田勝敬教授ほか	SPBS V9.67[1)]
MADr‐BMD[42]	Matthew W. Wheeler, A. John Bailer	MADr-BMD ver.1.0[1)]
BMD Express[43]	米国環境健康科学研究所 (NIEHS) 米国国家毒性プログラム (NTP)、カナダ保健省 (Health Canada)、EPA、Sciome LLC	BMD Express 2.2[1)]

1) 2019 年 2 月末時点

2) 2019 年 3 月 12 日時点

38 https://www.epa.gov/bmds

39 https://www.rivm.nl/en/Documents_and_publications/Scientific/Models/PROAST

40 https://shiny-efsa.openanalytics.eu/app/bmd

41 http://www.med.akita-u.ac.jp/~eisei/link.html

42 https://www.jstatsoft.org/article/view/v026i05

43 https://www.sciome.com/bmdexpress/

3.1　BMDS (EPA)

3.1.1　名称及びバージョン

開発元: EPA

名称及びバージョン情報: Benchmark Dose Software (BMDS) Version 3.1

下記 EPA ウェブサイト[44]より無償でダウンロードできる。

3.1.2　BMDS の概要

(1)　概要

BMDS は、EPA が行う汚染物質のリスク評価においてベンチマークドーズ (BMD) 法の活用を促進するためのツールとして開発されたソフトウェアである。2000 年に公開版 Version 1.2 が公表されて以降、更新が重ねられており、Version 3 としては 2018 年 9 月に Version 3.0 がリリースされた後、2019 年 2 月末、現行の最新版である Version 3.1 がリリースされた。Version 3.1 には約 20 のモデルが搭載され、二値データ、連続値データ、ネストしたデータ (発生毒性試験データ)、multiple tumor 解析に利用できる。Version 3 では BMDS として初めて、ベイズ推定に基づくパラメータ推定を行うモデル (ベイズ推定モデル) とモデル平均化機能が搭載された。Version 2.7 から Version 3.0 への更新における主な変更点と Version 3.0 から Version 3.1 への更新における主な変更点を以下に示す。

【Version 2.7 から Version 3.0 への主な変更点】

- アプリケーションの変更 (ソフトウェアからマクロ機能付き Microsoft Excel®への変更)
- 操作性の改善
  - ・1 つのワークシート上で解析条件の設定が可能となった。
  - ・複数のデータセット及び解析条件を用いた迅速な解析が可能となった。
- 既存モデルの更新
  - ・全ての連続値モデルに、ハイブリッド法による解析機能が追加された。
  - ・全ての連続値モデルに、連続値データの同一用量に対する反応量分布の仮定条件として、対数正規分布が追加された。
- 一部のモデルの削除 (詳細は (2)-2 に記載)
- ベイズ推定モデルの追加
  - ・二値データに対応するモデル (二値モデル) について、従来の、最尤推定 (頻度論) に基づくパラメータ推定を行うモデル (頻度論モデル) のほか、ベイズ推定に基づくパラメータ推定を行うベイズ推定モデルが追加された。

[44] https://www.epa.gov/bmds/benchmark-dose-software-bmds-version-31-download

- モデル平均化機能の追加
  ・ベイズ推定に基づくモデル平均化機能が追加された。
- モデリング結果の自動選択機能の追加
  ・EPA が推奨する、又はユーザーが設定したロジックに基づき、BMDL の採用が推奨される、モデリング結果を自動選択する機能が追加された。

【Version 3.0 から Version 3.1 への主な変更点】
- 2007 年以降リリースされた Microsoft Excel®に対応し互換性を拡張
- 操作性の改善
- プログラム上の不具合点の改善

上記のとおり、主要な機能の面では、Version 2.7 から Version 3.0 への更新に伴い、ベイズ推定モデル及びモデル平均化機能の追加を含む大幅な改訂がなされている。一方で、Version 3.0 と Version 3.1 間では大きな変更はない。以降の記載は、最新版の Version 3.1 に基づく。なお、モデルの名称は BMDS 3.1 と表記する。

(2)　主要な機能

(2)-1 利用可能なデータ

以下の 4 タイプのデータが解析可能。
- Continuous - summarized (連続値データ (要約))
- Continuous - individual (連続値データ (個体))
- Dichotomous (二値データ)
- Dichotomous - nested (ネストした二値データ)

(2)-2 収載する数理モデル

BMDS 3.1 に収載された数理モデルを表 3.1-1 に示す。
数理モデルに関して、Version 2.7 からの更新内容を以下に示す。

・二値モデルについて、従来の頻度論モデルに加え、ベイズ推定モデルが追加された。
・二値モデルについて、ベイズ推定に基づくモデル平均化機能が追加された。
・ネストした二値モデルのうち、Version 2.7 に搭載されていた以下のモデルが削除された。
- Rai and Van Ryzin model
- ToxicoDiffusion model
- Ten Berge model

なお、NCTR (National Center for Toxicological Research) nested dichotomous モデルは、

2019 年 2 月末現在では BMDS 3.1 には収載されていない (将来的に追加される予定)。

表3.1-1　BMDS 3.1に収載された数理モデル

データタイプ	数理モデル	略称	数式	パラメータ
Continuous (連続値データ)	Exponential Model	exp	$M2, \mu(dose) = a \times e^{-b \times dose}$ $M3, \mu(dose) = a \times e^{-(b \times dose)^d}$ $M4, \mu(dose) = a \times (c - (c-1) \times e^{-b \times dose}$ $M5, \mu(dose) = a \times (c - (c-1) \times e^{-(b \times dose)^d}$	a = background response (バックグラウンド時の反応量) (> 0) b = slope (傾き) (> 0) c = asymptote term (漸近項) increasing data (漸増傾向を示すデータの場合): c > 1 decreasing data (漸減傾向を示すデータの場合): 0 < c < 1 d = power (べき乗) (> 1)
	Hill Model	hil	$\mu(dose) = \gamma = \frac{v \times d^n}{K^n + d^n}$	γ = control (コントロール) (intercept (切片)) k = 最大半量変化 (half-maximal change) を生じる用量 n = power (べき乗) v = maximum change (最大変化量)
	Linear Model	lin	$\mu(dose) = \beta_0 + \beta_1 dose + \beta_2 dose^2 + \cdots + \beta_n dose^n$ Linear model は n = 1 とした polynomial model	$\beta_0 \ldots \beta_n$: 多項式係数 n: 多項式の次数 (ユーザーが正の整数を指定 (最大値＝21))
	Polynomial Model	ply		
	Power Model	pow	$\mu(dose) = \gamma + \beta \times (dose)^\delta$	γ = control (コントロール) (intercept (切片)) β = slope (傾き) δ = power (べき乗)

データタイプ	数理モデル	略称	数式	パラメータ
Dichotomous (二値データ)	Gamma Model	gam	$p(d) = \gamma + \frac{1-\gamma}{\Gamma(\alpha)} \int_0^{\beta d} t^{\alpha-1} \exp(-t)\, dt$	γ = background (バックグラウンド) (> 0, < 1) α = power (べき乗) (値の制限 (restriction)を行う場合: < 18, ≥ 1) β = slope (傾き) (restrictionを行う場合: ≥ 1)
	Logistic Model	log	$p(d) = \frac{1}{1+\exp[-\alpha-\beta d]}$	α = intercept (切片) β = slope (傾き) (restrictionを行う場合: >1)
	Log-Logistic Model	lnl	$p(d) = \gamma + \frac{1-\gamma}{1+\exp[-\alpha-\beta\log(d)]}$	γ = background (バックグラウンド) (> 0, < 1) α = power (べき乗) (restrictionを行う場合: < 18, ≥ 1) β = slope (傾き) (restrictionを行う場合: > 1)
	Probit Model	pro	$p(d) = \Phi(\alpha + \beta d)$, where $\phi(x) = \int_{-\infty}^{x} \phi(t)dt$ and $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$	α = intercept (切片) β = slope (傾き) (restrictionを行う場合: ≤ 18, ≥ 1)
	Log Probit Model	lnp	$p(d) = \gamma + (1-\gamma)\Phi[\alpha + \beta\log(d)]$	γ = background (バックグラウンド) (> 0, < 1) α = intercept (切片) β = slope (傾き) (restrictionを行う場合: ≤ 18, ≥ 1)
	Multistage Model	mst	$p(d) = \gamma + (1-\gamma)\left(1 - \exp\left[-\sum_{j=1}^{n} \beta_j dose^j\right]\right)$	γ = background (バックグラウンド) (> 0, < 1) β = dose coefficients (用量係数)
	Weibull Model	wei	$p(d) = \gamma + (1-\gamma)(1 - \exp[-\beta d^{\alpha}])$	γ = background (バックグラウンド) (> 0, < 1) α = power (べき乗) (restrictionを行う場合: < 18, ≥ 1) β = slope (傾き) (restrictionを行う場合: ≥ 1)

データタイプ	数理モデル	略称	数式	パラメータ
	Quantal Linear Model	qln	$\mu(dose) = \beta_0 + \beta_1 dose + \beta_2 dose^2 + \cdots + \beta_n dose^n$ Linear model は n = 1 とした polynomial model	$\beta_0 \ldots \beta_n$: 多項式係数 n: 多項式の次数 (ユーザーが正の整数を指定 (最大値＝21))
	Dichotomous Hill Model	dhl	$p(\mathrm{dose}) = \frac{g + (v - vg)}{1 + \exp(-\alpha - \beta \log(\mathrm{dose}))}$	g = background (バックグラウンド) (≥ 0, < 1) v = モデルにより予測される反応の最大確率 (0 < v ≤ 1) α = intercept (切片) β = slope (傾き) (restrictionを行う場合: ≤ 18, ≥ 1)
Dichotomous-Nested (ネストしたニ値データ)	NCTR (National Center for Toxicological Research) Model*	nctr	$p(d) = 1 - e^{\{-(\alpha+\theta_1(r_{ij}-r_m)-(\beta=\theta_2(r_{ij}-r_m)\times dose^{\rho}\}}$	α = intercept (切片) (≥ 0) ρ = power (べき乗) (≥ 0, restrictionを行う場合: ≥ 1) β = slope (傾き) (≥ 0) θ_1 = first coefficient for the litter specific covariate (腹固有の共変量に関する第1の係数) θ_2 = second coefficient** for the litter specific covariate (腹固有の共変量に関する第2の係数) $\phi_1, \ldots, \phi_g$ = inter-litter correlation coefficients (腹間の相関係数)
	Nested Logistic Model	nln	$p(d) = \alpha + \frac{\theta_1 r_{ij} + (1 - \alpha - \theta_1 r_{ij})}{(1 + e^{[-\beta - \theta_2 r_{ij} - \rho \times \ln(dose)]})}$ if $dose > 0$, and $\alpha + \theta_1 r_{ij}$ if $dose = 0$	α = intercept (切片) (≥ 0) ρ = power (べき乗) (≥ 0, restrictionを行う場合: ≥ 1) β= slope (傾き) (≥ 0)

データタイプ	数理モデル	略称	数式	パラメータ
				θ_1 = first coefficient for the litter specific covariate (腹固有の共変量に関する第1の係数) θ_2 = second coefficient** for the litter specific covariate (腹固有の共変量に関する第2の係数) $\phi_1,\ldots,\phi_g$ = inter-litter correlation coefficients (腹間の相関係数)
Dichotomous-Multi-tumor	Multi-Tumor (MS_Combo) Model	multi	$P(d) = 1 - e^{\{-(\beta_0+\beta_1 d+\beta_2 d^2+\cdots)\}}$	$\beta_0 = \sum \beta_{0i}$, $\beta_1 = \sum \beta_{1i}$, $\beta_2 = \sum \beta_{2i} \ldots.$ i = 1,…., t t: 検討対象とする腫瘍の数 βxj: 腫瘍 j のx 番目のパラメータ パラメータβのrestriction: 検出された又は特定のadverse directionに応じて、自動的に負の値をとらない (Non-negative) 、又は正の値をとらない (Non-positive)

*現時点ではBMDS 3.1に収載されていない。将来的に追加予定とされている。

**原文はfirst coefficientであるが、second coefficientの誤りと考えられる。

(2)-3　数理モデルの選択

各データタイプに応じて選択可能な数理モデルから、解析に用いる数理モデルを選択する。二値モデルについては、従来の頻度論モデルに加え、ベイズ推定モデルが追加された。

また、頻度論モデルにて最尤推定によるパラメータ推定を行う際は、パラメータの値の制限 (Restriction) を行うか否か (Restricted/Unrestricted) の設定が可能である (一部のモデル、パラメータについては固定) (詳細は 3.1.3. (4) 2) 表 3.1.-4 に記載)。

(2)-4　解析条件 (Modeling Options) の設定

解析に用いるモデルタイプに応じた解析条件 (BMR の種類 (BMR Type) と値 (BMRF)、信頼水準 (Confidence Level)、同一用量に対する反応量分布の仮定 (正規分布／対数正規分布) 等) を設定する (詳細は 3.1.3. (4) 3) 表 3.1.-5 に記載)。

BMDS 3.1 では、連続値データに対応するモデル (連続値モデル) の解析において、ハイブリッド法が設定できるようになった。また、同一用量に対する反応量分布の仮定に関して、従来は指数分布モデルのみ設定可能であった対数正規分布の仮定を、全ての連続値データで設定可能できるようになった。

また、旧バージョンと同様に、1 つの解析の中で複数モデルの同時解析が可能であり、さらにユーザーが選択したモデルに対して、ユーザーが独自に設定した”解析条件のセット”を適用することができるようになった。

(2)-5　モデリング結果の評価と推奨モデルの選定

Benchmark Dose Technical Guidance (EPA (2012)) で示された推奨決定ロジック (recommendation logic) 又はユーザーが事前に設定したユーザーロジック (user-defined decision logic) に従い、各モデリング結果を評価するとともに、推奨モデル (設定された評価ロジックに従い選択され、同モデルから導かれる BMDL の採用が推奨されるモデル) を自動的に選定する (必ずしも自動選択された推奨モデルを選定する必要はなく、最終的にはユーザーがモデルを選択する)。

BMDS 3.1におけるモデリング結果の評価は、以下の3分類である。

- Viable: 質の高いモデリング結果が得られたモデル。事前に設定したdecision logicに基づけば重大な欠陥はないが、注意を要する可能性はある。
- Questionable: 事前に設定したdecision logicに基づく重大な欠陥がある。
- Unusable: 使用不可。BMDやBMDL等の結果は推計されない。

"Logic"ワークシート (図 3.1-1) に、評価ロジックを構成する各評価項目について、その評価方法と基準値 (一部項目のみ) が示されており、同シート上で各評価項目の On/Off、基準値の変更が可能である。

主な評価項目を以下に示す。

1) 適合度

1)-1 評価項目の概要

連続値データ及び二値データについて、適合度検定のp 値及び残差に基づき、モデルのフィッティングの良さを評価する。評価基準を満たさない場合は重大な欠陥があると判断される。

2)-2 評価基準 (基準値はユーザーにより変更可能)

i) 適合度検定のp 値について、予め定義された臨界のp 値より大きい場合は、モデルはデータを適切に表しているといえる。EPAでは臨界値として通常は0.1を用いており、Multistage modelを発がん性データに適用する場合は0.05が用いられる場合がある。

ii) BMDに近い用量群におけるscaled residual value[45]の絶対値 ＜2

2) AIC[46]

2)-1 評価項目の概要

全てのデータタイプにおいてAICが算出されるか否かを評価する。また、使用可能なモデルの中から推奨モデルを選定する際にAIC値を利用する。

2)-2 評価基準

AICが算出されないモデルは使用不可と判断される。使用可能なモデルの中から推奨モデルを選定する際に、使用可能なモデルが3つ未満の場合、AIC値が最小となるモデルが選定される。

3) 分散性

3)-1 評価項目の概要

連続値データについて等分散性、不等分散性の検定を行い、p 値に基づき、推定対象の分散を適切にモデリングしているかを評価する。

3)-2 評価基準

等分散性の検定、不等分散性の検定で算出されるp 値について、デフォルトでは以下の評価基準が適用される (基準値はユーザーにより変更可能)。評価基準を満たさない場合は重大な欠陥がある (等分散/不等分散の仮定が適切でない) と判断される。

---

45 (観測値 － モデルによる推定値)／標準偏差で定義される。

46 赤池情報量規準。AIC = -2 log(L) + 2p (ここで log(L)はモデルの最大対数尤度、p はモデルに含まれる推定パラメータ数) で表され、類似のフィッティング手法をとる異なるモデルの比較に用いる。BMDS 3.1 では尤度が異なるモデルの比較を行なうために、正規化係数を用いて正規分布と対数正規分布の対数尤度を算出する。このため、Version 2.7 以前のバージョンの BMDS で算出された AIC と数値が異なる点に注意が必要である。

- 等分散性の検定のp 値 ＞0.05
- 不等分散性の検定p 値 ＞0.05

4) BMD/BMDL比

全てのデータタイプにおいてBMD/BMDL比を評価する。デフォルトでは、BMD/BMDL比20を超える場合は重大な欠陥あり、5を超える場合は注意が必要と判断される (基準値はユーザーにより変更可能)。

5) BMD、BMDLと用量範囲の比較

全てのデータタイプにおいて、BMD、BMDLが最小用量の10分の1を下回る場合は、重大な欠陥があると判断される。また、BMD、BMDLが最高用量を超える場合や最小用量の3分の1を下回る場合は、注意が必要と判断される (基準値はユーザーにより変更可能)。

6) 自由度 (degree of freedom)

全てのデータタイプにおいて、自由度 (degree of freedom) がゼロとなる場合、飽和モデル (saturated model) であり、重大な欠陥があると判断される。

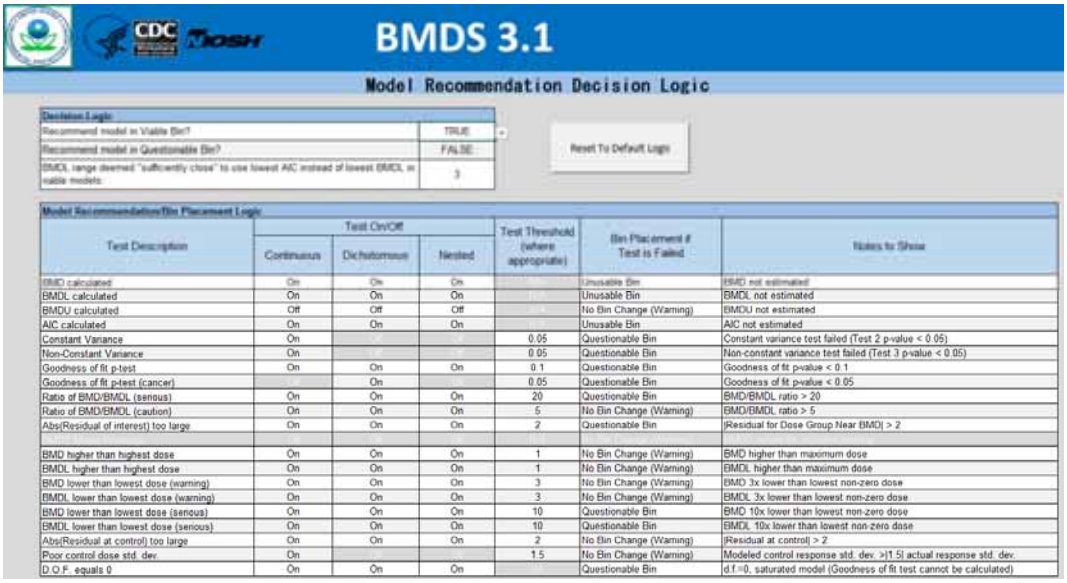


図3.1-1 Logicワークシート

また、評価ロジック (図3.1-2 にデフォルトで設定されている評価ロジックのフローチャートを示す) に基づき推奨モデルが選定される。

推奨モデルの選定基準

- ✓ Viableに分類されたモデルが3未満の場合、AICが最も小さいモデルのBMDLが推奨される。
- ✓ Viableに分類されたモデルが3以上の場合、それらのモデルから推定されるBMDLのうち最小のBMDLが推奨される。

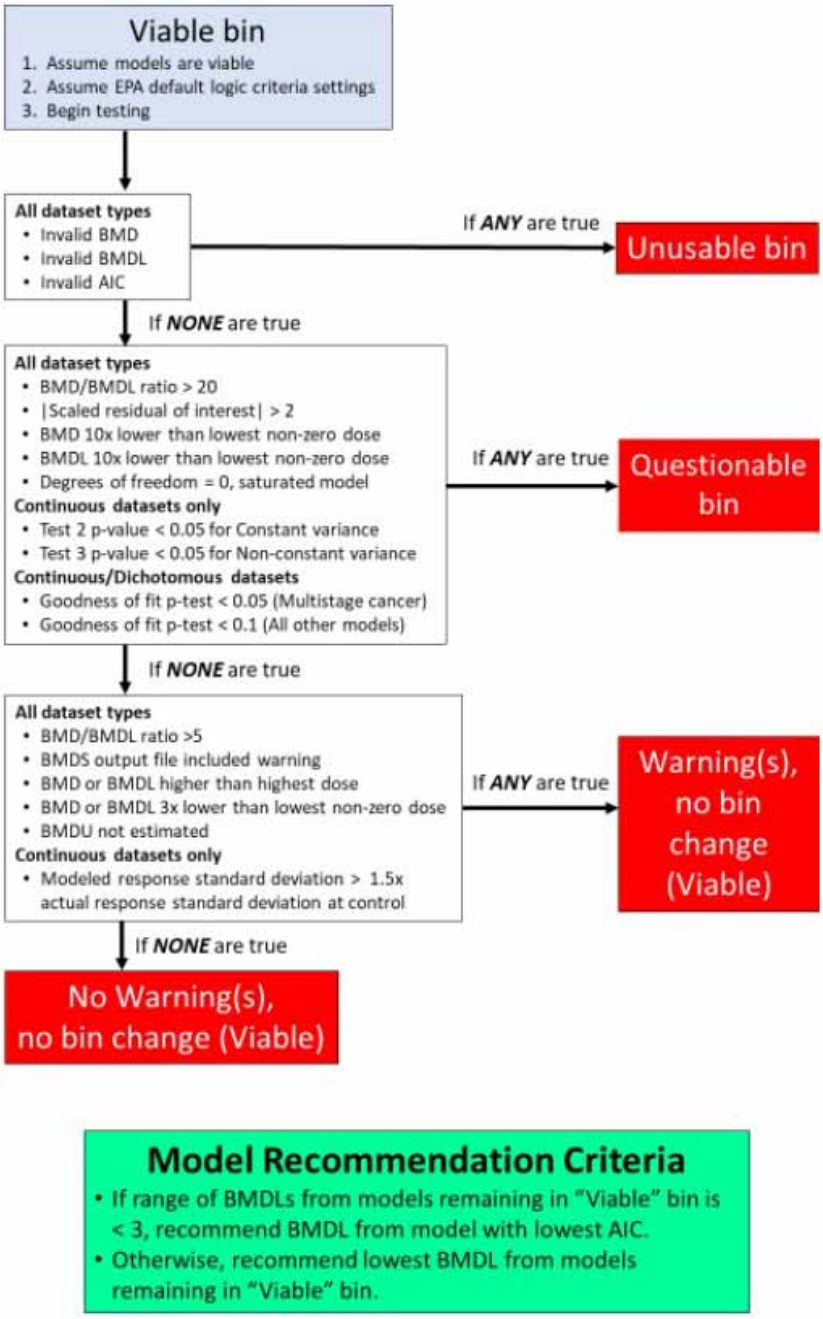


図3.1-2　デフォルトで設定されている評価ロジック (recommendation logic) を用いた推奨モデルの決定フロー

(BMDS Version 3.1ユーザーマニュアルより引用)

(2)-6　モデル平均化機能

BMDS 3.1 では、二値モデルを用いる解析において、ベイズ推定に基づくモデル平均化機能が追加された。同機能の実行を設定した場合、モデル平均化による BMD、BMDL、BMDU の算出が行なわれる。モデル平均化で用いられる数理モデルは、頻度論モデルと数式自体は同じだが、パラメータ等の事前情報 (prior information) を組み込みモデルフィッティング時にこの情報が利用される。ベイズ推定モデルの一覧を表 3.1-2 に示す。

【ベイズ推定に基づくモデル平均化の概要】

ベイズ統計学の考え方に基づき、モデル M とそのパラメータによって与えられるデータの生成メカニズムを定義することで推定が行われる。モデル M の事前確率を f(M)、モデルのパラメータの確率密度を f(θ)とすると、事前確率モデル π(M)は以下で表される。

$$\pi(M) = f(M)f(\theta|M)$$

単一のモデルでの推定の場合 f(M)=1 であり、データセットDが与えられたモデルMの推定はベイズの定理を用いて次式により定義される。

$$g(M|D) \propto \ell(D|M)\pi(M)$$

モデル M の事前確率を f(M)、f(θ) をモデルのパラメータの確率密度とすると、事前確率モデル π(M)は以下で表される。ここで、$0 \leq f(M) \leq 1$ であり、$\sum_{k=1}^{K} f_k(M_k) = 1$である。

データセットD及びモデルM_kについて、最大事後推定 (MAP) を用いて各モデルM_kのパラメータベクトル θ_k を推定し、BMDの事後確率密度、すなわち $g_k(BMD|M_k, D)$ を計算するとき、モデル平均化されたBMDの事後確率密度は以下で表される。

$$g_{ma}(BMD|\ D) = \sum_{k=1}^{9} \pi_k(M_k|\ D) g_k(BMD|\ M_k, D) \qquad (1)$$

ここで、π_kはデータセット D に対するモデルM_kの事後確率密度である。BMD と BMDL は、この事後確率密度π_kから計算される。

より具体的には、モデル平均化 BMD 推定値は、各モデルの BMD 推定値をそれぞれのモデルの事後確率密度π_kで重み付けして加重平均した値であり、BMDL 推定値は、適切に選択された確率レベルαに対する$100(1-\alpha)$パーセンタイルとみなされる。各モデルの重みとなる事後確率密度π_kは、ラプラス近似を用いて近似する。

表3.1-2　ベイズ推定モデル一覧

モデル名	数式	Restriction	事前確率[1)]	注記
Quantal linear	$p_1(\mathrm{dose}) = g + (1-g)(1-\exp[-\beta \mathrm{dose}])$	$\beta > 0$ $0 \le g \le 1$	$\log(\beta) \sim \mathrm{Normal}(0, 1)$ $\mathrm{logit}(\gamma) \sim \mathrm{Normal}(0, 2)$	Quantal linear modelとMultistage model では、それぞれの曲線の主要なパラメータを強調する目的で設定が異なる。
Multistage	$p_2(\mathrm{dose}) = g + (1-g)(1-\exp[-\sum_i \beta_i \mathrm{dose}^i])$	$\beta i > 0$ $0 \le g \le 1$	$\log(\beta_1) \sim \mathrm{Normal}(0, 0.25)$ $\log(\beta i) \sim \mathrm{Normal}(0, 1)\ i \ge 2$ $\mathrm{logit}(g) \sim \mathrm{Normal}(0, 2)$	βi パラメータの事前確率は、発がん性のtwo-hit model の二次項は正となるべきことを表す。 Multistage modelとQuantal linear modelでは、それぞれの曲線の主要なパラメータを強調する目的で設定が異なる。
Weibull	$p_3(\mathrm{dose}) = g + (1-g)(1-\exp[-\beta \mathrm{dose}^{\alpha}])$	$\beta > 0$ $\alpha > 0$ $0 \le g \le 1$	$\log(\beta) \sim \mathrm{Normal}(0, 1)$ $\log(\alpha) \sim \mathrm{Normal}(\log(2), 0.18)$ $\mathrm{logit}(g) \sim \mathrm{Normal}(0, 2)$	ここで、α に対する事前確率は0.01となるよう設定され、power パラメータは1 未満となる。これにより、supra-linear model が可能となるが、power パラメータが1を大きく下回る場合は多くのデータが必要となる。
Gamma	$p_4(\mathrm{dose}) = g + \frac{1-g}{\Gamma(\alpha)} \int_0^{\beta d} t^{\alpha-1} \exp(-t)\, \mathrm{dose}t$	$\beta > 0$ $\alpha > 0.2$ $0 \le g \le 1$	$\log(\beta) \sim \mathrm{Normal}(0, 1)$ $\log(\alpha) \sim \mathrm{Normal}(\log(2), 0.18)$ $\mathrm{logit}(g) \sim \mathrm{Normal}(0, 2)$	ここで、α に対する事前確率は低くなるように設計されており、power パラメータは1 未満とな

モデル名	数式	Restriction	事前確率[1)]	注記
				る。これにより、supra-linear model が可能となるが、power パラメータが1を大きく下回る場合は多くのデータが必要である。α は数値的な理由により0.2を超える値に制限される。
Dichotomous Hill	$p_5(\text{dose}) = g + \frac{v(1-g)}{1+\exp[-a-b\log(\text{dose})]}$	$0 \le g \le 1$ $0 \le v \le 1$ $-\infty < a < \infty$ $b > 0$	a ~ Normal(3, 3.3) log(b) ~ Normal(log(2), 0.5) logit(g) ~ Normal(-1, 2) logit(ν) ~ Normal(0, 3)	
Logistic	$p_6(\text{dose}) = \frac{1}{1+\exp[-\beta_0-\beta_1\text{dose}]}$	$-\infty < a < \infty$ $\beta_1 > 0$	β_0 ~ Normal(0, 2) log(β_e) ~ Normal(0, 1)	
Log-Logistic	$p_7(d\text{dose}) = g + \frac{1-g}{1+\exp[-\beta_0-\beta_1\log(\text{dose})]}$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	β_0 ~ Normal(0, 1) log(β_e) ~ Normal(log(2), 0.25) logit(g) ~ Normal(0, 2)	
e	$p_8(\text{dose}) = \Phi(\beta_0 + \beta_1\text{dose})$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	β_0 ~ Normal(0, 2) log(β_1) ~ Normal(0, 1)	
e	$p_9(\text{dose}) = g + (1-g)\Phi[\beta_0 + \beta_1\log(\text{dose})]$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	β_0 ~ Normal(0, 1) log(β_1) ~ Normal(log(2), 0.25) logit(g) ~ Normal(0, 2)	

1) $\text{logit}(\gamma) = \log\left(\frac{\gamma}{1-\gamma}\right)$

3.1.3　利用方法に関する情報

(1)　システム要件

デスクトップ版 Microsoft Excel® 2010、2013、2016 (32-bit 又は 64-bit)[47]

(2)　セットアップ方法

https://www.epa.gov/bmds/benchmark-dose-software-bmds-version-31-download より.zipファイルをダウンロードし、任意のフォルダに解凍する (Administrator 権限は不要)。"bmds3.xslm."を開き、BMDS 3.1を起動する。

(3)　インターフェース

BMDS 3.1による解析作業を行うファイル「Analysis Workbook」は以下の5つのワークシートからなる。

- Help
- Main
- Data
- Report Options
- Logic

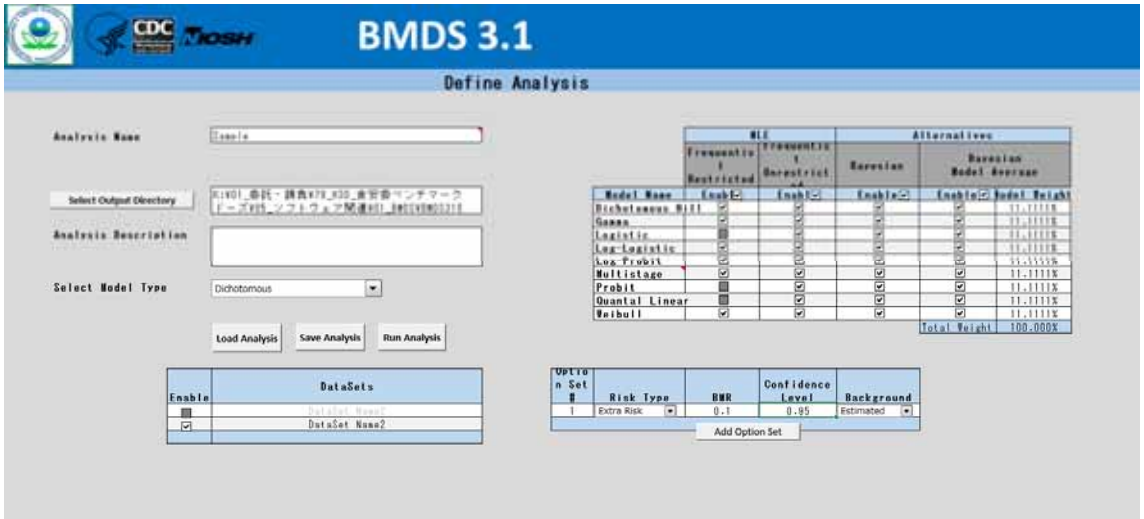


図3.1-3　Mainワークシート画面

[47] Version 3.0 では Microsoft Excel® 2016 (あるいはそれ以降) (32-bit) のみに対応していたが、Version 3.1 への更新により互換性が改良された。

(4)　操作の流れ

BMDS 3.1 の基本的な操作方法を以下に示す。

解析の名称、保存先の入力・指定

Mainワークシート上でAnalysis nameの入力、解析結果の出力ファイル (Result Workbook)とレポート (Word Report file) の保存先を指定する。

↓

データセットの入力

Data ワークシート上で、データタイプに応じた項目を入力する。

↓

解析条件の設定

Mainワークシート上でモデルタイプ及び解析条件を選択する。

選択したモデルタイプに応じた数理モデルのセット、解析条件が表示される。

- Continuous
- Dichotomous
- Dichotomous – Multi-tumor
- Dichotomous – Nested

↓

モデリングの実行、結果の確認

Mainワークシート上の”Run Analysis”をクリックし、モデリングを実行する。結果はデータセットごとにResult Workbookに保存される。

1) データセットの入力

データタイプ別に入力が必要な項目（入力項目）を表 3.1-3 に示す。また、連続値データ(要約) におけるデータセット入力画面を図 3.1-4 に示す。

表3.1-3　各データタイプの入力項目

データタイプ	入力項目
Continuous - summarized (連続値データ（要約))	- Dose (用量) - N (サンプル数) - Mean (平均値) - Std. Dev. (標準偏差)
Continuous - individual (連続値データ（個体))	- Dose (用量) - Response (反応量)
Dichotomous (二値データ)	- Dose (用量) - N (サンプル数) - Incidence (発現数)
Dichotomous - nested (ネストした二値データ)	- Dose (用量) - Litter Size (同腹児数) - Incidence (発現数) - Litter Specific Covariate (LSC) (腹固有の共変量)

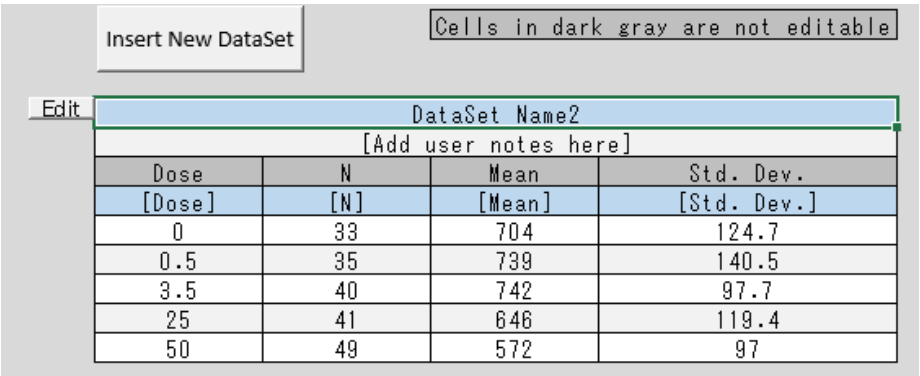


図3.1-4 連続値データ（要約）のデータセット入力画面

2)　解析条件の設定

2)-1　頻度論モデル

・Restriction の有無の設定

頻度論モデルにて最尤推定によるパラメータ推定を行う際は、Restriction の有無を設定する。各モデルの Restriction のデフォルト設定と Restricted/Unrestricted の設定の可否、パラメータの Restriction の内容を表 3.1-4 に示す。

表3.1-4　最尤推定によるパラメータ推定におけるRestrictionの設定

データタイプ	モデル	Frequentist Restricted	Frequentist Unrestricted	Restrictionの内容
Continuous (連続値データ)	Exponential	デフォルト	-	a = background response (> 0) b = slope (> 0) c = asymptote term: (> 1 for increasing data, 0 < c < 1 for decreasing data) d= power (> 1)
	Hill	デフォルト	設定可能	n= power (restrictionを行う場合: > 1)
	Linear	-	デフォルト	-
	Polynominal	デフォルト	設定可能	β = polynomial coefficients (restrictionを行う場合: "non-positive" 又は "non-negative")
	Power	デフォルト	設定可能	δ = power (restrictionを行う場合: > 1)
Dichotomous (二値データ)	Dichotomous Hill	設定可能	デフォルト	α = power (restrictionを行う場合: < 18, ≥ 1) β = slope (restrictionを行う場合: ≥ 1)
	Gamma	デフォルト	設定可能	α = power (restrictionを行う場合: < 18, ≥ 1) β = slope (restrictionを行う場合: ≥ 1)
	Logistic	-	デフォルト	-
	Log-Logistic	デフォルト	設定可能	α = power (restrictionを行う場合: < 18, ≥ 1) β = slope (restrictionを行う場合: > 1)
	Log-Probit	設定可能	デフォルト	β = slope (restrictionを行う場合: ≤ 18, ≥ 1)
	Mutistage	デフォルト	設定可能	β = dose coefficients (restrictionを行う場合: > 1)
	Probit	-	デフォルト	-

