

科学的意見書

リスク評価におけるベンチマークドース法の利用

科学委員会の指針

(Question No EFSA-Q-2005-232)

Adopted on 26 May 2009

専門家委員会メンバー

Susan Barlow, Andrew Chesson, John D. Collins, Albert Flynn, Anthony Hardy, Klaus-Dieter Jany, Ada Knaap, Harry Kuiper, John-Christian Larsen, David Lovell, Pierre Le Neindre, Jan Schans, Josef Schlatter, Vittorio Silano, Staffan Skerfving, Philippe Vannier

要 約

欧州食品安全機関（EFSA）の科学委員会および科学パネル（Scientific Panels）はリスク評価を行う際に、明快でかつ科学的妥当性のある利用すべき解析手法を持つことが必要であるとの考えに至り、これまで伝家宝刀の如く使用されてきた no-observed-adverse-effect level（NOAEL）法の代替として、ベンチマークドース（benchmark dose、BMD）法に関する現存する情報を評価・検討し、その上で BMD 法を EFSA は使用すべきか、またその使用が適切であるのはどのような状況下なのか、に関して勧告するよう科学委員会は求められた。さらに、実験研究に由来する量-反応データを解析する時に BMD 法をどのように用いるべきか、また観察的疫学研究データをこの方法にうまく適用する方法を見つけるよう依頼された。そして最後に、基準点（reference point）の導出に BMD 法を用いるとすると、何等かの不確実係数を選択する必要があるのかどうかについて忠告するよう求められた。

伝統的には、動物実験データを食品物質のリスク評価に使用する時、その物質が遺伝毒性や発がん性を持たないならば、当該物質の臨界影響を引起す際の NOAEL あるいは lowest-observed-adverse-effect level（LOAEL）が許容 1 日摂取量（ADI）のような健康指針値（health-based guidance value）を導出する際の基準点となる。しかしながら、この解析手法は、質的情報を利用することもあるが、量的方法で利用できるデータを有効活用していない。これと対照的に、BMD 法は、動物実験であれ、観察的疫学研究であれ得られた量-反応データを、よりうまく潜在的なリスクを特定し、かつ定量化するよう、徹底的に利用する。

これまで EFSA は、幾つかの科学 Panel が BMD 法を時折適用することはあったものの、系統的に使用するには至らなかった。その上、BMD 算出のための実験データは次から次へと提供されている。また EFSA の科学委員会は、過去の意見書の中で、遺伝毒性と発がん性の両方を有する物質の曝露マージン (margin of exposure) の推定に基準点を導出する BMD 法を利用するよう要望されていた。

科学委員会は、リスク評価における基準点を導出する方法として BMD 法と NOAEL 法の長所・短所を比較し、BMD 法は基準点の導出方法として 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。これは、BMD 法が利用可能な量-反応データを広範囲に使用するとともに、量-反応データにおける不確実性を数量化することを可能にするためである。併せて、BMD 法の使用は、あるベンチマーク反応値 (benchmark response) を用いる結果、より一貫性のある基準点を得ることができる。

BMD 解析用ソフト (software) が現在利用できる状態にある。その上、これは開発途上にあり、しかもモデル平均化や連続データの解析など、意義深い開発が遠からず期待されることに科学委員会は注目している。

また、OECD ガイドラインのような毒性試験指針が改訂される際、BMD 法の特有な性質にも配慮されるであろうと科学委員会は考えている。

BMD 法を用いて作られる健康指針値は、NOAEL 法で算出される値と同程度に保護的である、すなわち膨大な数のリスク評価を平均化したようなものでありと考える。したがって、現在適用されている不確実係数の初期設定値はそのまま適切であるとみなされるし、不確実係数をさらに追加することもない。

科学委員会は、過去に実施された評価の全てを BMD 法で再評価する必要があるとは考えていない。これは、BMD 法も NOAEL 法も似通った結果を出すからである。過去のリスク評価を再評価する必要があると考えられる (例えば、ヒトの曝露が ADI 値に接近しているような) 場合、BMD 法の適用は特に価値がある。

BMD 法は、農薬、添加物、汚染物質などの分類上の区分や由来と関係なく、食品中の全ての化学物質に適用可能である。BMD 法は、①NOAEL を同定することが難しい時、②遺伝毒性や発がん性を有する物質などで曝露マージンのための基準点を提供したい時、③観察的疫学データを用いて量-反応評価を行いたい時などに、特に価値がある。短期的には、以上述べてきたような状況で BMD 法を採用することを EFSA の科学パネルおよびユニットは強く推奨されよう。

長期的には、健康指標や曝露マージンを導出する際の基準点の決定にその選択方法として BMD 法が利用されることを科学委員会は期待している。そのようになるまで、その導入および EFSA での幅広い使用に関する実践上の配慮が必要であることを認識し、しかも、その適用に当たり、一定レベルの専門家の判断やモデル化の専門的技術が要求されていることを考え合わせると、量-反応のモデル化や当該ソフト使用に習熟するための訓練が科学パネルと EFSA ユニットの専門家に勧められる。その上で、科学委員会は、2 年毎に、EFSA の作業の中で BMD 法の履行状況、使用実態および受け入れ状況を検討したいと考えている。

Key words: ベンチマークドース、BMD、BMDL、BMR、NOAEL、量-反応モデル、有害危険因子の特定化、リスク評価、ヒトデータ、BMDS、PROAST

目 次

専門家委員会メンバー	1
要 約	1
目 次	3
EFSA から出された研究背景	4
EFSA からの諮問内容	4
謝 辞	5
評 価	
1. はじめに	6
2. 危険有害因子の確認－潜在的な臨界影響指標の選択	6
3. 危険有害因子の特定化における量－反応データの利用	7
3.1 NOAEL 法	7
3.2 BMD 法	8
3.3 NOAEL と BMD の解釈と特性	9
3.4 NOAEL と BMD 法、幾つかの例	11
3.5 BMR の選択	14
4. 危険有害因子/リスクの特定化の結果とリスクコミュニケーション	15
4.1 健康指針値の設定	15
4.2 遺伝毒性および発がん性がある物質のリスク評価	16
4.3 力価 (potency) の比較	16
4.4 確率論的なリスク評価	16
4.5 BMD と NOAEL、安全性の認識	17
5. BMD 法の適用指針	17
5.1 量－反応データの種類の特定	18
5.2 BMR の特定	18
5.3 量－反応モデルの候補の選択	18
5.4 モデルのあてはめと受け入れ	21
5.5 BMD の推定と基準点としての BMDL の確立	22
5.6 実例	22
5.7 ヒトの量－反応データにおける固有の問題	25
5.8 BMD 解析の報告	25
結論と勧告	28
文 献	29

EFSA から出された研究背景

食品中の化学物質のリスク評価に動物実験データが使用されねばならない時、伝統的には、ある物質の臨界影響の no-observed-adverse-effect level (NOAEL) や lowest-observed-adverse-effect level (LOAEL) がリスク評価における基準点となる。しかしながら、この解析手法は当該化学物質の量－反応曲線の形を直接的には考えていないので、使用可能な全情報が利用されていることになっていない。数字として表される NOAEL/LOAEL 値は、一方で批判的になっているように、臨界影響を同定する方法における感度とともに、用いられる用量間隔の選択やその選択された用量における対象動物数に依存する。

もし実験動物研究に由来する特定物質の量－反応曲線が広範囲に活用されるなら、ヒトで期待される曝露レベルにおける当該物質のリスク評価は潜在的なリスクをより良く特定し、かつ定量化できるようになる (FOSIE 2002)。ベンチマークドース (BMD) 法はこれを実行する手段を提供する (EPA 1995)。BMD 法は、低いが、測定可能な標的臓器影響を引き起こす (例えば、体重量いし臓器重量を 5% 減少させる、あるいは腎毒性の発症率を 10% 増加させる) 用量を推定する。当該推定用量の信頼下限値を計算することにより、そのデータの不確実性と変動性を考慮することになる。この信頼下限値を当初ベンチマークドースと呼んでいたが (Crump 1984)、後になって変わり、BMD 信頼下限値 (BMDL) として知られるようになった。

BMD 法は、疫学的研究から得られるヒトデータにも適用されている。リスク評価で疫学データを使って BMDL を計算する時、ヒト曝露の定量的推定にまつわる不確実性や交絡因子の影響などの付加的な要因に対しても考慮されねばならない。

食品に関する前 EC 科学委員会や欧州食品安全機関 (EFSA) は BMD 法をそれ程計画的に体系立てて使用していなかった。一方で、BMD を計算した実験データは次から次へと提出されている。EFSA は、BMDL を遺伝毒性および発がん性を有する物質の曝露マージンの推定のための基準点としたらよいと提案した—これは Joint FAO/WHO Expert Committee on Food Additives (JECFA 2006a) が 2005 年の JECFA 会議で使用した解析手法でもある。リスクが EFSA 科学委員会や科学パネルによって評価される時、使用する解析手法が明快でかつ科学的に妥当性のあるものであることが求められるのは当然である。その提案された任務は、まさに BMD 法の有用性に関する現存する情報を評価・検討することであり、また EFSA がこの解析手段を使用すべきかとともに、これが適切と考えられる状況とはどのようなものかについて忠告することである。さらに、BMDL を基準点として使用する時、適切な不確実係数を選択するための指針も必要となる。

EFSA からの諮問内容

この問題に関する自発的な取組に対する科学委員会の提案に続き、欧州食品安全機関 (EFSA) は科学委員会に以下の諮問を行う。

- 伝統的に用いられている NOAEL 値に対立するものとして、ベンチマークドース法の使用に伴う利点と限界について検討しなさい
- EFSA パネルおよび専門家グループが実施するリスク評価で、利用されるべきより好ましい解析方法についての意見をまとめなさい

- リスク評価法として BMD 法が最も相応しい解析手法であるとする、EFSA のパネルおよび専門家グループが実施する規制用リスク評価の中で、適切に、この解析方法を混乱なく導入できるようにする指針、それに、BMDL を出発点（すなわち、基準点）として使用する際に適切な不確実係数を選択する指針を提示しなさい

謝 辞

EFSA は、この意見書の作成にあたり、ベンチマークドースに関するワーキンググループの各委員— Ada Knaap (座長)、Alan Boobis、Hans Bouwmeester、Paul Brantom、Esben Budtz-Jørgensen、Lutz Edler、Claudia Fruijtier-Pöloth (2008 年 7 月まで同委員)、Philippe Grandjean、Inge Mangelsdorf、Douglas McGregor、Andrew Renwick、Ivonne Rietjens、Salomon Sand、Josef Schlatter、Wout Slob —に感謝申し上げます。

EFSA はまた David Bellinger 教授（ハーバード大学公衆衛生大学院環境保健学講座）が 5.7 章を寄稿して頂いたことに対し深謝したい。

評 価

1. はじめに

本書は主に実験研究で得られる量－反応データの解析を扱うが、観察的疫学研究から得られるデータへの適用も検討する。同様のアプローチは生態毒性研究にも適用できるかもしれないが、本書ではこれ以上考えない。毒性研究は試験物質の潜在的な有害影響を同定し、かつその特徴を明らかにするために行う。これらの研究で得られるデータ解析には、ヒトの健康リスク評価で出発点として利用される用量 (dose) を明らかにすることも含まれる。この目的のために使用される（もしくは導出される）用量は基準点 (reference point) と本書では呼ぶ。この用語は遺伝毒性および発がん性の両方を有する物質のリスク評価のための調和的アプローチに関する欧州食品安全機関科学委員会の意見書 (EFSA 2005) の中で既に使用されており、それゆえ米国環境保護庁のような他の組織で使用されている同義語、point of departure (PoD) よりも好まれる。

no-observed-adverse-effect level (NOAEL) は非遺伝毒性物質のリスク評価で広く用いられてきた基準点である。NOAEL はヒトのリスク評価の規制過程の中で使用された長い歴史を持ち、食品添加物や残留農薬の許容 1 日摂取量 (ADI) や汚染物質の耐容 1 日摂取量 (TDI) ないし耐容週間摂取量 (TWI) のような健康指針値 (health-based guidance) を推定するための慣例的な基準点である。NOAEL には幾つかの限界があることが知られている（後章を参照）ので、別の方法が近年検討されてきた。ベンチマークドース (BMD) 法は基準点を定義する別の方法として現れた。ここでは基準点を算出する方法として BMD 法を検討し、これまで基準点を決定するのに慣例的に用いられてきた NOAEL と比較する。