データタイプ	モデル	Frequentist Restricted	Frequentist Unrestricted	Restrictionの内容
	Quantal Linear	-	デフォルト	-
	Weibull	デフォルト	設定可能	α = power (restrictionを行う場合: < 18, ≥ 1) β = slope (restrictionを行う場合: ≥ 1)
Multi-tumor	Multi-Tumor (MS_Combo) Model	デフォルト	-	β: 検出された、又は特定のadverse directionに応じて、自動的に負の値をとらない (Non-negative) 、又は正の値をとらない (Non-positive)
Dichotomous-Nested (ネストした二値データ)	Nested Logistic	デフォルト	設定可能	ρ = power (≥ 0, restrictionを行う場合: ≥ 1)

・ベイズ推定モデル

現時点では、二値モデルでのみ利用可能である[48]。解析に用いるベイズ推定モデルを選択する (図 3.1-5)。

2)-2 ベイズ推定に基づくモデル平均化

現時点では、二値モデルのうちのベイズ推定モデルでのみ適用可能である[49]。解析に用いるベイズ推定モデルのうち、ベイズ推定に基づくモデル平均化に用いるモデルを選択し、Model Weight (0-100%) を設定する (デフォルトでは、選択したモデルの数に応じて均等に重み付けされる) (図 3.1-5)。

え 将来的に連続値モデルにも対応予定

49 将来的に連続値モデルにも対応予定

	MLE		Alternatives		
	Frequentist Restricted	Frequentist Unrestricted	Bayesian	Bayesian Model Average	
Model Name	Enab☐	Enabl☐	Enable ☑	Enable ☑	Model Weight
Dichotomous Hill	☐	☑	☑	☑	11.1111%
Gamma	☑	☐	☑	☑	11.1111%
Logistic	■	☑	☑	☑	11.1111%
Log-Logistic	☑	☐	☑	☑	11.1111%
Log-Probit	☐	☑	☑	☑	11.1111%
Multistage	☑	☐	☑	☑	11.1111%
Probit	■	☑	☑	☑	11.1111%
Quantal Linear	■	☐	☑	☑	11.1111%
Weibull	☑	☐	☑	☑	11.1111%
				Total Weight	100.000%

図3.1-5　二値データの場合のMLE/Bayesian/Bayesian Model Average設定画面

3)　解析条件の設定

選択したモデルタイプに応じた解析条件を設定する。

各モデルタイプの解析条件の設定項目を表 3.1-5 に示す。

表3.1-5　各モデルタイプの解析条件の設定項目

モデルタイプ	項目	選択 (入力) 値	備考
Continuous (連続値データ)	BMR Type	Relative Deviation / Absolute Deviation / Standard Deviation / Point / Hybrid	デフォルトはRelative Deviation
	BMRF	BMRファクター 0から1の値	
	Tail Probability	ユーザー指定の値	BMR TypeでHybridを選択した場合に有害影響を定義するカットオフ値として入力 デフォルトは0.01
	Confidence Level	0から1の値	EPAによる推奨値: 0.95
	Distribution	Normal / Log-normal	全ての連続値モデルについて、同一用量に対する反応量分布の仮定に関して正規分布

モデルタイプ	項目	選択 (入力) 値	備考
			か対数正規分布かを選択可能
	Variance	Constant / Non-constant	
	Polynomial Restrictions	Automatic / Non-negative / Non-positive	β値の制限
Dichotomous (二値データ)	BMR Type	Extra Risk / Added Risk	
	BMR	0から1の値	EPAによる推奨値: 0.1
	Confidence Level	0から1の値	EPAによる推奨値: 0.95
	Background	Estimated / Zero / 特定の値	ゼロまたは特定の値 明確な根拠がない場合は Estimatedを推奨
Multi-Tumor (MS_Combo)	BMR Type	Extra Risk / Added Risk	
	BMR	0から1の値	EPAによる推奨値: 0.1
	Confidence Level	0から1の値	EPAによる推奨値: 0.95
	Background	Estimated / Zero / ユーザー指定の値	ゼロまたは特定の値 明確な根拠がない場合は Estimatedを推奨
Dichotomous-Nested (ネストした二値データ)	BMR Type	Extra Risk / Added Risk	
	BMR	0から1の値	EPAによる推奨値: 0.1
	Confidence Level	0から1の値	EPAによる推奨値: 0.95
	Litter Specific Covariate	Overall Mean / Control Group Mean	
	Background	Estimated / Zero / ユーザー指定の値	ゼロまたは特定の値である明確な根拠がない場合は Estimatedを推奨
	Bootstrap Iterations	1000 (固定)	
	Bootstrap Seed	Automatic / User specified	

4)　モデリング結果の出力、記録

モデリング結果は解析条件とともに Result Workbook に出力される。デフォルトではエクセルシートへの出力となるが、ワードファイルへの出力を設定することもできる。

モデリングの結果、Result Workbook に出力される内容の例を以下に示す。

4)-1　連続値データ (要約) を用いた解析例

- i) Summary　結果表
- ii) Summary　グラフ
- iii) Recommended Model (Linear Model) 結果表
- iv) Recommended Model (Linear Model)　グラフ

4)-2　二値データを用いた解析例

- i) Summary　結果表
- ii) Summary　グラフ
- iii) Recommended Model (Log-Logistic Model) 結果表
- iv) Recommended Model (Log-Logistic Model) グラフ
- v) モデル平均化　結果表

4)-1　連続値データ (要約) を用いた解析例

i) Summary　結果表

Summary

BMDS 3.1

DataSet Name1			
[Add user notes here]			
Dose	N	Mean	Std. Dev.
[Dose]	[N]	[Mean]	[Std. Dev.]
0	33	704	124.7
0.5	35	739	140.5
3.5	40	742	97.7
25	41	646	119.4
50	49	572	97

BMDS3.1

Analysis Description

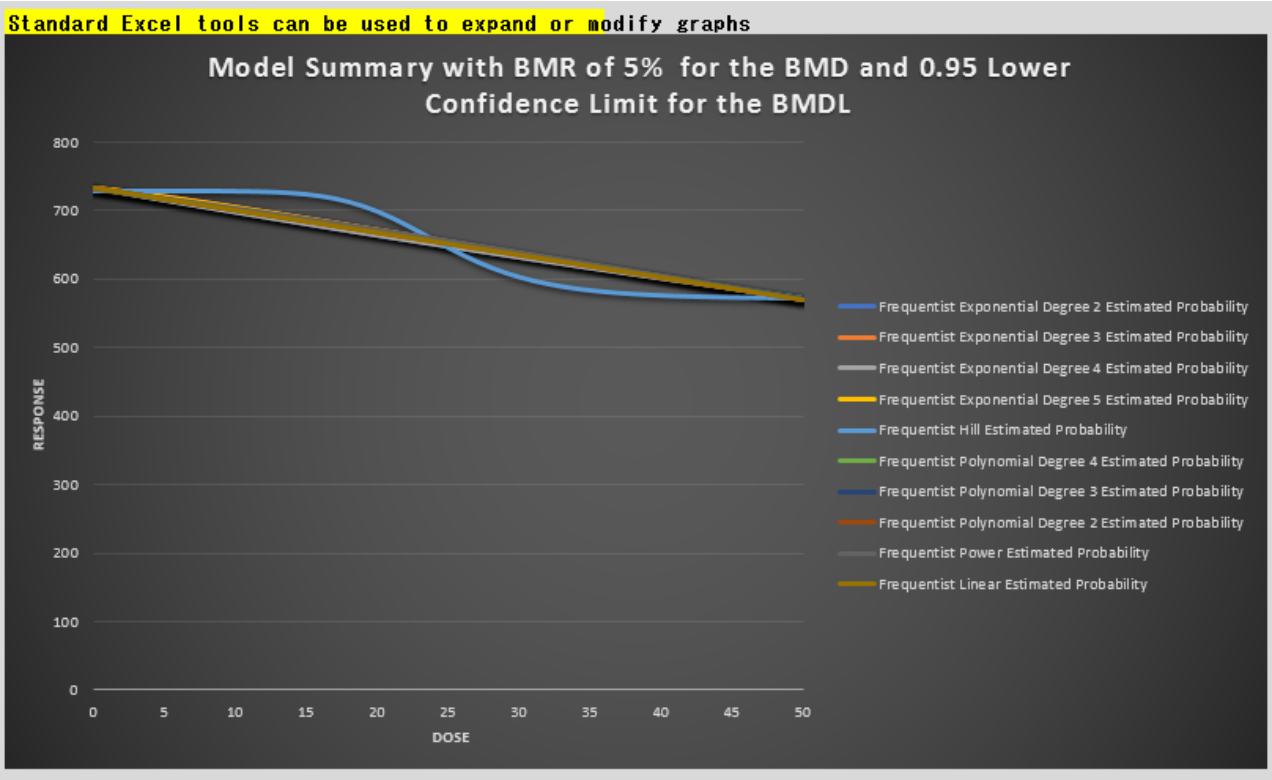
Color Key
Recommended frequentist model
Model averaging

Option set #1　roll right to see summary plot

Model	Analysis Type	Restriction	RiskType	BMRF	BMD	BMDL	BMDU	Test 4 P-Value	AIC	Unnormalized Log Posterior Probability	Scaled Residual for Dose Group near BMD	Scaled Residual for Control Dose Group	BMDS Recommendation	BMDS Recommendation Notes
Exponential 2 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	25.81787	20.62567	33.999554	0.2933852	2446.360882	-	-0.11739404	-1.481849637	Viable - Alternate	
Exponential 3 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	27.90527	20.86255	37.996812	0.1805689	2448.064831	-	-0.458749753	-1.418743401	Viable - Alternate	
Exponential 4 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	25.81787	16.72595	33.999554	0.2933852	2446.360882	-	-0.118003758	-1.481887144	Viable - Alternate	
Exponential 5 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	27.71606	16.95227	38.009285	0.0636752	2450.08032	-	-0.455045753	-1.447266322	Questionable	Goodness of fit p-value < 0.1
Hill (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	25.40131	19.15772	36.028144	0.1238565	2449.011731	-	-0.000957285	-1.278285508	Viable - Alternate	
Polynomial Degree 4 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	27.27051	22.34013	39.054452	0.3004955	2446.302165	-	-0.316538121	-1.423814582	Viable - Alternate	
Polynomial Degree 3 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	27.27051	22.34013	38.541419	0.3004958	2446.302163	-	-0.317112497	-1.424496111	Viable - Alternate	
Polynomial Degree 2 (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	27.2644	22.3402	38.336765	0.1603479	2448.302165	-	-0.316681886	-1.424493695	Viable - Alternate	
Power (CV)	frequentist	Restricted	Hybrid - extra risk	0.05	28.81165	22.40728	38.367303	0.1687288	2448.200271	-	-0.520860324	-1.336758184	Viable - Alternate	
Linear (CV)	**frequentist**	**nrestricted**	**Hybrid - extra ris**	**0.05**	**27.2644**	**22.3402**	**34.98436**	**0.300496**	**2448.30216**	-	**-0.316750073**	**-1.424421487**	**Viable - Recommended**	**Lowest AIC**

4)-1　連続値データ (要約) を用いた解析例

ii) Summary　グラフ



4)-1　連続値データ (要約) を用いた解析例

iii) Recommended Model (Linear Model) 結果表

Return to Summary

User Input

Info	
Model	frequentist Linear v1.0
Dataset Name	DataSet Name1
User notes	[Add user notes here]
Dose-Response Model	P[dose] = g + b1*dose
Variance Model	Var[i] = alpha

Model Options	
BMR Type	Hybrid - extra risk
BMRF	0.05
Tail Probability	0.01
Confidence Level	0.95
Distribution Type	Normal
Variance Type	Constant

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Mean]
Total # of Observatio	5
Adverse Direction	Automatic

Model Results

Benchmark Dose	
BMD	27.2644043
BMDL	22.34019554
BMDU	34.96435931
AIC	2446.302161
Test 4 P-value	0.300496022
D.O.F.	3

Model Parameters	
# of Parameters	3
Variable	Estimate
g	732.4729686
beta1	-3.231704272
alpha	13185.67616

Goodness of Fit

Dose	Size	Estimated Mean	Observed Mean	Estimated Std Dev	Observed Std Dev	Scaled Residual
0	33	732.4729686	704	114.8289	124.7	-1.424442149
0.5	35	730.8571164	739	114.8289	140.5	0.419528089
3.5	40	721.1620036	742	114.8289	97.7	1.147716832
25	41	651.6803618	646	114.8289	119.4	-0.31675007
50	49	570.887755	572	114.8289	97	0.067802748

Likelihoods of Interest

Model	Log Likelihood*	# of Parameters	AIC
A1	-1218.320673	6	2448.64135
A2	-1214.272956	10	2448.54591
A3	-1218.320673	6	2448.64135
fitted	-1220.15108	3	2446.30216
R	-1248.240273	2	2500.48055

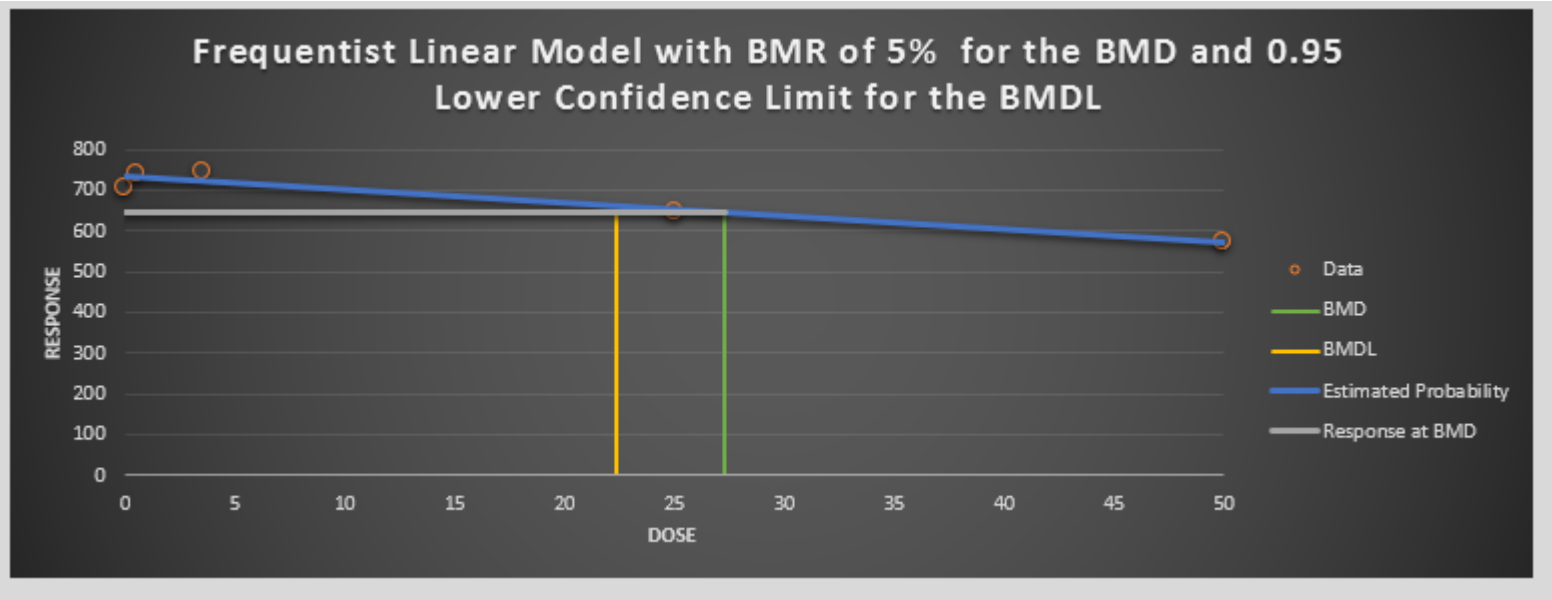
* Includes additive constant of -181.94983. This constant was not included in the LL derivat

Tests of Interest

Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	67.93463476	8	<0.0001
2	8.095434113	4	0.08814422
3	8.095434113	4	0.08814422
4	3.660815156	3	0.30049602

4)-1　連続値データ (要約) を用いた解析例

iv) Recommended Model (Linear Model) グラフ



4)-2　二値データを用いた解析例

i) Summary　結果表

Summary

BMDS 3.1

DataSet Name2		
[Add user notes here]		
Dose	N	Incidence
[Dose]	[N]	[Incidence]
0	20	0
0.5	18	2
3.5	19	5
25	20	6
50	18	12

Sample
Analysis Description

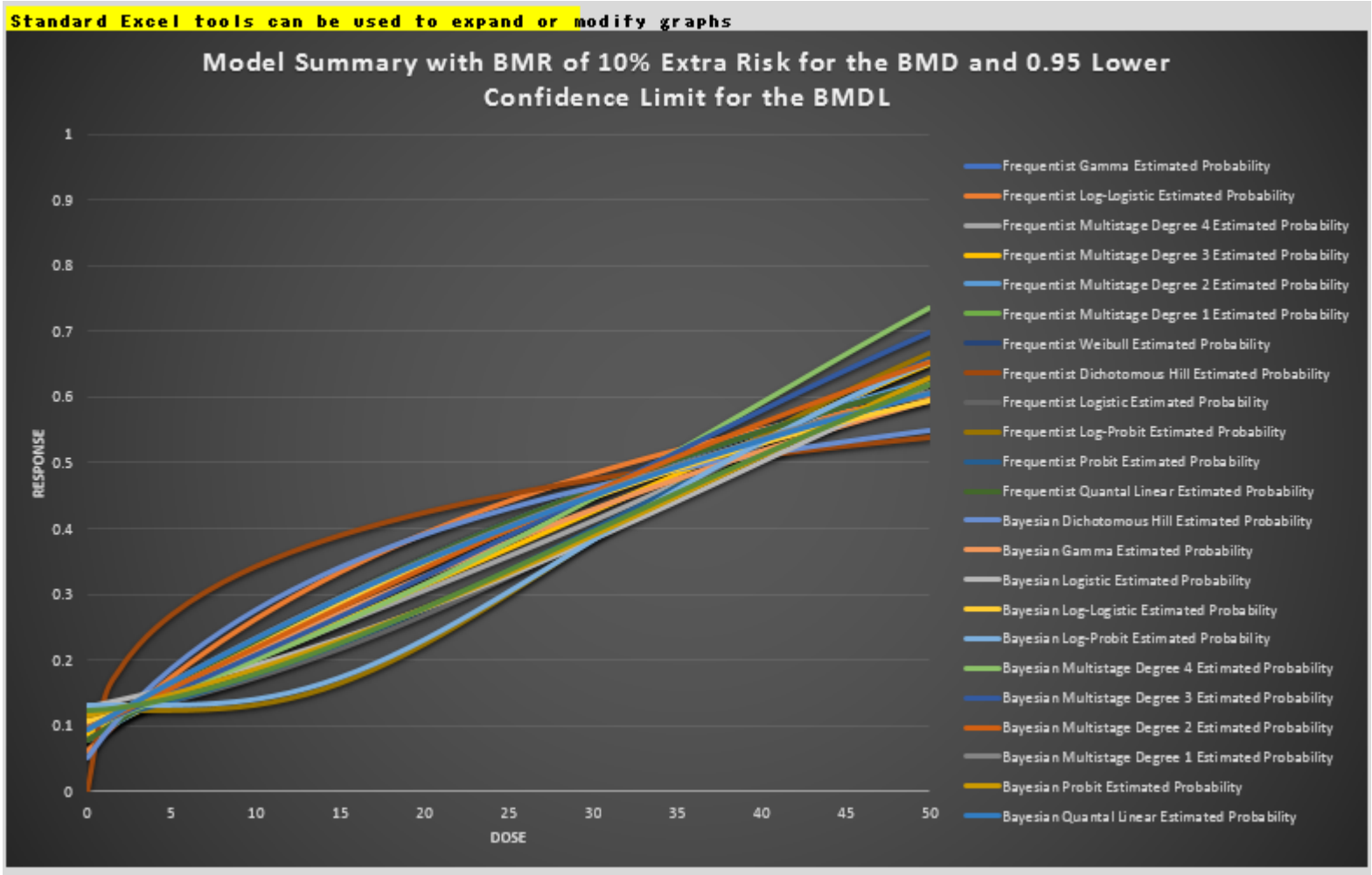
Color Key
Recommended frequentist model
Model averaging

Option set #1　　right to see summary p

Model	Analysis Type	Restriction	RiskType	BMRF	BMD	BMDL	BMDU	P Value	AIC	Unnormalized Log Posterior Probability	Scaled Residual for Dose Group near BMD	Scaled Residual for Control Dose Group	BMDS Recommendation	BMDS Recommendation Notes
Gamma	frequentist	Restricted	Extra Risk	0.1	5.911484	3.848187	28.237125	0.1240947	92.60995974	-	1.670817984	-1.293416048	Viable - Alternate	
Log-Logistic	frequentist	Restricted	Extra Risk	0.1	4.09014	2.16997	20.10302	0.121339	92.5290610	-	1.409774117	-1.152967073	Viable - Recommended	Lowest BMDL
Multistage Degree 4	frequentist	Restricted	Extra Risk	0.1	7.950677	3.941481	27.965365	0.069961	94.288978	-	1.742247023	-1.387937631	Questionable	Goodness of fit p-value < 0.1
Multistage Degree 3	frequentist	Restricted	Extra Risk	0.1	7.792313	3.902532	26.068947	0.0648968	94.41991749	-	1.747979984	-1.38014954	Questionable	Goodness of fit p-value < 0.1
Multistage Degree 2	frequentist	Restricted	Extra Risk	0.1	6.239516	3.849629	20.738041	0.0568506	94.6056744	-	1.691684536	-1.311754899	Questionable	Goodness of fit p-value < 0.1
Multistage Degree 1	frequentist	Restricted	Extra Risk	0.1	5.911483	3.848183	10.288415	0.1240947	92.60995974	-	1.670818011	-1.293415999	Viable - Alternate	
Weibull	frequentist	Restricted	Extra Risk	0.1	5.911484	3.848187	28.909143	0.1240947	92.60995974	-	1.670817895	-1.293416094	Viable - Alternate	
Dichotomous Hill	frequentist	Unrestricted	Extra Risk	0.1	0.46364	0.010753	2.0366776	0.0761914	93.0494066	-	0.106602351	-0.000552118	Questionable	Goodness of fit p-value < 0.1 BMD/BMDL ratio > 20 BMD/BMDL ratio > 5 BMDL 3x lower than lowest non-zero dose
Logistic	frequentist	Unrestricted	Extra Risk	0.1	12.73385	9.552159	17.155962	0.1321403	92.93039872	-	1.771605189	-1.551025655	Viable - Alternate	
Log-Probit	frequentist	Unrestricted	Extra Risk	0.1	19.02079	0	33.041797	0.0428974	95.81083601	-	-0.003495556	-1.672912397	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated Goodness of fit p-value < 0.1 Goodness of fit p-value < 0.05
Probit	frequentist	Unrestricted	Extra Risk	0.1	11.8138	8.976948	16.059809	0.1315969	92.87668272	-	1.778819231	-1.527764149	Viable - Alternate	
Quantal Linear	frequentist	Unrestricted	Extra Risk	0.1	5.911482	3.848186	10.287918	0.1240945	92.60995974	-	1.670819047	-1.293415348	Viable - Alternate	
Dichotomous Hill	bayesian	-	Extra Risk	0.1	3.14327	0.3381	22.826033	-	-	-49.80665797	1.117282953	-1.181930978		
Gamma	bayesian	-	Extra Risk	0.1	8.319187	2.056297	18.625049	-	-	-49.70681117	1.545086876	-1.539227395		
Logistic	bayesian	-	Extra Risk	0.1	12.99548	9.654102	18.647386	-	-	-48.79006834	1.396701704	-1.721449275		
Log-Logistic	bayesian	-	Extra Risk	0.1	8.157944	0.65189	24.103417	-	-	-48.40168032	1.588506567	-1.541085852		
Log-Probit	bayesian	-	Extra Risk	0.1	19.01864	6.663315	29.304345	-	-	-49.81002833	-0.038021924	-1.738624836		
Multistage Degree 4	bayesian	-	Extra Risk	0.1	8.769319	5.451538	13.96862	-	-	-52.86908935	1.672753167	-1.469922205		
Multistage Degree 3	bayesian	-	Extra Risk	0.1	8.201883	5.146077	13.085326	-	-	-49.95658807	1.635674335	-1.465244715		
Multistage Degree 2	bayesian	-	Extra Risk	0.1	7.395939	4.742638	11.768213	-	-	-49.2236973	1.573778407	-1.456103421		
Multistage Degree 1	bayesian	-	Extra Risk	0.1	6.311989	4.25946	10.054309	-	-	-47.17483787	1.46239372	-1.436985612		
Probit	bayesian	-	Extra Risk	0.1	12.10353	9.117713	17.080155	-	-	-48.72158876	1.572568325	-1.622611077		
Quantal Linear	bayesian	-	Extra Risk	0.1	6.334361	4.122462	10.951145	-	-	-47.92491982	1.461997028	-1.438783914		
Weibull	bayesian	-	Extra Risk	0.1	13.52067	2.663826	25.629422	-	-	-49.13431986	1.661846572	-1.671923557		
Model Average	bayesian MA	-	Extra Risk	0.1	9.372649	1.706954	21.579966	-	-	-	-	-		

4)-2　二値データを用いた解析例

ii) Summary　グラフ



4)-2　二値データを用いた解析例

iii) Recommended Model (Log-Logistic Model) 結果表

Return to Summary

User Input

Info	
Model	frequentist Log-Logistic v1.0
Dataset Name	DataSet Name2
User notes	[Add user notes here]
Dose-Response Model	P[dose] = g+(1-g)/[1+exp(-a-b*Log(dose))]

Model Options	
Risk Type	Extra Risk
BMR	0.1
Confidence Level	0.95
Background	Estimated

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Incidence]
Total # of Observatio	5

Model Results

Benchmark Dose	
BMD	4.09014385
BMDL	2.168674346
BMDU	28.10301546
AIC	92.52906185
P-value	0.121338576
D.O.F.	3
Chi^2	5.807920693

Model Parameters	
# of Parameters	3
Variable	Estimate
g	0.062324255
a	-3.605804718
b	1

Goodness of Fit

Dose	Estimated Probability	Expected	Observed	Size	Scaled Residual
0	0.062324255	1.246485107	0	20	-1.152968
0.5	0.07488983	1.348016947	2	18	0.5838374
3.5	0.143737297	2.731008641	5	19	1.4837741
25	0.44157361	8.831472203	6	20	-1.275008
50	0.602389778	10.84301601	12	18	0.5572159

Analysis of Deviance

Model	Log Likelihood	# of Parameters	Deviance	Test d.f.	P Value
Full Model	-40.90386726	0	-	-	-
Fitted Model	-44.26453092	2	6.72132732	3	0.0813309
Reduced Model	-54.75174214	1	27.6957497	4	<0.0001

4)-2　二値データを用いた解析例

iv) Recommended Model (Log-Logistic Model) グラフ

4)-2　二値データを用いた解析例

v) モデル平均化　結果表

Return to Summary

User Input

Info	
Model	Bayesian Model Averaging v1.0
Dataset Name	DataSet Name2
User notes	[Add user notes here]

Model Options	
Risk Type	Extra Risk
BMR	0.1
Confidence Level	0.95
Background	Estimated

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Incidence]
Total # of Observatio	5

Model Results

Benchmark Dose	
BMD	9.372649307
BMDL	1.706954464
BMDU	21.57996595

MA - Individual Models

Model	Posterior Probability	BMD	BMDL	BMDU
Dichotomous Hill	0.043073552	3.143270314	0.33809959	22.82603
Gamma	0.047596345	8.319187164	2.05629691	18.62505
Logistic	0.11904467	12.99547851	9.65410173	18.64739
Log-Logistic	0.175543492	8.157943934	0.65188962	24.10342
Log-Probit	0.042928624	19.01864111	6.663315	29.30435
Multistage	0.077159125	7.395938784	4.74263765	11.76821
Probit	0.12748241	12.10353151	9.11771283	17.08016
Quantal Linear	0.282798795	6.334196776	4.12247516	10.95112
Weibull	0.084372987	13.52066994	2.6638262	25.62942

3.2 PROAST (RIVM)

3.2.1 名称及びバージョン

最新バージョン：PROAST version 65.5 (2019 年 2 月末時点現在)[50]

3.2.2 PROAST の概要

(1) 概要

PROAST は、オランダ国立公衆衛生環境研究所 (RIVM) によって開発された用量反応データを統計解析するためのソフトウェアである。

PROAST には R パッケージと web 版のアプリケーション[51]がある。さらに、R パッケージには Menu バージョンと GUI (グラフィカルユーザーインターフェイス) バージョンの 2 つが収載されている。

Web 版のアプリケーションは PROAST 65.2 が使用されている。web 版のアプリケーションは、R パッケージのすべての機能が含まれるわけではないが、コンピュータへの R や R パッケージ等のインストールが不要であり、毒性データの基本的な用量反応解析を実行できるため、PROAST をすぐに利用したい場合に有用である。Web 版のアプリケーションのマニュアル「PROAST MANUAL WebApp」が RIVM のホームページからダウンロードできる。本調査ではより新しいバージョンであり、多くの機能を有する R パッケージを対象とする。

GUI バージョンは、version 38.9 から収載され、操作性が改善されたものである。GUI バージョンはユーザーフレンドリーであるが、詳細な解析には、解析のオプションが多い Menu バージョンが適しており、利用の目的によって使い分ける必要がある。

(2) 主要な機能

(2)-1 解析可能なデータタイプ及び使用される数理モデル

PROAST の主要な機能として用量反応モデリング、ヒト健康リスク評価のための BMD 導出、生態毒性リスク評価のための影響を及ぼす可能性のある濃度の導出がある。

PROAST (Menu バージョン又は GUI バージョン) に用量反応データを入力し、数理モデルをフィッティングさせることにより用量反応曲線が推定される。GUI バージョンで解析可能なデータセットは連続値データ (要約、個体) 及び二値データであり、Menu バージョンでは上記に加え順序カテゴリーデータ、クラスター化したデータの解析が可能である。

PROAST に収載されている数理モデルを表 3.2-1 に示す。連続値データで利用する数理

---

[50] https://www.rivm.nl/documenten/proast655

[51] Web 版のアプリケーション https://proastweb.rivm.nl/

モデルは指数分布 (Exp) モデル及び Hill モデルであり、二値データでは Logistic、Probit、Log-logistic、Log-Probit、Weibull、Gammma、Two-stage である。さらに二値データでは Latent variable models (LVMs、潜在変数モデル) も利用可能である。Menu バージョンでは上記の数理モデルのセットを用いた自動解析のほか、単一モデルによる解析が可能である。

表 3.2-1　PROAST に収載されている数理モデル

数理モデル	式
Null	$y = a$
Full	$y = \text{group mean}$
Logistic	$y = \dfrac{1}{1 + \exp(-a - bx)}$
Probit	$y = pnorm((x - a) \cdot b)$
Log-logistic	$y = a + \dfrac{1 - a}{1 + \exp\left(c \cdot \log\left(\frac{b}{x}\right)\right)}$
Log-probit	$y = a + (1 - a) \cdot pnorm\left(c \cdot \log\left(\frac{x}{b}\right)\right)$
Weibull	$y = a + (1 - a)\left(1 - \exp\left(-\left(\frac{x}{b}\right)^c\right)\right)$
Gamma	$y = pgamma(bx; c)$
Two-stage	$y = a + (1 - a)\left(1 - \exp\left(-\frac{x}{b} - c\left(\frac{x}{b}\right)^2\right)\right)$
Exp model 3	$y = a \cdot \exp(bx^d)$
Exp model 5	$y = a \cdot (c - (c - 1)\exp(-bx^d))$
Hill model 3	$y = a \cdot \left(1 - \dfrac{x^d}{b^d + x^d}\right)$

数理モデル	式
Hill model 5	$y = a \cdot \left(1 + (c-1)\frac{x^d}{b^d + x^d}\right)$

(2)-2　BMR の設定

BMR は、Menu バージョンでは任意の値を入力することができ、GUI バージョンではデフォルト値による予測がおこなわれる。BMR のデフォルト値は連続値データでは 5% (入力値としては 0.05) であり、これは生物学的な影響に関する測定値がバックグラウンド反応に比して平均値で 5%変化したことを意味する。連続値データでは、同一用量に対する反応量の分布が対数正規分布であると仮定する。二値データの BMR のデフォルト値は extra risk (過剰リスク) 10% (入力値としては 0.1)であり、Menu バージョンでは additional risk (追加リスク) も選択可能である。

(2)-3　モデリング結果の評価と信頼区間の決定

モデリング結果の評価は、デフォルトは AIC に基づき、AIC 間の 2 単位差を臨界値と見なしている (ユーザーによる変更が可能)。Menu バージョンでは AIC の他に対数尤度も選択できる。信頼区間の決定にはプロファイル尤度法又はブートストラップ法 (Menu バージョンのみ) が用いられる。ブートストラップの反復回数は任意の値を入力する必要がある (Menu バージョンのマニュアルには、200 回で合理的かつ正確な回答が得られるとの記載がある)。

(2)-4　Menu バージョンと GUI バージョンの機能の違い

Menu バージョンでのみ利用できる機能について以下に示す。

・順序カテゴリーデータ、クラスター化したデータの解析
・収載された数理モデルのセットを用いる自動解析に加え、単一モデルを用いる解析が可能
・数理モデルのフィッティングの確認に、AIC に加え対数尤度を選択可能
・数理モデルの形状パラメータに共変量を含めることが可能
・ (異なるパラメータに関連する) 複数の共変量を含めることが可能
・連続値データ応答変数に関して他の変換方法を選択可能 (平方根変換又は変換しない)
・多数のエンドポイントを一度に解析することが可能

・対照群に対する投与群のペアワイズ比較
・パラメータの初期値の調整
・パラメータの制限の調整
・数理モデル内で共変量以外の別の因子を使用し異なるマークでプロットする
・定量下限を左側打ち切り値と設定するのと同様に、データ上限として右側打ち切り値を設定できる
・x 及び y 軸のデータに誤差/変動がある場合、二乗和ではなく積の合計を最小化する
・推定するパラメータ数が多い場合は計算時間を短縮するため、フィッティングアルゴリズムの条件を緩和する
・近似モデルの任意のパラメータに対する信頼区間を計算するためのオプション
・RPF (相対力価係数) 又は同等のパラメータに対する信頼区間を導出するためのオプション
・結合データに数理モデルをフィッティングさせる際に、全用量範囲にわたり用量追加がある場合とない場合を明確にする
・時間の関数として反応を分析し (例：体重) 、用量反応解析でその曲線を特徴付けるパラメータ (例：成長曲線) を選択する
・サブグループ毎に、フィッティングさせた用量反応曲線をプロット
・異なるサブグループにおける用量反応曲線を、同じバックグラウンド反応量にスケーリングしたプロットの作成
・ブートストラップによる BMD の信頼区間の計算

(3) 旧バージョンからの変更点

PROAST の旧バージョンからの重要な変更点としては以下の 2 点である。

・選択又は許容される数理モデルは、(尤度比検定を用いた) 対数尤度ではなく、AIC に基づいている。現時点で EFSA は、PROAST のデフォルト値でもある AIC 間の 2 単位差を臨界値と見なしている。PROAST の Menu バージョンでは、本臨界値を変更することが可能である。
・(連続値データの場合) ネストした指数分布モデル又は Hill モデルから「最小」モデルを選定する方法が取り下げられた。代わりに、2 つの数理モデル (モデル 3 及び 5、表 3.2-1 参照) のみが考慮され、AIC の値が小さい数理モデルが、BMD の信頼区間を計算するための数理モデルとして選択される。

(4) BMDS との相違点

PROAST のマニュアルによると、PROAST は EPA が開発した BMDS ソフトウェアと一致させることを目的としている。一方で、マニュアルに記載された PROAST と BMDS との重要な違いは、以下の通りである。

・同一用量に対する反応量の分布について、BMDS は正規分布を連続値データのデフォルトとして用い[52]、PROAST は対数正規分布をデフォルトとして用いる。
・BMDS は大半の非連続データ用の数理モデルで形状パラメータの値に制限を課すのに対し、PROAST は制限しない。

[52] BMDS2.7 までは Exponential model のみ対数正規分布の仮定を選択できたが、最新の BMDS3.1 では、連続値データの解析に用いる全ての数理モデルについて、同一用量に対する反応量分布の仮定条件として対数正規分布を選択できる機能が追加された。

3.2.3　利用方法に関する情報

(1)　セットアップ方法

PROAST の R パッケージを RIVM のホームページからダウンロードする[53]。

(2)　操作の流れ

PROAST のプログラムは R 言語による解析であるため、解析を実施する前に、統計解析ソフトの R (又は S-Plus) をインストールする必要がある。R ソフトウェアは無料でダウンロードできる[54]。

数理モデルセットをフィッティングさせる基本的な解析の流れは、Menu バージョンと GUI バージョンで共通である。

まず、PROAST を起動し、データファイルを呼び出す。続いて、データタイプの指定、数理モデルの選択 (セット/単一モデル) を行う。その後、各変数、共変量、BMR の値を指定し、数理モデルのフィッティング、モデリング結果の出力を行う。

(2)-1　GUI バージョン

GUI バージョンの操作については、PROAST の GUI マニュアルに詳細が記載されている。

解析できるデータタイプは、連続値データ、二値データであり、ネストしたデータ (発生毒性試験における同腹効果等) も解析可能である。

連続値データを用いる解析では、指数分布モデルと Hill モデルで解析が実行される。

二値データを用いる解析では、EFSA が推奨する 7 つの数理モデルと、2 つの Latent variable models (LVMs) (潜在変数モデル) による解析が解析可能である。

1)　データファイルの作成

解析に用いるデータファイルを作成する (図 3.2-1)。データはエクセル形式で作成し、テキスト形式 (タブ区切り) に変換する。列は、用量、反応量及びその他の関連因子で構成される。行は、個々の動物 (又はその他の試験単位) のデータ、又は用量群毎の要約データである。要約データの場合は、列に群の個体数の情報が必要である。連続値データでかつ要約データの場合は、用量群毎の平均値と標準偏差 (SD) (又は標準誤差 (SEM)) の情報が必要となる。

---

[53] https://www.rivm.nl/documenten/proast655 (2019 年 1 月現在)

[54] www.cran.r-project.org

sample data_Continuous - メモ帳

ファイル(F)　編集(E)　書式(O)　表示(V)　ヘルプ(H)

dose	sex	mean.bw	sd.bw	n.bw
0	1	704	124.7	33
0.5	1	739	140.5	35
3.5	1	742	97.7	40
25	1	646	119.4	41
50	1	572	97	49
0	2	496	105.7	37
0.5	2	477	132.6	33
3.5	2	480	106.8	32
25	2	402	106.8	27
50	2	361	81.1	20

図 3.2-1　解析するデータファイルの一例 (連続値データ (要約))

2)　PROAST GUI の起動

ダウンロードした R ソフトウェアを開くと、Console ウィンドウが立ち上がる。同ウィンドウのパッケージタブから PROAST パッケージをロードした後 (図 3.2-2)、Console ウィンドウにコマンド「> g.proast()」を入力すると PROAST GUI が表示される (図 3.2-3)。PROAST GUI は R ウィンドウとは独立して配置され、いつでもアクセスできるようになっている。その後の作業は基本的に PROAST GUI 画面上で操作する。

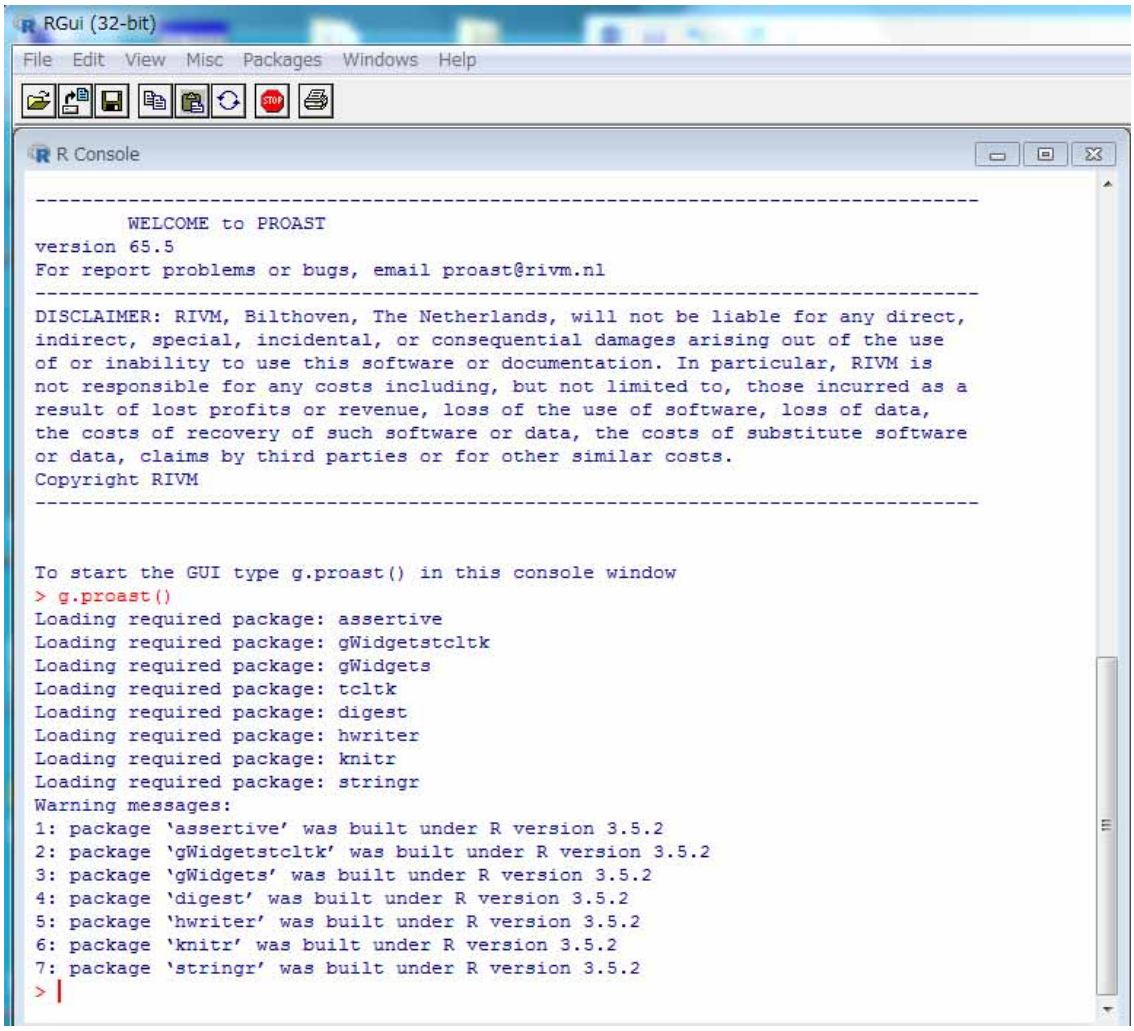


図 3.2-2　PROAST パッケージの呼び出し

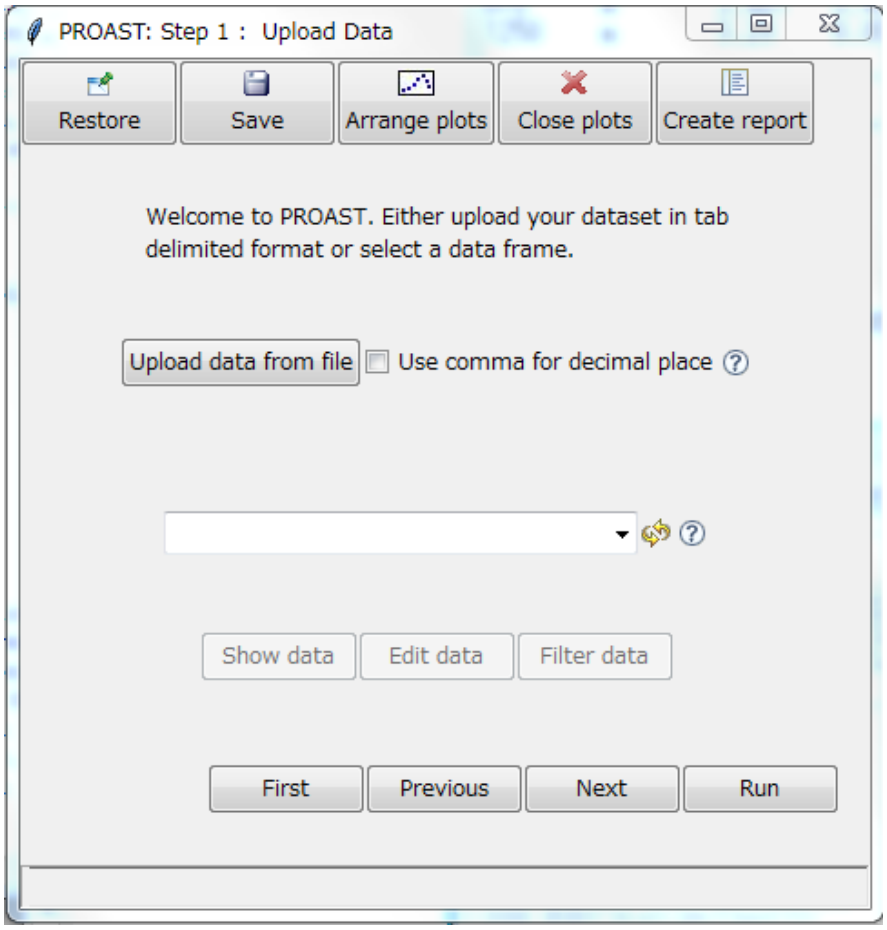


図 3.2-3　PROAST GUI の画面

3)　データセットの呼び出しと解析の設定

「upload data from file」を選択し 1) で作成したデータファイルをアップロードする。「show data」をクリックするとデータが確認できる。PROAST GUI の画面に従い、データの種類 (二値データ/連続値データ、個体データ/要約データ、同腹効果の有無) を選択し、解析データ (データのタイプ、標準偏差 (SD)/標準誤差 (SEM) (要約データの場合)、個体数 (要約データの場合)、反応データ、共変量の有無等) の列を指定する。その後 BMR (デフォルト値は二値データの場合 10%、連続値データの場合 5%) を指定し数理モデルを選択する。「run」をクリックし解析を実行する。最終的なプロットのフォーマットを変更したい場合は、「Advanced plotting options」を選択する。

4)　モデリング結果の出力、記録

連続値データの場合は 2 つの数理モデ (Exponential model, Hill model) のフィッティングの結果がプロットとして表示される。各プロットの右側には表 3.2-2 に示す凡例が表示さ

れる。プロット上の印 (△) は各用量における反応量の幾何学平均を示す。水平方向の破線は BMR を表し、垂直方向の線は近似曲線の BMD を表す。

二値データの場合は 8 つの数理モデルのフィッティングの結果がプロットとして表示される。垂直方向のバーで示される範囲は、信頼区間を表す。また、共変量を設定した場合は、同一プロット内に複数の用量反応曲線が表示される。

レポートを作成する場合は、Create report を選択し、レポートを表示させる。モデリングの結果、出力される内容の例を、入力したデータセットと共に以下に示す。

4)-1 連続値データ (要約) を用いた解析例 (性別を共変量とした例)

- i) データセット
- ii) 表示されるプロット
- iii) Console ウィンドウに表示される結果
- iv) 出力レポート

4)-2 連続値データ (個体) を用いた解析例

- i) データセット
- ii) 表示されるプロット
- iii) Console ウィンドウに表示される結果
- iv) 出力レポート

4)-3 二値データを用いた解析例

- i) データセット
- ii) 表示されるプロット
- iii) Console ウィンドウに表示される結果
- iv) 出力レポート

表 3.2-2　連続値データ (要約及び個体) で出力されるプロットの凡例

凡例	内容
version	PROAST のバージョン
loglik	最良のフィッティングに関連する対数尤度の値
var	群内の分散 (自然対数応答に関連)
a	近似モデルに従った a (background response (バックグラウンド反応)) の値
CED	臨界影響用量 (Critical Effect Dose) 又はそれと同等の、平均反応が指定パーセント変化する BMD
c	近似モデルによる c (maximum response) の値
d	近似モデルによる d (steepness) の値
CES	臨界影響サイズ (Critical Effect Size)。平均反応の変化率に対する BMR
CEDL	CED の信頼区間の下限値
CEDU	CED の信頼区間の上限値
b	近似モデルに従った b (background response (バックグラウンド反応)) の値
conv	近似アルゴリズムは収束したか？ 1 の場合は Yes、0 の場合は No
scaling factor on x	GUI バージョンでは常に 1 (Menu バージョンでのみこの値を調整可)
dtype	データタイプ
covariate	共変量として使用されたデータ

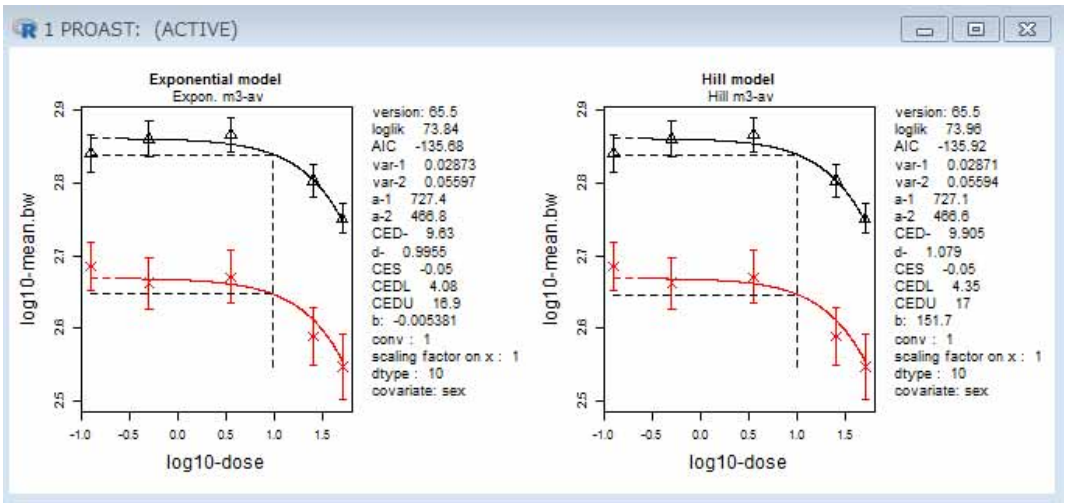
4)-1 連続値データ (要約) を用いた解析例

i) データセット

sample data_Continuous.txt

	dose	sex	mean.bw	sd.bw	n.bw
1	0	1	704	124.7	33
2	0.5	1	739	140.5	35
3	3.5	1	742	97.7	40
4	25	1	646	119.4	41
5	50	1	572	97	49
6	0	2	496	105.7	37
7	0.5	2	477	132.6	33
8	3.5	2	480	106.8	32
9	25	2	402	106.8	27
10	50	2	361	81.1	20

ii) 表示されるプロット



iii) Console ウィンドウに表示される結果

```
 -----------  CES =  0.05  ---------------------------------
The lowest BMDL and highest BMDU from exponential and Hill models are:
  covar.txt bmdl.lowest bmdu.highest
1       all        4.08           17
 ---------------------------------------------------
calculations started at: Wed Mar 13 13:47:10 2019
and ended at: Wed Mar 13 13:47:12 2019
------------------------- The analysis is complete. -------------------------
```

iv) 出力レポート

Benchmark dose analysis for continuous summary data0313.txt

This report was generated by `hayashi-tae` on `2019-03-13` using PROAST version `65.5` and `R version 3.5.1 (2018-07-02)`.

Dose

The dose variable was `'dose'`.

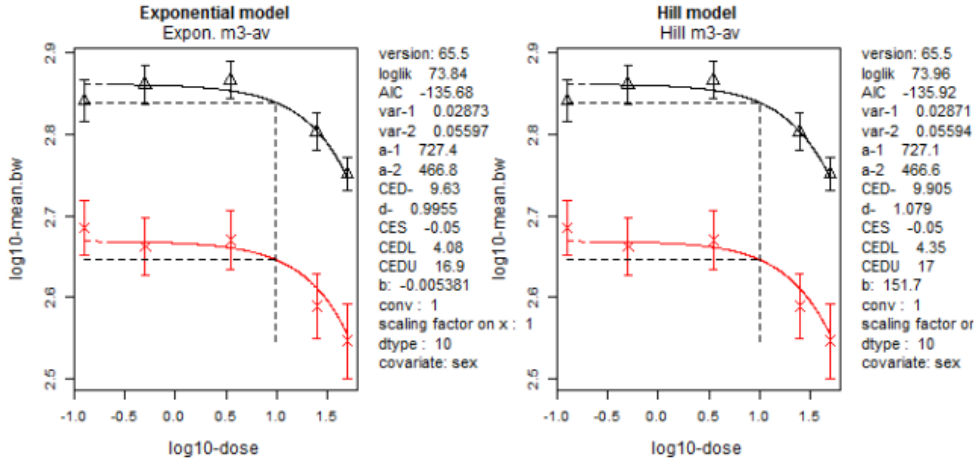
Response

The response variable was `'mean.bw'`.

Covariate

The covariate was 'sex' with levels '1', '2'.

Fitted Models



Exponential model

The chosen exponential model was `Expon. m3-av`.

Hill model

The chosen Hill model was `Hill m3-av`.

Benchmark dose

	Covariate	Lowest BMDL	Highest BMDU
1	All subgroups	4.08	17

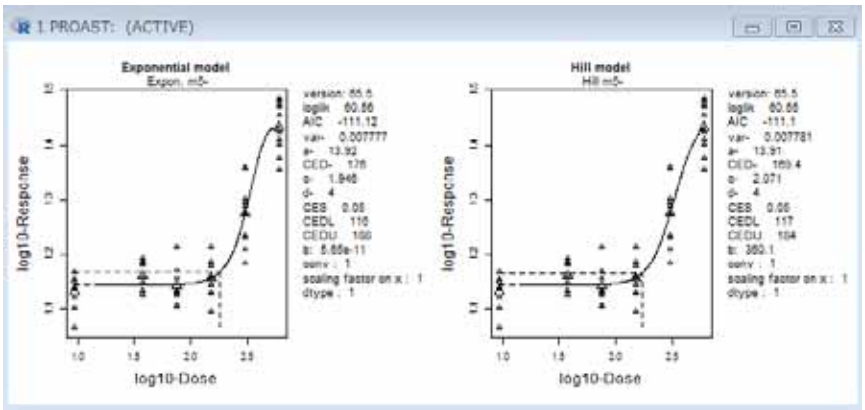
4)-2 連続値データ (個体) を用いた解析例

i) データセット (一部)

sample.data_C...

	Dose	Response
1	0.0	14.71
2	0.0	13.19
3	0.0	14.11
4	0.0	14.12
5	0.0	13.80
6	0.0	11.59
7	0.0	12.65
8	0.0	14.21
9	0.0	13.69
10	0.0	13.21
11	37.5	15.22
12	37.5	14.07
13	37.5	15.18
14	37.5	15.31
15	37.5	15.61
16	37.5	14.38
17	37.5	14.10
18	37.5	13.38
19	37.5	14.08
20	37.5	13.60
21	75.0	14.03
22	75.0	12.75
23	75.0	13.34
24	75.0	12.81
25	75.0	13.35
26	75.0	16.32
27	75.0	14.80

ii) 表示されるプロット



iii) Console ウィンドウに表示される結果

```
 -----------  CES =  0.05  ---------------------------------
The lowest BMDL and highest BMDU from exponential and Hill models are:
  covar.txt bmdl.lowest bmdu.highest
1       all         116          188
 ----------------------------------------------------
```

iv) 出力レポート

Benchmark dose analysis for sample.data_Continu_indiv.txt

This report was generated by `hayashi-tae` on `2019-03-13` using PROAST version `65.5` and `R version 3.5.1 (2018-07-02)`.

Dose

The dose variable was `'Dose'`.

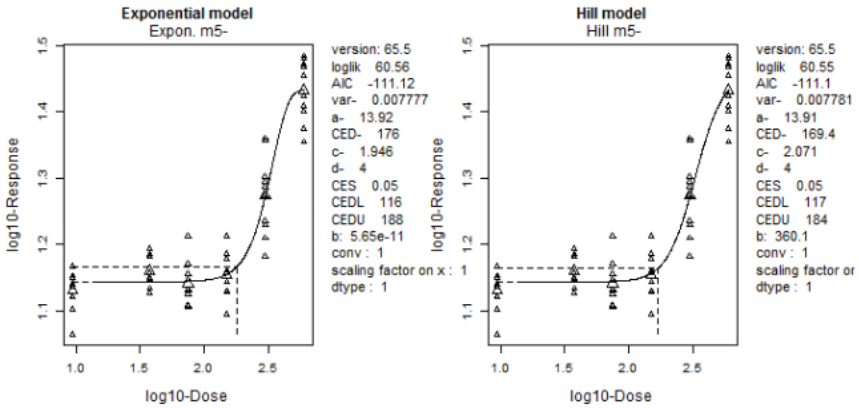
Response

The response variable was `'Response'`.

Covariate

There was no covariate.

Fitted Models



Exponential model

The chosen exponential model was `Expon. m5-`.

Hill model

The chosen Hill model was `Hill m5-`.

Benchmark dose

	Covariate	Lowest BMDL	Highest BMDU
1	All data	116	188

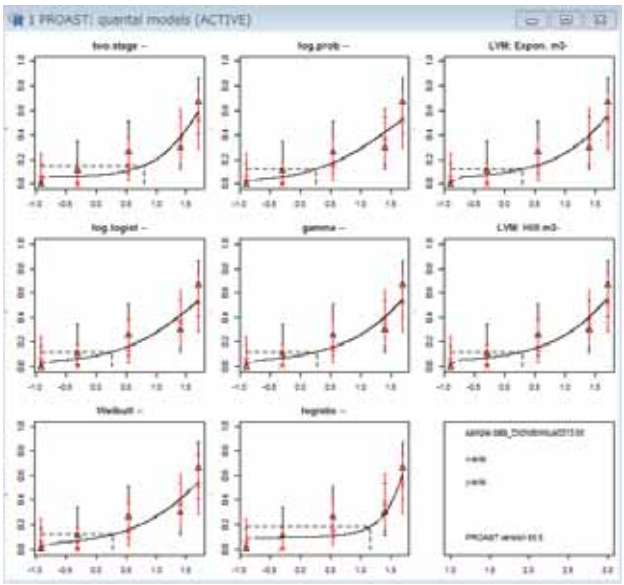
4)-3 二値データを用いた解析例

i) データセット

sample data_Dichotomous0313.txt

	dose	sex	response	group.size
1	0.0	1	0	20
2	0.5	1	2	18
3	3.5	1	5	19
4	25.0	1	6	20
5	50.0	1	12	18
6	0.0	2	1	20
7	0.5	2	0	19
8	3.5	2	3	20
9	25.0	2	7	19
10	50.0	2	10	19

ii) 表示されるプロット



iii) Console ウィンドウに表示される結果

```
response
             model No.par  loglik    AIC accepted  BMDL BMDU  BMD conv
1             null      1 -105.72 213.44             NA   NA   NA
2             full      5  -82.05 174.10             NA   NA   NA
3        two.stage      3  -84.20 174.40       no    NA   NA  6.3  yes
4       log.logist      3  -83.14 172.28      yes 0.438 6.72  1.9  yes
5          Weibull      3  -83.03 172.06      yes 0.401 7.21  1.9  yes
6         log.prob      3  -83.15 172.30      yes 0.484 5.92  1.8  yes
7            gamma      3  -82.98 171.96      yes 0.352 7.00  1.9  yes
8         logistic      2  -85.66 175.32       no    NA   NA 14.0  yes
9           probit      2  -85.45 174.90       no    NA   NA 13.0  yes
10 LVM: Expon. m3-      3  -83.01 172.02      yes 0.670 8.34  2.0  yes
11   LVM: Hill m3-      3  -83.05 172.10      yes 0.491 8.16  2.0  yes
BMR: 0.1 extra risk
constraint: no
no litter effects
critical AIC value: 2
PROAST version: 65.5


            lowest BMDL    highest BMDU    nr. of models
response          0.352           8.34                6


calculations started at: Wed Mar 13 14:30:53 2019
and ended at: Wed Mar 13 14:30:57 2019
------------------------- The analysis is complete. -------------------------
```

iv) 出力レポート

Benchmark dose analysis for sample data_Dichotomous0313.txt

This report was generated by **hayashi-tae** on 2019-03-13 using PROAST version 65.5 and R version 3.5.1 (2018-07-02).

Dose

The dose variable was '**dose**'.

Response

The response variable was '**response**'.

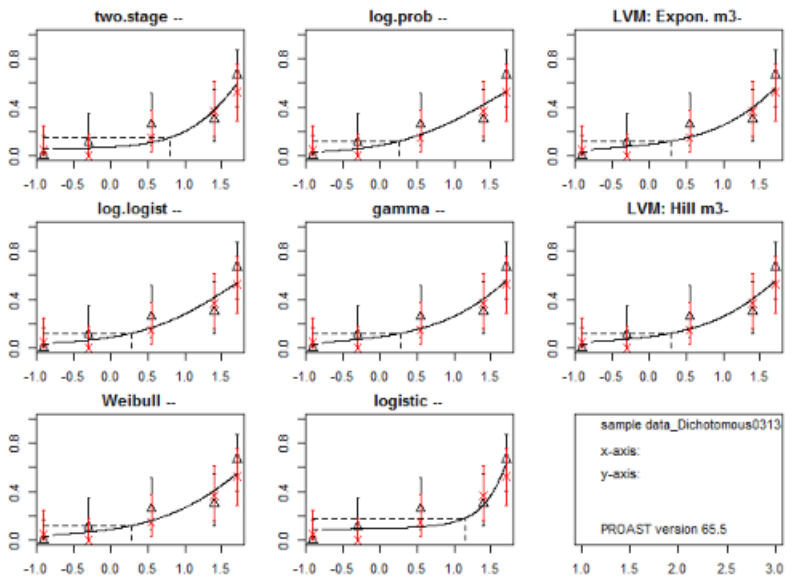
Covariate

The covariate was 'sex' with levels '1', '2'.

Fitted Models

model	No.par	loglik	AIC	accepted	BMDL	BMDU	BMD	sens.subgr	conv
null	1	-105.72	213.44						
full	10	-78.98	177.96						
two.stage	3	-84.20	174.40	no			6.3	--	yes
log.logist	3	-83.14	172.28	yes	0.438	6.72	1.9	--	yes
Weibull	3	-83.03	172.06	yes	0.401	7.21	1.9	--	yes
log.prob	3	-83.15	172.30	yes	0.484	5.92	1.8	--	yes
gamma	3	-82.98	171.96	yes	0.352	7.00	1.9	--	yes
logistic	2	-85.66	175.32	no			14.0	--	yes
LVM: Expon. m3-	3	-83.01	172.02	yes	0.670	8.34	2.0	1	yes
LVM: Hill m3-	3	-83.05	172.10	yes	0.491	8.16	2.0	1	yes

```
BMR: 0.1 extra risk
constraint: no
P-value GoF: 0.05
no litter effects
0
```



Benchmark dose

Lowest BMDL	Highest BMDU
0.352	8.34

(2)-2　Menu バージョン

Menu バージョンの操作については PROAST マニュアル (Menu バージョン) に詳細が記載されている。Menu バージョンでは、数理モデルのフィッティングを行う前に、図 3.2-4 の Change Setting (data specification) にある設定条件の設定ができる。また、メインメニュー (Main menu) からも設定可能である。

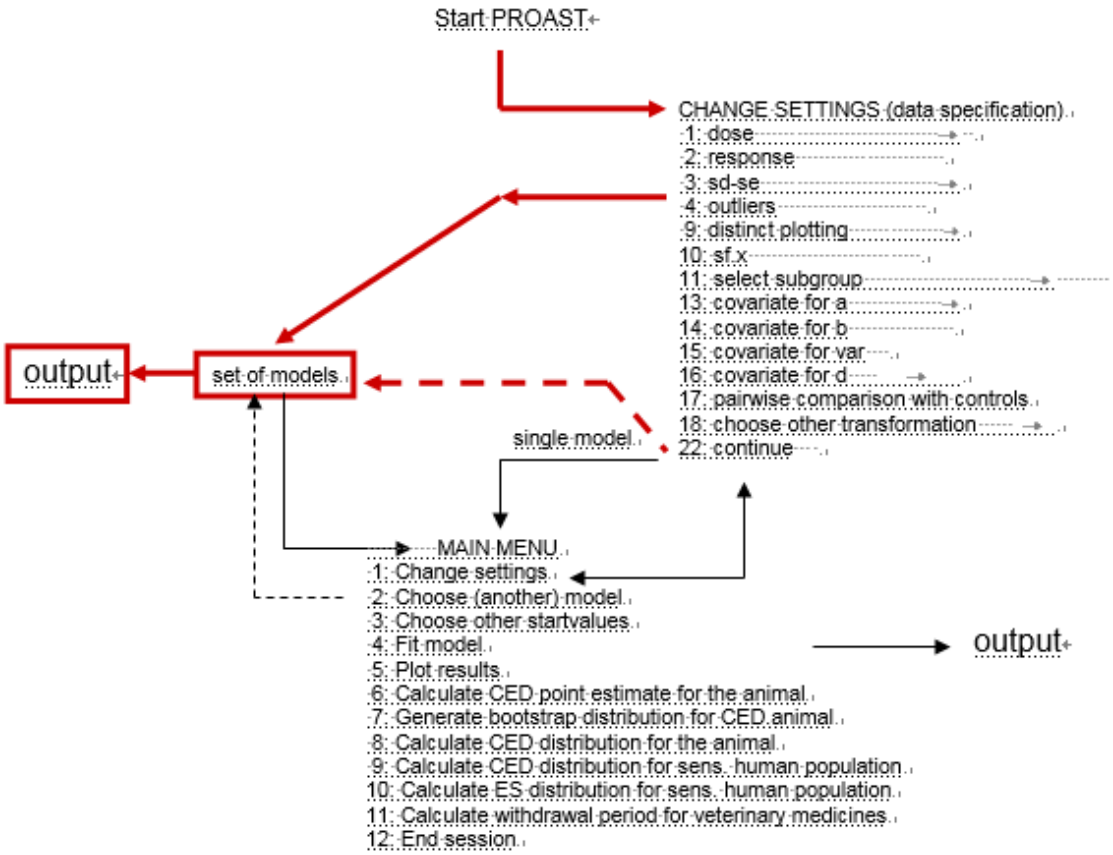


図3.2-4　Menuバージョンの解析の流れ
(PROASTマニュアル (Menuバージョン) より引用))

1)　データファイルの作成

GUI バージョンと同様の方法でデータファイルを作成する。

2)　データセットの呼び出し

R ソフトウェアを開くと、Console ウィンドウが立ち上がる。同ウィンドウのパッケージタブから PROAST パッケージをロードする (ここまでは GUI バージョンと同様)。

Menu バージョンは、Console ウィンドウ上にコマンドを入力することにより操作する。

まず、Console ウィンドウにコマンド「getwd()」を入力し、作業ディレクトリ[55]を表示させる。表示されたディレクトリが、解析に用いるデータセットを保存したディレクトリと異なる場合、解析に用いるデータセットを作業ディレクトリに移動する、又は、作業ディレクトリを変更する必要がある。作業ディレクトリを変更する場合は、R ウィンドウの左上隅にある「File」をクリックし、「Change dir ...」を選択し、データセットを保存したフォルダを参照する。参照が適切に行われたか否かは、再度コマンド「getwd()」を入力することにより確認できる。

Console ウィンドウにコマンド「1“データファイルのファイル名”.txt)」を入力し、ファイルを呼び出す (図 3.2-5)。呼び出しが可能なデータファイルについては Console ウィンドウにコマンド「ls()」を入力することで確認できる。なお、コマンド「f.proast()」に、解析結果のファイル名(「f.proast (,“解析結果のファイル名“)」) を入力することにより、過去に実施した解析結果を表示し、過去の解析を再開することも可能である。

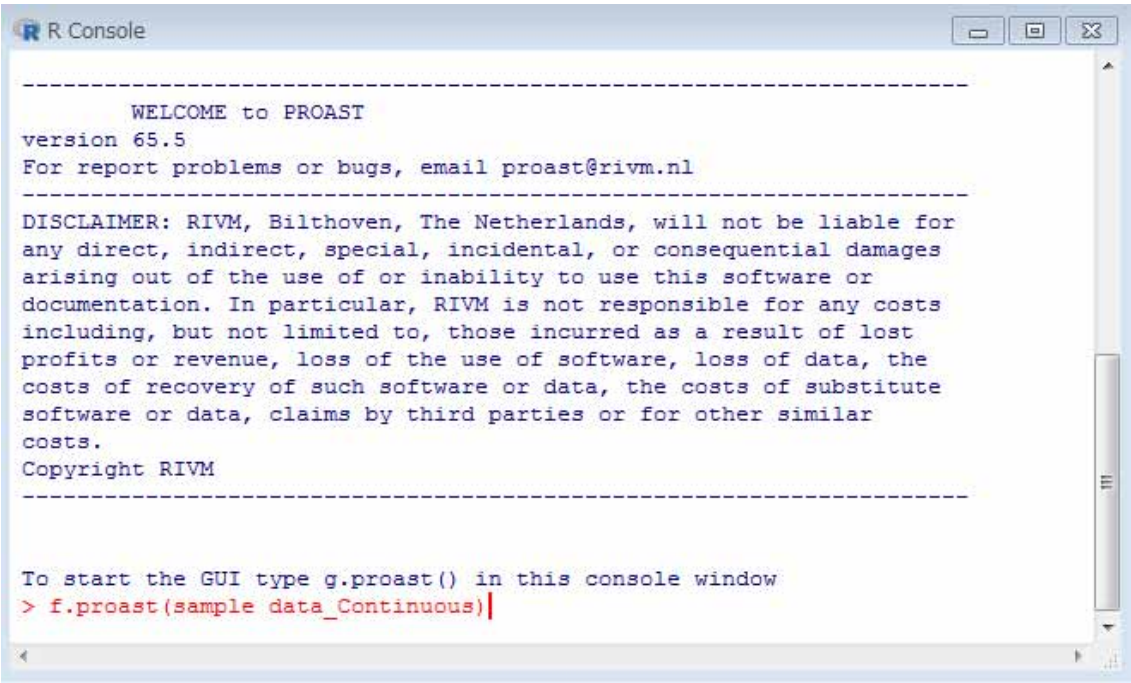


図 3.2-5　ファイルの呼び出し

[55]ファイルからデータを呼び出したり、解析結果をファイルで出力するコンピュータ上のフォルダのこと。

3) 解析の設定

データファイルを呼び出すと、基本的な解析条件の設定に関する質問が Console ウィンドウに順次表示される。解析を開始した際に表示される最初の質問は、以下の通りである。解析に用いるデータタイプを選択する。

What type of response data do you want to consider?