EFSA (2005) や Joint FAO/WHO Expert Committee on Food Additives (JECFA 2006a) は、遺伝毒性および発ガン性をともに有する物質の曝露マージン [訳者注: ‘margin of exposure’ は米国環境保護庁が定義した、腫瘍バイオアッセイ、疫学研究、生物学的マーカーなどから導出される用量の比、すなわち用量－反応関係における 10% 反応率の量を実際の（あるいは見積もられた）ヒトの曝露量で割った値を指す] を計算するため、その基準点を導出する BMD 法の利用を提案してきた。米国環境保護庁は、非遺伝毒性物質の reference dose (*RfD*) の設定や、発ガン物質の基準点や cancer slope factors を計算するために、BMD 法を長年使用してきた（尤も、NOAEL も前者の目的で個別的には利用している）。WHO、European Medicines Agency、EFSA などは個別の事例で BMD 法をリスク評価の一部として利用していたが、非発がん性物質の基準点を設定するために幅広く BMD 法を利用しているところは米国以外ではない。

本書の 1～3 章は、NOAEL と BMD 法の基本となる概念を幾つかの実例を交えて論じる。これらの章では、科学委員会がなぜ BMD 法をより強力な解析手法であると考えたのか述べる。4 章では、危険有害因子 (hazard) / リスクの特定化 (characterisation) とリスクコミュニケーションのために、BMD 法の使用がもたらす潜在的影響について論じ、5 章では BMD 法を実践の場でどのように適用できるのかを要約し、またその適用における幾つかの勧告を記す。Appendix では、ソフト (software) の使用方法とその出力結果をどのように解釈すべきかについて実践的な指針を示す。

2. 危険有害因子の確認－潜在的な臨界影響指標の選択

毒性研究は、物質によって生じる有害影響を確認し、その測定された有害影響に対する量－反応関係を明示するために計画される。ヒトの量－反応データが利用可能な場合もあるが、多くのリスク評価は動物実験に依存している。危険有害因子の特定化とはヒトの健康に関連する可能な臨界影響指標 (critical endpoint) を見分けることである。重要な点は、観察された影響が用量依存関係を有するかどうかを検討することである。伝統的には、これは（2群間の比較からなる統計学的評価などの）慣例的な統計手法とともに、目測 (visual inspection) で行われる。代わりになるべきものとして、BMD 法のような量－反応関係のモデル化手法が利用できる。

臨界影響の選択が統計学的手続きによってのみで決められることがあってはならない。さらなる毒性的議論は、十分に蓄積された毒性学データの評価の中で行われことが重要となる。BMD 法を利用したからと言って、反応データの批判的評価やヒトの健康影響への適切な評価が不要になるわけではない。

この第一段階の結果は、次章以下で述べるように、より詳細に解析されるべき可能な臨界影響指標を特定することである。

3. 危険有害因子の特定化における量－反応データの利用

量－反応関係の性質が危険有害因子の特定化に際して詳細に調べられる。多くの場合、毒性影響を明らかにする過程における第一段階は、実験条件下にある試験動物ではっきりと認知できる有害健康影響のない用量を確認することである。次に、種間および種内差、改善の研究における特性、欠損データのような不確実性や多様性を考慮しつつも、はっきりと認知できる有害健康影響は見られないと自信を持って言えるヒトの摂取量を算出するために、毒性研究から得られる基準点が用いられる。

リスク評価で危険有害因子を特定するために動物毒性研究の中で、一定の範囲の用量を用いることが必要となる。動物実験に用いられる小さな標本数の中で、観察が可能となる十分に大きな影響を引き起こす用量が求められる。また、量－反応関係の中で低用量部分の情報を提供することも必要である。

実験的バラツキや生物学的変動は反応の測定に影響する。その結果、各々の用量における反応平均値には統計的誤差が含まれる。従って、データにまつわる統計的誤差によって不適切な生物学的結論が導き出されないようにするために、量－反応データは統計学的方法で解析される必要がある。近年、そのような基準点を導き出すために2つの統計方法（すなわち、NOAEL 法と BMD 法）がある。本章はこれら2つの方法を概説し、各々の方法の長所と限界を比べる。

3.1 NOAEL 法

NOAEL 法は閾値を持つと考えられる全ての毒性影響に適用可能である。NOAEL は次のようにして導かれる

- － 各々の有害影響／影響指標に対して、専門家の意見とともに、対照群と各々の処理レベルを比較する統計学的検定を用いて、影響が検出されない最も高い実験用量を見つける
- － 幾つか検出された有害影響における NOAEL の中で最も低い NOAEL（すなわち、臨界影響における NOAEL）を当該研究における NOAEL とする。

すなわち、NOAEL は特定の実験で有害影響が観察されない最も高い用量である。数字で表した NOAEL 値は、研究が考案された時の用量レベルの選択や、有害影響を検出する研究自体の能力に依存す

る。検出力の低い研究（例えば、小さい標本数）や感度の低い方法は比較的大きな影響しか検出できず、より高い NOAEL になってしまう。もし全ての用量レベルで有意な影響があるなら、当該研究で用いられた最も低い用量が lowest-observed-adverse-effect level (LOAEL) と設定されることになる。

NOAEL の確定は必ずしも純粋に統計学に基づく決定であるとは限らない、と考えておくべきである。これには利点と欠点がある。評価者がデータによって十分に裏打ちされない NOAEL を却下することもあるし、異なる決定を下す可能性もありうる。NOAEL を確定する際に考慮する事項として、一貫性のある量－反応関係があるのか、また NOAEL の近傍で量－反応曲線が急峻な変化をするかどうかが含まれる。量－反応関係が低濃度域で単調でない状況では、バックグラウンド反応と統計的に有意な差が認められても、ある用量群が NOAEL 同定の際に無視される可能性がある。同様に、反応の大きさが小さい場合、たとえ統計的に有意であっても、ある評価者はそれを有害と見なさず、より高い用量を NOAEL とするかもしれない。逆に、LOAEL 以下の用量群の反応が小さく、有意でない増加を示していても、その量－反応関係が非常に急峻である場合、その NOAEL はもう一段階低い用量（つまり LOAEL より 2 段階低い用量群）として同定されるかもしれない。このような決定は専門家の判断を必要とするが、異なる評価者が異なる決定をする可能性があるというのも事実である。後者の例として、動物飼料に使用される添加剤物質に関する EFSA パネル (EFSA 2009a) と JECFA (2006b) は動物用医薬品である Ractopamine の残留物評価で異なる決定を過去に行った。

3.2 BMD 法

BMD 法はあらゆる毒性影響に適用できる。ある種の影響指標に対する総合的な量－反応関係の形を推定するのに、量－反応データの全てを利用する。その BMD は、推定された量－反応曲線から導かれ、ベンチマーク反応値 (benchmark response, 以下 BMR) と呼ばれるある特定の反応変化 (= 異常反応) をもたらす時の曝露量レベルを指す (3.5 章参照)。また、BMDL (benchmark dose lower confidence limit) は BMD のとり得る信頼区間の下限值であり、この値が基準点として通常用いられる。

BMD 法の基本概念を図 1 に示す。この図では、仮に BMR を 5% として計算する時の BMDL を表し、BMDL₀₅ はその異常反応率が 5% 以下となると思われるときの用量と解釈される。“思われる (likely)” という用語は統計的な信頼性レベルで定義され、通常 95% の信頼性である。

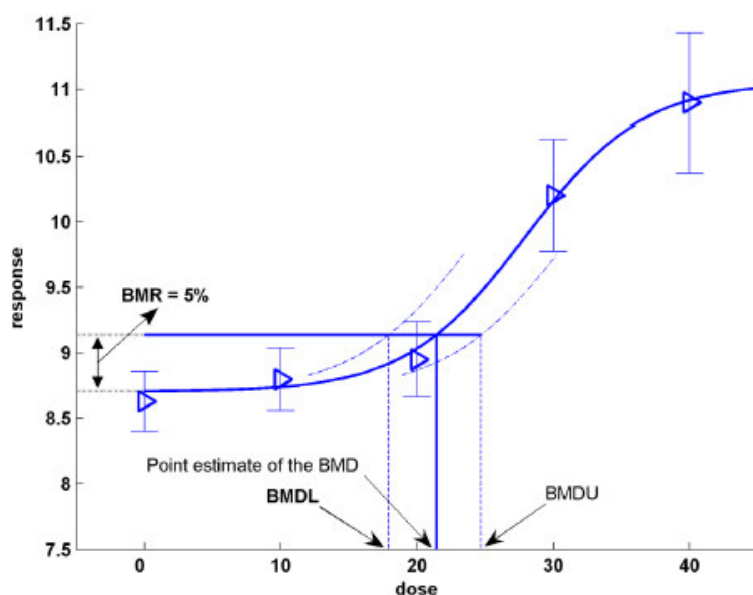


図 1. 仮想上の連続データを用いて描かれた BMD 法の基本概念 — 観察された平均反応率 (△) がその信頼区間とともにプロットされている。実線はこれら △ 点に最も当てはまる量－反応モデルである。この曲線から BMD (点推定) が決定され、その値は通常 BMR で表される (すなわち、低いと測定可能な変化をもたらす) 反応に対応する曝露量として定義される。点線は、用量関数 (function) として表した際の、影響の大きさの 95% 信頼上限と下限を各々示している。こ

これらの点線と横青線との交点が BMD の下限値および上限値となり、各々 BMDL と BMDU で表される。BMR は観察された平均バックグラウンド反応に関連する変化ではなく、適合したモデルから予測されるバックグラウンド反応に関連する変化として定義されていることを忘れないで頂きたい。この区別は重要であり、通常、あてはめられた曲線は観察されたバックグラウンド反応に必ずしも一致しないためである（そのため、この BMR を観察されたバックグラウンド反応に加えても、BMD 値での量－反応曲線との正確な交点を示さないであろう）。図では、BMD はバックグラウンド値と比較して 5% の反応変化（BMR=5%）と対応している。あてはめられた曲線の推定バックグラウンド反応の値は 8.7 であることから、その 5% 分の増加を加えると、 $9.14 (=8.7+0.05 \cdot 8.7)$ に等しい。つまり、21.50 の BMD₀₅ は反応 9.14 における水平線とあてはめられた量－反応モデルとの交点から得られる。この例では、BMDL₀₅ は 18 となる。

ある特定の研究で BMDL を同定するために必要なステップは

- － 低い測定可能な反応レベルの明確化（例えば、バックグラウンド反応と比べて 5% または 10% の反応増加あるいは減少）。これは BMR と呼ばれる。
- － 量－反応モデルをあてはめ、統計的規準に従ってデータを記述するモデルから BMD と BMDL を算出し、各々の有害影響/影響指標に対する BMDL 値を得る（5.3 章と 5.4 章を参照）
- － 各々の潜在的な臨界影響に対する BMDL を選択する（5.5 章を参照）
- － 異なる潜在的な臨界影響の或る範囲の BMDL 値から、ある総合的研究の BMDL（つまり、研究の重要な BMDL）を得る（5.5 章を参照）

原則として、BMD 法は関連する研究で測定される全ての影響指標に適用される。その臨界影響は、その後 NOAEL 法一すなわち、最も低い BMDL を示す影響指標と同様の方法で選択され、さらなる毒性的議論が、まさに NOAEL 法の場合と同様に、行われる。

NOAEL 法では、対照群データのバラツキが大きい、定量的測定法の検出力が悪い、標本数が小さいなどのデータは高い NOAEL 値を生じるので、可能性のある基準点としての NOAEL を算出する際に当該データセットを採用するかどうかの決断は重要である。データの判定はそれゆえ専門家の判断に委ねられる。対照的に、BMD 法はそれ自体、当該データの全特性を考慮することによって、データ品質を一定に保ったままで量的に評価する。データが比較的貧弱ないし情報価値のない場合、そのデータセットから生じる BMDL は低くなる傾向がある。しかし、その BMDL 値の意味は定義された時のままであり、関連する影響の大きさ（effect size）は使用する BMR よりも大きくなりそうにない用量レベルを映し出す。

にもかかわらず、データがあまりに貧弱で、可能な基準点としてその BMDL を使用するには忍びないということはあるし、そのようなデータセットは捨てる必要があるかもしれない。これは BMD の信頼区間が広いとか、モデルを変える度に BMDL が大きく変わる時に決断されるだろう。この問題に関する指針は 5.5 章に示す。

最もよく知られた BMD のソフトは米国環境保護庁で開発された Benchmark Dose Software (BMDS) と RIVM が開発した PROAST である。

3.3 NOAEL と BMD の解釈と特性

NOAEL は通常バックグラウンド反応と比べて反応に統計学的有意差が見られない用量レベルである。これは NOAEL が、当該研究の中で影響が検出できないくらい小さい用量であることを反映しており、それゆえ NOAEL 値での影響の大きさは恐らく未知のままである。これについての洞察を得る直接の方法

は、対照群と NOAEL 用量群の間で観察された反応変化の信頼区間を計算することである。

NOAEL は必ずしも“影響がない”用量ではないが、時々そのように解釈される。しかし、3.5 章で議論される動物実験が示すように、NOAEL 値の影響の大きさは、多くの研究の平均として、非連続的な値をとる反応では 10%、また連続的な値をとる反応では 5%に近い。ここで、非連続的な値をとる反応の影響の大きさは過剰リスク (extra risk) と表現される。過剰リスクは、反応頻度 (追加リスク additional risk) を対照集団の影響のない分画 ($100 - [\text{バックグラウンド反応}]$, %) で割った絶対的变化として定義される (例えば、追加リスクが 8.5% でバックグラウンド反応が 15% であれば、過剰リスクは $8.5/(100-15) = 10\%$ となる)。連続的な値をとる反応では、影響の大きさは反応平均値のパーセント変化として示される (例えば、平均体重が 100 g から 95 g に減った場合、影響の大きさは 5% となる)。個々の NOAEL では、統計的に有意でない影響の大きさはこれらの値より小さい (あるいは大きい) かもしれない。最近評価された物質における NOAEL 値での影響の上限値が計算され、表 1 に示されている。しかし、そのような付随的な情報は NOAEL 法の目下の適用において報告されていない。BMD 法では、その可能な影響の大きさ (つまり、BMR) はその定義上当然のこととして知られている。

Table 1: Illustrations of upper bounds⁷ of effect at NOAELs related to 10 compounds evaluated previously by JMPR or EFSA.