1: continuous, individual data
2: binary
3: ordinal
4: quantal
5: clustered continuous, individual data
6: clustered quantal
7: continuous, summary data
8: clustered continuous, summary data
9: quantal, CxT
10: other

データタイプの選択後は、PROAST からの質問に回答していく形で、基本的な解析条件の設定を行う (図 3.2-6)。連続値データ及び二値データの解析時に表示される基本的な解析条件を表 3.2-3 及び表 3.2-4 に示す。

また、Menu バージョンではそのほかにも、多岐にわたる解析条件の設定が「Change Settings」から可能である (図 3.2-7、表 3.2-5 及び表 3.2-6)。Change Settings は、基本的な解析条件の設定の過程やメインメニュー (Main menu) においてコマンド選択することで移行可能となる。Main menu の設定画面及び操作内容について図 3.2-8 及び表 3.2-7 に示す。

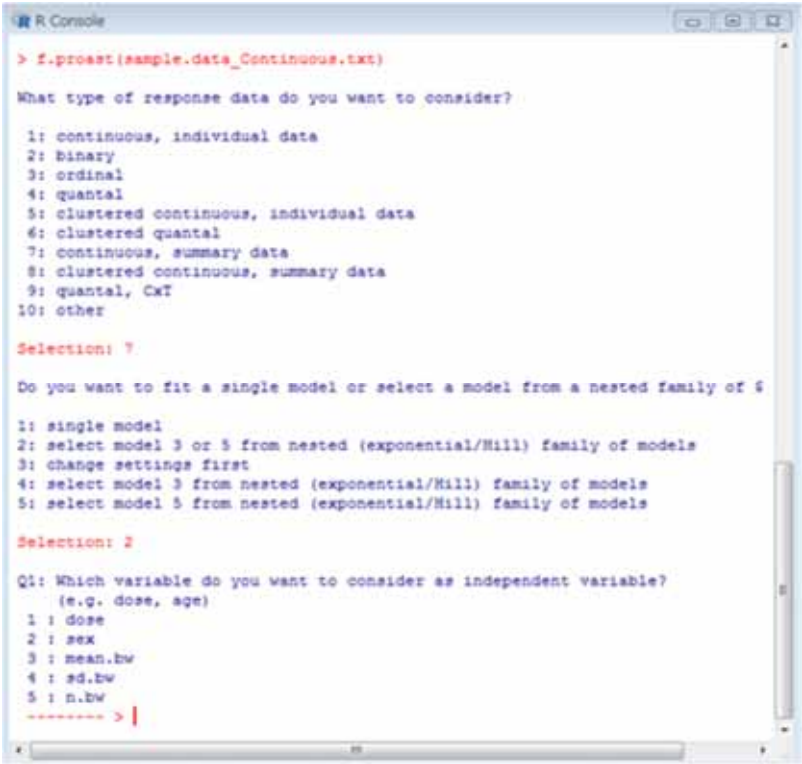


図 3.2-6　基本的な解析条件の設定画面

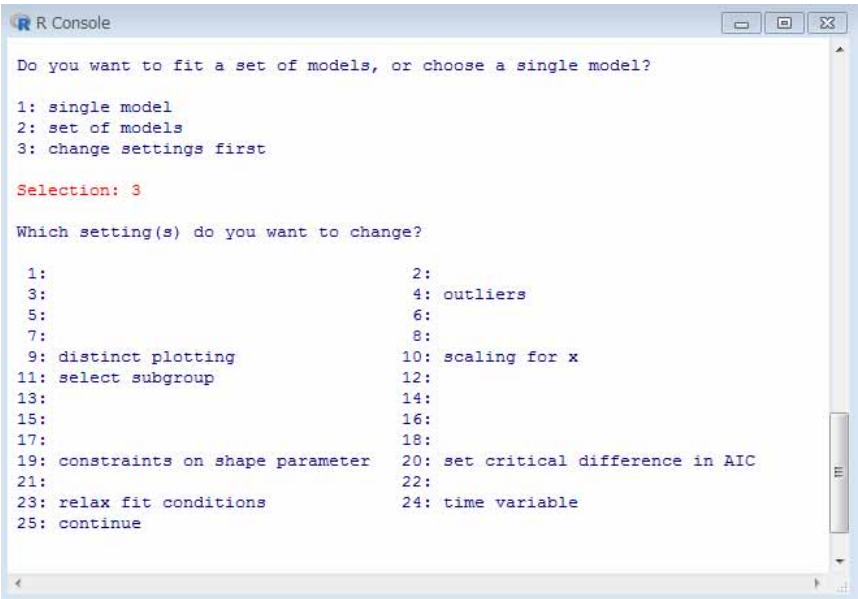


図 3.2-7　基本的な解析条件の設定の過程で **Change Settings** を選択した場合の解析条件の設定画面 **(**連続値データの例**)**

表 3.2-3　連続値データにおける基本的な解析条件

番号	設定項目	選択/入力項目
1	**データタイプの選択** (What type of response data do you want to consider?)	1: continuous, individual data、 2: binary、 3: ordinal、 4: quantal、 5: clustered continuous, individual data、 6: clustered quantal、 7: continuous, summary data、 8: clustered continuous, summary data、 9: quantal, CxT、10: other
2	**数理モデルの選択** (Do you want to fit a single model or select a model from a nested family of models?)　フィッティングさせる単一又は一連の数理モデルを選択できる。選択肢 1 (単一モデル) を選択することで、選択肢 2、4、5 と比べユーザーが解析を行う際の自由度と選択肢が増す。一連の数理モデルを一度にフィッティングさせる自動解析の方が速く作業できる。選択肢のうち 3 を選択すると、各種条件（表 3.2-5 参照）の変更ができる。	1: single model、2: select model 3 or 5 from nested (exponential/Hill) family of models、3: change settings first、4: select model 3 from nested (exponential/Hill) family of models、5: select model 5 from nested (exponential/Hill) family of models
3	**独立変数の指定** (Which variable do you want to consider as independent variable?)　独立変数、すなわち x 軸にプロットする変数の列を選択する。	該当項目の番号を選択する
4	**反応データの指定** (Give number(s) of the response(s) you want to analyse)	該当項目の番号を選択する
5	**分散値 (SD 又は SE) の指定** (要約データの場合) (Give number(s) of the variation statistic associated to these endpoints)	(要約データの場合)該当項目の番号を選択する
6	**群サイズの指定** (要約データの場合) (Give number(s) of the groups size associated to these endpoints)	(要約データの場合)該当項目の番号を選択する
7	**共変量の指定** (Give number of factor serving as potential covariate (e.g. sex))	(共変量がある場合は) 該当項目の番号を選択する
8	**BMR の入力** (Give value for CES (always positive))	任意の値を入力（5%であれば 0.05）

表 3.2-4　二値データにおける基本的な解析条件

番号	設定項目	選択/入力項目
1	**データタイプの選択** (What type of response data do you want to consider?)	1: continuous, individual data、2: binary、3: ordinal、4: quantal、5: clustered continuous, individual data、6: clustered quantal、7: continuous, summary data、 8: clustered continuous, summary data、9: quantal, CxT、10: other
2	**数理モデルの選択** (Do you want to fit a set of models, or choose a single model?) フィッティングさせる単一又は一連の数理モデルを選択できる。選択肢 1 (単一モデル) を選択することで、選択肢 2、4、5 と比べユーザーが解析を行う際の自由度と選択肢が増す。一連の数理モデルを一度にフィッティングさせる自動解析の方が速く作業できる。選択肢のうち 3 を選択すると、各種条件（表 3.2-6 参照）の変更ができる。	1: single model、2: set of models、3: change settings first
3	**独立変数の指定** (Which variable do you want to consider as independent variable?) 独立変数、すなわち x 軸にプロットする変数の列を選択する。	該当項目の番号を選択する
4	**反応データの指定** (Which response(s) you want to analyse)	該当項目の番号を選択する
5	**群サイズの指定** (Enter column number(s) with the associated sample sizes)	該当項目の番号を選択する
6	**共変量の指定** (Give number of factor serving as potential covariate (e.g. sex))	(共変量がある場合は) 該当項目の番号を選択する
7	**BMR のタイプの指定** (What type of Benchmark response do you want to consider?) 過剰リスクの場合は 3 を選択	1: ED50、2: Additional risk, i.e. P[BMD] - P[0]、3: Extra risk, i.e. (P[BMD]-P[0])/(1-P[0])、4: CED for latent variable

番号	設定項目	選択/入力項目
8	**BMR** の入力 (Give value for the BMR)	任意の値を入力 (10%であれば 0.1)
9	モデル平均化の実施/非実施の指定 (Do you want to calculate the BMD confidence interval by model averaging?)	1: no、2: yes
10	潜在変数モデルに関する指定（**9** でモデル平均化の実施を選択した場合） (Do you want to include the latent variable models in model averaging?)	1: no、2: yes
11	ブートストラップ反復回数の設定（**9** でモデル平均化の実施を選択した場合） (give number of bootstrap runs for calculating BMD confidence interval (e.g. 200 or more)) モデル平均化で行われるブートストラップの反復回数を入力する。マニュアルでは、通常、200 回で合理的かつ正確な回答が得られるとされ、また、ブートストラップの反復回数が多いほど計算結果はより正確になるが、計算にはより多くの時間がかかるとされている。 一連の数理モデルを用いて通常通りフィッティングが行われた後、モデル平均化が開始される。平均化で得られたデータの新しいプロットが表示され、同プロットに、各ブートストラップによって得られたモデル平均化曲線を表す新しい破線で書かれた曲線が追加される。	任意の値を入力

表 3.2-5　連続値データの Change Settings から設定できる項目

番号	設定項目	内容
1	independent variable	ユーザーが独立変数、すなわち x 軸にプロットする変数とする変数の列を選択することができる。
2	response variable	解析の対象となる反応変数の列を選択する。
4	outliers	あるデータポイントが外れ値であり、解析に用いることが適切でない場合に、当該データポイントを除外することができる。
5	denominator	同じデータセットに含まれる体重のデータで除す必要のある臓器重量のデータがある場合などに利用できる。
6	detection limit	当該検出限界値を入力する。
7	small value	全ての反応データに追加する小さい値を入力することができる。
9	distinct plotting	データを分割して表示することができる。例として、データセットに性別のデータがある場合、これらのデータが体系的に異なるか否かを目視で判断するため、両者を別々にプロットできる。その他、異なるばく露期間や異なる化合物の間等でデータ比較する場合もある。
10	scaling factor on X	X 軸の数値 (用量) が入力した値 (scaling factor) で除される。例えば、1 を入力した場合、元の値が使用される。解析における数値上の問題を防ぐ上で、用量の値をスケール調整することが有用な場合がある。例えば、最大用量が 5,000 (mg/kg) であった場合、scaling factor1,000 を入力することで、解析時の最大用量の値を 5 (g/kg) とすることができる。
11	subgroup	データ分析に当たって制限（解析対照から除外）したい要因（subgroup）を指定できる。
13	covariate with respect to parameter a	モデルのフィッティングに当たりパラメータ a の共変量として含めたい因子について、表示されたリスト中の対応する番号を入力する。
14	covariate with respect to parameter b	モデルのフィッティングに当たりパラメータ b の共変量として含めたい因子について、表示されたリスト中の対応する番号を入力する。指定した共変量の各レベルに対して、異なるパラメータ b の値が推定される。
15	covariate with respect	モデルのフィッティングに当たりパラメータ ver の共変量として含めたい因子について、表示されたリスト中の対

	to parameter var	応する番号を入力する。指定した共変量の各レベルに対して、異なる分散パラメータの値が推定される。分散パラメータ（var）は、実測値の（自然）対数に関連する。
16	covariate with respect to parameter d*	パラメータ d が異なるサブグループ間で等しいと仮定可能かどうか調べる際に使用する場合がある。その場合は、用量反応（曲線）が、対数スケールで平行の位置関係にある。
17	covariate with respect to parameter c	特定の用途、例えば定数 c の仮定が合理的であるかを確認する等の場合において、パラメータ c に対する共変量を設定することが有用な場合がある。
18	pairwise comparison of dose groups	yes と回答することで、対照群に対する用量群のペアワイズ比較の結果が得られる。
19	other transformation	デフォルト (対数変換) から、無変換又は平方根変換を選択できる。
20	critical difference in AICs	EFSA で推奨する AIC の有意な差 (2 単位) を変更できる。
21	value for right censoring	反応変数の上限値としての右側打ち切り値を設定できる。
22	type of regression	x と y の両方でエラーが起こるデータには、二乗和よりも積和を最小化する方がより適切である。
23	relax fit conditions	多数のパラメータを推定する場合や、複数のサブグループの信頼区間を計算する場合、計算時間が長くなることがある。最適化アルゴリズム (R 関数の nlminb) において、より緩やかな条件を選択することにより、計算時間を削減できる。ただし、状況によってはフィッティング結果が不正確なものとなる場合がある。
24	time variable	2 つの独立変数 (CxT) とフィッティングさせたい場合、時間を表す列の番号を入力する。

*PROAST マニュアル (Menu バージョン) には通常、リスク評価者は変更しないと記載されている。

表 3.2-6　二値データの Change Settings から設定できる項目*

番号	設定項目	内容
15	covariate with respect to parameter theta	パラメータ theta に関して共変量として働く因子の数を与える。
19	finite slope at zero	原点付近の傾きを制限し BMDS 計算結果と比較することが有用な場合がある。
20	critical difference in AICs among models	AIC の有意な差として 2 単位が推奨される。ただし、デフォルトの AIC 差がフィッティングさせたモデルにおいて許容できない場合は変更 (増加) できる。

*連続値データと共通の項目については表 3.2-5 を参照

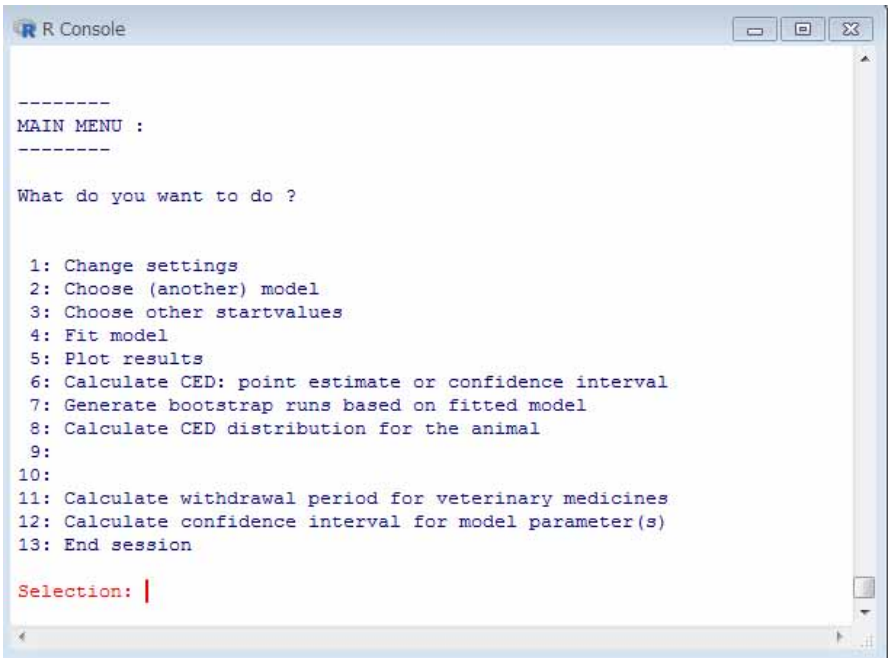


図 3.2-8　Main menu の解析の設定画面

表 3.2-7　Main menu における操作内容

操作*	内容
Option 1: Change settings	設定変更 (表 3.2-5 及び表 3.2-6 参照)
Option 2: Choose (another) model	フィッティングに用いる特定の数理モデルを指定する。 【連続値データ】 モデル 38 又は 39 を選択することにより、自動モデルフィッティングに進む 【二値データ、クラスター化したデータの場合】 フィッティングさせるモデルとして、単一の古典的モデル (classical models)もしくは単一の潜在変数モデル、又は一連のモデルセットを選択する (選択肢 3 は欠番)。 1: classical models 2: latent variable models 4: set of models 【順序カテゴリーデータ】 連続値データのモデルのリストからのサブセットである、31 モデルのリストが表示される
Option 3: Choose other startvalues	パラメータ推計を行う際の開始値を変更する。 フィッティングさせた数理モデルとデータとの一致が良好で

操作*	内容
	ない場合、推計開始時のパラメータの値を変更することにより、より良好な値が得られる可能性がある。
Option 4: Fit model	数理モデルのフィッティングを行う。
Option 5: Plot results	用量反応データとフィッティングさせた数理モデルのプロットの表示形式を変更する。設定項目は以下の通り。 【連続値データの場合】 1: x-y、2: log(x)-y、3: x-log(y)、4: log(x)-log(y)、5: sqrt(x)-y、6: sqrt(x)-log(y)、7: x-sqrt(y) 8: log(x)-sqrt(y)、9: qq-plot of residuals、10: residuals against dose、11: residuals against model prediction、12: change plot limits、13: shift datapoints at same dose、14: use current plot for identifying points、15: back to main menu 【二値データの場合】 1: y vs. x、2: y vs. log(x)、3: log(y) vs. x、4: log(y) vs. log(x)、5: arcsin.sqrt(y) vs. x、6: arcsin.sqrt(y) vs. log(x)、7: latent variable vs. x、8: latent variable vs. log-x、9: log(latent variable) vs. x、10: log(latent variable) vs. log-x、11: qqplot for betadistr、12: change plot limits、13: back to main menu
Option 6: Calculate BMD/CED point estimate or confidence interval	プロファイル尤度法に基づき BMD/CED 値の推定又は信頼区間の算出を行う。CED 及び CES (critical effect size) は以下の式により定義される。 $CES = \frac{f(CED) - f(0)}{f(0)} = \frac{f(CED)}{f(0)} - 1$
Option 7: Generate bootstrap runs based on fitted model	フィッティングした数理モデルに基づいてブートストラップを実行する。CED 周辺の不確実性分布を作成し定量化できる。CED 周辺の信頼区間を評価するための Option 6 (プロファイル尤度法) の代替法としても用いることができる。 クラスター化した連続値データの場合、CED の信頼区間をプロファイル尤度法で計算できないため、ブートストラップ法で評価する。
Option 8: Calculate CED distribution for the animal	ブートストラップ解析を実行後、ブートストラップ CED のヒストグラムを作成し、そのヒストグラムに基づいて動物試験の CED の信頼区間を推定できる。
Option 11: Calculate withdrawal period for	食肉中の動物用医薬品濃度の時間当たりの濃度減少率から休薬期間を計算するなど、特殊な用途で利用する場合の設定。

操作*	内容
veterinary medicines	その目的は、あるマトリクス中の濃度が、試験動物の 95%で参照値を下回る時点を、その 95%信頼区間とともに推計すること。
Option 12: Calculate confidence interval for model parameter(s)	プロファイル尤度法に基づき、数理モデルのフィッティングにおいて推定されたパラメータの信頼区間を計算できる。

* Option 9: Calculate CED distribution for sens.及び Option 10: Calculate ES distribution for sens. human population については現バージョンの解析画面では表示されない

4) モデリング結果の出力、記録

解析の設定が完了すると、解析が開始され計算過程と結果が表示される。モデリングの結果、出力される内容の例を、入力したデータセットと共に以下に示す。また、連続値データ及び二値データのプロットの右側に表示される凡例を表 3.2-8 及び表 3.2-9 に示す。

4)-1 連続値データ (要約) を用いた解析例 (性別を共変量とした例)

- i) データセット
- ii) 表示されるプロット
- iii) Console ウィンドウに表示される結果

4)-2 連続値データ (個体) を用いた解析例

- i) データセット
- ii) 表示されるプロット
- iii) Console ウィンドウに表示される結果

4)-3 二値データを用いた解析例

- i) データセット
- ii) 表示されるプロット
- iii) Console ウィンドウに表示される結果
- iv) モデル平均化に基づくブートストラップ曲線

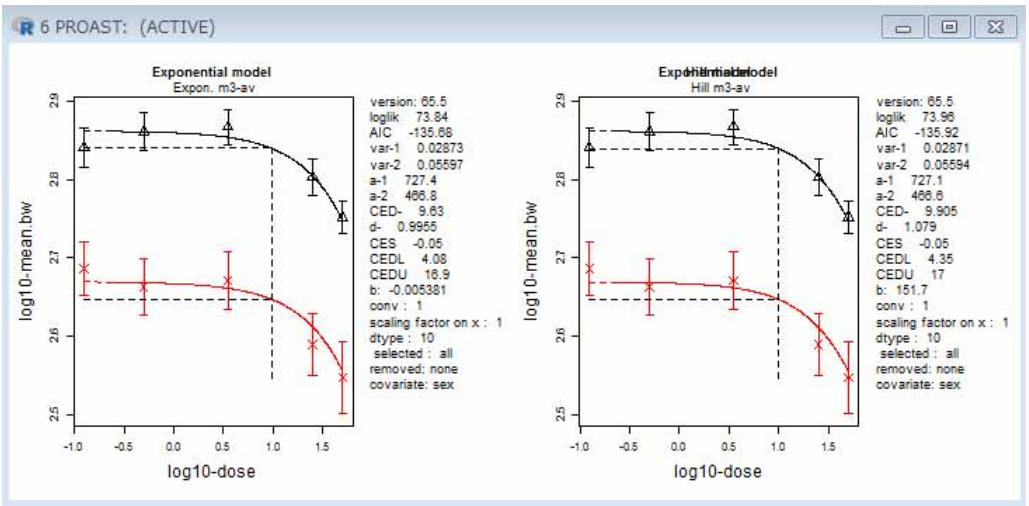
4)-1 連続値データ (要約) を用いた解析例

i) データセット

ファイル(F)　編集(E)　書式(O)　表示(V)　ヘルプ(H)

dose	sex	mean.bw	sd.bw	n
0	1	704	124.7	33
0.5	1	739	140.5	35
3.5	1	742	97.7	40
25	1	646	119.4	41
50	1	572	97	49
0	2	496	105.7	37
0.5	2	477	132.6	33
3.5	2	480	106.8	32
25	2	402	106.8	27
50	2	361	81.1	20

ii) 表示されるプロット



iii) Console ウィンドウに表示される結果

```
response:  mean.bw
ANALYSIS WITH EXPONENTIAL MODELS
model     converged       npar     loglik          aic
 full            1        11       67.63   -113.26
 full-v          1        12       77.28   -130.56
 null-v          1        3        -56.45          118.9
 null-av         1        4        29.21   -50.42
 Expon. m3-v     1        5        -32.07          74.14
 Expon. m3-av    1        6        73.84   -135.68
 Expon. m3-abv   1        7        74.46   -134.92
 Expon. m5-av    1        7        74.79   -135.58
 Expon. m5-abv   1        8        75.42   -134.84
---------------------------------------------------
Best model with covariates is: Expon. m3-av
However Expon. m5-av is a reasonable model as well
and so is Expon. m3-abv
and so is Expon. m5-abv
---------------------------------------------------

-----------------------------------------------------
selected model: Expon. m3-av
 estimate for  var-1 :  0.02873
  estimate for  var-2 :  0.05597
  estimate for  a-1 :  727.4
  estimate for  a-2 :  466.8
  estimate for  CED- :  9.63
  estimate for  d- :  0.9955
-----------------------------------------------------

 calculating confidence intervals ....

the CED (in orig. units) and the 90 % confidence interval is:
 9.6
 4.08
 16.9

response:  mean.bw
ANALYSIS WITH HILL MODELS
model     converged       npar     loglik          aic
 Hill m3-av      1        6        73.96   -135.92
 Hill m3-abv     1        7        74.61   -135.22
 Hill m5-av      1        7        74.78   -135.56
 Hill m5-abv     1        8        75.54   -135.08
---------------------------------------------------
Best model with covariates is: Hill m3-av
However Hill m5-av is a reasonable model as well
and so is Hill m3-abv
and so is Hill m5-abv
---------------------------------------------------

-----------------------------------------------------
selected model: Hill m3-av
 estimate for  var-1 :  0.02871
  estimate for  var-2 :  0.05594
  estimate for  a-1 :  727.1
  estimate for  a-2 :  466.6
  estimate for  CED- :  9.905
  estimate for  d- :  1.079
-----------------------------------------------------

 calculating confidence intervals ....

the CED (in orig. units) and the 90 % confidence interval is:
 9.9
 4.35
 17


 -----------  CES =  0.05  --------------------------------
The lowest BMDL and highest BMDU from exponential and Hill models are:
  covar.txt bmdl.lowest bmdu.highest
1       all        4.08           17
 ---------------------------------------------------- end of analysis for response:  mean.bw
```

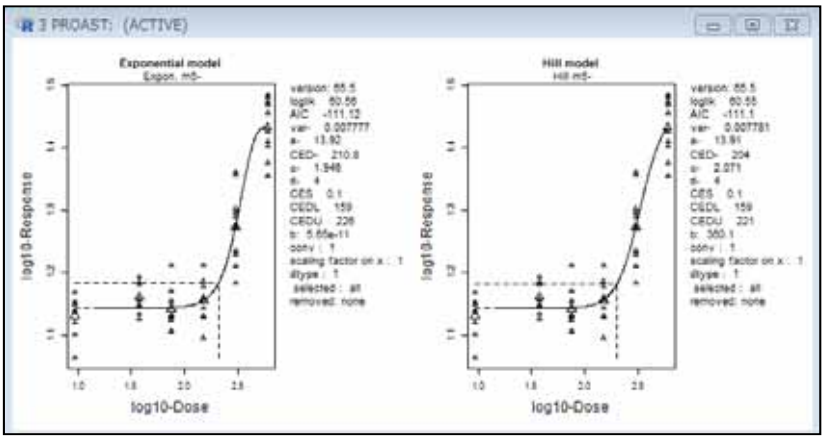
4)-2 連続値データ (個体) の表示例

i) データセット (一部分)

ファイル(F)　編集(E)

Dose	Response
0	14.71
0	13.19
0	14.11
0	14.12
0	13.8
0	11.59
0	12.65
0	14.21
0	13.69
0	13.21
37.5	15.22
37.5	14.07
37.5	15.18
37.5	15.31
37.5	15.61
37.5	14.38
37.5	14.1
37.5	13.38
37.5	14.08
37.5	13.6
75	14.03
75	12.75
75	13.34
75	12.81
75	13.35
75	16.32
75	14.8
75	14.3
75	13.53
75	13.54
150	15.36
150	15.06
150	14.35

ii) 表示されるプロット



iii)　Console ウィンドウに表示される結果

```
response:  Response
ANALYSIS WITH EXPONENTIAL MODELS
model     converged       npar    loglik          aic
 full             1       7       62.22   -110.44
 null-   1        2       -5.03   14.06
 Expon. m3-       1       4       56.94   -105.88
 Expon. m5-       1       5       60.56   -111.12
-------------------------------------------------------
selected model: Expon. m5-
 estimate for  var- :  0.007777
  estimate for  a- :  13.92
  estimate for  CED- :  176
  estimate for  c- :  1.946
  estimate for  d- :  4
-------------------------------------------------------

The Grubb outlier for this sample size is: 3.19966

No outliers detected

 calculating confidence intervals ....

the CED (in orig. units) and the 90 % confidence interval is:
 180
 116
 188


response:  Response
ANALYSIS WITH HILL MODELS
model     converged       npar    loglik          aic
 Hill m3-         1       4       56.07   -104.14
 Hill m5-         1       5       60.55   -111.1
-------------------------------------------------------
selected model: Hill m5-
 estimate for  var- :  0.007781
  estimate for  a- :  13.91
  estimate for  CED- :  169.4
  estimate for  c- :  2.071
  estimate for  d- :  4
-------------------------------------------------------

The Grubb outlier for this sample size is: 3.19966

No outliers detected

 calculating confidence intervals ....

the CED (in orig. units) and the 90 % confidence interval is:
 170
 117
 184


 -----------  CES =  0.05  ---------------------------------
The lowest BMDL and highest BMDU from exponential and Hill models are:

  covar.txt bmdl.lowest bmdu.highest
1       all         116          188
 ------------------------------------------------------ end of analysis for response:  Response
```

表 3.2-8　連続値データで出力されるプロットの凡例

凡例	内容
version	PROAST のバージョン
loglik	最良のフィッティングに関連する対数尤度の値
AIC	赤池情報量規準
var	標準偏差/標準誤差
a-f and a-m	雌雄のパラメータ a の値 (MLEs)
CED	臨界影響用量 (Critical Effect Dose)
d	近似モデルによる d (steepness) の値
CES	臨界影響サイズ (Critical Effect Size)
CEDL and CEDU	プロファイル尤度法により得られた CED の信頼区間の下限値 ブートストラップ法により得られた値の場合は CEDL.bt 及び CEDU.bt と表示される
b	近似モデルに従った b (background response (バックグラウンド反応)) の値
conv	近似アルゴリズムは収束したか？ 1 の場合は Yes、0 の場合は No
scaling factor on x	change settings 後に入力された値。データが分割された要因
dtype	データタイプ
selected	選択されたプラスレベル因子 (Factor plus levels that were selected)
removed	除外されたプラスレベル因子 (Factor plus levels that were removed)
covariate	共変量として使用された因子
fact1	パラメータ a に対する共変量として使用された因子
fact2	パラメータ b に対する共変量として使用された因子
fact3	パラメータ c に対する共変量として使用された因子
fact4	パラメータ d に対する共変量として使用された因子
fact5	パラメータ e に対する共変量として使用された因子

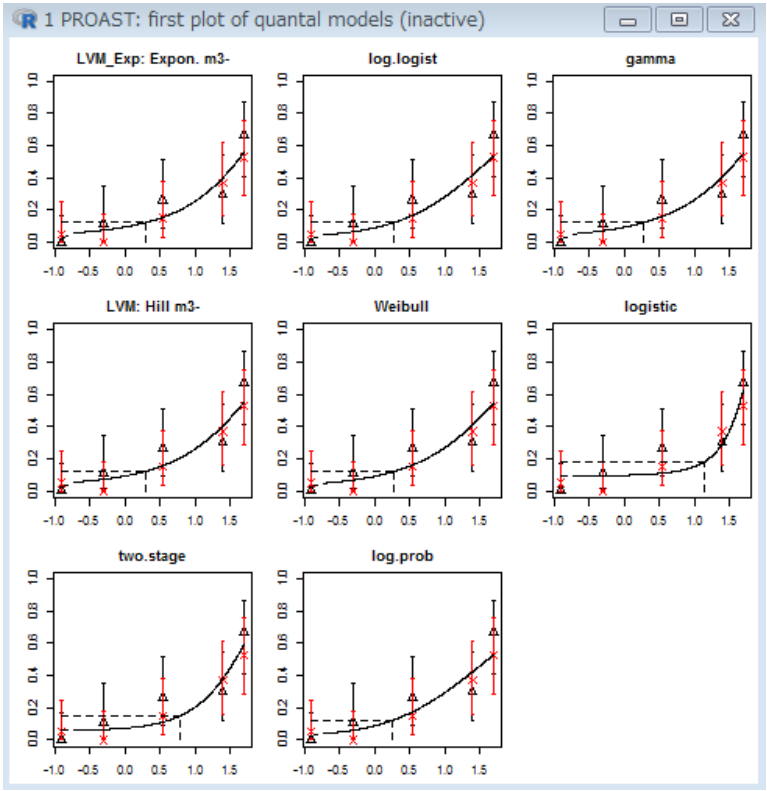
4)-3 二値データの表示例

i) データセット

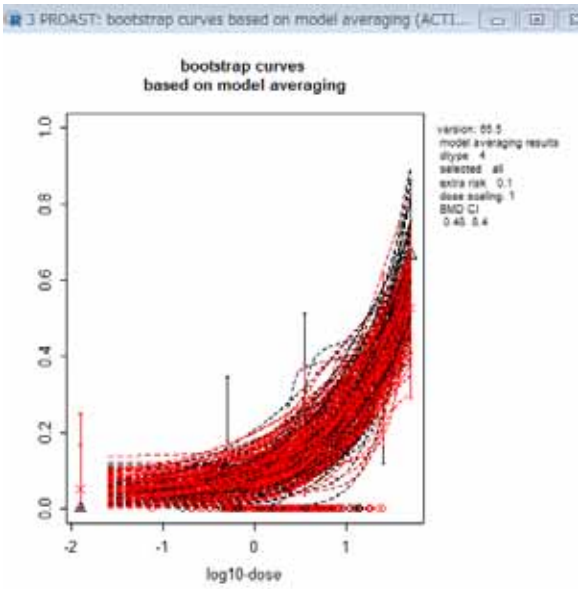
ファイル(F)　編集(E)　書式(O)　表示(V)　ヘルプ(H)

dose	sex	response	group.size
0	1	0	20
0.5	1	2	18
3.5	1	5	19
25	1	6	20
50	1	12	18
0	2	1	20
0.5	2	0	19
3.5	2	3	20
25	2	7	19
50	2	10	19

ii) 表示されるプロット



iii) モデル平均化に基づくブートストラップ曲線



iv) Console ウィンドウの結果表示 Console ウィンドウに表示される結果

```
Calculating BMD confidence intervals for all models

              model No.par  loglik    AIC accepted  BMDL BMDU  BMD sens.subgr conv
1              null      1 -105.72 213.44              NA   NA   NA
2              full     10  -78.98 177.96              NA   NA   NA
3         two.stage      3  -84.20 174.40        no    NA   NA  6.3              yes
4       log.logist       3  -83.14 172.28       yes 0.438 6.72  1.9              yes
5           Weibull      3  -83.03 172.06       yes 0.401 7.21  1.9              yes
6          log.prob      3  -83.15 172.30       yes 0.484 5.92  1.8              yes
7             gamma      3  -82.98 171.96       yes 0.352 7.00  1.9              yes
8          logistic      2  -85.66 175.32        no    NA   NA 14.0              yes
9  LVM: Expon. m3-       3  -83.01 172.02       yes 0.670 8.34  2.0           1  yes
10   LVM: Hill m3-       3  -83.05 172.10       yes 0.491 8.16  2.0           1  yes
BMR: 0.1 extra risk
constraint: no
no litter effects
critical AIC value: 2
PROAST version: 65.5


     lowest BMDL     highest BMDU     nr. of models
          0.352            8.34                 6


 The weights used in model averaging are:
        model weight
1  two.stage 0.0489
2 log.logist 0.1413
3    Weibull 0.1577
4   log.prob 0.1399
5      gamma 0.1658
6   logistic 0.0309
7        EXP 0.1609
8       HILL 0.1546
Calculating model-average BMD confidence interval, this may take some time ...
run   1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21$
ATTENTION: no convergence was reached for the following models:
 two.stage
  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  8$
ATTENTION: no convergence was reached for the following models:
 two.stage
  123  124  125  126  127  128  129  130  131  132  133  134  135  136  137  13$
ATTENTION: no convergence was reached for the following models:
 two.stage
  157  158  159  160  161  162  163  164  165  166  167  168  169  170  171  17$

duration of bootstrap calculations:
[1] "Wed Mar 13 17:01:59 2019"
[1] "Wed Mar 13 17:12:01 2019"

The model-average BMD confidence interval is:
  subgroup Vlower.ma Vupper.ma
1        1     0.476      8.36
2        2     0.486      9.05
```

表 **3.2-9**　二値データで出力されるプロットの凡例

凡例	内容
version	PROAST のバージョン
model	A: 古典的モデル (classical models) B: 潜在変数モデル モデル平均化の結果
dtype	データタイプ
ces.ans	BMR のタイプ (50%有効量 (ED50)、追加リスク (additional risk) 過剰リスク (extra risk))
dose scaling	用量単位
BMD	ベンチマークドーズ
CI	信頼区間

3.3 Benchmark Dose Modelling (Efsa)

3.3.1 名称及びバージョン

BMD 分析用の EFSA Web ツール

(基礎となる計算は R-package PROAST ver. 66.24 で実行 (2019 年 3 月 12 日時点))

3.3.2 Benchmark Dose Modelling の概要

(1) 概要

Benchmark Dose Modelling は、EFSA、RIVM と OpenAnalytics の共同プロジェクトにより作成され 2017 年から公開された無償の web アプリケーションである (https://shiny-efsa.openanalytics.eu/app/bmd)。利用に際しては、利用者の EFSA アカウントを作成する必要がある。アカウントは EFSA ShinyProxy に必要事項を入力、登録することにより作成できる。

(2) 主要な機能

(2)-1 利用可能なデータ及び使用される数理モデル

主要な機能としては、連続値データ (要約、個体)、二値データ、2 進データ、順序カテゴリーデータを対象とした BMDL、BMDU、BMD の算出であり、BMDL の計算は PROAST を使用している (3.2 PROAST 参照)。モデル平均化については、ver.65.7 では MADr-BMD (Wheeler & Bailer (2007)) と同様に二値データに対してのみ対応していたが、現時点 (2019 年 3 月) の ver. 66.24 では、連続値データ、二値データ及び 2 進データに対応している (順序カテゴリーデータは未対応)。

収載されている数理モデルを表 3.3-1 に示す。

表 3.3-1　**Benchmark Dose Modelling** に収載されている数理モデル

モデル	パラメータ数	式	データ			
			連続値	二値	2進	順序カテゴリー
Null	1	$y = a$	○	○	○	○
Full	no. of groups	$y = \text{group mean}$	○	○	○	○
Logistic	2	$y = \frac{1}{1 + \exp(-a - bx)}$		○	○	
Probit	2	$y = pnorm((x - a) \cdot b)$		○	○	
Log-logistic	3	$y = a + \frac{1 - a}{1 + \exp\left(c \cdot \log\left(\frac{b}{x}\right)\right)}$		○	○	
Log-probit	3	$y = a + (1 - a) \cdot pnorm\left(c \cdot \log\left(\frac{x}{b}\right)\right)$		○	○	
Weibull	3	$y = a + (1 - a)\left(1 - \exp\left(-\left(\frac{x}{b}\right)^c\right)\right)$		○	○	
Gamma	3	$y = pgamma(bx; c)$		○	○	
Two-stage	3	$y = a + (1 - a)\left(1 - \exp\left(-\frac{x}{b} - c\left(\frac{x}{b}\right)^2\right)\right)$		○	○	
Exp model 3	3	$y = a \cdot \exp(bx^d)$	○	○	○	○
Exp model 5	4	$y = a \cdot (c - (c - 1)\exp(-bx^d))$	○	○	○	○
Hill model 3	3	$y = a \cdot \left(1 - \frac{x^d}{b^d + x^d}\right)$	○	○	○	○

モデル	パラメータ数	式	データ			
			連続値	二値	2進	順序カテゴリー
Hill model 5	4	$y = a \cdot \left(1 + (c-1)\frac{x^d}{b^d + x^d}\right)$	○	○	○	○
Inverse Exponential	4	$y = a \cdot (1 + (c-1)\exp(-bx^{-d}))$	○			○
Log-Normal Family	4	$y = a \cdot (1 + (c-1)\Phi(\ln b + d\ln x))$	○			○

(2)-2　BMR の設定

数理モデルのフィッティング及びパラメータ推定には頻度論に基づく統計学的アプローチ (すなわち最尤推定) が使用される。全ての数理モデルで BMD を計算後、AIC に基づいて AIC 間の 2 単位差を臨界値とし、すなわち最小の AIC に 2 を加えた数値以下の AIC を示す数理モデルが選択される (AIC ≤ AICMin + 2 の数理モデル)。AIC の臨界値はユーザーによる変更が可能であり、選択された数理モデルについてのみ BMDL、BMDU が算出される。なお、モデル平均化の場合は、全ての数理モデルが用いられ、AIC を基にした重み付けが行われる。連続値データについては同一用量に対する反応量分布を対数正規分布であると仮定している。

3.3.3　利用方法に関する情報

(1)　セットアップ方法

操作方法に関する資料（操作マニュアル、ガイダンス等）として EFSA から公表されている「Manual for BMD Modeling」がある。また、本 web アプリケーション操作画面の右上にある About ボタンを押すことにより同マニュアルが表示され、印刷、ダウンロードすることができる。[56]

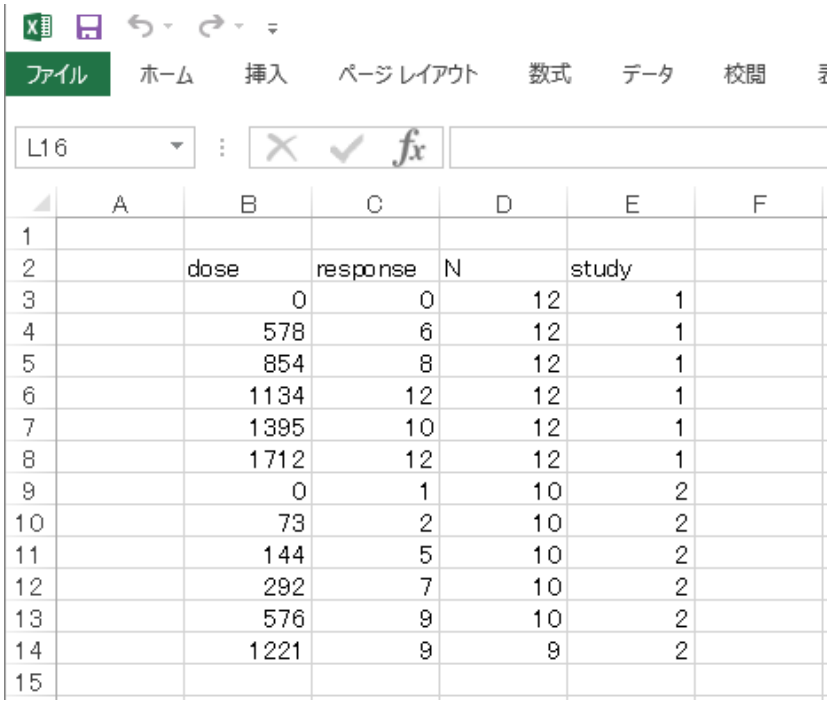
[56] https://shiny-efsa.openanalytics.eu/app/bmd

(2)　操作の流れ

操作の概要を以下に示す。

1)　データファイルの作成

エクセルで作成し、保存する際に CSV ファイル又はテキストファイルの形式を選択する (タブ区切り等) (図 3.3-1)。



	A	B	C	D	E	F
1						
2		dose	response	N	study	
3		0	0	12	1	
4		578	6	12	1	
5		854	8	12	1	
6		1134	12	12	1	
7		1395	10	12	1	
8		1712	12	12	1	
9		0	1	10	2	
10		73	2	10	2	
11		144	5	10	2	
12		292	7	10	2	
13		576	9	10	2	
14		1221	9	9	2	
15						

図 3.3-1　入力データの例

2)　Web アプリケーション

タブページを使用して、以下の項目を切り替えることができる (図 3.3-2)。

- Data (データ)：分析に使用するデータの指定
- Fit Models (数理モデルのフィッティング)：PROAST を使用して数理モデルをフィッティングし、可能な場合はモデル平均化を実行
- Advanced Plotting: 最終的なプロットのフォーマットの変更

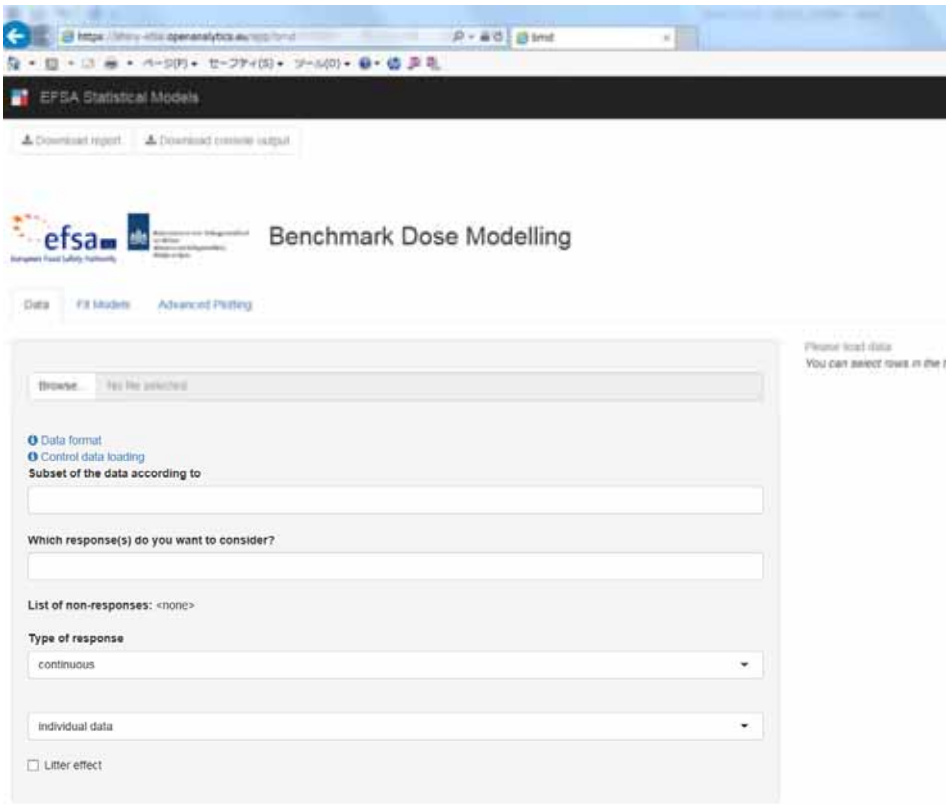


図 3.3-2　Web アプリケーションの開始画面

3)　データファイルの読み込み

Data タブページを開き、Browse ボタンを押し、読み込むデータファイルを指定する (図 3.3-3)。

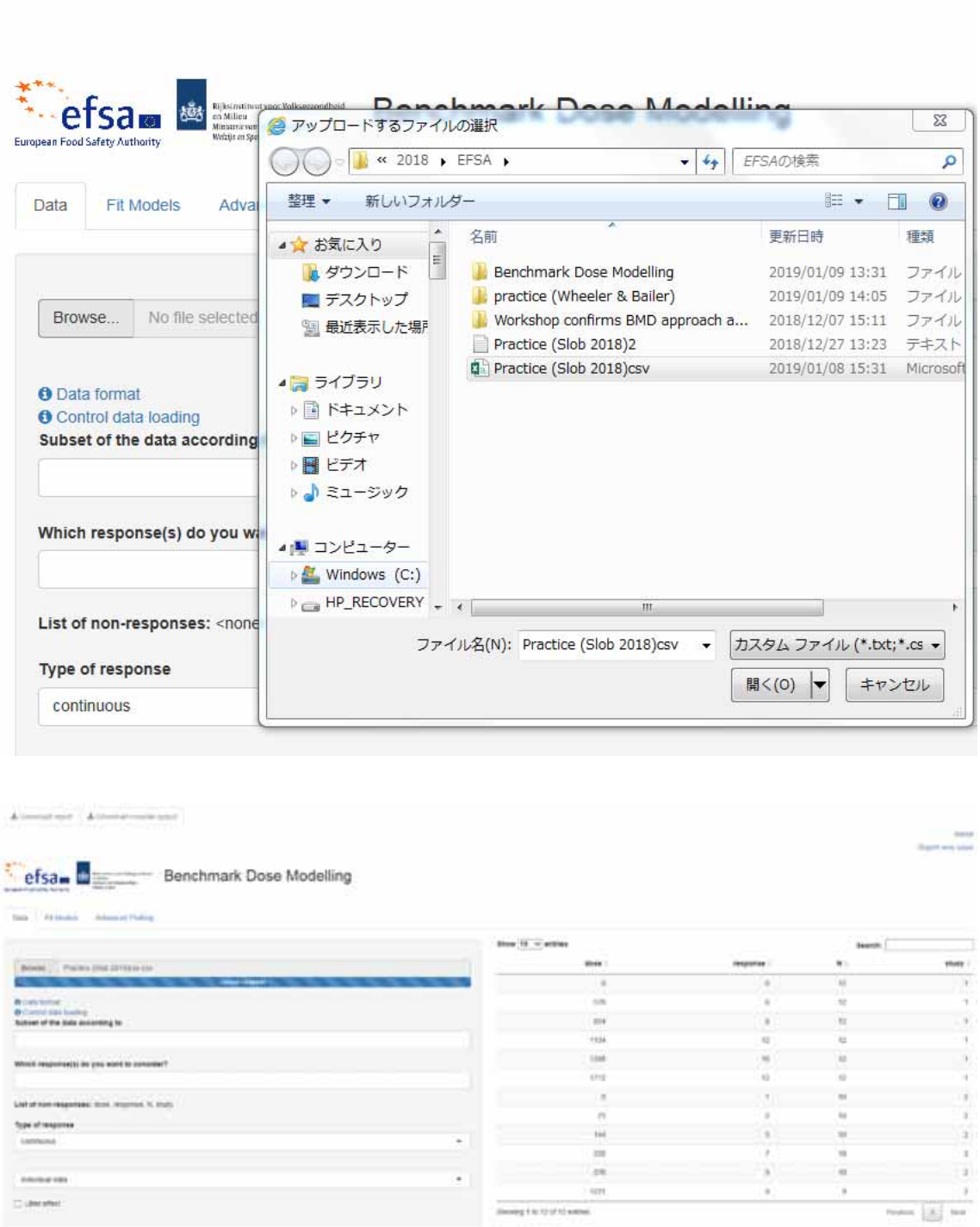


図 3.3-3　データファイルの読み込み画面

4)　データの設定

Data タブページを開き、必要項目を設定する (図 3.3-4)

Subset of the data according to (データのサブセットの設定) では、解析対象のデータセットをサブセットに分割する場合に用いるデータを指定することができる (サブセットを使用しない場合は空欄とする)。全てのデータセットをサブセットに分割して詳細に指定する必要はなく、通常は用量にあたるデータ（列）を指定し、解析に用いる用量を指定する (通常は全ての用量を指定するが、特定の用量で死亡が多くみられた場合等でその用量を解析から除外した方が良いと考えられる場合等では用量の指定を考慮する)。Which response(s) do you want to consider? (解析対象とする反応の設定) では、解析に用いる Response Variable (反応変数) となるデータ (列) を指定する。Type of response（反応変数の設定）では、Response のタイプ (連続値データ (要約又は個体)/二値データ/2 進データ/順序カテゴリーデータ) を指定する

Litter effect （同腹効果の有無の設定）では同腹効果が考えられる場合に指定する。

図 3.3-4　データの設定画面

5)　データ変数の設定

Fit Models タブページを開き、必要項目を入力する。

Data Variables (データの変数) の設定画面で Independent variable (独立変数、例えば、用量)、Response variable(s) (反応変数、例えば、平均体重) 、Sample size (サンプルサイズ、例えば n 数)となるデータ (列) を指定する。なお、連続値データの要約データの場合、Variation statistic (反応変数のばらつきに関する指標) のタイプ (標準偏差/標準誤差)とデータ(列) を選択する。

また、用いるデータファイルに共変量 (従属変数と最も関心のある独立変数との関係に影響を与える二次的な変数) が含まれる場合、例えば、雌雄両方のデータが同一データファイルに含まれる場合等は、必要に応じて Covariate(s) (共変量) とするデータ（列）を指定する。

6) 解析条件の設定

Fit Models タブページを開き、必要項目に入力する。

(ア)二値データ及び 2 進データ

平均化実施の有無 (デフォルト：実施)、ブートストラップの反復回数 (モデル平均化から得られる BMD の信頼区間を計算するためのブートストラップの反復回数、デフォルト：200)、AIC 間の臨界値 (BMD の推定に用いるモデルの選択基準である AIC の最小モデルと比較した AIC 値の差、デフォルト：2)、Benchmark metric (ベンチマーク判定基準) ：ED50／Additional risk (追加リスク) ／Extra risk (過剰リスク(デフォルト))、過剰リスクや追加リスクの BMR (デフォルト：0.1)、BMD 信頼区間の信頼水準 (デフォルト：0.9)) を変更する。

(イ)連続値データ (要約、個体)

平均化実施の有無 (デフォルト：実施)、ブートストラップの反復回数 (デフォルト：200)、AIC 間の臨界値 (デフォルト：2)、臨界影響サイズ (CES：Critical Effect Size) (平均反応量からの変化率を基にした BMR、デフォルト：0.05)、BMD 信頼区間の信頼水準 (デフォルト：0.9)を変更する。

(ウ)順序カテゴリーデータ (2019 年 3 月 12 現在、モデル平均化は未対応)

AIC 最大値 (デフォルト：2)、臨界影響サイズ (CES、デフォルト：0.05)、BMD 信頼区間の信頼水準 (デフォルト：0.9)を変更する。

7) 解析の実行

Fit Models のタブページの左上の Fit Model (s) ボタンを押す (図 3.3-5)。

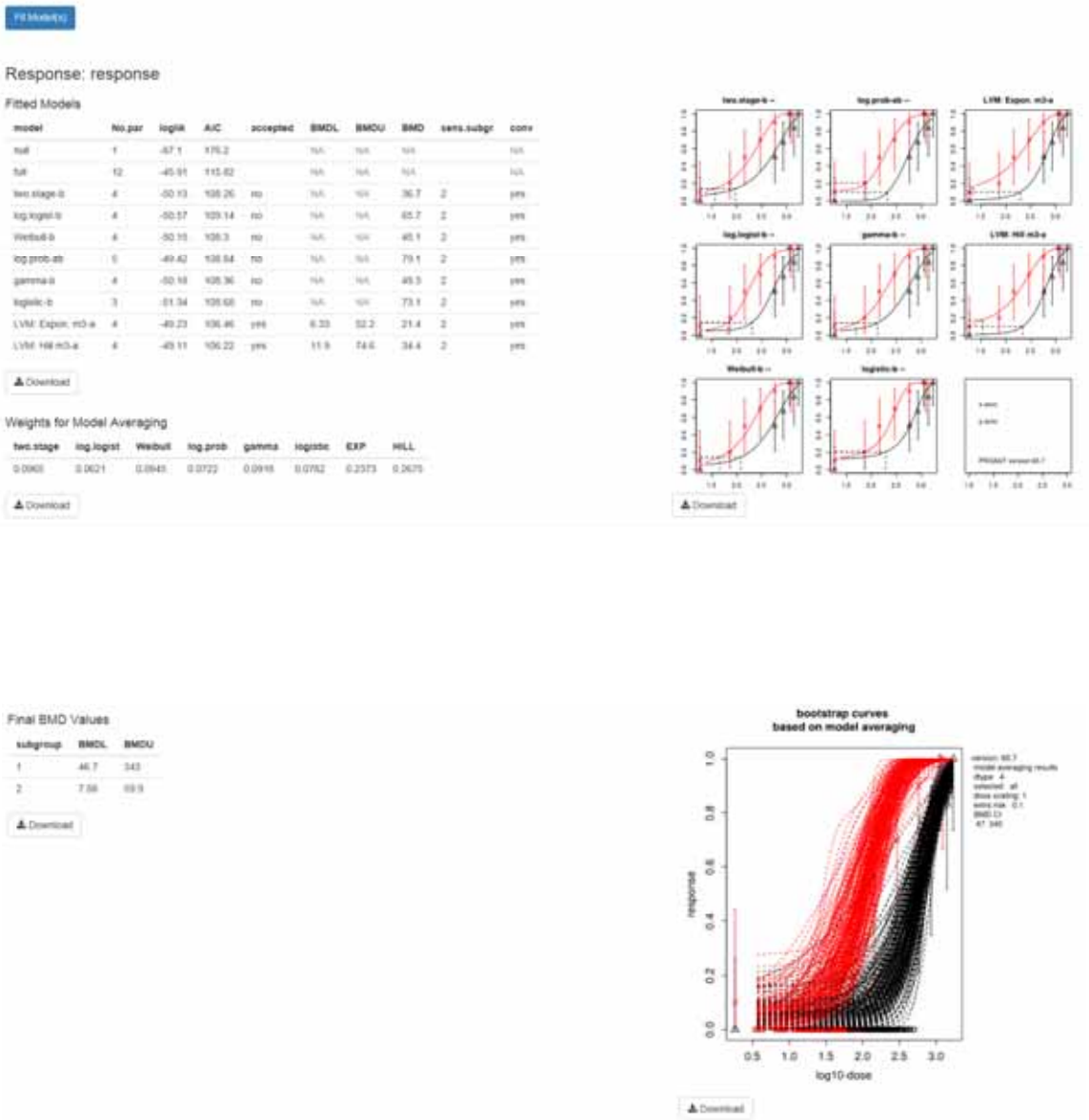


図 3.3-5　解析の実行画面 (二値データ)

Fit Models のタブページの左上の Download report を押すと要約された結果を含む Word 文書が得られる。

高度なプロット

Advanced Plotting のタブページ上で設定、表示を行う。

解析結果から、デフォルトプロットに加えてインタラクティブプロットと非インタラクティブプロットを作成することができる。インタラクティブプロットは対話型のプロット

でグラフにデータが埋め込まれており、利用者の操作に対して応答して、画面上で拡大縮小、マウスを置いた場所のデータの表示、選択したモデルのグラフの表示・非表示等が行える。Advanced Plotting のタブページ画面の左側では、プロットに使用されるデータを決定し、X 軸と Y 軸にプロットする変数の変更、X 軸と Y 軸のスケールの変更（変更無し、log10、平方根）等を行う。画面の右側ではプロットの外観のデフォルト設定をユーザーが変更することができる。

8) 解析結果の出力

解析結果の出力例（連続値データ及び二値データ）を以下に示す。

(ア) 連続値データ (要約)

Benchmark Dose Modeling: Report

European Food Safety Authority (EFSA)

1. Abstract

(Max. 300 words, no paragraph breaks; no tables, footnotes, graphs or figures. Note that the abstract should end with the copyright)

© European Food Safety Authority, 20YY

Key words: (max. seven key words)

Requestor: add requesting party

Question number: EFSA-Q-YYYY-NNNNN

Correspondence: xxx@efsa.europa.eu

Acknowledgements: [Scientific Committee OR EFSA] wishes to thank the following for the support provided to this scientific output: [staff members or others who made a contribution but are not eligible as authors]. The Panel [Scientific Committee OR EFSA] wishes to acknowledge all European competent institutions, Member State bodies and other organisations that provided data for this scientific output.

Suggested citation: EFSA (European Food Safety Authority), Individual authors [add names in the format Surname followed by Initial(s), Surname followed by Initial(s) and Surname followed by Initial(s)], 20YY. Title of the report. EFSA supporting publication 20YY:EN-NNNN. 10 pp. doi:10.2903/sp.efsa.20YY.EN-NNNN

ISSN: 2397-8325

© European Food Safety Authority, 20YY

Reproduction is authorised provided the source is acknowledged.

Reproduction of the images listed below is prohibited and permission must be sought directly from the copyright holder:

Figure 1: © Stockphoto; Figure 5: © WHO

2. Summary

The summary should not include tables, footnotes, graphs or pictures or references.

Table of Contents

Abstract

Summary

1. Data Description

2. Selection of the BMR

3. Software Used

4. Specification of Deviations from Default Assumptions

5. Results

6. Advanced Plots

7. Conclusions

Appendix

References

3. Data Description

The endpoint to be analyzed is: mean.bw.

Subset of the data is taken for dose, retaining value(s) 0, 0.5, 3.5, 25, 50.

Data used for analysis:

dose	mean.bw	sd.bw	n.bw	sex
0.0	704	124.7	33	1
0.5	739	140.5	35	1
3.5	742	97.7	40	1
25.0	646	119.4	41	1
50.0	572	97.0	49	1
0.0	496	105.7	37	2
0.5	477	132.6	33	2
3.5	480	106.8	32	2
25.0	402	106.8	27	2
50.0	361	81.1	20	2

Information pertaining to this endpoint.

4. Selection of the BMR

The BMR (benchmark response) used is a 5% change in mean response compared to the controls. The BMD (benchmark dose) is the dose corresponding with the BMR of interest.

A 90% confidence interval around the BMD will be estimated, the lower bound is reported by BMDL and the upper bound by BMDU.

5. Software Used

Results are obtained using the EFSA web-tool for BMD analysis, which uses the R-package PROAST, version 66.24, for the underlying calculations.

6. Specification of Deviations from Default Assumptions

General assumptions

Please motivate in detail assumptions made when deviating from the recommended defaults (e.g. gamma distributional assumption instead of log-normal, heteroscedasticity instead of homoscedasticity).

Dose-response models

Other models than the recommended ones that were fitted should be listed, with the respective description of reasons to include them.

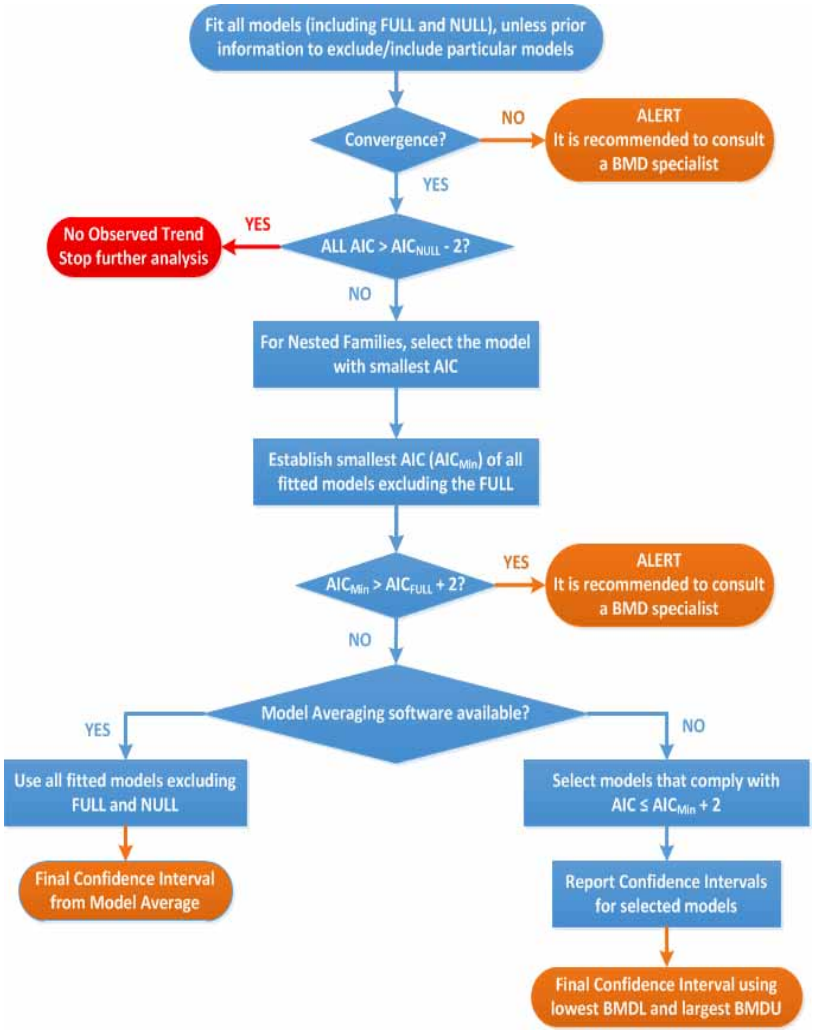
Default set of fitted models:

Model	Number of parameters	Formula
Null	1	$y = a$
Full	no. of groups	$y = \text{group mean}$
Exp model 3	3	$y = a \cdot \exp(bx^d)$
Exp model 5	4	$y = a \cdot (c - (c-1)\exp(-bx^d))$
Hill model 3	3	$y = a \cdot \left(1 - \frac{x^d}{b^d + x^d}\right)$
Hill model 5	4	$y = a \cdot \left(1 + (c-1)\frac{x^d}{b^d + x^d}\right)$
Inverse Exponential	4	$y = a \cdot (1 + (c-1)\exp(-bx^{-d}))$
Log-Normal Family	4	$y = a \cdot (1 + (c-1)\Phi(\ln b + d\ln x))$

As a covariate is included in the analysis, these models will also be fitted assuming that some of the parameters [background response parameter (a), potency parameter (BMD) and/or variance (var)] depend on the subgroup defined by the covariate. Therefore the number of parameters in each model might be larger than indicated in the table above.

Procedure for selection of BMDL

Description of any deviation from the procedure described in the flow chart to obtain the final BMD confidence interval.



Flowchart for selection of BMDL

7. Results

Response variable: mean.bw

Fitted Models

model	converged	loglik	npar	AIC
full model	yes	67.63	11	-113.26
full-v	yes	77.28	12	-130.56
null modelv	yes	-56.45	3	118.90
null model-a-v	yes	29.21	4	-50.42
Expon. m3-v	yes	-32.07	5	74.14

Expon. m3-av	yes	73.84	6	-135.68
Expon. m3-abv	yes	74.46	7	-134.92
Expon. m5-av	yes	74.78	7	-135.56
Expon. m5-abv	yes	75.44	8	-134.88
Hill m3-av	yes	73.84	6	-135.68
Hill m3-abv	yes	74.46	7	-134.92
Hill m5-av	yes	74.78	7	-135.56
Hill m5-abv	yes	75.54	8	-135.08
Inv.Expon. m3-av	yes	74.23	6	-136.46
Inv.Expon. m3-abv	yes	74.90	7	-135.80
Inv.Expon. m5-av	yes	74.79	7	-135.58
Inv.Expon. m5-abv	yes	75.60	8	-135.20
LN m3-av	yes	74.07	6	-136.14
LN m3-abv	yes	74.72	7	-135.44
LN m5-av	yes	74.79	7	-135.58
LN m5-abv	yes	75.59	8	-135.18

Estimated Model Parameters

EXP

estimate for var-1 : 0.02873

estimate for var-2 : 0.05597

estimate for a-1 : 727.4

estimate for a-2 : 466.8

estimate for CED- : 9.637

estimate for d- : 0.9975

HILL

estimate for var-1 : 0.02873

estimate for var-2 : 0.05597

estimate for a-1 : 727.4

estimate for a-2 : 466.8

estimate for CED- : 9.644

estimate for d- : 0.9994

INVEXP

estimate for var-1 : 0.02864

estimate for var-2 : 0.05592

estimate for a-1 : 725.8

estimate for a-2 : 465.8

estimate for CED- : 10.52

estimate for d- : 0.1789

LOGN

estimate for var-1 : 0.02868

estimate for var-2 : 0.05594

estimate for a-1 : 726.4

estimate for a-2 : 466.2

estimate for CED- : 10.17

estimate for d- : 0.3351

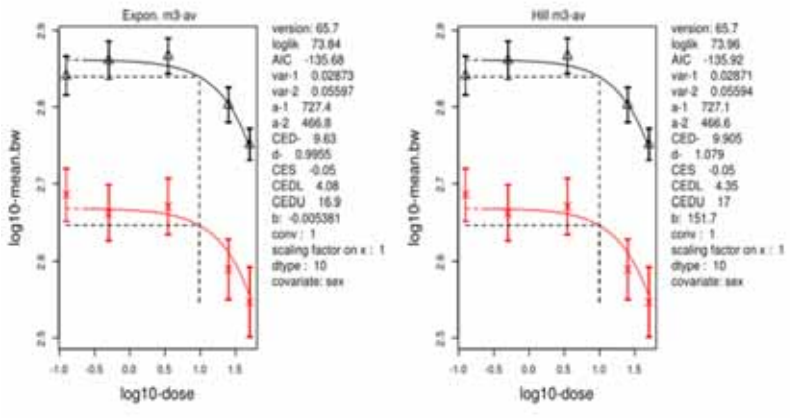
Final BMD Values

model	BMDL	BMDU	BMD
Expon. m3-av	4.08	16.9	9.63
Hill m3-av	4.35	17.0	9.90

Lowest BMDL and highest BMDU Values

subgroup	bmdl.lowest	bmdu.highest
all	4.08	17

Visualization



Plot for response 'mean.bw

Advanced Plots

No results available: If needed, please create advanced plots in the application.

Conclusions

The section should discuss the results for the different endpoints and, if applicable, specific issues such as:

- *Discuss if there were any alerts, and if so, how they well dealt with.*
- *Discuss any particular circumstances, if relevant for the final outcome of the BMD confidence interval.*

The BMD confidence interval of the critical endpoint (and the BMDL selected as reference point) should be reported and discussed.

Appendix

References

(イ) 二値データ

Benchmark Dose Modeling: Report

European Food Safety Authority (EFSA)

1. Abstract

(Max. 300 words, no paragraph breaks; no tables, footnotes, graphs or figures. Note that the abstract should end with the copyright)

© European Food Safety Authority, 20YY

Key words: (max. seven key words)

Requestor: add requesting party

Question number: EFSA-Q-YYYY-NNNNN

Correspondence: xxx@efsa.europa.eu

Acknowledgements: [Scientific Committee OR EFSA] wishes to thank the following for the support provided to this scientific output: [staff members or others who made a contribution but are not eligible as authors]. The Panel [Scientific Committee OR EFSA] wishes to acknowledge all European competent institutions, Member State bodies and other organisations that provided data for this scientific output.

Suggested citation: EFSA (European Food Safety Authority), Individual authors [add names in the format Surname followed by Initial(s), Surname followed by Initial(s) and Surname followed by Initial(s)], 20YY. Title of the report. EFSA supporting publication 20YY:EN-NNNN. 10 pp. doi:10.2903/sp.efsa.20YY.EN-NNNN

ISSN: 2397-8325

© European Food Safety Authority, 20YY

Reproduction is authorised provided the source is acknowledged.

Reproduction of the images listed below is prohibited and permission must be sought directly from the copyright holder:

Figure 1: © Stockphoto; Figure 5: © WHO

2. Summary

The summary should not include tables, footnotes, graphs or pictures or references.

3. Table of Contents

Abstract

Summary

1. Data Description

2. Selection of the BMR

3. Software Used

4. Specification of Deviations from Default Assumptions

5. Results

6. Advanced Plots

7. Conclusions

Appendix

References

3. Data Description

The endpoint to be analyzed is: response.

Subset of the data is taken for dose, retaining value(s) 0, 578, 854, 1134, 1395, 1712, 73, 144, 292, 576, 1221.

Data used for analysis:

dose	response	N	study
0	0	12	1
578	6	12	1
854	8	12	1
1134	12	12	1
1395	10	12	1
1712	12	12	1
0	1	10	2
73	2	10	2
144	5	10	2
292	7	10	2
576	9	10	2
1221	9	9	2

Information pertaining to this endpoint.

4. Selection of the BMR

The BMR (benchmark response) used is an extra risk of 10% compared to the controls.

When the specified BMR deviates from the default value, the rationale behind the choice made should be described.

The BMD (benchmark dose) is the dose corresponding with the BMR of interest.

A 90% confidence interval around the BMD will be estimated, the lower bound is reported by BMDL and the upper bound by BMDU.

5. Software Used

Results are obtained using the EFSA web-tool for BMD analysis, which uses the R-package PROAST, version 66.24, for the underlying calculations.

6. Specification of Deviations from Default Assumptions

General assumptions

Please motivate in detail assumptions made when deviating from the recommended defaults (e.g. gamma distributional assumption instead of log-normal, heteroscedasticity instead of homoscedasticity).

Dose-response models

Other models than the recommended ones that were fitted should be listed, with the respective description of reasons to include them.

Default set of fitted models:

Model	Number of parameters	Formula
Null	1	$y = a$
Full	no. of groups	$y = \text{group mean}$
Logistic	2	$y = \frac{1}{1 + \exp(-a - bx)}$
Probit	2	$y = pnorm((x - a) \cdot b)$

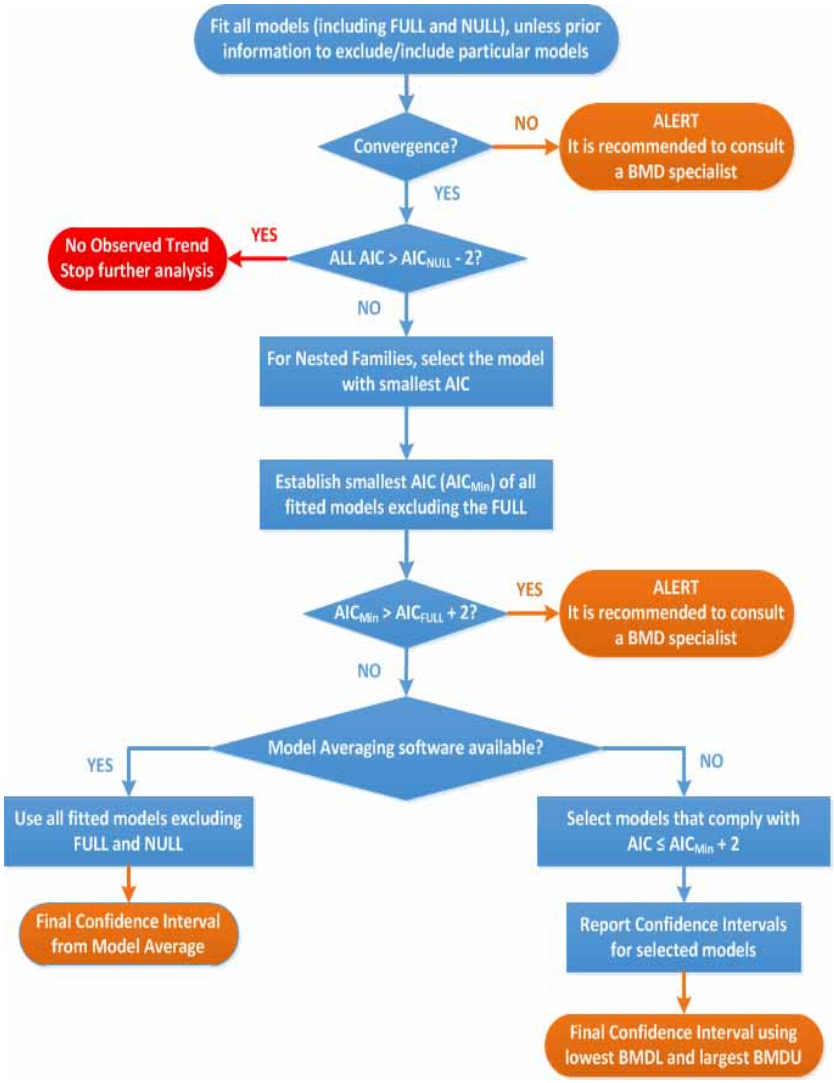
Model	Number of parameters	Formula
Log-logistic	3	$y = a + \frac{1-a}{1+\exp\left(c \cdot \log\left(\frac{b}{x}\right)\right)}$
Log-probit	3	$y = a + (1-a) \cdot pnorm\left(c \cdot \log\left(\frac{x}{b}\right)\right)$
Weibull	3	$y = a + (1-a)\left(1 - \exp\left(-\left(\frac{x}{b}\right)^c\right)\right)$
Gamma	3	$y = pgamma(bx; c)$
Two-stage	3	$y = a + (1-a)\left(1 - \exp\left(-\frac{x}{b} - c\left(\frac{x}{b}\right)^2\right)\right)$
Exp model 3	3	$y = a \cdot \exp(bx^d)$
Exp model 5	4	$y = a \cdot (c - (c-1)\exp(-bx^d))$
Hill model 3	3	$y = a \cdot \left(1 - \frac{x^d}{b^d + x^d}\right)$
Hill model 5	4	$y = a \cdot \left(1 + (c-1)\frac{x^d}{b^d + x^d}\right)$

For the Exp and Hill family, we fit models with 3 and 4 parameters as listed in the table. The 3-parameter model is selected if the difference in AIC is smaller than 5, otherwise the 4-parameter model is selected.

As a covariate is included in the analysis, these models will also be fitted assuming that some of the parameters [background response parameter (a), potency parameter (BMD) and/or variance (var)] depend on the subgroup defined by the covariate. Therefore the number of parameters in each model might be larger than indicated in the table above.

Procedure for selection of BMDL

Description of any deviation from the procedure described in the flow chart to obtain the final BMD confidence interval.