Compound (source +year)	Endpoint	Quantal data	Continuous data	References
		Upper bound extra risk (%) ⁸	Upper bound percent change (%) ⁹	
Thiodicarb (JMPR 2000)	splenic extramedullary haematopoiesis	21		www.inchem.org/documents/jmpr/jmpmono/v00pr09.htm
Carbaryl (JMPR 2001)	vascular tumours	15		www.inchem.org/documents/jmpr/jmpmono/2001pr02.htm
Spinosad (JMPR 2001)	thyroid epithelial cell vacuolation	2.7		www.inchem.org/documents/jmpr/jmpmono/2001pr12.htm
Flutolanil (JMPR 2002)	erythrocyte volume fraction haemoglobin concentration mean corpuscular haemoglobin decreased cellular elements in the spleen	30	9 9.7 3	www.inchem.org/documents/jmpr/jmpmono/2002pr07.htm
Metalaxyl (JMPR 2002)	serum alkaline phosphatase activity serum AST		260 100	www.inchem.org/documents/jmpr/jmpmono/2002pr09.htm
Cyprodinil (JMPR 2003)	spongiosis hepatitis	5.1		www.inchem.org/documents/jmpr/jmpmono/v2003pr03.htm
Famoxadone (JMPR 2003)	cataracts microscopic lenticular degeneration	29 29		www.inchem.org/documents/jmpr/jmpmono/v2003pr05.htm
Tributyltin (EFSA 2004)	testis weight		9.1	www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178620762916.htm
Fumonisin (EFSA 2005)	nephrosis	8.6		www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178620807204.htm
Deoxynivalenol (EFSA 2004)	body weight		10.5	www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178620763160.htm
Ethyl lauroyl arginate (EFSA 2007)	white blood cell counts		23	www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1178622334379.htm

⁷ As calculated by the Scientific Committee.

⁸ Two-sided 90%-confidence interval for extra risk was calculated by the likelihood profile method

⁹ Two-sided 90%-confidence interval was calculated for the difference on log-scale, and then transformed back, resulting in the confidence interval for percent change (see Slob 2002 for further statistical assumptions).

幾つかの化合物に対して、科学委員会は表 1 に要約している影響の大きさの上限値を確定した。非連続的な値をとる影響指標では、過剰リスクに関係する上限値は約 3 から 30%の間で変わる。ある場合に、NOAEL 値における過剰リスクが 10% (これは非連続データで推奨されている BMR である) を超えることもありうる (5.4 章を参照)。同様に、表 1 の限られた化合物では、連続的な値をとる影響指標における NOAEL 値の影響の大きさが 3%と小さい上限値のこともある。しかし、多くの場合、それは 10%台であり、連続データ時の BMR の勧告値 5% (5.4 章を参照) と比べて高い。表 1 の一つは、変動の大きい臨床化学パラメーターであり、影響の上限値は 260%と高かった。

BMD 法もある統計学的方法を用いるのであるが、データの一部を用いて 2 群間の比較を行う代わりに、データセット一式の情報をを用いる。また、NOAEL 法は事前に設定した用量に限られるが、BMD 法は設定された用量の間を補間することも可能である。このため、NOAEL は既設の用量のみ表し、かつその対応する潜在的な影響の大きさは通常計算されないが、BMDL は、BMDL として計算される値の‘定義済

みの影響の大きさ’ といつも対応する。したがって、BMDL 値は BMR によって定義される用量における影響の上限値を明確に示すことによって、NOAEL よりも多くの情報を提供することになる。

BMD 法が本来もっている意義は計算された BMD の不確実性の評価である（それは BMD のまわりの信頼区間を示すことによって反映される）。この不確実性の評価は通常 NOAEL 法では行われない。これは、データが制限されている場合や大きくばらついている場合に、特に重要な相違点である。NOAEL はそのような場合でもしばしば導出されるが、得られた値の不確実性は通常示されない。

リスク評価を行うための NOAEL 法のデータ要件は研究デザインの国際的に認められたガイドライン（例えば、化学物質試験のための OECD ガイドライン）に組み入れられている。しかし、データの有用性はこれら研究デザイン（例えば、用量群の数や群の大きさ）に関する全体的な側面だけでなく、実際に選択される用量や観察される反応の変動などの研究の質の面にも依存する。NOAEL 法におけるデータの有用性は研究デザインなどの生得的な考えにかなりの程度依存するが、BMD 法はこれまで議論してきたようにこれらの因子に殆ど拘束されない。その上、当該データセットの特異性を考えながらデータの質を評価することで（例えば、データのバラツキや量－反応の情報）、さらに前に行く。このように、データセットが基準点の導出に耐えられるかどうかに関して、より情報に基づく決定が可能となる。BMDL は（下限値を推定することによって）当該データセットの質にまで十分配慮しているので、データの質が NOAEL の場合と比べて重大な問題にならないだろう。

研究デザインに関する現行の国際的ガイドラインは、NOAEL 法を想定して考えて練り上げられたが、BMD 法を適用することに対して何等異議を唱えていない。しかしながら、現行のガイドラインは BMD 法に対して最適とは言えない。これらのガイドラインは、例えば OECD Test Guideline Programme の枠内で改訂されるので、BMD 法に適したデータを提供することが必要となる。

3.4 NOAEL と BMD 法、いくつかの例

本章は量－反応評価のために NOAEL と BMD 法の計算例を示す。1 番目と 2 番目の例題は、毒性研究における実在の量－反応データが連続および非連続の場合であり、各々 NOAEL 法と BMD 法を説明するのに使う。3 番目の例題はヒトにおける量－反応データである。

Table 2: Pair-wise comparison of dose groups, data from Fig. 2

Dose	N	Geometric mean	ES (%)	Lower 95%-confidence bound (one sided) of ES (%)**	Upper 95%-confidence bound (one sided) of ES (%)**	t-statistic	p-value
0	5	7.16					
75	5	7.35	2.6	-2.0	7.4	0.99	0.17
300	5	7.72	7.8	3.0	12.9	2.88	0.006
600	4*	8.19	14.4	9.0	20.1	4.85	< 0.001

GM = geometric mean

ES = effect size (in percent change compared to response at dose zero)

* = one animal died at the top dose

** Two-sided 90%-confidence interval was calculated for the difference on log-scale, and then transformed back, resulting in the confidence interval for percent change (see Slob 2002 for further statistical assumptions)

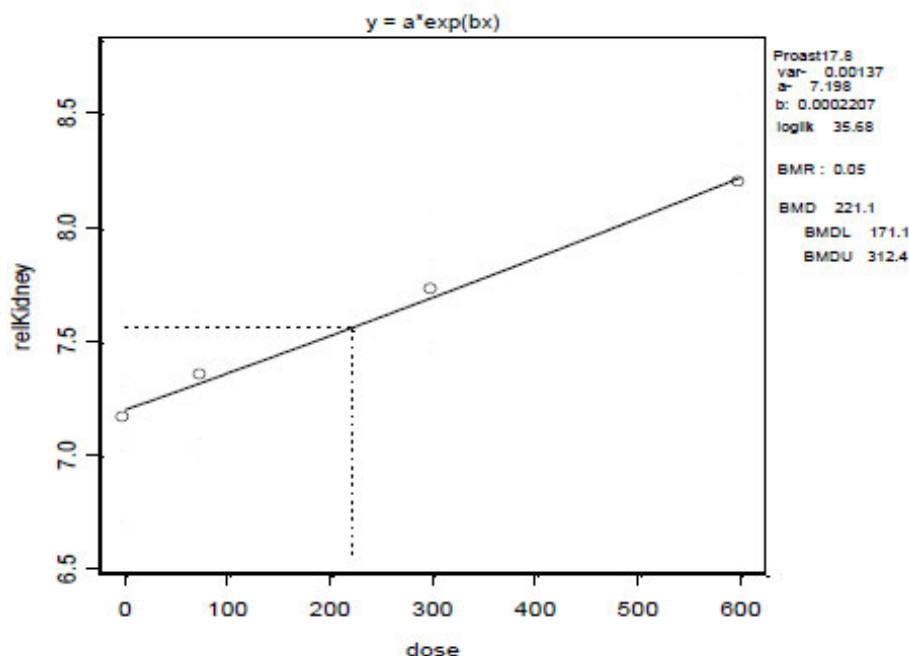
例 1、連続的な量－反応データ

この例は相対腎重量に関するもので、連続反応データである。連続データでは各々の対象で数字として示される値が測定される。連続反応の BMR は影響における程度ないし重篤度の測定と解釈されるが、非連続データの BMR は発症率の変化を反映する（例 2 参照）。

NOAEL 法と BMD 法との相違点を示すために、これら 2 方法を当該データセットに適用してみよう。NOAEL 法において各々の用量群を対照群の反応と比較すると、低用量 75 mg/kg 群の影響は有意でないが、用量 300 mg/kg 群と 600 mg/kg 群の影響（反応）は、表 2 の最後列（*P* 値）に示したように、統計的に $P < 0.05$ で有意に異なる。統計的に有意な結果の規準に照らすと、75 mg/kg が NOAEL とみなされるだろう。にもかかわらず、この用量で起こる影響（反応）の 95%信頼上限値（片側）は相対腎重量で 7.4% の増加である（表 2）。

同じデータセットで BMD 法を説明するために、そのデータを量-反応モデル ($y = a \cdot \exp(b \cdot x)$) にあてはめ、相対腎重量が 5% の増加を表す BMR を用いた（図 2 参照）。このモデルの解析結果は BMR = 5% で BMD₀₅ が 221 mg/kg となり、BMDL₀₅ は 171 mg/kg となる（図 2 の説明文参照）。

図 2. 各々の用量群に動物 5 匹を用いたときの相対腎重量 (mg/kg BW) — 大きな丸は各群の幾何平均を示す。実線は適合させた量-反応モデルを示し、破線は BMR = 5% のときの BMD を示す。本書を準備する時点では BMDs ではこの図を作成できなかった。



このデータセットでは BMDL₀₅ は NOAEL より高い（各々、171 mg/kg と 75 mg/kg）。にもかかわらず、BMDL₀₅ における影響の大きさは（95%の信頼度で）5%より小さいと述べることができるが、2 群間データの比較（表 2 参照）では NOAEL 値の影響の大きさは（こちらも 95%の信頼度で）7.4%より小さいとしか言えない。換言すると、BMDL で算出される基準点の信頼度は NOAEL で算出されるものよりも大きい [訳者注: 区間推定の範囲が広いほど、信頼度は低いとみなされる]。すなわち、NOAEL 法でバックグラウンド反応と個々の用量群との間で比較するよりも、BMD 法で全データセットを解析する方がデータ情報をより上手く活用していることを示す。

例 2：非連続反応データ

非連続反応データでは、BMR はバックグラウンド発症率からのある特定増加（specified increase）と定義される。図 3 に示す例では、10% の BMR（過剰リスク）が用いられる。