Flowchart for selection of BMDL

7. Results

Response variable: response

Fitted Models

model	No.par	loglik	AIC	accepted	BMDL	BMDU	BMD	sens.subgr	conv
null	1	-87.10	176.20		NA	NA	NA		NA
full	12	-45.91	115.82		NA	NA	NA		NA

model	No.par	loglik	AIC	accepted	BMDL	BMDU	BMD	sens.subgr	conv
two.stage-b	4	-50.13	108.26	no	NA	NA	36.7	2	yes
log.logist-b	4	-50.57	109.14	no	NA	NA	65.7	2	yes
Weibull-b	4	-50.15	108.30	no	NA	NA	45.1	2	yes
log.prob-ab	5	-49.42	108.84	no	NA	NA	79.1	2	yes
gamma-b	4	-50.18	108.36	no	NA	NA	49.3	2	yes
logistic-b	3	-51.34	108.68	no	NA	NA	73.1	2	yes
LVM: Expon. m3-a	4	-49.23	106.46	yes	6.343	52.32	21.4	2	yes
LVM: Hill m3-a	4	-49.11	106.22	yes	11.90	74.6	34.4	2	yes

Estimated Model Parameters

two.stage

estimate for a- : 0.04506

estimate for BMD-1 : 94.14

estimate for BMD-2 : 36.73

estimate for c : 0.4479

log.logist

estimate for a- : 0.04977

estimate for BMD-1 : 200.9

estimate for BMD-2 : 65.67

estimate for c : 2.239

Weibull

estimate for a- : 0.0459

estimate for BMD-1 : 119.2

estimate for BMD-2 : 45.1

estimate for c : 1.279

log.prob

estimate for a-1 : 1e-06

estimate for a-2 : 0.121

estimate for BMD-1 : 209.7

estimate for BMD-2 : 79.11

estimate for c : 1.389

gamma

estimate for a- : 0.04606

estimate for BMD-1 : 132.4

estimate for BMD-2 : 49.26

estimate for cc : 1.556

logistic

estimate for a- : -2.035

estimate for BMD-1 : 195.2

estimate for BMD-2 : 73.13

EXP

estimate for a-1 : 2.016

estimate for CED1 : 194.3

estimate for CED2 : 21.37

estimate for d- : 0.5325

estimate for th : 0

estimate for sigma : 0.25

HILL

estimate for a-1 : 2.029

estimate for CED1 : 216.6

estimate for CED2 : 34.44

estimate for d- : 0.7636

estimate for th : 0

estimate for sigma : 0.25

Weights for Model Averaging

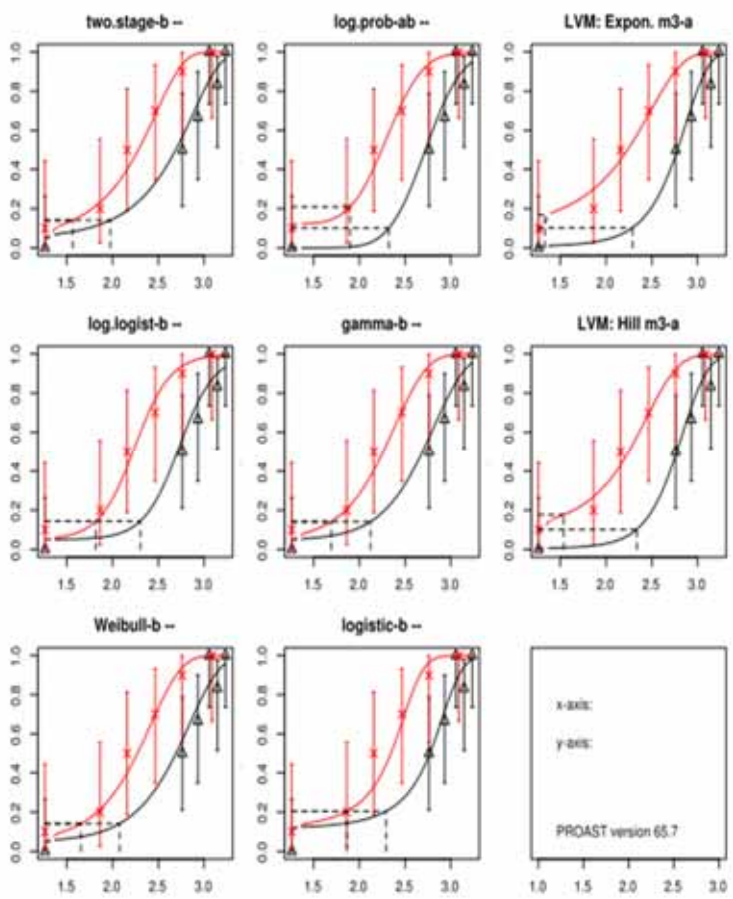
two.stage	log.logist	Weibull	log.prob	gamma	logistic	EXP	HILL
0.1	0.06	0.09	0.07	0.09	0.08	0.24	0.27

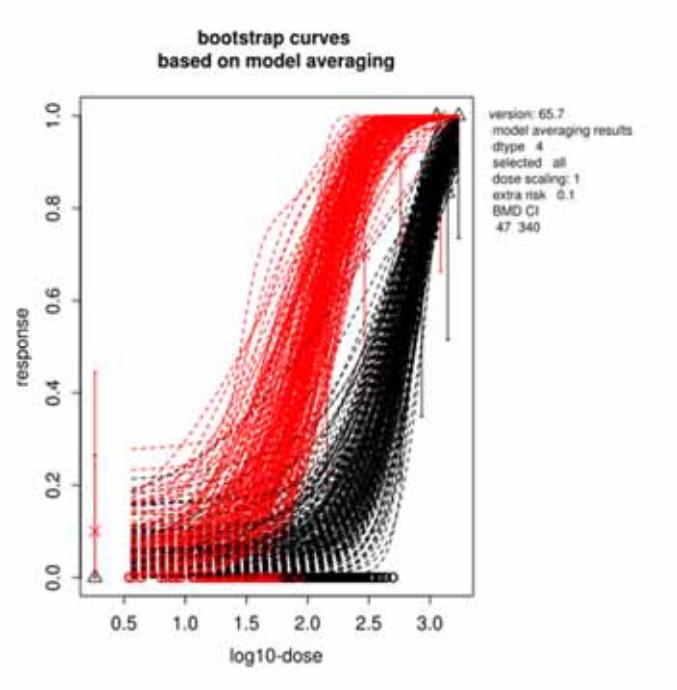
Final BMD Values

subgroup	BMDL	BMDU

1	46.70	343.0
2	7.88	59.9

Confidence intervals for the BMD are based on 200 bootstrap data sets.Visualization





8. Advanced Plots

No results available: If needed, please create advanced plots in the application.

9. Conclusions

The section should discuss the results for the different endpoints and, if applicable, specific issues such as:

- *Discuss if there were any alerts, and if so, how they well dealt with.*
- *Discuss any particular circumstances, if relevant for the final outcome of the BMD confidence interval.*

The BMD confidence interval of the critical endpoint (and the BMDL selected as reference point) should be reported and discussed.

3.4　SPBS (秋田大学)

3.4.1　名称及びバージョン

SPBS V9.67 (2014 年 1 月 10 日改訂版) (2019 年 2 月末時点)

Microsoft Windows XP/Vista/7 で稼動

3.4.2　SPBS の概要

(1)　概要

SPBS は、村田勝敬教授 (秋田大学) によって 1986 年に初代が開発された。無償で入手できる教育用統計ソフトウェアである (図 3.4-1)。

図 3.4-1　ダウンロード画面

(2)　主要な機能

(2)-1　利用可能なデータ及び使用されるモデル

主要な機能は、連続値データ (個体) を対象とした、BMD 及び BMD の信頼区間の下限

値 (BMDL) 算出である。収載されているモデルは、一次回帰モデル、べき関数モデル、log (X+1) 変換モデルであり、手動で選択する。

(2)-2　BMR の設定

BMR は 0.02、0.05、0.10 の 3 種類から選択し、同一用量に対する反応量分布が正規分布と仮定している。分析方法は Budtz-Jørgensen ら (2001)[57]の方法に基づくハイブリッド法である。この手法は、非ばく露集団において設定された異常率 P0 に対して、ばく露集団における α % (BMR) の異常増加をもたらすばく露濃度を BMD と定義し、この BMD の 95%信頼区間の下限値を BMDL として求める手法である。分析方法の概要を以下に示す。

非ばく露群の異常率 P_0、有意な影響の反応レベル BMR、ばく露指標 d、影響指標 μ(d)、用量反応関係 $\mu(d) = b_0+b_1g(d)$とすると、$b_1>0$ はばく露を受けると影響指数が増加することを示し、$b_1<0$ はばく露を受けると影響指数が減少することを示す。

① $b_1>0$ (ばく露を受けると影響指標が増加する) 場合：

$P_0 = 1-\Phi([C-b_0]/\sigma)$ 及び $BMD = g^{-1}\{[\Phi^{-1}(1-P_0)-\Phi^{-1}(1-P_0-BMR)]\ \sigma/b_1\}$から計算

② $b_1<0$ (ばく露を受けると影響指標が減少する) 場合：

$P_0 = \Phi([C-b_0]/\sigma)$ 及び $BMD = g^{-1}\{[\Phi^{-1}(P_0)-\Phi^{-1}(P_0+BMR)]\ \sigma/b_1\}$から計算

($g(d)= d^k$ はべき関数、Φ は正規累積分布関数、C はカットオフ値、b_0 は期待反応関数の定数値、σ は集団の標準偏差値を示す。また K は≥1)

$P_0 = \Phi([C-b_0]/\sigma)$ は Crump (1995)[58] が考案した式で、非ばく露群の P_0=5%を入力すると、カットオフ値として非ばく露群の 95%上限値が算出される。

非ばく露群において異常率を P_0 とした場合に定義される影響指標のカットオフ値が、臨床上の基準値に近いか否かでモデルあるいは対象集団の妥当性を判断する。

例えば、ばく露によって血圧が上昇し、カットオフ値が 140 mmHg と算出された場合、これを臨床上の基準値と比較しその妥当性を判断する。

---

[57] Budtz-Jørgensen E, Keiding N and Grandjean P: Benchmark dose calculation from epidemiological data. Biometrics Vol. 57, No. 3 (Sep., 2001), pp. 698-706.

[58] Crump KS. Calculation of benchmark doses from continuous data. Risk Anal 1995; 15: 78-89.

3.1.1.　3.4.3　利用方法に関する情報

(1)　セットアップ方法

秋田大学のホームページからダウンロードできる[59]。

上記ホームページの教育用統計ソフトウェア SPBS V9.67 をクリックし、圧縮ファイルの解凍方法を参考に、SPBS (圧縮ファイル)、NBtest (圧縮ファイル) をダウンロードし解凍することにより利用できる。

操作方法に関する資料 (操作マニュアル、ガイダンス等)

・ダウンロードした spbs フォルダ内の spbshelp ファイルに操作方法が記載されている。

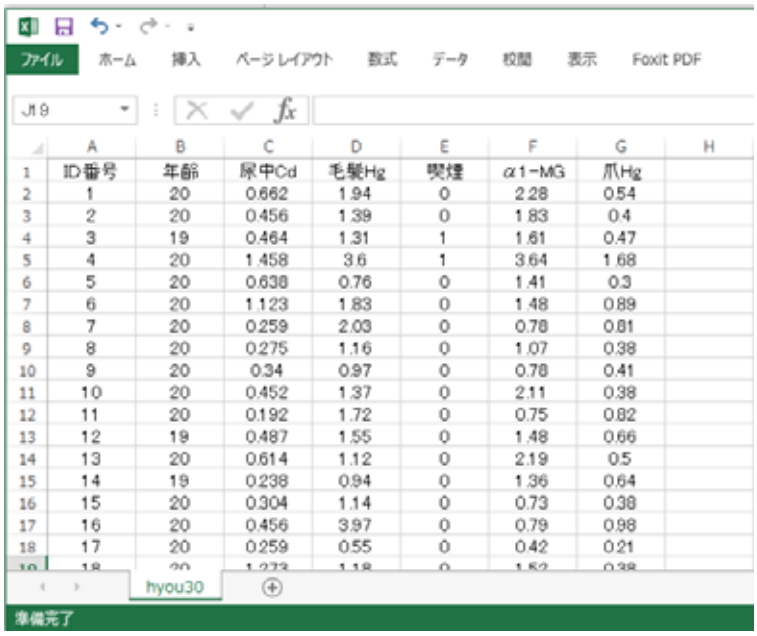
・ダウンロードした spbs フォルダ内の電卓マークの spbs ファイルのショートカットを作成し、デスクトップに移動させる。

(2)　操作の流れ

操作の流れを以下に示す。

1)　入力データの作成

エクセルで作成し、保存する際に CSV ファイルの形式で選択する(タブ区切り等) (図 3.4-2)。



	A	B	C	D	E	F	G
1	ID番号	年齢	尿中Cd	毛髪Hg	喫煙	α1-MG	爪Hg
2	1	20	0.662	1.94	0	2.28	0.54
3	2	20	0.456	1.39	0	1.83	0.4
4	3	19	0.464	1.31	1	1.61	0.47
5	4	20	1.458	3.6	1	3.64	1.68
6	5	20	0.638	0.76	0	1.41	0.3
7	6	20	1.123	1.83	0	1.48	0.89
8	7	20	0.259	2.03	0	0.78	0.81
9	8	20	0.275	1.16	0	1.07	0.38
10	9	20	0.34	0.97	0	0.78	0.41
11	10	20	0.452	1.37	0	2.11	0.38
12	11	20	0.192	1.72	0	0.75	0.82
13	12	19	0.487	1.55	0	1.48	0.66
14	13	20	0.614	1.12	0	2.19	0.5
15	14	19	0.238	0.94	0	1.36	0.64
16	15	20	0.304	1.14	0	0.73	0.38
17	16	20	0.456	3.97	0	0.79	0.98
18	17	20	0.259	0.55	0	0.42	0.21

図 3.4-2　入力データの例

[59] http://www.med.akita-u.ac.jp/~eisei/link.html

2)　プログラムの起動

Spbsspbs (計算機のアイコン) をダブルクリックし起動する (図 3.4-3)。左上の Program(D)を選び、プルダウンして Several tests→Benchmark dose Calculation を選択する。

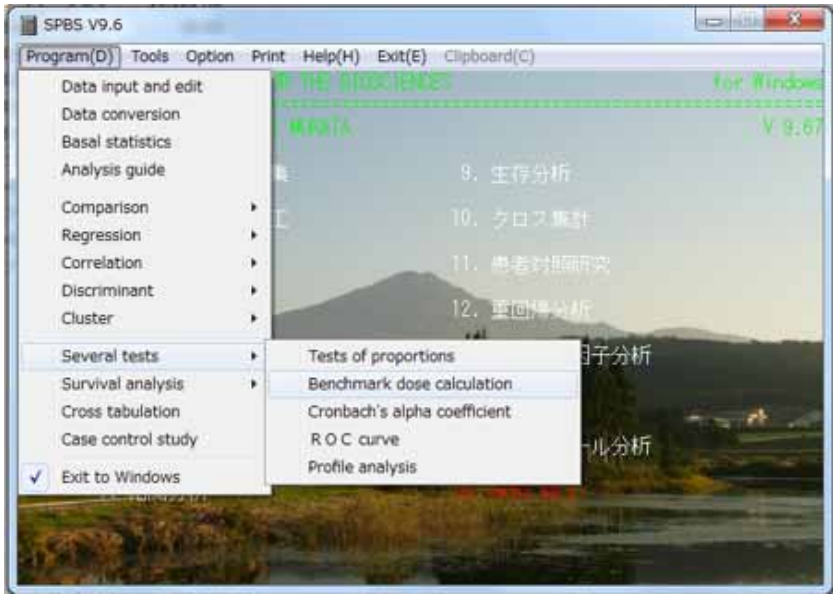


図 3.4-3　プログラムの起動画面

3)　分析するデータファイルの選択、P_0 設定、べき関数のフィッティング

①　分析するデータファイルの選択

・Benchmark dose Calculation 画面上 (図 3.4-4) で File 確認ボタンを押し、解析に用いる用量反応データファイルを選択し、確認する。(図 3.4-5、図 3.4-6)

・ファイル名にスペースが含まれている場合や、ファイル名称が半角 12 文字以上の場合は読み込めない場合がある旨が表示される

図 3.4-4　Benchmark dose Calculation 画面

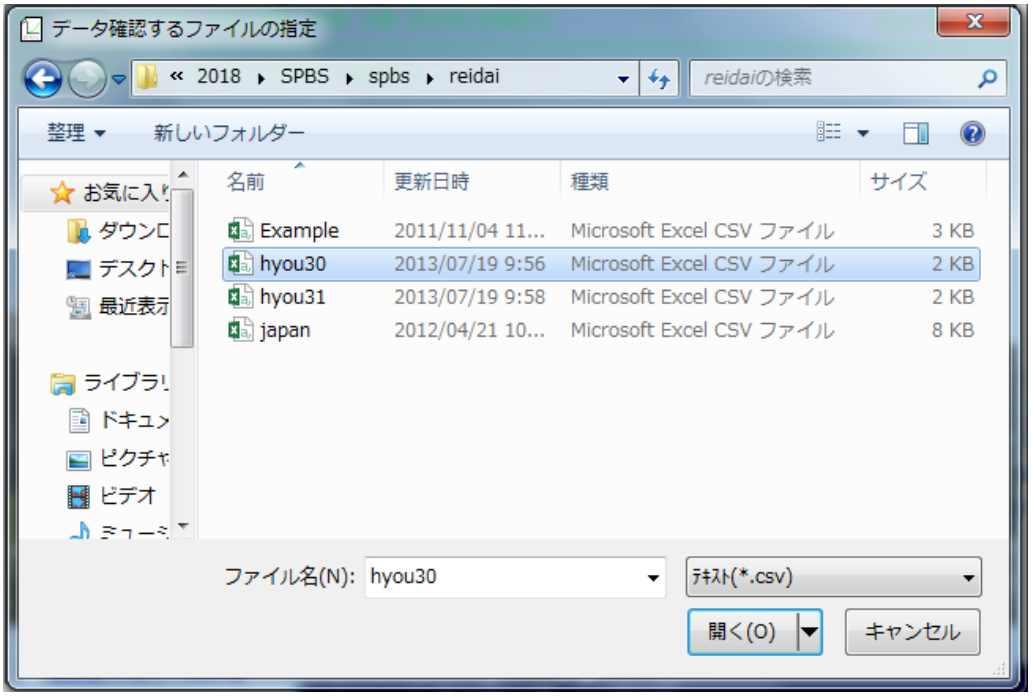


図 3.4-5　解析に用いる用量反応データファイルの選択

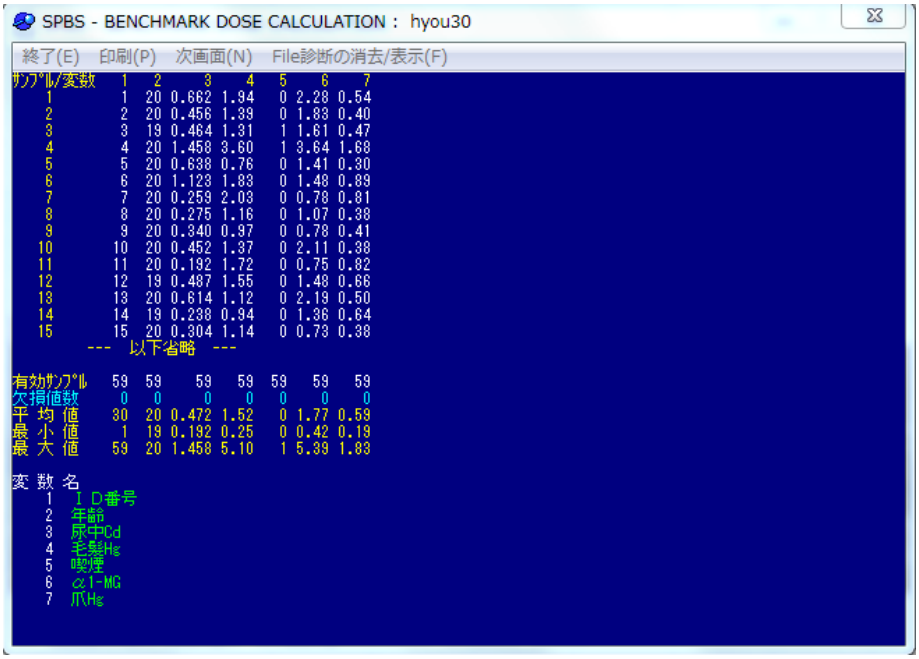


図 3.4-6　解析に用いる用量反応データファイルの確認

② 非ばく露集団の異常率の設定

Benchmark dose Calculation 画面上 (図 3.4-4) で $P_0 = 0.05$、$P_0=0.10$ のいずれかを選択する。 (通常 $P_0=0.05$ を選択)

③ ばく露群における用量のべき関数によるフィッティング

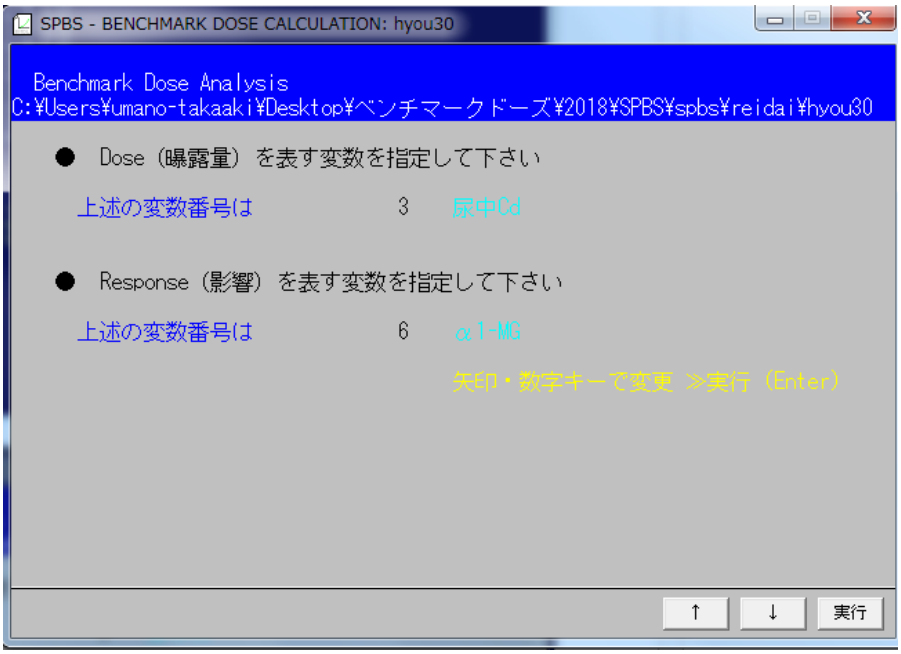
Benchmark dose Calculation 画面上 (図 3.4-4) でばく露量の log(X+1)変換にチェックを入れて実行する。実行ボタンを押し、再度、解析に用いる用量反応データファイルを選択する。

4) データ変数の設定及び解析

各変数の定義付けを行なう。「Dose (ばく露量) を表す変数」を指定し実行ボタンを押す。次に「Response (影響) を表す変数」を指定し実行ボタンを押す (図 3.4-7)。

共変量の設定については、交絡変数の指定を行う画面で設定を行う (図 3.4-8)。

設定に際しては、過去の論文を確認し、共変量として調整した方が良いことが確認された場合は使用し、確認されなかった場合は、一旦共変量として加えて、解析結果「量-影響関係」表の変数名と t 値に対する P 値が有意であれば使用する (この判断は、試行錯誤的に行う) 。一方、統計学的に有意でない場合でも、“Response (影響)” 指標と関連することが研究報告等で知られている場合には含めることが望ましい。



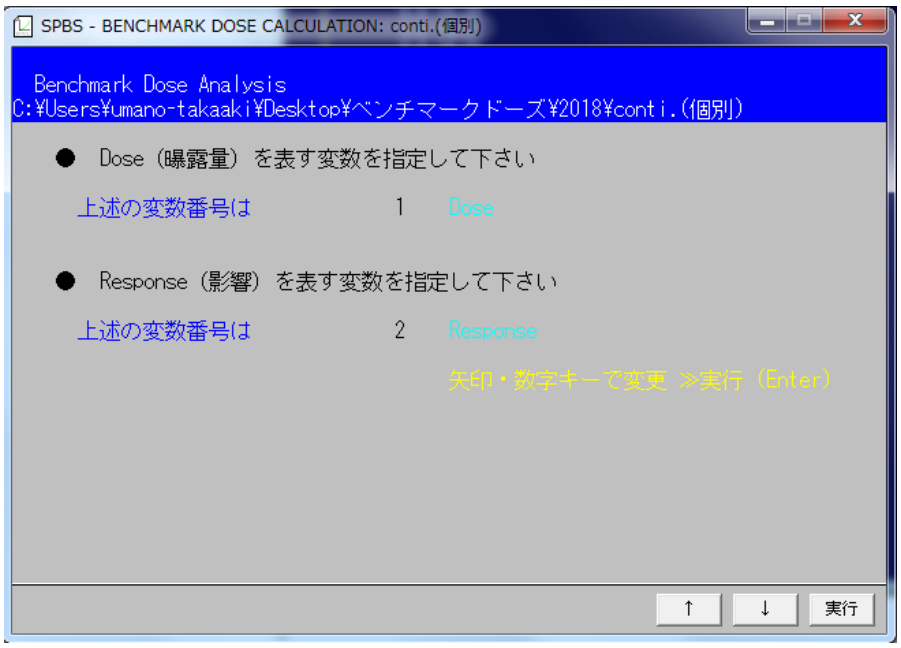


図 3.4-7　データ変数の設定（ばく露量を表す変数、影響を表す変数）の設定場面

図 3.4-8　交絡変数の設定画面

5)　数理モデルのフィッティングの確認

目視で確認する。また尤度比 (-2log L) 検定の結果 (P 値) が有意でない数理モデルは原則として使用しない。

6)　最適な数理モデルの選択

Response (影響) データを対数変換するか否かを選択する。 (最初は対数変換せずにモデル化し目視で対数変換するか否かを検討する。影響を表す変数が正規分布しないで対数正規分布する場合は、同変数を対数変換して解析をする。しかし、このプログラムでは影響を表す変数及びばく露量を表す変数を同時に対数変換することは出来ないことに注意が必要)

Dose (ばく露量) データの変換の実施の有無について、Dose (ばく露量) の log (X+1) はフィットする例があまりないため、一次回帰モデル、べき関数モデルを優先的に検討する。それぞれの数理モデルに関して、目視でモデリング結果の図をみながら、フィットする数理モデルを試行錯誤的に探索する必要がある。

7)　解析結果の出力

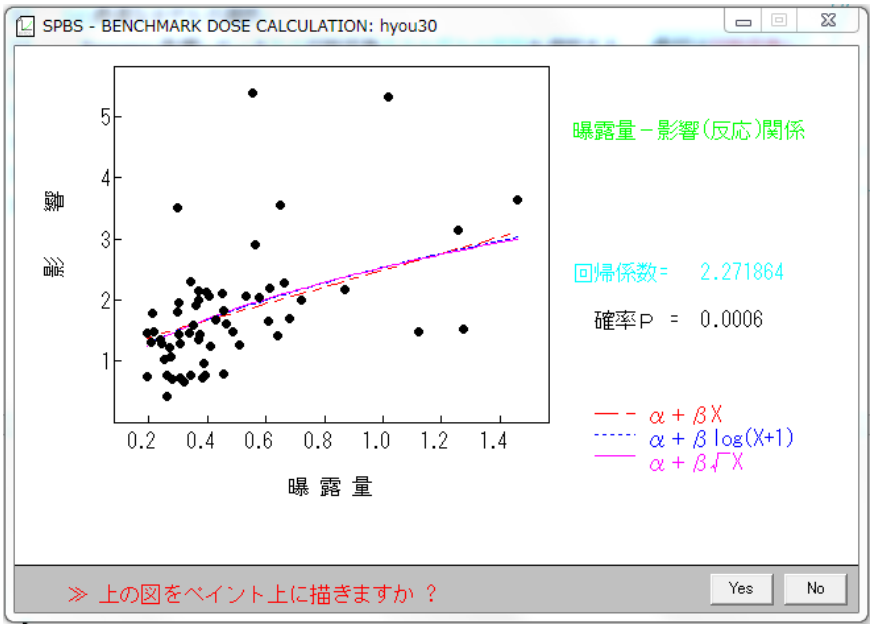
解析に使用した用量反応データ (表 3.4-1) 及び解析結果の出力例を以下に示す。

表 3.4-1　使用した用量反応データ (個体データ)

ID 番号	年齢	尿中 Cd	毛髪 Hg	喫煙	α1-MG	爪 Hg
1	20	0.662	1.94	0	2.28	0.54
2	20	0.456	1.39	0	1.83	0.4
3	19	0.464	1.31	1	1.61	0.47
4	20	1.458	3.6	1	3.64	1.68
5	20	0.638	0.76	0	1.41	0.3
6	20	1.123	1.83	0	1.48	0.89
7	20	0.259	2.03	0	0.78	0.81
8	20	0.275	1.16	0	1.07	0.38
9	20	0.34	0.97	0	0.78	0.41
10	20	0.452	1.37	0	2.11	0.38
11	20	0.192	1.72	0	0.75	0.82
12	19	0.487	1.55	0	1.48	0.66
13	20	0.614	1.12	0	2.19	0.5
14	19	0.238	0.94	0	1.36	0.64
15	20	0.304	1.14	0	0.73	0.38
16	20	0.456	3.97	0	0.79	0.98
17	20	0.259	0.55	0	0.42	0.21
18	20	1.273	1.18	0	1.52	0.38
19	20	0.41	1.44	0	1.24	0.46
20	20	0.389	1.06	0	0.97	0.39
21	20	0.214	2.05	0	1.49	0.71
22	19	0.406	1.16	0	2.07	0.44
23	20	0.575	3.54	0	2.05	1.02
24	20	0.307	1.74	0	1.29	0.94
25	20	0.561	1.29	0	2.9	0.41
26	20	0.252	1.89	0	1.03	0.58
27	20	0.268	1.96	0	1.22	0.73
28	20	0.678	1.52	0	1.7	0.47
29	20	0.278	1.43	0	0.7	0.49

ID 番号	年齢	尿中 Cd	毛髪 Hg	喫煙	α1-MG	爪 Hg
30	20	0.395	1.73	0	2.12	0.9
31	20	0.336	0.98	0	1.46	0.3
32	20	0.554	5.1	1	5.39	1.83
33	19	0.302	0.74	0	1.96	0.22
34	20	0.722	2.39	0	1.99	0.72
35	20	0.344	0.59	0	2.31	0.38
36	20	0.297	0.72	1	1.81	0.31
37	20	0.51	1.34	0	1.27	0.45
38	20	0.36	0.78	0	1.92	0.33
39	20	1.259	1.08	0	3.15	0.35
40	20	0.352	3.44	1	1.6	0.83
41	19	0.868	1.57	1	2.17	0.72
42	20	0.371	1.5	1	2	0.65
43	20	1.017	1.7	0	5.31	1.08
44	20	0.299	1.32	0	3.51	0.58
45	20	0.245	1.39	0	1.28	0.52
46	19	0.533	2.21	0	2.07	0.75
47	20	0.3	0.49	0	1.44	0.19
48	20	0.391	0.69	0	0.78	0.22
49	20	0.43	1.63	0	1.67	0.47
50	20	0.609	0.6	0	1.65	0.34
51	20	0.318	1.53	0	0.66	0.75
52	20	0.209	0.25	0	1.32	0.26
53	20	0.192	1.13	0	1.46	0.49
54	20	0.382	0.76	0	0.72	0.51
55	20	0.647	1.14	1	3.56	0.4
56	20	0.37	2.49	0	2.14	1.06
57	20	0.368	0.39	0	1.36	0.58
58	20	0.213	0.82	1	1.79	0.29
59	20	0.374	1.3	0	1.43	0.74

(SPBS ソフトのダウンロードフォルダ内にあるデータから引用)



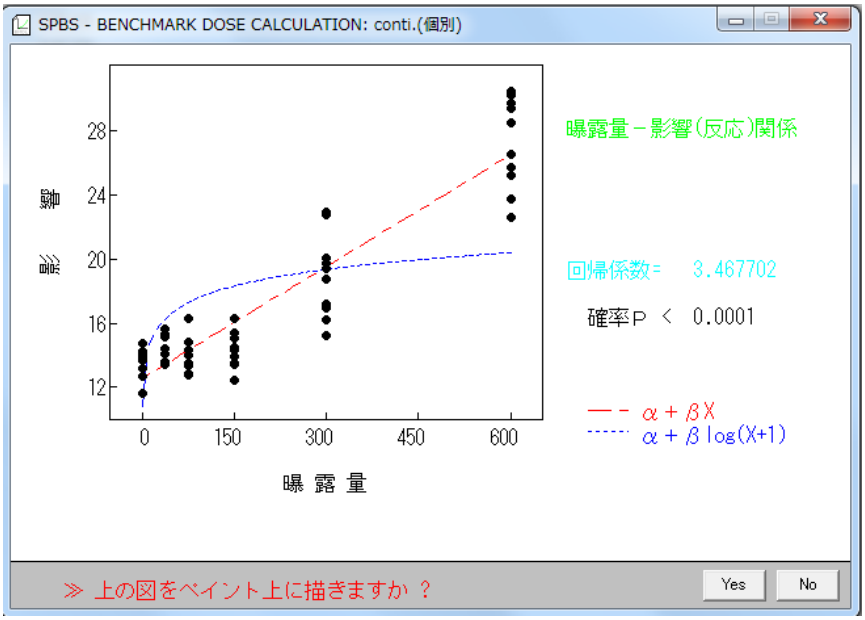


図 3.4-9　一次回帰モデル、ばく露量の log (X+1) モデル及びべき関数モデルによる用量反応データ

① 一次回帰モデルによる解析結果 (基本的に画面の表示内容を転載)

■■ 基本データ ■■　　解析の症例数 = 59

変数	平均値	標準偏差	最小値	最大値	変 数 名	
1	0.472	0.277	0.192	1.458	尿中 Cd	⇐ ばく露変数
2	1.515	0.907	0.250	5.100	毛髪 Hg	⇐ 交絡変数
3	0.588	0.319	0.190	1.830	爪 Hg	⇐ 交絡変数
4	1.766	0.983	0.420	5.390	α1-MG	⇐影響変数

◇ 相関行列 [▲: P 値　▼: 相関係数]

	1	2	3	4
1		0.248	0.316	0.498
2	.0580		0.857	0.368
3	.0148	<.0001		0.471
4	<.0001	.0042	.0002	

量－影響関係　(回帰式の F 値　10.322　P<0.0001)

回帰係数	標準回帰係数	標準誤差	t 値	変数名
1.366	0.38542	0.4033	3.387 (P=0.0013)	尿中 Cd
−0.1022	−0.10403	0.2061	−0.496 (P=0.6219)	毛髪 Hg
0.4309	0.43843	0.2105	2.047 (P=0.0454)	爪 Hg
1.121				定数項

Cutoff point = 2.535　⇐ この値の妥当性を検討する

○ Benchmark Dose Calculation

Po=0.05　BMR=0.02　BMD= 0.106　(BMDL= 0.069)

Po=0.05　BMR=0.05　BMD= 0.229　(BMDL= 0.148)　⇐ 通常この値を使用

Po=0.05　BMR=0.10　BMD= 0.383　(BMDL= 0.248)

全尤値（定数項のみの対数尤度：共変量によらない対数尤度）　171.43
偏尤値（回帰モデルの対数尤度：共変量からの最大対数尤度）　145.08
モデルの尤度比検定(-2 LOG L)　26.35　（自由度 ＝3）　P<0.0001

② ばく露量の log (X+1) モデルによる解析結果(基本的に画面の表示内容を転載)

■■ 基本データ ■■　解析の症例数 ＝ 59

変数	平均値	標準偏差	最小値	最大値	変 数 名	
1	0.162	0.072	0.076	0.391	尿中 Cd	⇐ ばく露変数
2	1.515	0.907	0.250	5.100	毛髪 Hg	⇐ 交絡変数
3	0.588	0.319	0.190	1.830	爪 Hg	⇐ 交絡変数
4	1.766	0.983	0.420	5.390	α1-MG	⇐影響変数

注）ばく露データは log(X+1)変換値

◇ 相関行列［▲: Ｐ値　▼: 相関係数］

	1	2	3	4
1		0.256	0.306	0.511
2	.0505		0.857	0.368
3	.0183	<.0001		0.471
4	<.0001	.0042	.0002	

量－影響関係　(回帰式のＦ値　10.986　P<0.0001)

回帰係数	標準回帰係数	標準誤差	t 値	変数名
5.500	0.40409	1.525	3.607 (P=0.0007)	尿中 Cd
−0.1242	−0.12641	0.2036	−0.610 (P=0.5442)	毛髪 Hg
0.4477	0.45557	0.2068	2.165 (P=0.0347)	爪 Hg
0.8775				定数項

Cutoff point = 2.277

○ Benchmark Dose Calculation [log(X+1)変換値から算出]

Po=0.05 BMR=0.02 BMD= 0.062 (BMDL= 0.041)
Po=0.05 BMR=0.05 BMD= 0.138 (BMDL= 0.089) ⇐ 通常この値を使用
Po=0.05 BMR=0.10 BMD= 0.242 (BMDL= 0.154)

全尤値（定数項のみの対数尤度：共変量によらない対数尤度） 171.43
偏尤値（回帰モデルの対数尤度：共変量からの最大対数尤度） 143.73
モデルの尤度比検定(-2 LOG L) 27.70 （自由度 ＝3） P<0.0001

③ ばく露量のべき関数モデルによる解析結果(基本的に画面の表示内容を転載)

■■ 基本データ ■■ 解析の症例数 = 59

変数	平均値	標準偏差	最小値	最大値	変 数 名	
1	0.664	0.176	0.438	1.207	尿中 Cd	⇐ ばく露変数
2	1.515	0.907	0.250	5.100	毛髪 Hg	⇐ 交絡変数
3	0.588	0.319	0.190	1.830	爪 Hg	⇐ 交絡変数
4	1.766	0.983	0.420	5.390	α1-MG	⇐影響変数

注）ばく露データは√X 変換値

◇ 相関行列［▲: Ｐ値 ▼: 相関係数］

	1	2	3	4
1		0.259	0.303	0.513
2	.0475		0.857	0.368
3	.0195	<.0001		0.471
4	<.0001	.0042	.0002	

量－影響関係 （回帰式のＦ値 11.129 P<0.0001）

回帰係数	標準回帰係数	標準誤差	t 値	変数名
2.272	0.40779	0.6219	3.653 (P=0.0006)	尿中 Cd
−0.1330	−0.13536	0.2031	−0.655 (P=0.5152)	毛髪 Hg
0.4553	0.46329	0.2058	2.212 (P=0.0311)	爪 Hg
0.2563				定数項

Cutoff point = 1.654

○ Benchmark Dose Calculation [√X 変換値から算出]

Po=0.05 BMR=0.02 BMD= 0.004 (BMDL= 0.002)
Po=0.05 BMR=0.05 BMD= 0.018 (BMDL= 0.008) ⇦ 通常この値を使用
Po=0.05 BMR=0.10 BMD= 0.052 (BMDL= 0.023)

全尤値（定数項のみの対数尤度：共変量によらない対数尤度） 171.43
偏尤値（回帰モデルの対数尤度：共変量からの最大対数尤度） 143.45
モデルの尤度比検定(-2 LOG L) 27.99 （自由度 ＝3） P<0.0001

3.5　MADr-BMD

3.5.1　名称及びバージョン

MADr-BMD ver.1.0 (2019 年 2 月末時点)

3.5.2　MADr-BMD の概要

(1)　概要

MADr-BMD は、Matthew W. Wheeler (CDC/NIOSH) と A. John Bailer (マイアミ大学と CDC/NIOSH) によって 2007 年に開発された、無償で入手できるソフトウェアである。本ソフトウェアは、Microsoft Windows オペレーションシステムにおけるシェルコマンドプロンプト60上で操作する他のプログラムを要せず単体で動作可能なプログラムである。操作は、マウスではなくキーボードでコマンドと呼ばれる命令文を入力して行う。

(2)　主要な機能

主要な機能は、二値の用量反応データを対象とした数理モデル平均化手法による BMD 及び BMD の信頼区間の下限値 (BMDL) の算出である。収載されている数理モデルを表 3.5-1 に示す。数理モデルの選択は手動で行う (単一モデルの選択も可能)。BMR は 1%、5%、10%等の任意の値を入力できる。モデル性能の評価法は AIC、BIC、KIC から選択する (その他 BICB、AICB、KICB も選択可能)。また、信頼区間の推定はブートストラップ法、パラメータ推定は最尤推定による。MADr-BMD に収載されている数理モデルを表 3.5-1 に示す。

表3.5-1　MADr-BMDに収載されている数理モデル

数理モデル	式	パラメータ
Logistic	$\pi_1(d) = \frac{1}{1 + exp[-(\alpha + \beta \times d)]}$	
Log-logistic	$\pi_2(d) = \gamma + \frac{(1-\gamma)}{1 + exp[-(\alpha + \beta \times d)]}$	$0 \leq \gamma \leq 1$ $\beta \geq 0.5$

[60] コマンドと呼ばれる命令文を用いて Windows の操作や設定を行うためのツール

数理モデル	式	パラメータ
Gamma	$\pi_3(d) = \gamma + \frac{1-\gamma}{\Gamma(\alpha)} \int_0^{\beta d} t^{\alpha-1} e^{-t} dt$	$0 \le \gamma \le 1$ $\alpha \ge 1, \beta \ge 0.5$
Multistage	$\pi_4(d) = \gamma + (1-\gamma)(1 - exp(-\theta_1 d - \theta_2 d^2 \ldots))$	$0 \le \gamma \le 1$ $\theta_1 \ge 0, \theta_2 \ge 0, \ldots$
Probit	$\pi_5(d) = \Phi(\alpha + \beta d)$	
Log-probit	$\pi_6(d) = \gamma + (1-\gamma)\Phi(\alpha + \beta ln(d))$	$0 \le \gamma \le 1, \beta \ge 0$
Quantal Linear	$\pi_7(d) = \gamma + (1-\gamma)(1 - exp(-\beta d))$	$0 \le \gamma \le 1, \beta \ge 0$
Quantal quadratic	$\pi_8(d) = \gamma + (1-\gamma)(1 - exp(-\beta d^2))$	$0 \le \gamma \le 1, \beta \ge 0$
Weibull	$\pi_9(d) = \gamma + (1-\gamma)(1 - exp(-\beta d^\alpha))$	$0 \le \gamma \le 1$ $\beta \ge 0.5$

(Wheeler and Bailer (2008)から引用)

3.5.3　利用方法に関する情報

(1)　セットアップ方法

1)　MADr-BMD ソフト等のダウンロード

米国 EPA のサイトから入手できる[61]

また、操作方法に関する資料 (操作マニュアル、ガイダンス等) の入手方法は以下のとおりである。

・操作方法は EPA の上記のサイトに記載されている。

・統計学的なアプローチ、使用例、設定例については、Wheeler and Bailer (2008) の論文から入手できる。

・ソースコードや解析に必要な入力ファイルの例 (input と表示されているテキストファ

[61] https://www.epa.gov/tsca-screening-tools/model-averaging-dichotomous-response-benchmark-dose-madr-bmd-tool

イル) も上記のサイトから入手できる。

2) madrbmd ソフトと入力ファイルの書き込み可能なフォルダへの移動

書き込み可能なフォルダに、ダウンロードした madrbmd ソフトと入力ファイルを移動する。

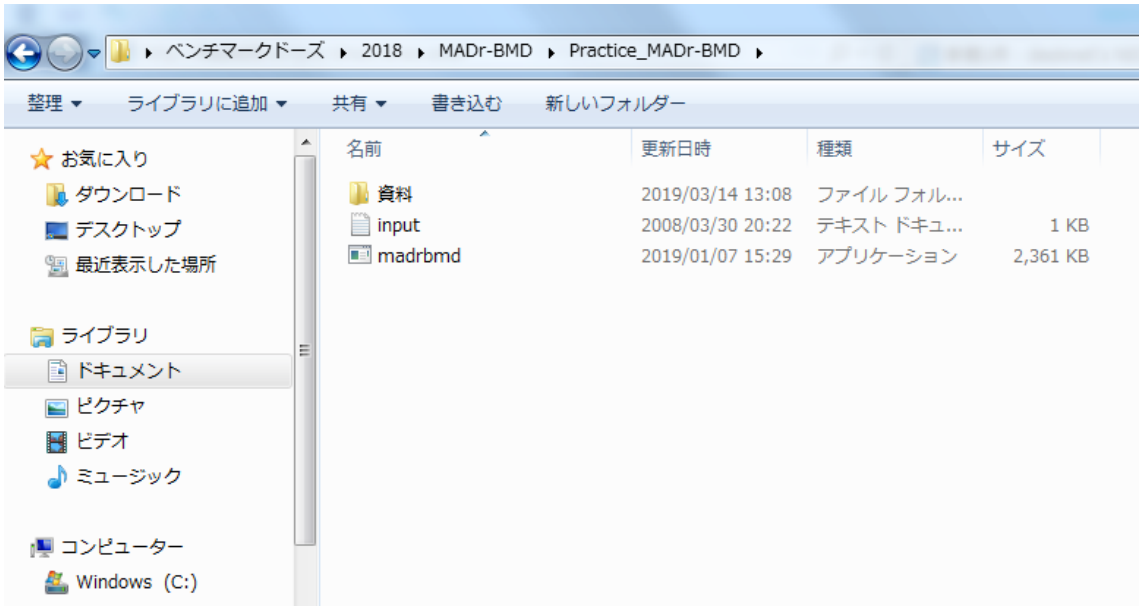


図 3.5-1　ダウンロードした madrbmd ソフトと入力ファイルの保管場所の例

(2)　操作の流れ

MSDr-BMD では、入力ファイル上で、解析する用量反応データとともに、解析に当たっての各種条件を入力、指定する必要がある。

操作の流れを以下に示す。

1) データの入力

入力ファイルの入力例及び設定について以下に示す (図 3.5-2、表 3.5-2)。

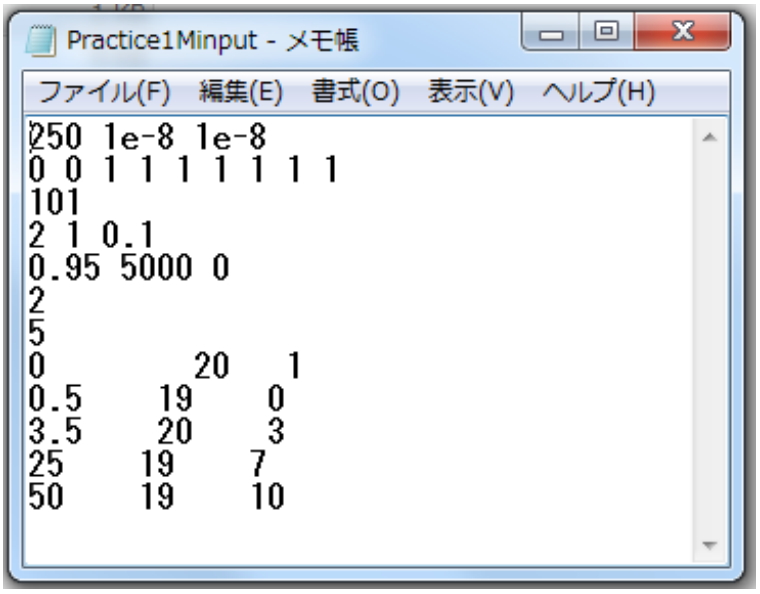


図 3.5-2　入力ファイルの入力例

表 3.5-2　入力ファイルで設定する解析条件

250　1e-8　1e-8	最大反復回数、関連する収束、一般的な収束
0 0 1 1 1 1 1 1 1	モデル平均化 (MA) に含める数理モデル(1＝ 含める、0＝含めない) 左から次の数理モデルを示している(Quantal-linear、Quantal-quadratic、Multistage、Logistic、Probit、Weibull、Log-probit、Log-logistic、Gamma) (単一モデルのみをフィッティングしたい場合は、当該モデルの順番の位置に 1 と入力し、その他の数理モデルは 0 を入力する。)
101	ランダムシード (0 を用いると現在の時刻を使用)
2　1　0.1 ①　②　③	①平均化の際に重み付けとして用いる指標 (1 ＝ BIC、2 ＝ AIC、3＝KIC、4＝BICB、5＝AICB、6＝KICB) ②BMR のタイプ(1＝ 追加リスク(added risk)、2＝ 過剰リスク(extra risk)) ③BMR の設定値
0.955000　0	①BMD の信頼区間の信頼水準（BMD の信頼区間の下限値(BMDL)算出に反映される）

	②ブートストラップサンプリング数 ③ブートストラップサンプルの出力 (0＝しない、1＝ する)
2	Multistage の次数
5	データの行数
0 20 1 0.5 19 0 3.5 20 3 25 19 7 50 19 10	データの入力項目 1 列目：用量、2 列目：個体数（n 数)、3 列目：発生数

実際にデータの入力、解析時の条件を設定する場合は、ダウンロードした入力ファイル(input と表示されているテキストファイル) を基に入力し、入力ファイルを“任意の名称”.txt として保存する。この際、ファイル名にスペースは使用しない。

2) データの解析

① コマンドプロンプトの起動

使用するパソコンの OS に応じた方法でコマンドプロンプトウィンドウを開く (図 3.5-3) (例：Windows 7 の場合は、スタートメニューを開き、一番下のエントリーに “cmd” と入力、Enter キーを押す) 。

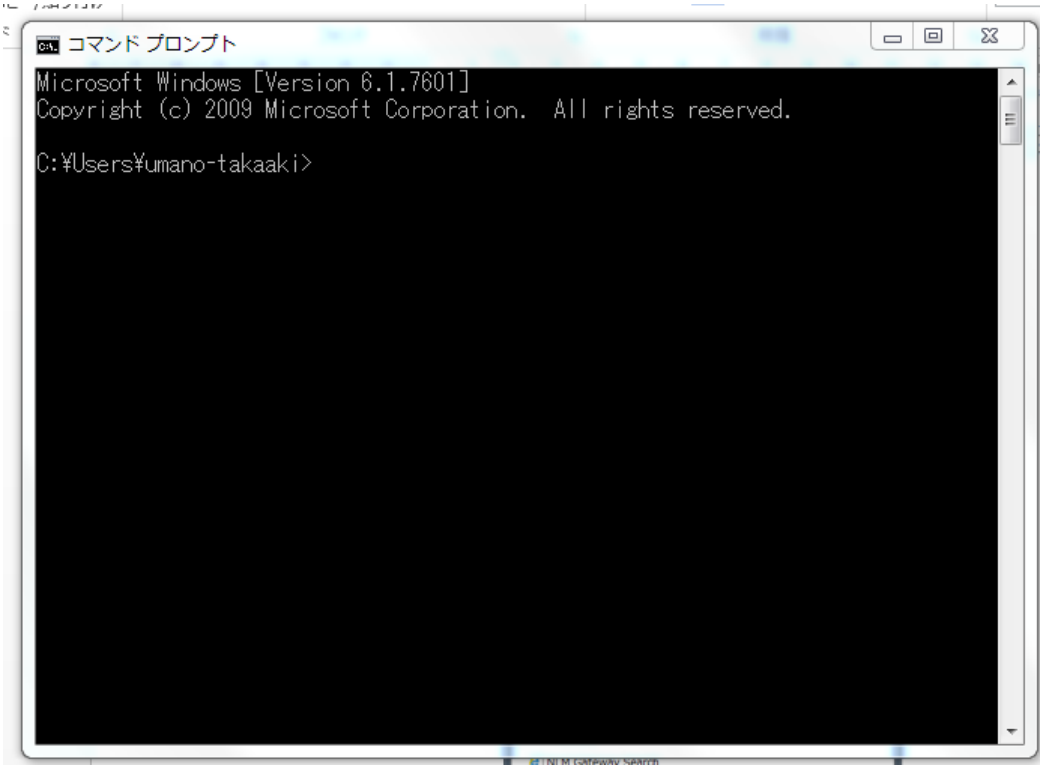


図 3.5-3　コマンドプロンプトウィンドウ

② ディレクトリの変更と確認

・ディレクトリの変更 (図 3.5-4)

madrbmd ソフトと解析に用いるデータ入力ファイルが保存されているディレクトリに変更するため、“cd コマンド”を用いて変更する。

コマンドプロンプトウィンドウに表示されているディレクトリ名の後に、cd+半角スペース+「 madrbmd ソフトとデータ入力ファイルが保存されているディレクトリ名」を入力

ディレクトリの変更例：保存先が Desktop¥ベンチマークドーズ¥2018¥MADr-BMD¥Practice_MADr-BMD の場合

C:¥Users¥umano-takaaki>の後に、cd C:¥Users¥umano-takaaki¥Desktop¥ベンチマークドーズ¥2018¥MADr-BMD¥Practice_MADr-BMD と入力し Enter

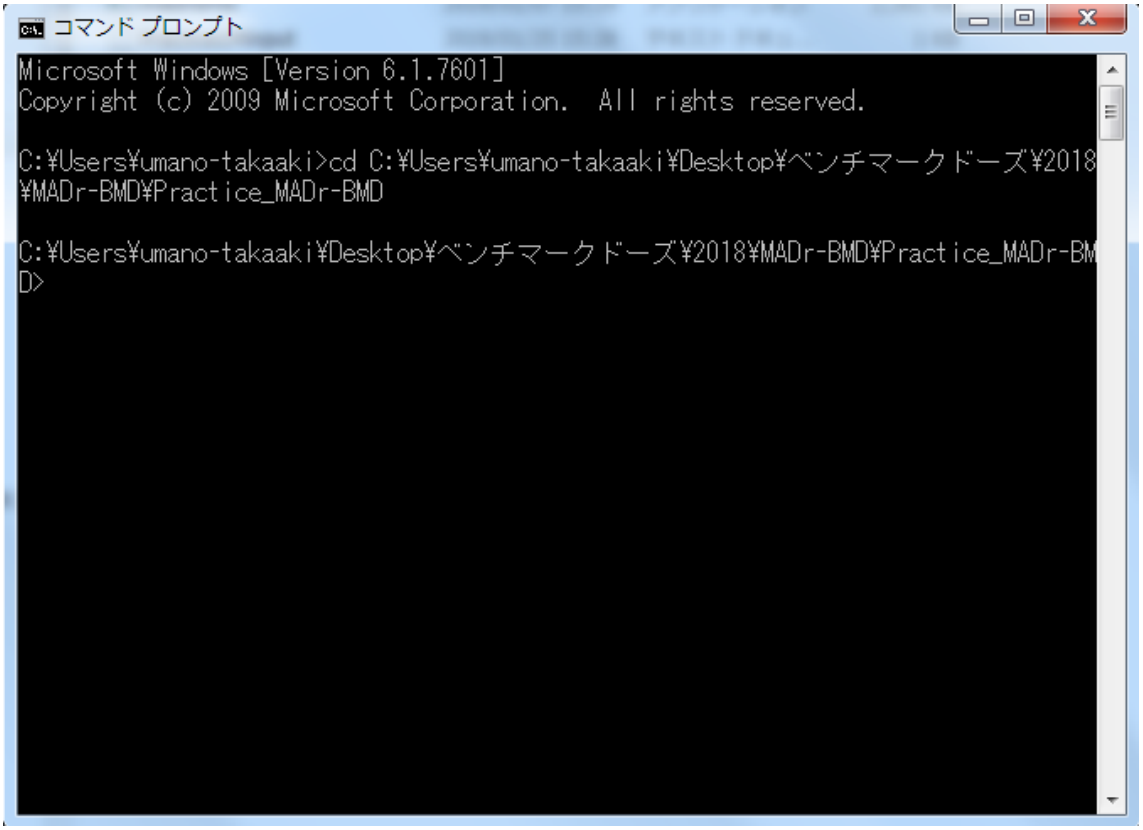


図 3.5-4　ディレクトリの変更操作画面

・ディレクトリの変更の確認

madrbmd ソフトとデータ入力ファイルは同じディレクトリにあることが必要なため、操作中のディレクトリにあることを次の操作により確認する。

“dir コマンド”を用いてディレクトリの変更を確認する。“dir”と入力するとディレクトリにあるファイルが表示され確認できる。

dir コマンドの例：

C:¥Userrs¥umano-takaaki>cd C:¥Users¥umano-takaaki¥Desktop¥ベンチマークドーズ¥2018¥MADr-BMD¥Practice_MADr-BMD>の後に dir を入力

図 3.5-5 に示すように、解析対象のデータである Practice1Minput.txt ファイル、解析ソフトである madrbmd.exe ファイルがこのディレクトリにあることが確認できる。

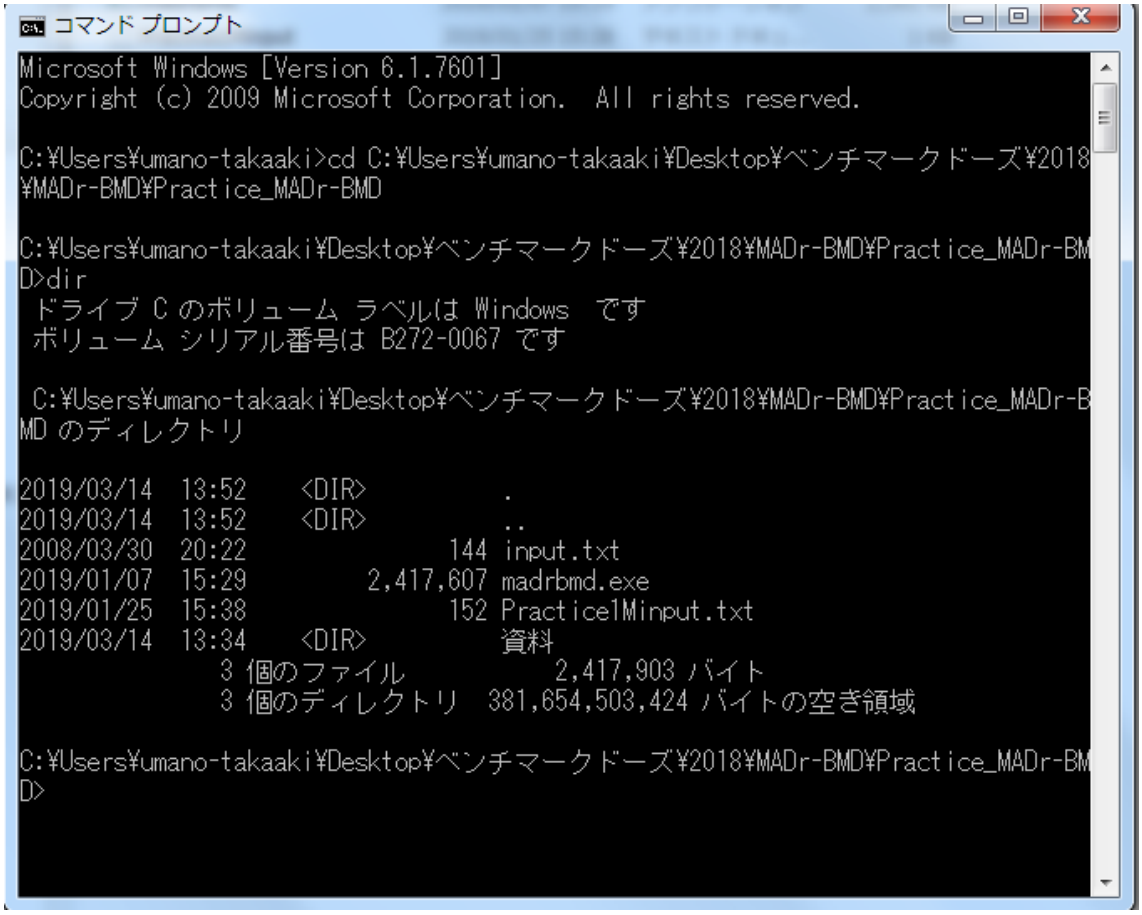


図 3.5-5　ディレクトリの変更の確認画面

③　MADr-BMD ソフトの起動・解析

解析に用いる入力データの読み込みは、次のように madrbmd+半角スペース+入力ファイル名.txt と入力する。例えば、「madrbmd Practice1Minput .txt」と入力すると、入力ファイルとして Practice1Minput .txt を使用するようにプログラムに指示がされる。数十秒程度で、解析結果がコマンドプロンプトウィンドウに表示される。

入力データの読み込み、解析例 (図 3.5-6)：

入力データが madrbmd Practice1Minput.txt の場合、

C:¥Userrs¥umano-takaaki>cd C:¥Users¥umano-takaaki¥Desktop¥ベンチマークドーズ¥2018¥MADr-BMD¥Practice_MADr-BMD> の後に、「madrbmd Practice1Minput.txt」を入力

```
コマンド プロンプト
C:¥Users¥umano-takaaki¥Desktop¥ベンチマークドーズ¥2018¥MADr-BMD¥Practice_MADr-BM
D>madrbmd Practice1Minput.txt
MABMD VERSION 1.0
Thu Mar 14 15:34:03 2019

This program's results are given "as is", without warranty, either expressed
or implied by the National Institute for Occupational Safety and Health.


INPUT DATA
Dose            Count          Observed
---------------------------------------------------------------------------
0.000000           20              1
0.500000           19              0
3.500000           20              3
25.000000          19              7
50.000000          19             10
---------------------------------------------------------------------------


Model Fit Statistics
---------------------------------------------------------------------------
Model                  Weight  -2log(L)          AIC           BIC
---------------------------------------------------------------------------
Multistage             0.162     78.87          84.87         92.59
Logistic               0.111     81.63          85.63         90.78
Probit                 0.137     81.21          85.21         90.36
Weibull                0.162     78.87          84.87         92.59
Log-Probit             0.068     80.61          86.61         94.33
Log-Logistic           0.199     78.45          84.45         92.18
Gamma                  0.162     78.87          84.87         92.59
---------------------------------------------------------------------------



-----------------------------------------
'Average-Model' Benchmark Dose Estimate
-----------------------------------------
Nominally Specified Confidence Level:0.950
Weighting Criterion: AIC
BMD Calculation: Added Risk
BMR: 0.200000
BMD: 16.991535812414
BMDL(BCa):7.163513952322
BMDL(Percentile):11.500404572231
Acceleration: 0.030539
Bootstrap Resamples: 5000
Random Seed: 101
-----------------------------------------
```

図 3.5-6　解析結果画面

3)　解析結果の出力

ソフト起動時に、madrbmd 解析に用いる入力ファイル.txt> ユーザー指定の出力ファイル名.txt とすると、解析結果がユーザー指定の出力ファイル名の付いたテキストとして、同じディレクトリに自動で保存される。(この際、出力ファイル名は必ず独立したファイル名とするように注意する)(図 3.5-7)

例：

C:¥Userrs¥umano-takaaki>cd C:¥Users¥umano-takaaki¥Desktop¥ベンチマークドーズ¥2018¥MADr-BMD¥Practice_MADr-BMD> madrbmd Practice1Minput.txt> Practice1Moutput.txt

なお、出力ファイルを設定しなかった場合、結果が表示されているコマンドプロンプトウィンドウを右クリックして全てを選択した後コピーし、ワードファイル等にペーストすることで、解析結果を出力することもできる。

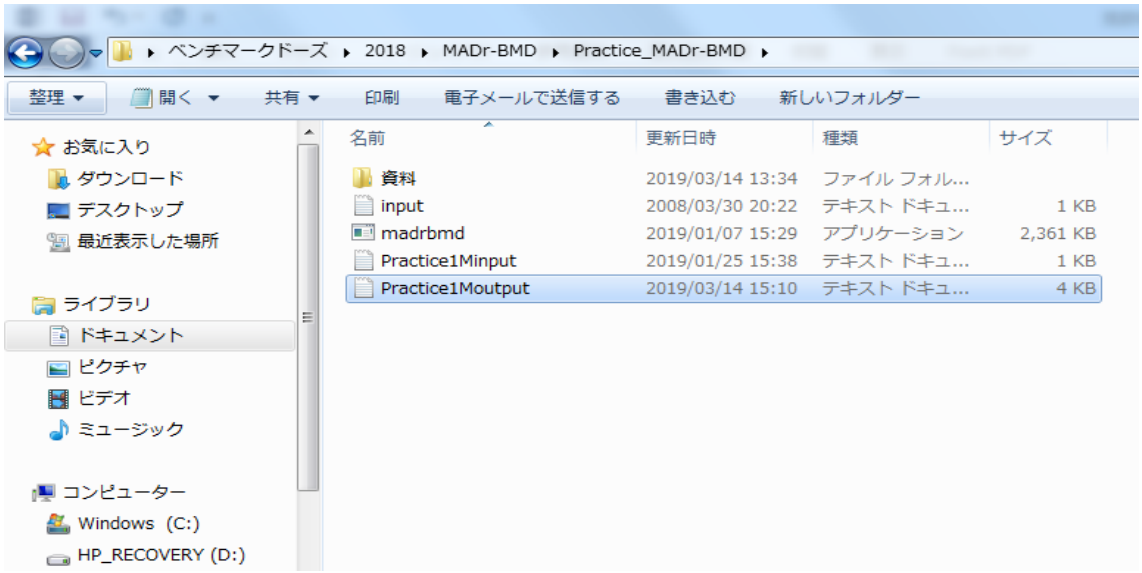


図 3.5-7　出力ファイルの自動保存

解析結果の出力例を以下に示す。

MABMD VERSION 1.0
Thu Mar 14 17:00:07 2019

This program's results are given "as is", without warranty, either expressed or implied by the National Institute for Occupational Safety and Health.

INPUT DATA

Dose	Count	Observed
0.000000	20	1
0.500000	19	0
3.500000	20	3
25.000000	19	7
50.000000	19	10

Model Fit Statistics

Model	Weight	-2log(L)	AIC	BIC
Multistage	0.162	78.87	84.87	92.59
Logistic	0.111	81.63	85.63	90.78
Probit	0.137	81.21	85.21	90.36
Weibull	0.162	78.87	84.87	92.59
Log-Probit	0.068	80.61	86.61	94.33
Log-Logistic	0.199	78.45	84.45	92.18
Gamma	0.162	78.87	84.87	92.59

'Average-Model' Benchmark Dose Estimate

Nominally Specified Confidence Level: 0.950

Weighting Criterion: AIC

BMD Calculation: Added Risk

BMR: 0.100000

BMD: 8.212328013576

BMDL (BCa):3.041899204839

BMDL (Percentile):5.574802713355

Acceleration: 0.027111

Bootstrap Resamples: 5000

Random Seed: 101

'Average-Model' Goodness of Fit Test

Test Statistic : 2.135780

Bootstrap P-Value: 0.444600

MODEL: Multistage, 2-degree polynomial:

Parameters	Estimate	StdErr
GAMMA:	3.852992e-002	3.188900e-002
BETA(1):	1.573000e-002	4.292851e-003
BETA(2):	0.000000e+000	N/A

Optimization Succeeded

MODEL: Logistic:

Parameters	Estimate	StdErr
ALPHA:	-2.460499	0.446256

BETA: 0.056080 0.013585

Optimization Succeeded

MODEL: Probit:

Parameters	Estimate	StdErr
ALPHA:	-1.455329	0.234799
BETA:	0.033341	0.007740

Optimization Succeeded

MODEL: Weibull:

Parameters	Estimate	StdErr
GAMMA:	0.038530	0.031884
ALPHA:	1.000000	N/A
BETA:	0.015730	0.004293

Optimization Succeeded

MODEL: Log-Probit:

Parameters	Estimate	StdErr
GAMMA:	0.064740004794	0.033896
ALPHA:	-3.794321	0.245232
BETA:	1.000000	N/A

Optimization Succeeded

MODEL: Log-Logistic:

```
Parameters        Estimate          StdErr
-----------------------------------------------------------
GAMMA:            0.032534          0.029626
ALPHA:            -3.806055         0.348835
BETA:             1.000000          N/A
Optimization Succeeded
-----------------------------------------------------------

MODEL: Gamma:
-----------------------------------------------------------
Parameters        Estimate          StdErr
-----------------------------------------------------------
GAMMA:            0.038530          0.031891
ALPHA:            1.000000          N/A
BETA:             0.015730          0.004293
Optimization Succeeded
-----------------------------------------------------------
```

引用

Matthew W. Wheeler, A. John Bailer (2008) Model Averaging Software for Dichotomous Dose Response Risk Estimation. Journal of Statistical Software. Vol. 26, Issue 5. https://www.jstatsoft.org/article/view/v026i05

3.6 BMDExpress

3.6.1 名称及びバージョン

最新バージョン：BMDExpress 2.2 (2019 年 2 月末時点)

2018 年 10 月 30 日にリリースされた最新バージョン 2.2 には Linux 版、Mac 版、Windows 64bit 版が存在する。但し、Linux 版及び Mac 版では、BMDExpress プロジェクトファイルの保存形式である.bm2 ファイルの閲覧や既存の BMD 解析における機能分類のみ可能で、BMD 解析や一方向 ANOVA フィルタリングは実行できない。

3.6.2 BMDExpress の概要

(1) 概要

BMDExpress は、毒性学、医薬品設計、リスクアセスメント及びトランスレーショナルリサーチ[62]の分野で増加しているトランスクリプトミクスによる用量反応データ[63]の解析を目的に開発された無償のデスクトップアプリケーションである。このアプリケーションでは、トキシコゲノミクスデータからの BMD の推計と、ユーザーが定義したカテゴリー (シグナル経路、遺伝子オントロジー (GO) 等) に基づく機能分類解析を組み合わせたワークフローが提供される。その開発は米国の国立環境衛生科学研究所 (NIEHS)、カナダ保健省、米国 EPA 及び Sciome 社によって行われ、ソフトウェアの設計や実装は Sciome 社が担当した。

BMDExpress では、個々の遺伝子プローブや関連する統計量の用量反応データを近似したパラメータ曲線と、視覚化された一連の分析結果が得られる。トランスクリプトミクスは、ガイドラインに従った毒性試験と比べ、リソースが集約されていないが、BMDExpress アプローチによってトランスクリプトミクスデータを用いた用量反応モデリングが可能になってきた。リスクアセスメントで考慮する必要がある生物学的因子の適用範囲が拡がったと認知されるようになった。

BMDExpress を用いた解析の特徴として、下記の解析フロー (特に機能分類解析モジュール) が挙げられる。その他、視覚化による解析効率の向上も本ソフトの特徴といえる。

BMDExpress で採用されている全ての用量反応曲線のフィッティングモデルは、米国 EPA の BMDS ソフトウェア[64]に収載されている連続値に対応したパラメトリックモデル

---

[62] 大学等における基礎研究の優れた成果を次世代の革新的な診断・治療法の開発につなげるための橋渡し研究。

[63] 薬物等を投与又は適用した場合の各用量におけるトランスクリプトーム (特定の条件下において細胞中に存在する全ての mRNA (又は一次転写産物) の総体) を DNA マイクロアレイ等の技術で網羅的に解析したデータ。

[64] https://www.epa.gov/bmds

(continuous parametric models) である。個々のプローブ[65]や遺伝子レベルのモデル化プロセスは、米国 EPA が毒性評価 (US EPA, 2012) において推奨しているものと概ね一致する。

(2)　主要な機能

(2)-1　プラットフォーム及びアノテーション

BMDExpress では、現在までに各メーカーのプラットフォーム (Affymetrix：20 種、Agilent：8 種、Illumina：4 種及び BioSpyder：3 種) と複数の RNA-Seq アノテーション[66] (RefSeq：5 種及び Ensembl：5 種) 合計 40 種がサポートされている (表 3.6-1) 。

また、GO アノテーションは最新のものが反映され、パスウェイ分類[67]は REACTOME の定義を採用している。アノテーションはカスタムで定義することも可能である。

表 3.6-1　BMDExpress でサポートされているマイクロアレイ及び次世代シークエンスプラットフォーム

ID	名称	プロバイダー	Species
GPL1319	Zebrafish Genome Array	Affymetrix	*Danio rerio*
GPL1322	Drosophila Genome 2.0 Array	Affymetrix	*Drosophila melanogaster*
GPL72	Drosophila Genome Array	Affymetrix	*Drosophila melanogaster*
GPL13158	HT HG-U133+ PM Array Plate	Affymetrix	*Homo sapiens*
GPL570	Human Genome U133 Plus 2.0 Array	Affymetrix	*Homo sapiens*
GPL571	Human Genome U133A 2.0 Array	Affymetrix	*Homo sapiens*
GPL96	Human Genome U133A Array	Affymetrix	*Homo sapiens*
GPL341	Rat Expression 230A Array	Affymetrix	*Homo sapiens*
GPL11180	HT MG-430 PM Array Plate	Affymetrix	*Mus musculus*
GPL16985	HT RG-230 PM Array Plate	Affymetrix	*Mus musculus*
GPL201	Human HG-Focus Target Array	Affymetrix	*Mus musculus*
GPL1261	Mouse Genome 430 2.0 Array	Affymetrix	*Mus musculus*
GPL8321	Mouse Genome 430A 2.0 Array	Affymetrix	*Mus musculus*
GPL81	Murine Genome U74A 2.0 Array	Affymetrix	*Mus musculus*
GPL32	Murine Genome U74A Array	Affymetrix	*Mus musculus*
GPL339	Mouse Expression 430A Array	Affymetrix	*Rattus*

[65] DNA マイクロアレイ（多数の DNA 断片をプラスチック等の基板上に配置した分析器具）の基板上に固定された遺伝子断片。プローブの塩基配列が相補関係にある mRNA と水素結合を形成して 2 本鎖を形成 (ハイブリダイゼーション) する性質を利用し、mRNA の発現の有無や増減を把握できる。

[66] RNA の塩基配列に対して付与される、公式の名称や遺伝子の機能の説明、データベース上の ID 等の「注釈」。

[67] 生体内での遺伝子やタンパク質、化合物の分子間相互作用を「経路」として表現した「パスウェイ」を、その機能等に応じて分類したもの。

ID	名称	プロバイダー	Species
			norvegicus
GPL340	Mouse Expression 430B Array	Affymetrix	*Rattus norvegicus*
GPL342	Rat Expression 230B Array	Affymetrix	*Rattus norvegicus*
GPL1355	Rat Genome 230 2.0 Array	Affymetrix	*Rattus norvegicus*
GPL85	Rat Genome U34 Array	Affymetrix	*Rattus norvegicus*
GPL7300	018972 D. melanogaster Oligo Microarray 4x44K G2519F	Agilent	*Drosophila melanogaster*
GPL6480	014850 Whole Human Genome Microarray 4x44K G4112F	Agilent	*Homo sapiens*
GPL14746	028282 Whole Rat Genome Microarray 4x44K v3	Agilent	*Homo sapiens*
GPL7202	014868 Whole Mouse Genome Microarray 4x44K G4122F	Agilent	*Mus musculus*
GPL11202	026655 Whole Mouse Genome Microarray 4x44K v2	Agilent	*Mus musculus*
GPL21163	074809 SurePrint G3 Mouse GE v2 8x60K Microarray	Agilent	*Mus musculus*
GPL7294	014879 Whole Rat Genome Microarray 4x44K G4131F	Agilent	*Rattus norvegicus*
GPL14664	026437 D. rerio (Zebrafish) Oligo Microarray V3	Agilent	*Rattus norvegicus*
EPA_BSP_Whole Transcriptome_170424	EPA_BSP_WholeTranscriptome_170424	BioSpyder	*Homo sapiens*
S1500_Human	S1500_Human	BioSpyder	*Homo sapiens*
S1500_Rat	S1500_Rat	BioSpyder	*Rattus norvegicus*
danRer10_Ensembl	danRer10_Ensembl	Ensembl	*Danio rerio*
dm6_Ensembl	dm6_Ensembl	Ensembl	*Drosophila melanogaster*
hg19_Ensembl	hg19_Ensembl	Ensembl	*Homo sapiens*
mm10_Ensembl	mm10_Ensembl	Ensembl	*Mus musculus*
rn6_Ensembl	rn6_Ensembl	Ensembl	*Rattus norvegicus*
GPL10558	HumanHT-12 V4.0 expression beadchip	Illumina	*Homo sapiens*
GPL6885	MouseRef-8 v2.0 expression beadchip	Illumina	*Mus musculus*
GPL6887	MouseWG-6 v2.0 expression beadchip	Illumina	*Mus musculus*
GPL6101	ratRef-12 v1.0 expression beadchip	Illumina	*Rattus norvegicus*
danRer10_RefSeq	danRer10_RefSeq	RefSeq	*Danio rerio*