ここで、中等用量と低用量の発症率はバックグラウンド反応の発症率と比べた時有意差が認められな

かった。すなわち、450 mg/kg の中等用量がこの研究の影響指標における NOAEL となる。この場合、バックグラウンド反応との 2 群間比較を行うと、NOAEL 値における影響の大きさに対して非常に大きな 95%信頼上限値（片側）を生じる（すなわち、過剰リスクの値は 47%前後となる）。

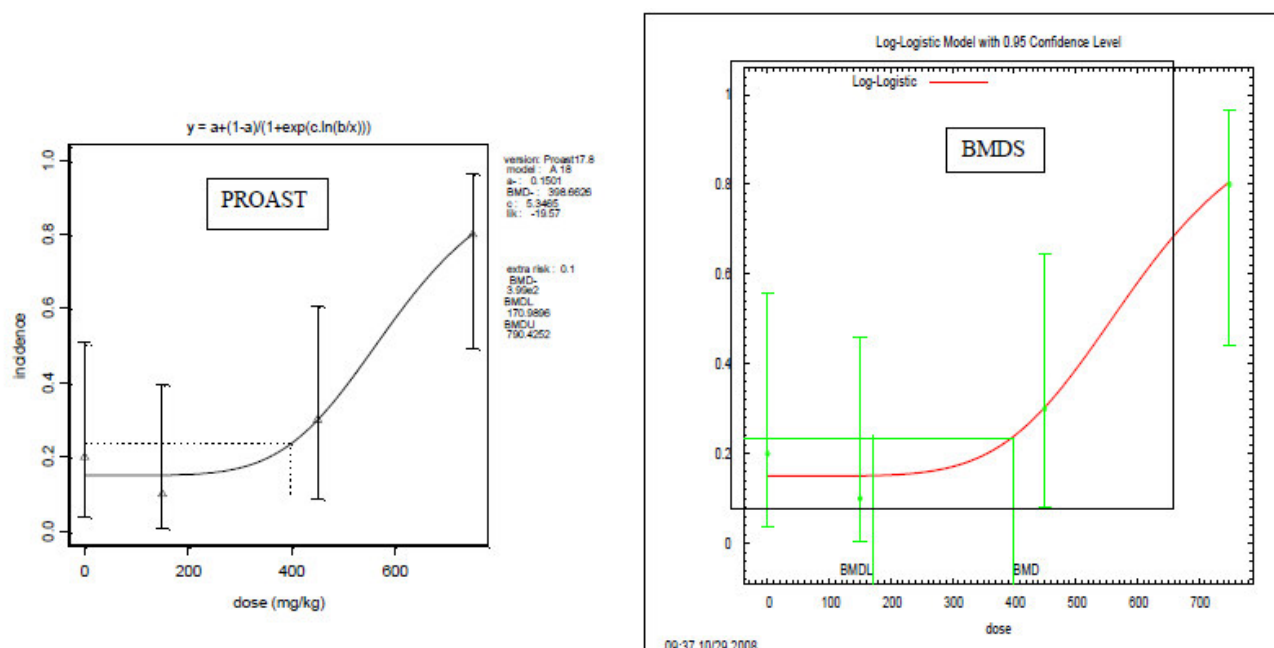


図 3. PROAST と BMDS ソフトで得られた非連続データの解析 — 各々の用量群で 10 匹の動物が使用された毒性研究における影響された動物の割合。データに量－反応モデルがあてはめられ（実線）、破線はゼロ用量の反応と比べて 10%の過剰リスク時の BMR を示す。PROAST と BMDS と対数ロジスティックモデルがあてはめられた（表 3 参照）。示された図はソフトが生成したグラフである。

対数ロジスティックモデルを用いて量－反応データをモデル化すると、図 3 のように BMD_{10} が 399 mg/kg、 $BMDL_{10}$ が 171 mg/kg となり、NOAEL よりもかなり低い。

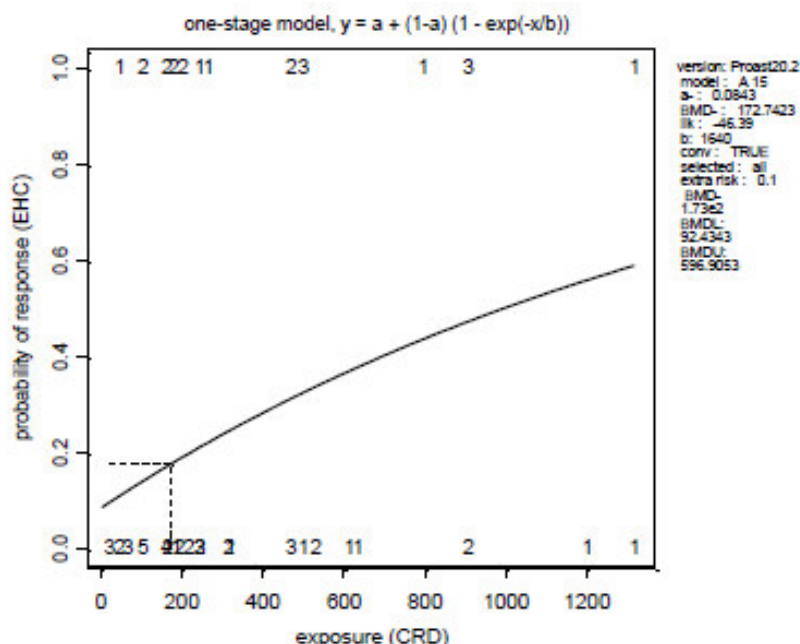
$BMDL$ 値での関連する影響は（95%の信頼度で）10%以上にならないと述べることを BMD の方法論は許可するが、これは、NOAEL 用量群とバックグラウンド反応の 2 群間比較によって計算される（NOAEL 値での）影響の上限値 47%前後よりかなり低い。これは BMD 法がデータから入手できる情報をより上手く活用しているということを再び示している。

例 3：ヒトの量－反応関係データ

ヒトの量－反応データの解析は、交絡因子の存在や曝露推定値における不正確さがあるために、概して動物実験で得られる典型的な量－反応データより複雑である。ここに示す例ではそのような複雑なものとは扱わず、起こりうるヒトデータの一例、すなわち非常に小さな曝露集団の例を示す。ここでは、曝露レベルが各々の対象者で測定されている。その時、曝露量を限られた人数の用量群にカテゴリー化するならば NOAEL 法は適用できるが、それは情報の損失に繋がる。対照的に、 BMD 法は図 4 に示すように、カテゴリー化せずに適用することができる。この例では、全ての対象者が目－手共同運動について正常（= 0）あるいは異常（= 1）とスコア化されている。このようなタイプの観察に目測で量－反応関係を見出すことは難しいが、量－反応モデルをこれらデータにあてはめ、用量に関連する反応が存在する

ことを証明できる。この例では、あてはめたモデルと関連する曲線は所与の曝露レベルで反応するヒトの確率を表す。あてはめたモデルは、あてはめられた水平線と比べ、統計的に有意な改善を示し、統計的に有意な曝露影響の存在を示唆した。BMD 法はこの曲線を過剰リスクが 10% となる曝露レベルを推定するために使用された (5.7 章を参照)。

図 4. ヒトにおける個々の曝露量と量-反応データの BMD 解析 — 個々の作業員（プロットされた数字は当該曝露レベルにおける対象者数を示す）において観察された目-手共同運動が曝露関数として得点化される（0=正常、1=異常）。量-反応モデルがこれらのデータにあてはめられている。BMD₁₀（破線）は 173 であり、BMDL₁₀ は 92 である。10% の BMR が使用された。



この例では、さらに BMD 法が対照群の存在しない状況でも適用できることを示している。すなわち、バックグラウンド反応レベルはあてはめた量-反応曲線によって推定される。

3.5 BMR の選択

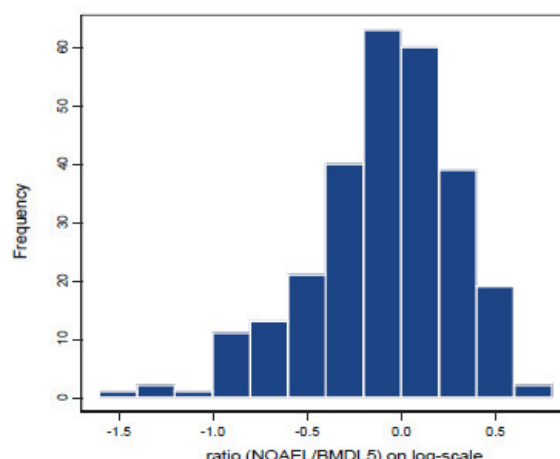
BMD 法を用いて基準点を算出するためには BMR に特定の値を選択する必要がある。ここで、BMR の初期設定値の選択に際して何等かの正当性が求められる。

理想的には、BMR は無視できる、あるいは有害でない影響の大きさを反映するだろう。しかし、実践上の制約として、選択されるべき BMR は、観察値の範囲を超える外挿によって BMD を推定することがないように、あまり小さくすべきでない。と言うのは、BMDL は使用するモデルに大いに依存するからである。

当初、BMD 法は主に非連続データに適用され、1%、5%、ないし 10% の過剰リスクの BMR 値が動物実験データで使用されるよう提案されていた (Crump 1984; EPA 1995)。5% の BMR を使用することはあるデータには実用的であるが、全ての毒性データに必ずしも実用的という訳でない。それに、BMDL はより低い BMR 値を用いた際の量-反応モデルの選択に実際依存するので (Sand *et al*, 2002)、非連続データの場合には BMR=10% の方がより適切であるように思える。平均すると、BMDL₁₀ は致死データの NOAEL に近いと報告されている (Fowles *et al*, 1999)。同様に、発達毒性領域の非連続データに対する BMDL は、平均的に、5% や 1% の BMR よりも 10% の BMR の NOAEL 平均値に近いことがわかった。しかし、この場合 BMDL₁₀ は NOAEL 値より平均して 1/2 程度であった (Allen *et al*, 1994)。1% 程度の BMR がヒト研究に使用されており、もし対象者数が非常に大きいならば正当化されるかもしれない (実際、1% の BMR がヒト研究における Aflatoxin の BMD/BMDL の算出に用いられた—EFSA 2007)。

連続データでは、5%の反応レベルがしばしば観察値の範囲内にあり、この BMR はモデル依存性がそれほど強くない BMD と BMDL の推定値を提供するのである (Woutersen *et al*, 2001; Piersma *et al*, 2000; Sand *et al*, 2006)。米国国家毒性プログラム (NTP) 研究 (Bokkers & Solb 2007) の多くを再解析すると、BMDL₀₅ が同じデータセットに由来する NOAEL に平均として近いことが示され (図 5 参照)、ほとんどのデータセットで一桁以内の違いであった。同様の観察が胎児重量の研究でも認められた (Kavlock *et al*, 1995)。

図 5. ラットおよびマウス研究における同じ量-反応データから算出した 395 の NOAEL/BMDL₀₅ 比の頻度グラフ



科学委員会は、結論として、動物実験由来の BMR の初期設定値は非連続データでは 10%、連続データでは 5%を使用することを提案する。これまでに述べてきたように、この BMR の初期設定値は統計的あるいは毒性学的考察に基づいて修正されるかもしれない。

4. 危険有害因子/リスクの特定化の結果とリスクコミュニケーション

前章までに、基準点の導出方法を説明する中で BMD 法を紹介してきた。危険有害因子の特定化の次の段階は、遺伝毒性と発がん性の両方の性質をもつ物質に対する曝露マージンの設定、それに食品添加物および残留農薬の ADI や汚染物質の TDI あるいは TWI のような遺伝毒性も発がん性もない物質の健康指針値の設定をすることである。

4.1 健康指針値の設定

基準点から ADI/TDI を導出する時、不確実係数が NOAEL に適用される (WHO 1987)。BMDL を基準点として使用する時、より大きなあるいはさらなる不確実係数が相応しいかもしれないと示唆されてきた。そこで言われている議論は、NOAEL と異なり、BMDL は“影響がない”用量を反映していないということである。この議論は「NOAEL は全く影響がないことと関連する」という誤った仮定に基づいている。3.5 章で検討したように、BMR の初期設定値は“平均すると BMDL が NOAEL と一致する”レベルである。さらに 3.3 章で示したように、NOAEL 値で現れる影響の可能な大きさは BMD/BMDL に関連するある特定影響の大きさ (BMR) よりも大きくなりうる。これらを考え合わせると、通常適用されている以上の追加的な不確実係数は必要でない。BMDL から導出される健康指針値は、NOAEL から導出されるものと同じくらいに保護的であると予想される。結論として、NOAEL に現在適用されている不確実係数の初期設定値は BMDL に対しても同等に適用可能であろう。

ある研究では、対照群の反応と統計的あるいは生物学的に有意に異なる低い用量 (LOAEL) で影響が見られるかもしれない。NOAEL 法の中では、LOAEL は伝統的に追加的な不確実係数で除される。これに対して、BMD 法では所与のデータから望ましい BMR を用いて BMDL が導出されており、そのような追加的な不確実係数は必要ないだろう。もし量-反応データが望むレベルでの BMR を可としないなら

ば (3.5 章を参照)、より高い反応を担保する BMR を選択するか、あるいは BMDL に追加的な不確実係数を課す必要がでてくる。

4.2 遺伝毒性および発がん性がある物質のリスク評価

科学委員会 (EFSA 2005) は、考えられる選択肢から、曝露マージン法が遺伝毒性および発がん性をもつ物質のリスク評価の中で最も適切な方法であるとの結論を出した。当委員会は基準点として BMDL₁₀ を使用するよう、すなわち曝露マージンの分子として BMDL₁₀ が指名されるべき、と提案した。

4.3 力価 (potency) の比較

異なる物質あるいは異なる曝露条件下における同一物質の力価を比較する際、同じ大きさの影響／反応を生じさせる用量に関する情報を必要とする。BMD 法は、適用された用量の間を補間することができるので、そのような解析に相応しい道具と言える。同じ理由で、BMD 法は共通の毒性作用をもつ混合物中における個々の物質の毒性等価因子 (toxic equivalency factors) を導き出すのにも適している。実際、異なる有機リン化合物の相対力価推定値を提示するために BMD 法が使用されてきた。また、BMD 法を用いて得られる相対力価推定値は行動分析様式 (mode of action analysis) で使用する NOAEL よりも適切である (Boobis *et al*, 2006)。Plant Protection products and their Residues 委員会は、相対力価推定値を決めるためには BMDL より BMD の方が好まれるとする意見書を最近発表した (EFSA 2008)。

BMD 法は量－反応に関与する性、曝露期間、他の化学物質の同時曝露など、共変量 (交絡変数) の影響を解析するのに適している。例えば、“性” 因子を量－反応解析の中に共変量として含める。そうすれば、男子と女子が有意に異なる量－反応関係を示すかどうか評価できるのである。一方の性が高方よりも化学物質に対して感受性が高いという所見はリスク評価にとって有用な情報となるかもしれない。その上、量－反応関係が男子と女子、あるいは他の小集団間で類似していることが判れば、これら 2 群の量－反応データを足し合わせることも可能となる。これは解析の全標本数を効果的に増すことに繋がるし、ある場合には適用される用量の総数が効果的に増えることに繋がる。同様に、似通った特徴をもつ異なる研究からの量－反応データは、各研究を共変量として、連結することができ、これら全てのデータを再活用することが可能となる (Murata *et al*, 2002)。

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、動物実験の解析では今のところ限られている。後者は 5.7 章でさらに検討する。

4.4 確率論的なリスク評価

リスク評価における確率論的方法が、曝露評価 (例えば、Gibney & van der Voet 2003; Fryer *et al*, 2006; Tressou *et al*, 2004) および危険有害因子の特定化 (例えば、Baird *et al*, 1996; Swartout *et al*, 1998) の両者に関して、大いなる注目を浴びている。BMD の不確実性は分布の形状の中で定量化されるので (Slob & Pieters 1998)、BMD 法は確率論的な危険有害因子の特定化と合致する。さらに、その BMD 法の背後にある量－反応関係のモデル化は、ある種の曝露レベル (例えば、当該集団における現在の曝露量) が仮定されるなら、ヒト集団における潜在的な健康影響の大きさを推定する手段となりうる。これは、例えば Mycotoxin Deoxynivalenol の解析に用いられた (Pieters *et al*, 2004)。

4.5 BMD と NOAEL、安全性の認識

BMD は影響レベルと理解されているために、BMD 法の導入はリスク管理者、政治家、消費者団体、一般大衆との意志疎通の中で問題を引き起こすかもしれないと言われてきた。一方、NOAEL は影響がないレベルと時々不正確に理解されている。しかし、4.1 章で説明したように、リスク評価における BMDL の使用は基本的方法や前提を根本的に変えるものではない。

BMD 法を支持する意見の一つは、BMDL はデータの性質を NOAEL よりも上手く扱っているので、BMD 法はその結論に対し、いかなる場合にも、より高い信頼性を提供するということである。これは全ての過去のデータの再評価が必要であるということを意図していない。何故なら、4.1 章で述べたように、NOAEL と BMDL は平均すると似通っていると考えられるからである。推定 1 日摂取量と健康指針値（例えば、ADI）の間に大きな隔たりがある状況では、再検討などは必ずしも必要でない。実際の推定 1 日摂取量が健康指針値に近いか超えるような物質に対しては、BMD 法を用いたデータの再評価により、洗練されたリスク評価が可能となるかもしれない。

リスク評価の中で不確実性を発生する多くの原因があることを認めなければならないし、量－反応関係のモデル化はその原因の一つとなる。所与のリスク評価に BMD 法を適用するに相応しい利点を課す時、不確実性の原因、その大きさ、その評価に相応しい影響に対して何等かの配慮が払われるべきである。そのような情報は、BMD 法によって賦与されるに相応しい改善がリスク評価に重大な変化をもたらすかどうかを決める助けになろう。前述のことは、BMD 法を使用する唯一の目的が基準点の信頼性を改善することにある時、意味を持つ。

さらに、健康指針値は BMD 法を基礎としている。そして、量－反応曲線に由来する全てのデータを考慮する時、健康指針値を超える状況で、BMD 法はそのリスクを定量化する最良の根拠を提供し、かくして、それはリスクコミュニケーションのより良い礎となるのである。

最後に、ADI や TDI のような健康指針値は「認識可能な健康リスクを持たないでヒトが一生涯、毎日曝露するレベル」と定義されていることをはっきりと意識することが大切であり、健康指針値が NOAEL の代わりに BMDL から算出されてもこの定義は変わらないのである。

5. BMD 法の適用指針

この章は、量－反応データから BMD および BMDL をどのように導くかについて全体像を述べるとともに、なされるべき特殊な選択について勧告する。現在利用可能なソフトは、コンピューターに関する詳細な専門知識がなくても、BMD 法を用いることができる。しかしながら、本書の中で述べてきた BMD 法の概念を正しく理解することが、結果を正しく解釈する上で、必要条件となる。

BMD 法を適用する手順は以下のステップに要約される。個々のステップを以下に説明する。

1. 量－反応データの種類の特定
2. BMR の特定
3. 量－反応モデル候補の選択
4. 受け入れ可能なモデルの同定
5. BMD の推定と基準点としての BMDL の確立

5.1 量－反応データの種類の特定

量－反応関係のない影響指標は通常 BMD の算出に使用しない。影響指標を無視するという決定は当該データの目測によって、あるいは上述のステップ 3 で統計的に有意な変動傾向がないという決定がされる場合であっても、公式には BMD 解析にそれらを含めることによって行う。

反応データには様々な種類があり、連続量、非連続量、順序変数を含む。データの種類の見極めは（元の統計学的分布の前提など）統計的根拠として重要であるが、BMR を解釈するためにも重要である。連続量と非連続量についての BMR の解釈は 3.4 章（例 1 と 2）を参照しなさい。順序変数は連続データと非連続データの間型のデータとみなされるかもしれない——例えば、組織病理学的観察で重症度分類（極めて軽症、軽症、中等症など）が個人に割り当てられた場合などがこれに相当する。順序変数は非連続データに変換することは可能であるが、これは情報の損失を意味するので推奨できない。

個人の連続データからなる観察値は、願わくは BMD 解析に投入する方向で取扱うべきであろう。個人データでなく要約データのみが利用可能な時は、各々の処置群の平均値、標準偏差（または標準誤差）、サンプル数を用いて BMD 解析を行う。個人データから計算する時と比べ、要約データを使用すると若干異なる結果を導くことがある。非連続データの時は、各々の用量群のサンプル数と影響を受けた人数が必要となる。

5.2 BMR の特定

非連続データの BMR は、バックグラウンド反応と比較したときの病変/得点化された反応の発症率の増加として定義される。これを行う一般的な手順は、3.4 章に記したように、追加リスクあるいは過剰リスクのいずれかによる。非連続データで推奨される初期設定値は 10% の過剰リスクの BMR である。

連続データの BMR は様々な方法で定義される。ここで推奨される方法は、予測されるバックグラウンド反応と比較するときの反応変数における平均値のパーセント変化として定義することである。推奨される初期設定値は 5% の BMR（例えば、赤血球数の 5% 減少）である。

ある種の連続データ（例えば、認知スコア）における BMR は、非連続データと類似した方法で、すなわち異常反応を示す対象者数が 5% 増加する点と定義するのが最良の方法である。この種の BMR は所謂ハイブリッド法と称される方法である（Crump 2002）。

5.3 量－反応モデルの候補の選択

同じ程度にデータに上手く適合する異なるモデルであっても、得られる BMD と BMDL は、モデルの不確実性を反映して、同じにならないことがある。不確実性のこの面を考慮するために、同じデータセットに様々なモデルをあてはめてみる必要がある。本章は、あてはめモデル、特に BMD ソフトパッケージ（例えば、BMDS や PROAST）の中で現在利用可能なモデルの候補を詮索する。このソフトは一連のデータセットを解析するために適切かつ柔軟なモデルを多数用意している。もし他のソフトを使用する時は、同じ組合せの候補モデルを適用することが推奨される。

これらのモデルは表 3 に要約してある。これらの幾つかは一連の繰り込まれたモデルの一部として挙げられている。すなわち、（パラメーターを殆ど持たない）より単純なモデルは（多数のパラメーターを持つ）より複雑なモデルに至るまで拡張することができる。表 3 は指数型分布族モデルの繰り込み構造を示している。

ある族の繰り込み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ー数（すなわち、必要以上でもなくかつ不足でもない）を持つのかを決められることである。例えば、S字状の量－反応関係は少なくとも3つのパラメーター（反応が最大の半分に達する位置を表すパラメーター、最大反応の大きさを表すパラメーター、残りは最小と最大反応の間の曲線の傾斜を表すパラメーター）のモデルが必要となるが、量－反応関係がほぼ線形であるならば2つのパラメーターのモデルで十分であろう。最初（わずかなパラメーターを持つ）単純モデルから始め、その後モデルにパラメーターを追加し、当該あてはめに対して有意な改善が認められるかどうか確認しながら、統計的に最適なモデルを見つける。そうすれば、追加のパラメーターを持つモデルが受け入れられる。この過程はあてはめに改善が見られなくなるまで続けられ、やがて最適数のパラメーターを持つモデルが選択される。パラメーターの追加によってあてはめが改善するか否かの決定は尤度比検定による有意性検定を基にする（5.4章を参照）。

Table 3: Recommended models for use in the BMD approach⁴⁾.

Model	Number of model parameters	Model expression response (y) as function of dose (x)	Constraints
<i>Continuous data</i>			
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 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	$y = a [1 + (c-1)x^d/(b^d + x^d)]$	$a > 0, b > 0, c > 0, d > 1$
<i>Quantal data²⁾</i>			
Logistic	2	$y = 1 / (1 + \exp(-a - bx))$	$b > 0$
Probit	2	$y = \text{CumlNorm}(a + bx)$	$b > 0$
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	$y = a + (1-a) \text{CumlNorm}(\log(x/b) / c)$	$0 \leq a \leq 1, b > 0, c > 0$
Weibull	3	$y = a + (1-a) \exp(-(x/b)^c)$	$0 \leq a \leq 1, b > 0, c > 1$
Gamma	3	$y = a + (1-a) \text{CumGam}(bx^c)$	$0 \leq a \leq 1, b > 0, c > 1$
Linearized multistage (LMS) family³⁾			
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 : unknown parameters that are estimated by fitting the model to the data.