ID	名称	プロバイダー	Species
dm6_RefSeq	dm6_RefSeq	RefSeq	*Drosophila melanogaster*
hg19_RefSeq	hg19_RefSeq	RefSeq	*Homo sapiens*
mm10_RefSeq	mm10_RefSeq	RefSeq	*Mus musculus*
rn6_RefSeq	rn6_RefSeq	RefSeq	*Rattus norvegicus*

(2)-2　フィルタリングと可視化

ANOVA、ウィリアムズの両側傾向検定 (two-sided William's trend test)、ORIOGEN (Order Restricted Inference for Ordered Gene Expression) や変動倍率フィルタリング[68]によって統計学的に有意ではない発現変化を示すプローブを除去する。各手法の概要とデフォルト設定値を表 3.6-2 にまとめた。

なお、BMDExpress には、リアルタイムで表示データのフィルタリングを行うためのユーザーインターフェースが用意されている。パラメータに様々な制限を加えた場合に、結果を可視化情報とデータの両方からリアルタイムで確認できる。

表 3.6-2　統計学的フィルタリング手法の概要とデフォルト設定値

手法	概要	デフォルト設定値
ANOVA 検定	「異なる用量での反応が全て同じである」という帰無仮説の検定。対立仮説は「応答がすべて同じではなく、応答の変化の方向に制限がない」。	P-Value Cutoff：0.05 (0.1、0.01 もデフォルトの選択肢に含む) Multiple Testing Correction：False Filter Out Control Genes：True
ウィリアムズの両側傾向検定	「用量反応がない」という帰無仮説と「単調に増加又は減少する」という対立仮説の比較。等張回帰 (データに対する単調応答にフィットするノンパラメトリック回帰) を使用し、用量特異的応答の推定値及び入手可能な試験統計量を得る。	
ORIOGEN	ノンパラメトリックを利用して重要な遺伝子を同時に識別し、さまざまな不等式パターンに従ってそれらをグループ化する。 BMDExpress 2 では、遺伝子の全体的な有意性の p 値は、別の用量反応プロファイルの和集合に対して用量反応なしの帰無仮説を検定することによって計算する。	P-Value Cutoff：0.05 (0.1、0.01 もデフォルトの選択肢に含む) Number of Initial Bootstrap Samples：500 Number of Maximum Bootstrap Samples：1000 Shrinkage Adjustment Percentile：5.0

[68] 遺伝子発現の膨大な変動データには、わずかな変動や統計学的に有意でない用量依存性変動も含まれる。そのため一般的に、トランスクリプトミクスデータを扱う際は予め設定した変動倍率 (例えば 2 倍) に満たない変動データを除去する。

Multiple Testing Correction：False
Filter Out Control Genes：True

(2)-3　BMDの算出

数理モデルのフィッティングの確認は米国 EPA の BMDS に実装されている手法[69]で行われる。

トランスクリプトミクスデータで得られる連続値モデルの場合、モデル式の項数は用量反応データの用量群の数によって自動的に決定される。なお、指数モデルを使用する場合には、負の値を生成しないようデータを予め正規化することが望ましい。

BMD の算出は、Linux 版、Mac 版、Windows 版のいずれの BMD Express でも可能であるが、BMDS モデル実行ファイルで採用されている浮動小数点演算[70]の影響により、OS 間で算出結果に若干、差が出る可能性がある。経験的に、プラットフォームから得られる結果の差異は限定的であるが、少数のプローブから得られる結果は大きな差となる可能性がある。なお、モデル実行ファイルの検証は Windows でのみ行われている。

BMDExpress で実装されているモデル式を下記に示す。

多項式モデル (Polynomial model)：

$\mu(dose) = \beta_0+\beta_{(1)} dose+ \beta_{(2)} dose^2+ \cdots + \beta_{(n)} dose^n$

(ここで n は多項式の次数)

線形モデル (Linear model)：

多項式モデルで n=1 とする場合

累乗モデル (Power model)：

$\mu(dose)=\gamma+\beta\ dose^\delta$

(ここで $0 < \gamma < 1$, $\beta \geq 0$, and $18 \geq \delta > 0$)

Hill モデル (Hill model)：

$\mu(dose)=\gamma+ (v\ dose^n)/(k^n+ dose^n)$

指数 2 (Exponential 2)：

$\mu(dose)=aexp(sign\ b*dose)$

---

[69] 3.6.3 (2)-3 項及び米国 EPA 作成の Benchmark Dose Technical Guidance の項を参照。

[70] コンピュータ演算における小数の扱い方の 1 つで、小数点の位置を固定せず、仮数、基数、指数の要素で表現することで有限桁数の 2 進数として扱う。例えば、「1.23」は「123×10^{-2}」と表すことができる。このとき、仮数は「123」、基数は「10」、指数は「-2」である。浮動小数点式を用いると、桁数の制約により誤差 (丸め誤差) が生じ、数値の精度が落ちることがある。さらに、演算時の処理オプションが OS によって異なるため、浮動小数点演算の結果は OS 間で異なることがある。

指数 3 (Exponential 3)：

μ(dose)=aexp(sign (b*dose)^d)

指数 4 (Exponential 4)：

μ(dose)=a*(c-(c-1)exp(-1 b*dose))

指数 5 (Exponential 5)：

μ(dose)=a*(c-(c-1)exp(-1 (b*dose)^d))

(2)-4　機能分類

プローブセットは、遺伝子に紐付けられた NCBI Entrez Gene ID に基づいて統合される。複数のプローブセットが単一遺伝子に関連する場合、BMD は平均されて Entrez ID に対応する単一の値が得られる。BMDExpress では、Entrez ID と各種機能分類の照合に基づいた遺伝子のカテゴリー分類が行われ、デフォルトでは下記の機能分類が採用されている。

・GO (生物学的プロセス／分子機能／細胞の構成要素／全て)
・シグナル経路 (Reactome)
・ユーザー定義カテゴリー

(2)-5　組み込まれているデータベース

BMDExpress には、表 3.6-3 に示すデータベース及びソフトウェア情報が組み込まれている。

表 3.6-3　BMDExpress に組み込まれているデータベース及びソフトウェア

名称	内容	URL
Reactome Pathway Database	シグナル経路	https://www.reactome.org/
Gene Ontology	遺伝子オントロジー	http://www.geneontology.org/
DONLP2 optimizer	非線形適合	https://www2.mathematik.tu-darmstadt.de/fbereiche/numerik/staff/spellucci/DONLP2/

(3)　今後の展開

今後、以下の機能の更新・リリースが計画されている。

・モデル平均化

・ノンパラメトリック用量反応モデリング
・自動レポート生成
・IVIVE (in vitro から in vivo への外挿) 機能
・データ閲覧ビューの更新
・より安定した接続マッピングを可能にする wAUC (加重曲線下面積) メトリクス
・GUI のパフォーマンス改善

その他の統計学的手法、バイオインフォマティクス的手法についても現在開発が進められており、今後ワークフローに導入される可能性がある。

3.6.3 利用方法に関する情報

(1) セットアップ方法

BMDExpress は Sciome 社のホームページからダウンロードできる[71]。

(2) 操作の流れ

BMDExpress による解析の流れを示す。図は下記 WEB サイトから引用した。操作の詳細も WEB サイトを参照のこと。解析画面例を図 3.6-1 に示す。

https://github.com/auerbachs/BMDExpress-2/wiki

図 3.6-1 解析画面例

[71] https://www.sciome.com/bmdexpress/ (2019 年 1 月現在)

(2)-1 データ入力

遺伝子発現データファイル (タブ区切り.txt ファイル) をインポートし、アノテーションを付与する。なお、BMDExpress では遺伝子発現データ以外の連続値データの用量反応モデリングも実行可能である。データファイルの読み込み画面を図 3.6-2 に示す。

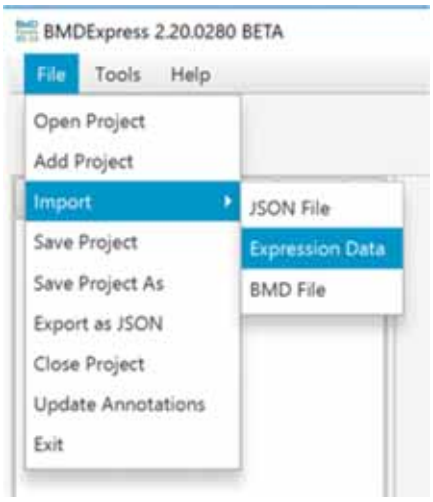


図 3.6-2　データファイルの読み込み画面

データファイルの読み込み後、アレイプラットフォームを指定する。データ取得に使用したプラットフォームがドロップダウンリストに含まれない場合は ”generic” を選択する。さらに、データの前処理に使用したログ変換の種類を選択する。データファイルのプラットフォーム及びログ変換の設定画面を図 3.6-3 に示す。

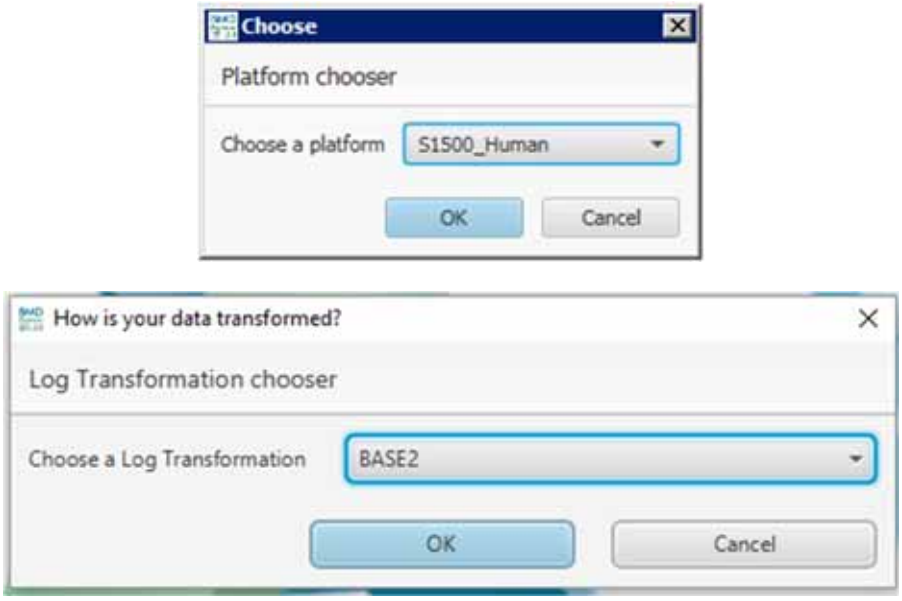


図 3.6-3　データファイルのプラットフォーム及びログ変換の設定画面

(2)-2 フィルタリング

変動倍率フィルタリングと ANOVA や傾向検定を組み合わせ、分析対象とするプローブの数を減らす (任意) 。データフィルタリングの設定画面を図 3.6-4 に示す。

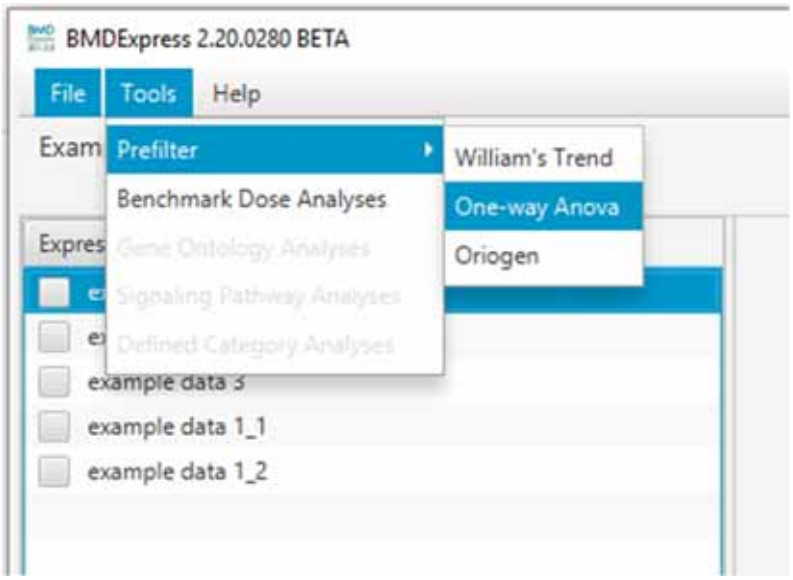


図 3.6-4　データフィルタリングの設定画面

一元配置分散分析 (One-Way ANOVA) 及びウィリアムズの両側傾向検定では以下のオプションを設定する。一元配置分散分析及びウィリアムズの両側傾向検定のオプション設定画面を図 3.6-5 に示す。

P-Value Cutoff：*p* 値に基づくフィルター。デフォルトでは 0.05 に設定。

Multiple Testing Correction：選択した *p* 値に偽陽性率 (False Discovery Rate；FDR) 補正を適用する。

Filter Out Control Genes：プラットフォームに特有の内部標準遺伝子を除去する。

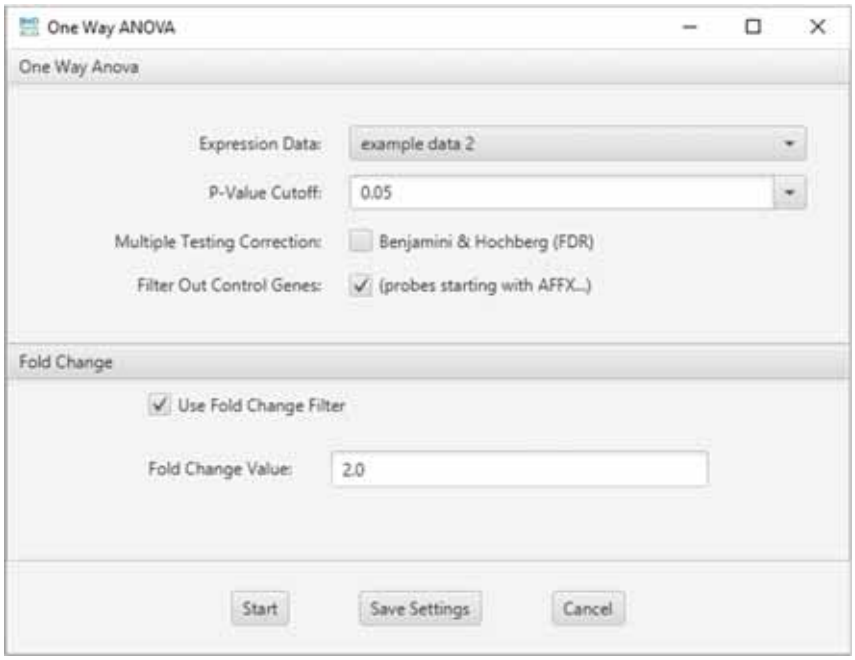


図 3.6-5　一元配置分散分析及びウィリアムズの両側傾向検定のオプション設定画面

ORIOGEN では以下のオプションを設定する。ORIOGEN のオプション設定画面を図 3.6-6 に示す。

P-Value Cutoff：*p* 値に基づくフィルター。デフォルトでは 0.05 に設定。フィルタリング結果画面を図 3.6-7 に示す。

Number of Initial Bootstrap Samples：

計算効率の最大化のため、ORIOGEN では設定の範囲内でサンプル数を徐々に増やしながらブートストラップで *p* 値を算出する。このオプションでは、全プローブの *p* 値の計算に用いるブートストラップの開始時のサンプル数を設定する。

Number of Maximum Bootstrap Samples：

全プローブの *p* 値の計算に用いるブートストラップの最大サンプル数。

Shrinkage Adjustment Percentile：

偽陽性の識別制御に用いる。デフォルト設定の 5.0 は、データセット内の全プローブの 5 パーセンタイル標準偏差を変動閾値に反映する。値を小さくすると、フィルターを通過するプローブの数が減少する可能性がある。

Multiple Testing Correction：選択した *p* 値に偽陽性率 (FDR) 補正を適用する。

Filter Out Control Genes：プラットフォームに特有の内部標準遺伝子を除去する。

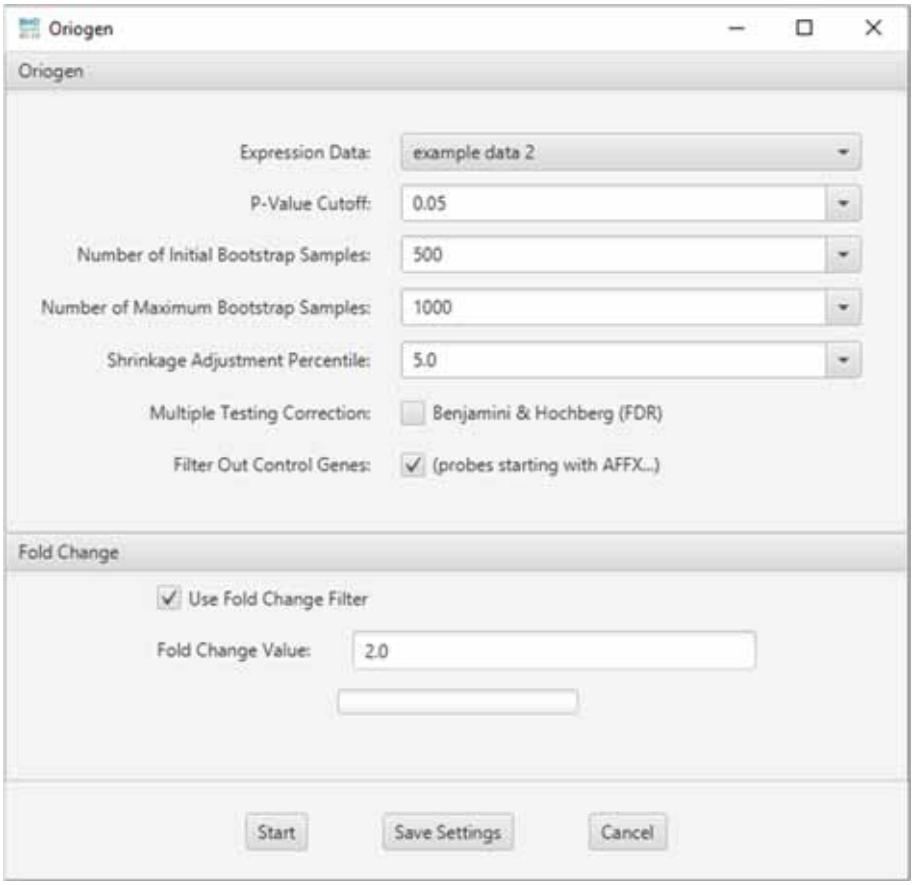


図 3.6-6　ORIOGEN のオプション設定画面

図 3.6-7　フィルタリング結果画面

(2)-3 曲線のフィッティング(Tools > Benchmark Dose Analysis)

数理モデルサブセットに基づき、曲線のフィッティングを計算する。各プローブに対して最もよくフィッティングする数理モデルの BMD とその信頼区間の上限値 (BMDU) 及び下限値 (BMDL) を算出する。BMD オプション設定画面を図 3.6-8 に示す。BMD データオプションは以下の通り設定する。

Continuous Models：

曲線フィッティングの数理モデルを単一又は組合せとして選択する。一部の数理モデルはデフォルトで選択される。また、用量反応データを構成する用量数によって利用可能な数理モデルは異なる。BMDExpress で実装されているモデルは 3.6.2 (2)-3 に記載。

Parameters：

モデルフィッティングは米国 EPA の BMDS に実装されている方法（アルゴリズム）で実行される。

Maximum Iterations：数理モデルの収束基準。

Confidence Level：BMD の信頼区間算出に用いる信頼水準。

Restrict Power：累乗モデルにのみ適用。1 以上の値を指定する。

Constant Variance：選択すると数理モデルに定数分散が適用される。

BMR Factor：BMD の定義に用いる標準偏差の値。Benchmark response (BMR) や critical effect size と同義。デフォルト値は EPA の推奨値でもある 1 (1 標準偏差に相当)。

Model Selection：

Best Poly Model Test：

- Nested Chi Square：ネストしたモデルに対して尤度比検定を行い線形モデルと多項式モデルのいずれかを選択した後、選択されたモデル、Hill モデル、累乗モデルの各モデル分布族間で AIC の比較を行い、AIC が最も低い数理モデルを選択する。
- Lowest AIC：AIC に基づく選択プロセスが実行される。

P-Value Cutoff：最良の線形/多項式モデルの選択における、ネストしたモデルに対して行う尤度比検定法 (カイ二乗検定) で用いる統計学的閾値 (有意水準)。

Flag Hill Model with ‘k’ Parameter < :

データにフィットする数理モデルの 1 つとして Hill モデルが選択されており、その’k’パラメータが最低陽性用量又はその 1/2 や 1/3 よりも小さい場合に、Hill モデルにフラグを立てる。Hill モデルでは非現実的な BMD 及び BMDL 値を算出することがあるため、このオプションが含まれる。

Multiple Threads：

Number of threads：同時に実行するモデルフィット計算の数。

Model Execution Timeout (secs)：個々のプローブに対する数理モデルの実行時間。デフォルトの設定は 600 秒。

BMD 結果画面を図 3.6-9 に示す。

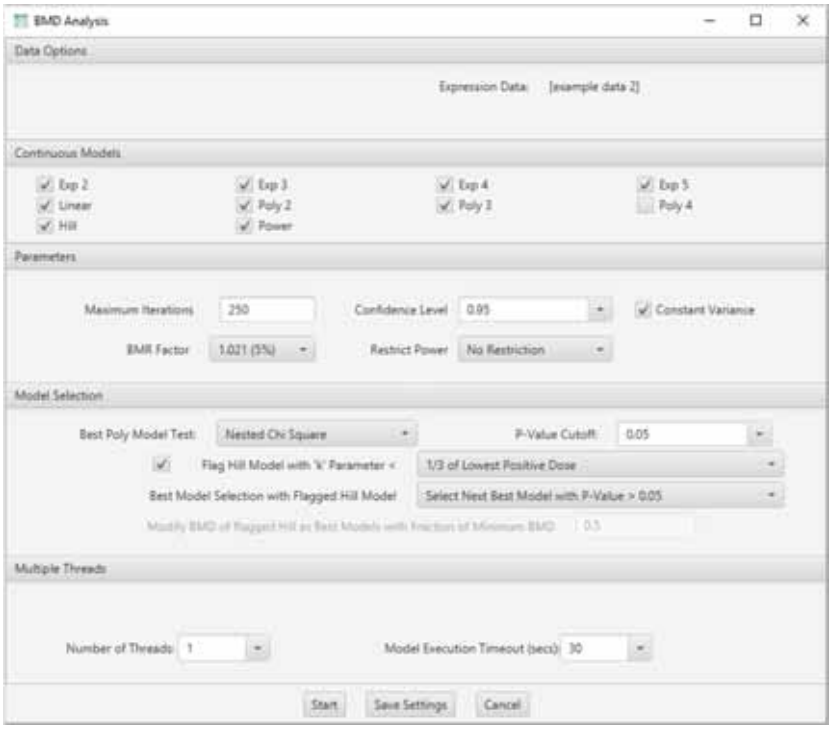


図 3.6-8　BMD オプション設定画面

図 3.6-9　BMD 結果画面

Curve Viewer で各プローブ ID について用量反応プロットとモデルフィット曲線を確認できる。プローブや数理モデルの切替えは、それぞれ「ID」「Model」のプルダウン選択で行う。Curve Viewer 画面を図 3.6-10 に示す。

図 3.6-10　Curve Viewer 画面

(2)-4 遺伝子セット BMD/BMDL/BMDU の算出

各プローブ/プローブセットを固有の遺伝子にマッピングし、遺伝子を GO やシグナリングパスウェイ等に基づいてカテゴリーに分類する。

データ選択エリアで目的の BMD 結果セットを選択して、次のいずれかを選択する。BMD 結果セットの選択画面を図 3.6-11 に示す。

Tools > Gene Ontology Analysis

Tools > Signaling Pathway Analysis

Tools > Defined Category Analysis

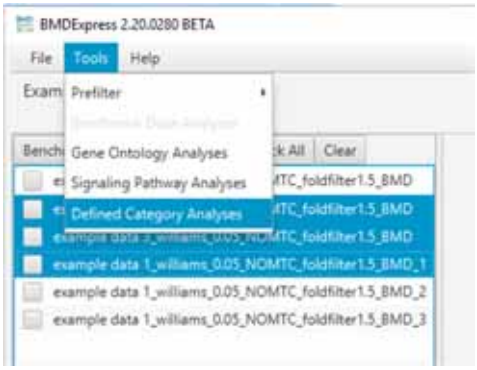


図 3.6-11　BMD 結果セットの選択画面

解析時のオプション (解析対象から除くデータの条件等) を設定し、各カテゴリーの全遺伝子のデータから、カテゴリー遺伝子セットの BMD/BMDL/BMDU の中心傾向とそれに関連する変動性を示す要約値を算出する。解析オプションの設定画面を図 3.6-12 に示す。

図 3.6-12　解析オプションの設定画面 (GO 解析の例)

解析結果は表形式での出力に加え、視覚化された形でも入手できる。図 3.6-13、図 3.6-14、図 3.6-15 に、その例を示す。

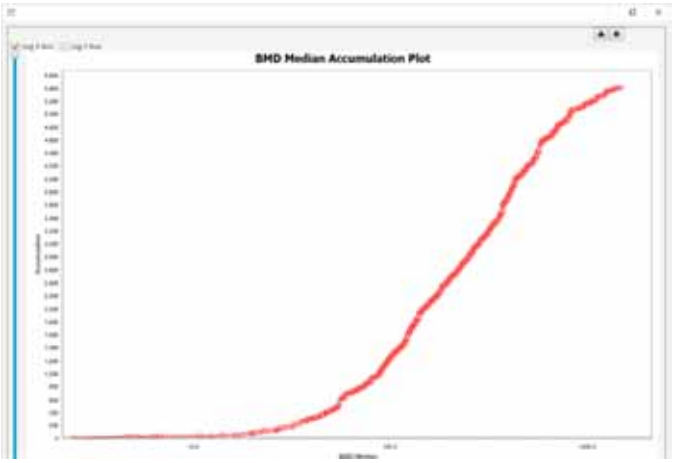


図 3.6-13　BMD 中央値の累積プロット

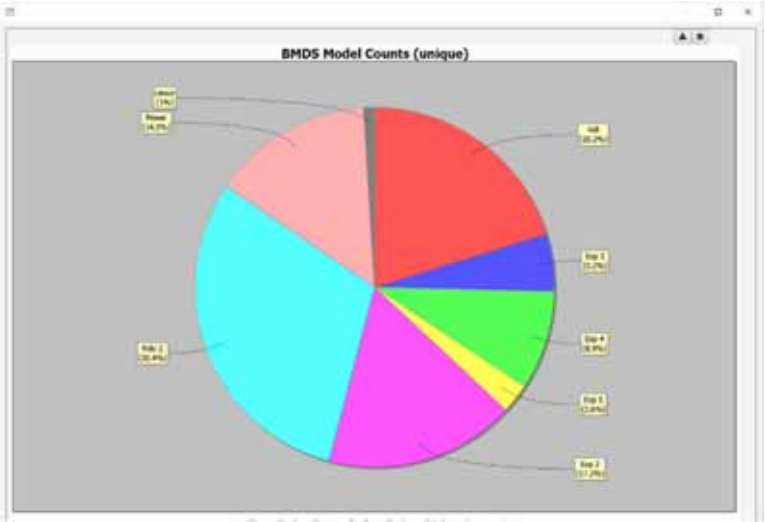


図 3.6-14　最適モデルの円グラフ

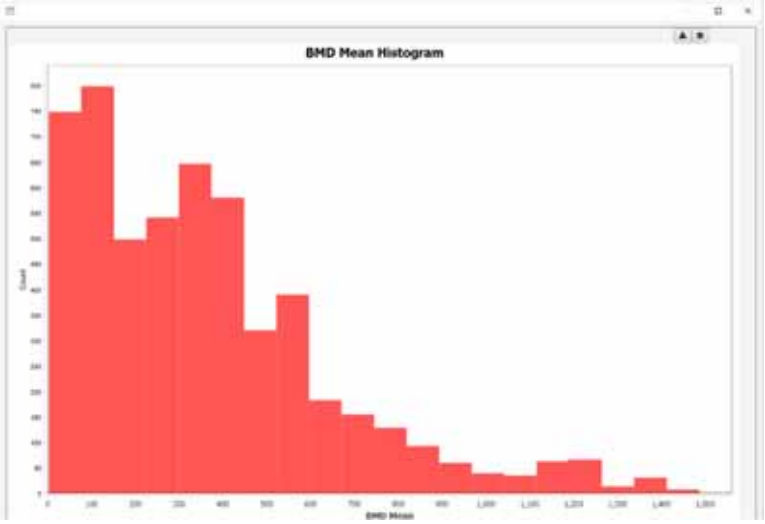


図 3.6-15　BMD 平均値のヒストグラム

3.7　ソフトウェアの主要な機能の比較

			BMDS		PROAST*	EFSA_Benchmark Dose Modelling	MADr-BMD	SPBS	BBMDS**
			Version 3.1	Version 2.7					
解析可能なデータセットの種類			連続値データ (要約、個体)、二値データ、二値データ (ネスト)	連続値データ (要約、個体)、二値データ、二値データ (ネスト)、一般データ (Generic Data)	連続値データ (要約、個体)、二値データ、順序カテゴリーデータ、クラスター化データ	連続値データ (要約、個体)、二値データ、2 進データ、順序カテゴリーデータ	二値データ	連続値データ (個体)	連続値データ、二値データ
モデルの種類	連続値		Exponential、Hill、Linear、Polynomial、Power	Exponential、Hill、Linear、Polynomial、Power	Exponential model 3、Exponential model 5、Hill model 3、Hill model 5	Exponential model 3、Exponential model 5、Hill model 3、Hill model 5、Inv.Expon. m3、Inv.Expon. m5、Logn. DR m3、Logn. DR m5		一次回帰モデル べき関数モデル log (X-1)変換モデル	Linear、Power、Michaelis-Menten model、Hill、Exponential model 2～5
	二値	頻度論	Gamma、Logistic、Log-Logistic、Probit Log Probit、Multistage Weibull、Quantal Linear、Dichotomous Hill 【ネストしたモデル】 Nested Logistic、NCTR*** 【Multi-Tumor モデル】 MS_Combo	Gamma、Logistic、Log-Logistic、Probit、Log Probit、Mutistage、Multistage-Cancer、Quantal Linear、Weibull Dichotomous Hill 【ネストしたモデル】 Nested Logistic、NCTR Rai and van Ryzin、ToxicoDiffusion、Ten Berge 【Multi-Tumor モデル】 MS_Combo	Logistic、Probit、Log-logistic、Log-Probit、Weibull、Gammma、Multistage、Latent variable models (LVMs)	Logistic、Probit、Log-logistic、Log-Probit、Weibull、Gammma、Multistage、Latent variable models (LVMs)	Logistic、Probit、Log-logistic、Log-Probit、Weibull、Gammma、Multistage、Quantal-linear、Quantal-quadratic		
		ベイス法	Gamma、Logistic、Log-Logistic、Probit、Log Probit、Multistage、Weibull、Quantal Linear、Dichotomous Hill						Quantal-linear、Probit、Logistic、Weibull、Multistage (2nd degree)、LogLogistic、LogProbit、Dichotomous Hill
パラメータ制限の可否			ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	デフォルトで設定されている	デフォルトで設定されている	無し	ユーザーによる設定が可能
パラメータ推定法			最尤推定/ベイズ法	最尤推定	最尤推定	最尤推定	最尤推定	最尤推定	マルコフ連鎖モンテカルロサンプリング (MCMC) を適用したベイズ法

		BMDS		PROAST*	EFSA_Benchmark Dose Modelling	MADr-BMD	SPBS	BBMDS**
		Version 3.1	Version 2.7					
連続値データが示す確率分布の初期設定		正規分布/対数正規分布の選択が可能	Exponential 以外のモデルは正規分布のみ (Exponential モデルのみ正規分布/対数正規分布の選択が可能)	対数正規分布 (対数変換の省略か、又は平方根変換を適用することが可能)	対数正規分布		正規分布 (対数正規分布の場合は影響指数を対数(底 10)変換して解析)	対数正規分布
連続値データの BMR 設定方法		中心傾向に基づく方法/ハイブリッド法を選択	中心傾向に基づく方法	中心傾向に基づく方法	中心傾向に基づく方法		ハイブリッド法	中心傾向に基づく方法/ハイブリッド法を選択
使用するモデルの選択		ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能
BMR 値		任意の値を入力 (EPA の基準: 0.1)	任意の値を入力 (EPA の基準: 0.1)	任意の値を入力 (デフォルトは連続値データ 5%、二値データで 10%)	任意の値を入力 (デフォルトは連続値データ 5%、二値データで 10%)	任意の値を入力 (デフォルト：0.1 (10%))	0.02、0.05、0.10 が自動で算出 (通常 0.05 を選択)	任意の値を入力 (連続値データ、二値データ共にデフォルトは 10%)
モデリング結果の評価		尤度比検定の p 値> 0.1 AIC が小さいモデルを推奨 BMD/BMDL 比< 20 (推奨モデルの自動選択機能有り)	尤度比検定の p 値> 0.1 AIC BMD/BMDL 比	AIC に基づく (AIC 間の 2 単位差を臨界値とする) 尤度比検定も選択可能	AIC に基づく（最小値を採用） モデル平均化：（AIC ≤ AICMin + 2 のモデルのみ平均化に用いる）	AIC、BIC、KIC から選択 (デフォルト：AIC) (その他 BICB、AICB、KICB もあるが有効ではない)	・非ばく露群における影響指標の cutoff 値と臨床上の基準値との近似性の確認 ・目視による確認 ・尤度比 (-2logL) 検定、P 値が有意なもの	事後予測 p 値（Posterior predictive p-value: PPP）：0.05 から 0.95 の範囲内 BIC モデルの重み近似法
モデル平均化機能		有り (二値データのみ) 各モデルの BMD 推定値を各モデルの事後確率密度で重み付けして加重平均する	無し	有り (二値データのみ) ブートストラップリサンプリング数：任意に変更 (デフォルト：200)	有り (二値データ/連続値データ) ブートストラップリサンプリング数：任意に変更 (デフォルト：200)	有り (二値データのみ) ブートストラップリサンプリング数：任意に変更 (デフォルト：5000)	無し	有り(二値データ/連続値データ)
信頼区間の設定方法	**基本的な設定方法**	プロファイル尤度法	プロファイル尤度法	プロファイル尤度法	プロファイル尤度法		-	MCMC を適用したベイズ法により算出した事後分布から両端（100-a）/2 を切り取って a%信用区間****とする 事後分布=（事前分布×尤度） / データの分布
	モデル平均化	設定された信頼水準 100(1−α)パーセンタイル値を BMDL として算出		ブートストラップ法	ブートストラップ法	ブートストラップ法		

*Menu バージョンについて記載

**ソフトウェアの調査対象ではないが、参考として記載

***現時点では収載されていない

AIC: Akaike information criterion, BIC: Bayesian information criterion, KIC: Kashyap's information criterion

****信頼区間は頻度主義の考えに基づき、信用区間はベイズ主義の考えに基づく

4. BMD 法に関連する論文等の収集

BMD 法に関する論文のうち以下の項目に該当する情報を収集した。検索は、データベース内でキーワードを用いることにより実施した。検索対象とするデータベースを以下に示す。

・検索したデータベース

PubMed、TOXLINE、BIOSIS Previews、Web of Science Core Collection 等

・収集対象とした文献

① BMD に関するガイダンスに引用されている論文(168 報)

② BMD 法で用いるモデル平均化手法を取り扱った論文 BBMD 法に関する論文、分位変動閾値法に関する論文 (47 報)

③ 2016 年以降に公表された BMD 法を取り扱った論文(上記①及び②に該当するものを除く) (100 報)

5. 資料の和訳

1.～3.の調査で収集した資料について、重要な資料を事務局との協議の上、和訳した。和訳対象とした文書を表 5 に示す。

表 5 和訳対象の文書

資料名	発行年	作成機関 作成者
BMD 法に関するガイダンス		
Update: use of the benchmark dose approach in risk assessment	2017	EFSA
The use of advanced risk assessment methods in answering various types of risk management questions	2006	RIVM
ソフトウェアの操作方法等に関する資料		
PROAST MANUAL Menu version versions 65.6	2018	RIVM
Joint project on Benchmark Dose modelling with RIVM	2018	Wout Slob
Benchmark dose software (BMDS) user manual version 2.6	2016	EPA
Benchmark dose software (BMDS) ver 3.0 user guide	2018	
Manual for BMD Modeling	2017	EFSA
User Manual and Technical Guidance for the Bayesian Benchmark Dose (BBMD) Analysis System Version 1.1.0.0	2016	Shao K and Shapiro AJ