連続データ

連続データでは、次に掲げる望ましい特性により、指数型分布族と Hill 族のモデルが推奨されている。

- ー モデルが常に増加あるいは減少を続ける
- ー モデルが最大反応で横ばいになる、あるいは閾値様の曲率を示す（図 6 参照）
- ー モデルが、実際遭遇するように、殆ど全ての量－反応データを表すのに十分な柔軟性を持つ
- ー モデルは、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

Hill 族は BMDS ソフトの繰り込みに含まれておらず、現在使用できない。

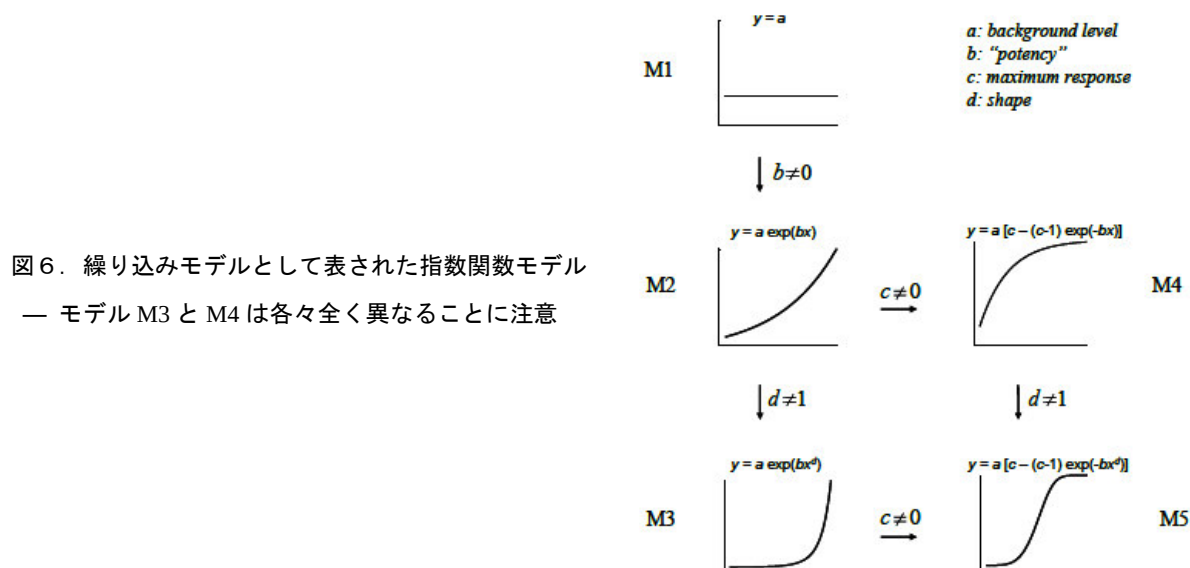


表 3 は指数型分布族と Hill 族の各々の構成要素を示す。

最適指数型分布モデルの BMDL と最適 Hill モデルの BMDL の違いは当該データセットにおけるモデル不確実性の程度の目安を表す（1 例として、図 7 と表 5 を参照のこと）。

たまに、量－反応データが、例えば高用量で体重増加が正から負の値に減少するように、負の値を含むように表されることがある。このような場合、追加のバックグラウンドパラメーターを持つモデルが必要となるかもしれない。代わりに、もし個々人の体重データが利用可能ならば、体重増加を差ではなく比（パーセント変化）として表現することもできる。

線形モデルはしばしば疫学データを十分に記述するし、それゆえ BMD の導出に用いられる。

より高次の多項式モデルは、統計解析では一般的に用いられているが、量－反応データには推奨されない。これは、幾つかの非現実的なモデル特性を有するからである。例えば、多項式モデルでは負の値を生じうるが、生物学的には通常ありえないものである。また、データが支持していない時ですら、単調でない形を示すこともある。

非連続データ

非連続データでの使用が推奨されている 7 モデルが表 3 に掲げられている。線形多段階モデル (LMS) のみが繰り込み族のモデルを形成している。したがって、この族およびこの (BMD 解析で使用されている) 構成要素から、最も相応しい 1 つの多段階モデルが選択されるべきである。

連続および非連続データの解析に対するモデル化の制約

望ましくない特性を持つモデルを避けるために、ある制約がモデルのパラメーターに課されている。例えば、連続反応は通常負にならないので、バックグラウンド反応パラメーター (a) は連続モデルの中

で正になるように制約が加えられる。非連続モデルでは0と1の間になるよう制約される（つまり、反応は0%と100%の間）。さらに、（多くの影響指標で生物学的に信じがたいと思われる）用量0で無限の傾きをとる曲線を避けるために、幾つかのモデルにある“形状”パラメーターはしばしば次のように制限されている：（連続データ）指数関数と Hill 族の M3 と M5 モデルでは $d > 1$

（非連続データ）対数 logistic、Weibull、Gamma モデルでは $c > 1$

5.4 モデルのあてはめと受け入れ

利用可能な BMD ソフトはモデルのあてはめに注意を払っている。それは、計算中の量-反応曲線を実データに限りなく近づけるように、モデルの未知パラメーター値を見つけることを指す。これは最適適合度（best fit）と呼ばれ、対数尤度を最大にすることによって達成される。

BMD 法は、BMD の統計的な最適推定値を見つける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もっともらしい値を全てを見つけることにある。それ故、その到達目標は唯一の最適なあてはめモデルを見つけることではなく、受け入れ可能なあてはめを持つモデルを見つけることである。

あてはめモデルの受け入れは2つの原則に基づく。第一の原則は、繰り込み族のモデルの中で様々な構成要素の対数尤度を尤度比検定で比較し、これはという構成要素を決定する。前述したように、わずかしかなパラメーターを持たない構成要素がそれ以上貧弱なあてはめを示さないのであれば、その構成要素の方が好まれる [訳注: 「モデルはより単純であればあるほど好まれる」という意味と思われる]。

第二の原則は、適合度検定で $P > 0.05$ であれば、あてはめモデルはその量-反応データに理にかなった説明を与えるということである。適合度検定は何種類もある。尤度比検定はここで推奨される選択肢である。もしこれが使用しているソフトパッケージで利用できないならば、ピアソン χ^2 検定などの別の検定を考慮すべきである。尤度比検定では、あてはめモデルに関連する対数尤度値がいわゆる“全モデル（full model）”の対数尤度値と比較され、検定される。全モデルは単純に各々の適用された用量で観察された反応（平均）からなる。これ故に、パラメーターの数は用量群の数に等しい。もしモデルの適合度が全モデルより有意に劣っていなければ、そのモデルは受け入れられることになる。

尤度比検定は、繰り込みモデルの追加パラメーターがあてはめを有意に改善させるかどうかを検定するのに使われる。この検定は、2つのモデルに関連する対数尤度値に-2を掛けた値（-2 Log L）が当該2モデル間のパラメーター数の差に等しい自由度をもった χ^2 分布に従うという事実に基づいている。表4はパラメーター数（自由度）が増加する時の、2つの対数尤度の間の有意水準差を示す。その詳細は Slob (2002) を参照しなさい。

Table 4: Critical difference in log-likelihood values in comparing two nested models.

df	Critical difference in log-likelihoods	df	Critical difference in log-likelihoods
1	1.92	11	9.84
2	3.00	12	10.51
3	3.91	13	11.18
4	4.74	14	11.84
5	5.54	15	12.50
6	6.30	16	13.15
7	7.03	17	13.79
8	7.75	18	14.43
9	8.46	19	15.07
10	9.15	20	15.71

When the difference exceeds the critical value, the fit of the model with more parameters is better at a 5%-significance level.

The number of degrees of freedom (df) is equal to the difference in the number of estimated parameters between both models.

データの目測で幾つかのモデルは観察された量－反応関係を適切に説明しているように見えても、いずれのモデルも適合度検定をパスしない場合もありうると心に留めておくことは重要である。そのような場合（仮定分布からの逸脱、観察値が必ずしも全て独立でない、非ランダムエラー等）、適合度検定の前提は打ち砕かれ、そのようなモデルを受け入れようと決定するときには高レベルの不確実性をその評価の中に認めざるを得なくなる。

5.5 BMD の推定と基準点としての BMDL の確立

BMD 解析は研究の中で測定された可能性のある臨界影響指標に対して実施されるべきである。

- ・可能性のある臨界影響指標に対して、モデルのセット（表 3）を適用し、受け入れられたモデルの各々で、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。初期設定の信頼レベルは 95%（片側）である。この手順はある一定範囲の BMDL を生むが、それはまさに使用したモデル間の差を反映する。
- ・各々の可能性のある臨界影響指標に対する総体的な BMDL となる、最も低い BMDL をこの範囲の中から決定する。つまり、これが各々の場合における最も無難な値である。
- ・これらの全ての影響指標から最も低い BMDL を決定する。それが当該研究を代表する BMDL である。

同じ影響指標の中で最も低い BMDL を選択することが無難である。しかし、“モデル加算平均”のようなより進化した方法が完全に開発され正当化されるまでの間、これを使うことが推奨される。

このように、BMDL は各種影響指標に対して決めることができる。もし BMD 解析の前に臨界影響指標を確立することができなかったなら、NOAEL 法の時と同様のやり方（すなわち、最も低い BMDL を生じる影響指標—その影響がヒトの健康に関連するならば—を選択する）で、BMD 解析の結果に基づいて決める。

実験データが量－反応関係について十分な情報を提供する場合、当該データにあてはまる各種モデルは類似した形を持ち、得られる BMDL の範囲も狭くなる。そのような BMDL 値はリスク評価目的に安全な基準を提供し、かつ健康指針値の確立ないし曝露マージン算出のための基準点を定義するのに用いることができる（第 4 章参照）。

ある状況で、量－反応関係が上手く定義できないことがある。例えば、連続した反応レベルの間に大きな隙間がある場合などである。そのようなデータセットでは、（これまでに議論してきた統計的規準に従う）データにあてはまる各種モデル（表 3 参照）が異なる形状を推定するかもしれないし、結果として得られる BMD 値の範囲も広くなるかもしれない。これらの BMDL 値は健康指針値を確立するための安全基準とならないだろう。得られた BMDL 値の範囲に基づいて、量－反応データの妥当性を判断する基準は未だ確立されていない。科学委員会は、一般的ルールとして、量－反応データは承認された異なるモデルからの BMDL の範囲が一桁を超えないようにすべきであると提案している。この値を超える時は、幾つかの選択肢が利用できるので、個々の対応（例えば、BMR を増加させる、モデルのセットを再評価する、モデル加算平均するなど）を考慮しなければならない。

5.6 実例

連続データ

量－反応データセットから BMD 解析を行った結果が表 5 および図 7 に示されている。この例題用データセットでは、用量レベルの数（6）が典型的な毒性研究におけるより多めである。指数型分布族と Hill 族の両モデルがその量－反応データセットに適用された。表は 2 族のモデルにおける全モデルおよび各々の構成要素から得られた対数尤度値を示す。この解析で、2 族のモデルとも M3 が最適モデルとして選択された（多くの場合、異なる族から選択されるモデルは同数のパラメーターを持つ）。これら 2 つのモデルの BMDL は各々 326 mg/kg、316 mg/kg となり、この差はモデルの不確実性を映し出す（その差は、用量群数が相対的に多い割に小さい）。かくして、この影響指標の BMDL は、最も低い BMDL である 316 mg/kg となる。

Table 5: Results of BMD analysis using two families of models.

Model ¹⁾	Number of model parameters	Log-likelihood		BMD ₀₅ ⁺ (mg/kg bw)		BMDL ₀₅ ⁺ (mg/kg bw)	
		Exponential	Hill	Exponential	Hill	Exponential	Hill
M1 (null model)	1	20.85					
└ M2	2	28.97	28.61				
└└ M3	3	33.35 **	33.39 **	488	479	326	316
└└ M4	3	28.97	28.61				
└└└ M5	4	33.56	33.56				
Full model	7	35.95					

* At BMR of 5%.

** Optimal model within this family according to log-likelihood criterion. It also passes the goodness-of-fit test, as the full model does not result in a significantly better fit.

¹⁾ See Figure 7 and Table 3 for an illustration of models M1 to M5, and their nested structure

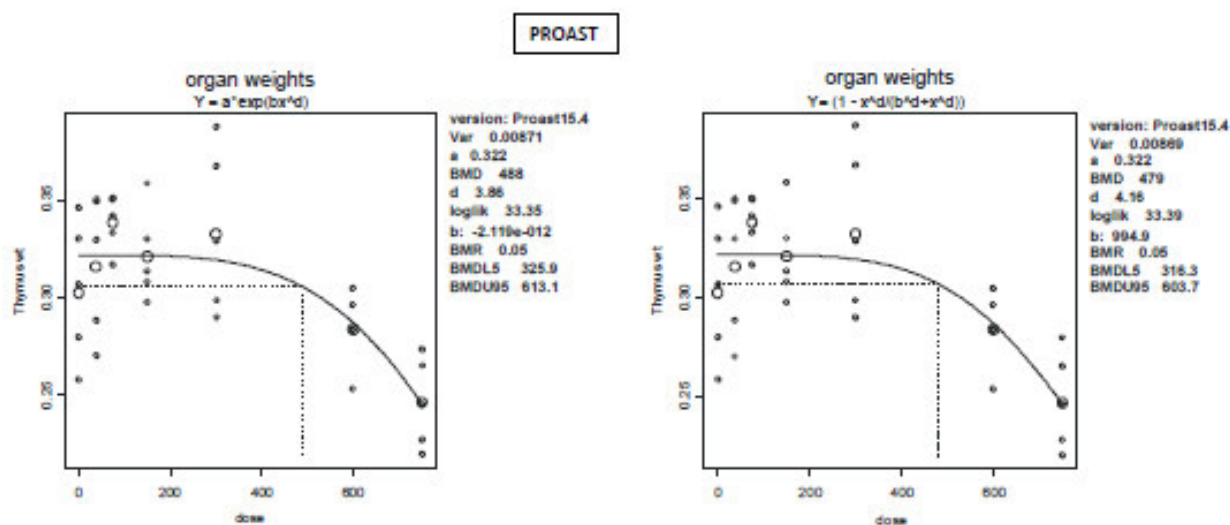


図 7. 指数型分布族および Hill 族のモデル M3 に適合させた曲線（表 3 を参照）

繰り込み族からの最適モデルの選択は幾つかの BMD ソフトによって実行できるが、使用者自身によっても行うこともできる。実例として、表 5 の 3 列目の対数尤度比を考えてみよう。どのモデルが他のモデルに繰り込まれるかに注意を払うと（図 6 参照）、以下の解析手順が続くだろう — M2 に関連する対数尤度は M1 より 8 単位以上高く、パラメーターも 1 つ多い。尤度比検定によると、対数尤度が 1.92 以上増加している時、モデルにパラメーターを加えると有意に良いあてはめを生じる（ $P=0.05$ ）。つまり、M2 は明らかに M1 より良い。同じ規準で、M3 は M2 より良いが、M4 は M2 ほど良くない（追加パラメ

ーター c は 0 と推定され、その場合 M4 は M2 と同じになる)。その上、M5 は M3 ほど良くない。したがって、M3 がこの繰り込み族の中で“最適”モデルとなる。最終的に、M3 は有意に改善されないし、しかも適合度検定をパスする。対数尤度値の要請された増加は追加 4 パラメーターで 4.74 となる。

Table 6: Results of BMD analysis of quantal data shown in Fig. 8.

model	number of model parameters	log-likelihood	accepted	BMD ₁₀	BMDL ₁₀
null model	1	-143.01	--		
LMS: one-stage	2	-56.8	no		
LMS: two-stage	3	-53.71	no		
logistic	2	-55.63	no		
probit	2	-60.40	no		
Weibull	3	-52.33	no		
log-logistic	3	-49.35	yes	3.28	2.45
log-probit	3	-49.56	yes	3.16	2.42
gamma	3	-50.96	yes	2.88	1.93
full model	4	-49.33	--		

BMR: 0.10

Constraint: no

P-value Goodness of Fit: <0.05 leads to model rejection. If the model passes the goodness-of-fit test, see section 5.4

非連続データ

表 6 と図 8 は非連続データにあてはめた各種モデルの結果である。受け入れられたモデルの決定は上で述べたような 3 つの規準に基づく。

- 全てのモデルが反応無しモデルよりも有意に良い。対数尤度値は概ね 50 単位以上であるが、2 パラメーターモデルで 1.92 単位 また 3 パラメーターモデルで 3.00 単位の増加であり、十分であろう（表 4 参照）
- (繰り込み) LMS モデルでは、対数尤度の差は 1.92 より大きく、2 段階モデルは 1 段階モデルよりも有意に良い。したがって、この場合 2 段階モデルが LMS モデルを代表する
- 7 モデル中の 3 モデルだけが適合度検定をパスした。それ以外のモデルは全モデルの対数尤度よりも有意に低い対数尤度を示している

受け入れられた 3 モデルの BMDL は 1.93 から 2.45 mg/kg であり、この小さな範囲はこれら量-反応データが BMDL₁₀ を導出するのに相応しいことを示している。

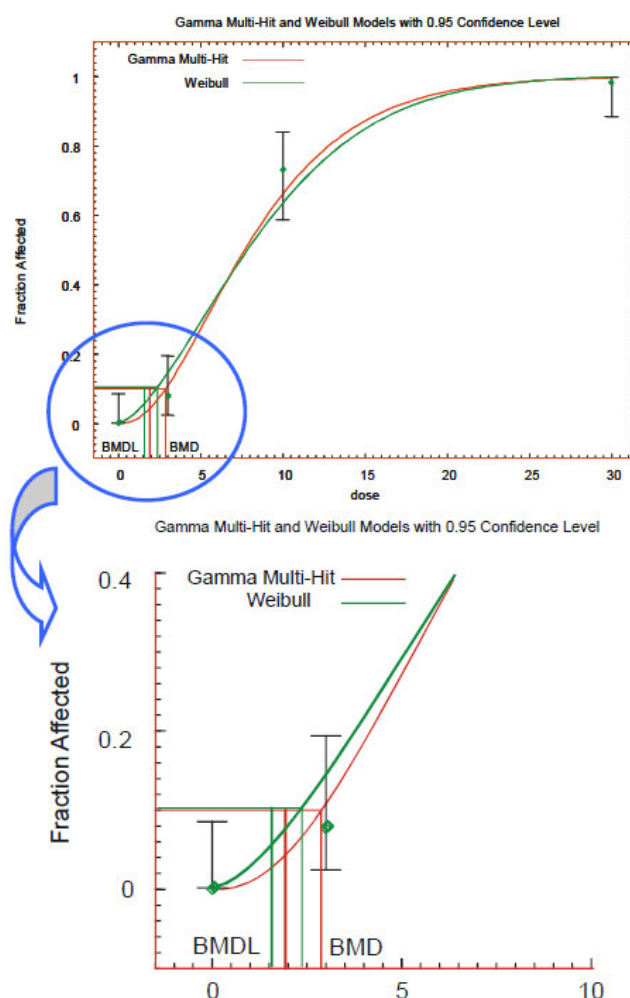


図 8. ガンマおよびワイブルの非連続モデルを適合させた例

5.7 ヒトの量－反応データにおける固有の問題

観察的疫学研究に由来する量－反応データは、幾つかの点で典型的な動物毒性データと異なる。BMD 算出に関連する主要な違いを以下に手短に示す。

曝露データは、よく定義された用量群数が減多に小さくなることはない。NOAEL 法を適用するためには、個人曝露データがカテゴリー化される必要があり、結果的に情報損失を伴う。例 3 で例証したように、BMD 法は個人曝露データを容易に利用することができる。

多くの実験研究と異なり、観察研究は非曝露対照群を含まず、例えば大気汚染物質や食品汚染物質などを、全ての対象者がある程度曝露されている。この場合でも、量－反応曲線のあてはめは必ずしも曝露ゼロの観察値を必要としていないので BMD 法は適用される。しかしゼロ曝露での反応は低用量外挿法によって推定される必要がある。それゆえ疫学研究に由来する BMD は強いモデル依存性を持っている (Budtz-Jørgensen *et al*, 2001)。

ヒトの曝露推定値はしばしば不正確であり、その不正確さを無視すると量－反応関係の評価を歪めてしまう恐れがあることを心に留めておくべきである。この問題に取り組み、かつ量－反応関係および BMD 値の偏りの少ない推定値を得るための統計的方法が開発されている (Budtz-Jørgensen *et al*, 2004; Budtz-Jørgensen 2007)。

ヒト研究の反応変数は、しばしば、その量－反応に干渉するかもしれない交絡因子の影響を受ける。交絡因子への配慮を上手く行わないと BMD を過小評価したり、過大評価したりする可能性がある。これは一般的に交絡因子を共変量として (例えば、重回帰モデルの中に) 入れることによって対処される。高齢者になると傷病に大いに罹りやすくなるように、影響の修飾は考慮されるべき付加的問題をもたらすかもしれない。その時、BMDL は最も脆弱な特定集団における反応を反映すべきである。影響の修飾はそれに関連する共変量を含めることによってモデル化することができる。交絡因子や影響の修飾に対する補正は BMDS ソフトでは実行できないが、PROAST ソフトで一部可能である。

これらの方法論上の問題があるにもかかわらず、BMD 解析は、多くの場合、観察的疫学研究で使用されている (例えば、EFSA 2009b)。

5.8 BMD 解析の報告

BMD 解析の結果は他者がその解析法を追従することができるような方法で報告されるべきである。

ある研究の BMD 解析を報告する時、解析された全ての影響指標に関する情報を提供する必要はなく、重要なものについてだけで良い。これらの影響指標がなぜ選択されたのかが話の中で明らかにされるべきである。

以下の情報が提供されるべきである。

A. BMD 解析で報告される影響指標についてのデータの要約表

非連続データでは、各用量レベルでの使われた動物の総数および反応した動物の数が用量レベル毎に表されるべきである。連続的な影響指標では、反応平均値、標準偏差 (あるいは標準誤差)、サンプル数を用量レベル毎に表すべきである。ヒトのデータでは、同様な情報を交絡因子の詳細とともに提供されるべきである

B. 選択された BMR 値とその根拠

C. 使用されたソフトとそのバージョン

D. モデルあてはめ手順の設定と前提

E. 使用されたモデルと対数尤度比の表、モデルのあてはめと受け入れに関する情報、および受け入れられたモデルから算出された BMD と BMDL

F. 少なくとも、重要な影響指標のデータにあてはめられたモデルのプロット（最も小さい BMDL を含む）

G. この特定研究の BMDL に関する結論

非連続データセットおよび連続データセットについて、BMD 解析の報告の仕方を以下に示す。いずれの場合にも、臨界影響を表すと考えられる 1 つの影響指標が解析される。読者は BMDs と PROAST ソフトの使用法の簡単な指針として Appendix を参照することができる。

BMD ソフトの使用に関する指針を提供しようとする努力がこの意見書の中で行われているが、そのようなソフトは未だ開発中であり、新バージョンが頻繁に公開されることを使用者は知っておくべきである。したがって、使用時に入手できるソフトのバージョンがここで参照するものと同じでないかもしれない。使用者は使っているソフトの各々のバージョンについての仕様書を調べるのが勧められる。使ったソフトのバージョンがいずれであれ、報告する文章構成は同じである。

例 1：連続データ

この例は PROAST ソフトを使用したものである。BMDs ソフトを使用した場合幾つかの初期設定値が本書での推奨のように変更される必要があることに注意しなければならない (Appendix 参照)。この例はマウスにおける Deoxynivalenol の 2 年研究に関連する。用量に関連した体重の減少が観察された。この影響指標が重要な影響であると推定された。次に示す BMD 解析は、EFSA の意見書に基づいて連続データセットの BMD 解析から得られた結果をどのように報告するかの例である。

A. データ（下表）

dose	body weight, group mean	SD	n	Sex
0	43.85	2.69	37	M
0.1	43.51	2.86	35	M
0.5	40.04	3.00	43	M
1.1	35.09	2.56	42	M

B. BMR：初期設定値（5%の変化）

C. 使用したソフト：PROAST Ver. 17.3

D. 追加の前提：なし

E. BMD 結果（下表）

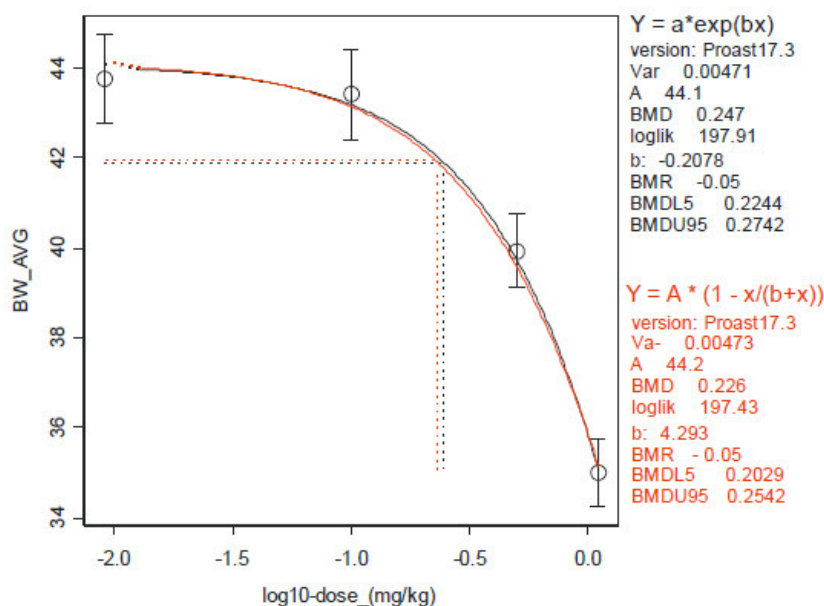
Body weight (males)

Model	N° of parameters (variance excluded)	Log-likelihood		BMD ₀₅ [*] (mg/kg)		BMDL ₀₅ [*] (mg/kg)	
		Exponential	Hill	Exponential	Hill	Exponential	Hill
M1 (no response)	1	119.03					
↳ M2**	2	197.91	197.43	0.25	0.22	0.23	0.20
↳ M3	3						
↳ M4	3						
↳ M5	4						
Full model	7	198.34					

^{*} At BMR of 5%.

^{**} Optimal model within this family according to log-likelihood criterion. It also passes the goodness-of-fit test, as the full model does not result in a significantly better fit. Consequently, there was no need to fit the higher models. M3, M4 and M5

F. 図 (右図)



G. 結論

指数型分布族と Hill 族の両方のモデルについて、M2 が選択された。関連する $BMDL_{05}$ は 0.23 と 0.20 mg/kg である。この研究の $BMDL_{05}$ は 0.20 mg/kg である

例 2：非連続データ

この例はラットを 2 年間追跡した研究であり、3 つの用量で化合物が動物に投与された。用量に関連した甲状腺上皮細胞の空胞化の変化が観察され、これらのデータが BMD 解析に使用された。

A. データ (右表)

Dose (mg/kg bw/day)	N° of animals showing thyroid epithelial vacuolization	N° of animals in dose group	sex
0	6	50	f
3	6	50	f
12	34	50	f
30	42	50	f

B. BMR: 初期設定値 (過剰リスク=0.10)

C. 使用したソフト: BMDS Ver. 2.0

D. 追加の前提: なし

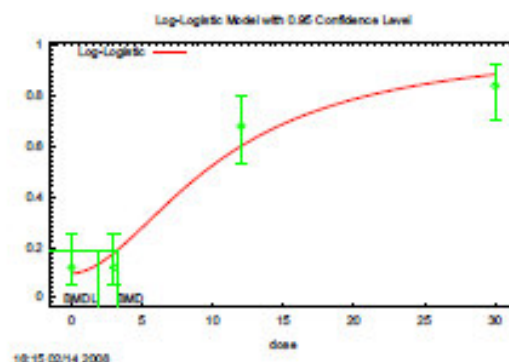
E. BMDの結果の表

Model used	N° of parameters	log-likelihood	accepted	BMD ₁₀	BMDL ₁₀
Null model	1	-137.19			
Gamma	3	-93.49	no		
Logistic	2	-97.23	no		
Log-logistic	2	-91.91	yes *	3.23	1.90
Log-probit	3	-91.86	yes *	3.31	2.34
LMS: one-stage ¹⁾	2	-94.10	no		
LMS: two-stage	3	-94.10	no		
Probit	2	-97.54	no		
Weibull	3	-93.77	no		
full model	4	-90.02			

* not significantly worse than full model (and significantly better than null model)

¹⁾ The one-stage model is identical to the quantal linear model as implemented in BMDS; note that in BMDS, this model is called "multistage" and the number of stages has to be defined by setting the degree of the polynomial in this model, e.g. 2 for a two-stage model.

F. 図 (右図)



G. 結論

2つのモデルが承認された。関連する $BMDL_{10}$ は 1.9 から 2.3 mg/kg BW の範囲だった。この研究の $BMDL_{10}$ は 1.9 mg/kg BW である

結論と勧告

BMD 法は NOAEL 法と比べ科学的により進歩した基準点算出の方法である、と科学委員会は結論する。何故なら、BMD 法は量-反応データを広範囲に利用しうるし、量-反応データの不確実性とバラツキを定量化することができるからである。BMD 法を利用することでより一貫性の高い、定めた BMR の、基準点を算出することができる。

BMD 解析のためのソフトは現在利用できる状況にあるが、ソフトは未だ開発途上にあり、中期的にはモデルの加算平均や連続データの解析などの重要な改善が加えられであろう。

科学委員会はまた、毒性試験ガイドライン（例えば、OECD ガイドライン）が改訂される時、BMD 法のある部分が考慮されると考えている。

BMD 法を用いて導出される健康指針値は、平均的な多くのリスク評価にわたって、NOAEL 法から導出されたものと同じくらい保護的であると期待される。したがって、現在用いられている不確実係数の初期設定値は同様に適用可能である。

科学委員会は、NOAEL 法を用いて行われた過去の全評価を BMD 法で再解析する必要があるとは考えていない。何故なら、これらの方法は平均すると似通った数値を算出しているからである。以前のリスク評価を再検討することが必要と考えられる場合、BMD 法の適用は大いに価値があるだろう。

BMD 法は食物中の全ての化学物質に（例えば、殺虫剤、添加物あるいは汚染物質の分類や由来に関係なく）適用できる。BMD 法が特別の価値を持つ例として、NOAEL の同定が不明瞭であるとか、遺伝毒性と発がん性を有する物質の曝露マージンや疫学的データの量-反応評価などで基準点が必要とされる場合などである。短期的には、科学パネルと EFSA ユニットは上述のような場合に BMD 法を採用することを強く勧める。

長期的には、BMD 法が健康指針値や曝露マージンを導出する際の基準点を決定する方法として利用されることを期待している。実際、BMD 法が導入され、幅広く利用されると EFSA は考えており、適用に際しては一定レベルの専門家の判断とモデル化の専門的知識が必要となる。このため、科学パネルおよび EFSA ユニットの専門家は量-反応のモデル化およびソフトの使用法に習熟しておかなければならない。科学委員会は、2 年毎に、EFSA の作業の中で BMD 法の履行状況、使用実態および受け入れ状況を検討したいと考えている。

文 献

- Allen BC, Kavlock RJ, Kimmel CA & Faustman EM (1994). Dose-response assessment for developmental toxicity. II. Comparison of generic benchmark dose estimates with no-observed-adverse-effects levels. *Fundam. Appl. Toxicol.* 23: 487-495.
- Baird SJS, Cohen JT, Graham JD, Shylakter AI & Evans JS (1996). Noncancer risk assessment: A probabilistic alternative to current practice. *Hum. Ecol. Risk Assess.* 2: 79-102.
- Bokkers BGH & Slob W (2007). Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach. *Crit. Rev. Toxicol.* 37: 353-377.
- Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D & Farland W (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit. Rev. Toxicol.* 36: 781-792.
- Budtz-Jørgensen E, Keiding N & Grandjean P (2001). Benchmark dose calculation from epidemiological data. *Biometrics* 57: 698-706.
- Budtz-Jørgensen E (2007). Estimation of the benchmark dose by structural equation models. *Biostatistics* (in press).
- Crump KS (1984). A new method for determining allowable daily intakes. *Fundam. Appl. Toxicol.* 4: 854-871.
- Crump K (2002). Critical issues in benchmark calculations from continuous data. *Crit. Rev. Toxicol.* 32: 133-153.
- EFSA (2005). Opinion of the scientific committee on a request from EFSA related to a harmonized approach for risk assessment of substances which are both genotoxic and carcinogenic. *The EFSA Journal* 282: 1-31.
- EFSA (2007). Opinion of the scientific panel on contaminants in the food chain [CONTAM] related to the potential increase of consumer health risk by a possible increase of the existing maximum levels for aflatoxins in almonds, hazelnuts and pistachios and derived products. *The EFSA Journal* 446: 1-127.
- EFSA (2008). Opinion of the scientific panel on plant protection products and their residues to evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health with a view to set MRLs for those pesticides in the frame of Regulation (EC) 396/2005. *The EFSA Journal* 704: 1-84.
- EFSA (2009a). Safety evaluation of ractopamine. Scientific opinion of the panel on additives and products or substances used in animal feed. *The EFSA Journal* 1041:1-52.
- EFSA (2009b). Cadmium in food. Scientific opinion of the panel on contaminants in the food chain. *The EFSA Journal* 980: 1-139.
- EPA (1995). The use of the benchmark dose approach in health risk assessment. EPA/630/R-94/007. Risk Assessment Forum, Washington DC.
- FOSIE (2002). Food safety in Europe: Risk assessment of chemicals in food and diet. *Food Chem. Toxicol.* 40: 2-3.
- Fowles JR, Alexeeff GV & Dodge D (1999). The use of benchmark dose methodology with acute inhalation lethality data. *Regul. Toxicol Pharmacol.* 29: 262-278.
- Fryer M, Collins CD, Ferrier H, Colville RN & Nieuwenhuijsen MJ (2006). Human exposure modeling for chemical risk assessment: A review of current approaches and research and policy implications. *Environ. Sci. Policy* 9: 261-274.
- Gibney MJ & van der Voet H (2003). Introduction to the Monte Carlo project and the approach to the validation of probabilistic models of dietary exposure to selected food chemicals. *Food Addit. Contam.* 20 (suppl 1): S1-7.
- Joint FAO/WHO Expert committee on Food Additives – JECFA (2006a). Sixty-fourth meeting, WHO/IPCS safety evaluation of certain contaminants in food. WHO Food Additives Series 55.
- Joint FAO/WHO Expert committee on Food Additives – JECFA (2006b). WHO Technical Report Series 939. Evaluation of certain veterinary drug residues in food. 66th report of the Joint FAO/WHO Expert Committee on

food additives.

- Kavlock RJ, Allen BC, Faustman EM & Kimmel CA (1995). Dose-response assessments for developmental toxicity IV. Benchmark doses for fetal weight changes. *Fundam. Appl. Toxicol.* 26: 211-222.
- Murata K, Budtz-Jørgensen E & Grandjean P (2002). Benchmark dose calculations for methylmercury-associated delays on evoked potential latencies in children. *Risk Anal.* 22: 465-474.
- Piersma AH, Verhoef A, te Biesebeek JD, Pieters MN & Slob W (2000). Developmental toxicity of butyl benzyl phthalate in the rat using a multiple dose study design. *Reprod. Toxicol.* 14: 417-425.
- Pieters MN, Bakker M & Slob W (2004). Reduced intake of deoxynivalenol in the Netherlands: A risk assessment update. *Toxicol. Lett.* 153: 145-153.
- Sand S, Falk Filipsson A & Victorin K (2002). Evaluation of the benchmark dose method for dichotomous data: Model dependence and model selection. *Regul. Toxicol. Pharmacol.* 36: 184-197.
- Sand S, von Rosen D, Victorin K & Falk Filipsson A (2006). Identification of a critical dose level for risk assessment: developments in benchmark dose analysis of continuous endpoints. *Toxicol Sci.* 90: 241-251.
- Slob W & Pieters MN (1998). A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: General framework. *Risk Anal.* 18: 787-798.
- Slob W (2002). Dose-response modeling of continuous endpoints. *Toxicol. Sci.* 66: 298-312.
- Swartout JC, Price PS, Dourson ML, Carlson-Lynch HL & Keenan RE (1998). A probabilistic framework for the reference dose (Probabilistic RfD). *Risk Anal.* 18: 271-282.
- Tressou J, Leblanc JC, Feinberg M & Bertail P (2004). Statistical methodology to evaluate food exposure to a contaminant and influence of sanitary limits: Application to ochratoxin A. *Regul. Toxicol. Pharmacol.* 40: 252-263.
- Wheeler MW & Bailer AJ (2007). Properties of model-averaged BMDLs: A study of model averaging in dichotomous response risk estimation. *Risk Anal.* 27: 659-670.
- Wheeler MW & Bailer AJ (2008). Model averaging software for dichotomous dose response risk estimation. *J. Stat. Software* 26: 1-15.
- WHO (1987). Principles for the safety assessment of food additives and contaminants in food. Environmental Health Criteria 70. WHO/IPCS.
- Woutersen RA, Jonker D, Stevenson H, te Biesebeek JD & Slob W (2001). The BMD approach applied to a 28-day toxicity study with Rhodorsil Silane in rats: the impact of increasing the number of dose groups. *Food Chem. Toxicol.* 39: 697-707.

< 翻訳協力者—嶽石美和子、苅田香苗、熊田洋高、岩田豊人、佐藤 洋 >

<2010.12.23 翻訳修正責任者 村田勝敬 >

本翻訳は、食品安全委員会食品健康影響評価技術研究「用量反応性評価におけるベンチマークドース法の適用に関する研究（課題番号－1007）」（主任研究者 広瀬明彦@国立医薬品食品衛生研究所）の研究補助を受けた