

ベンチマークドース（BMD）法の最近の動向について

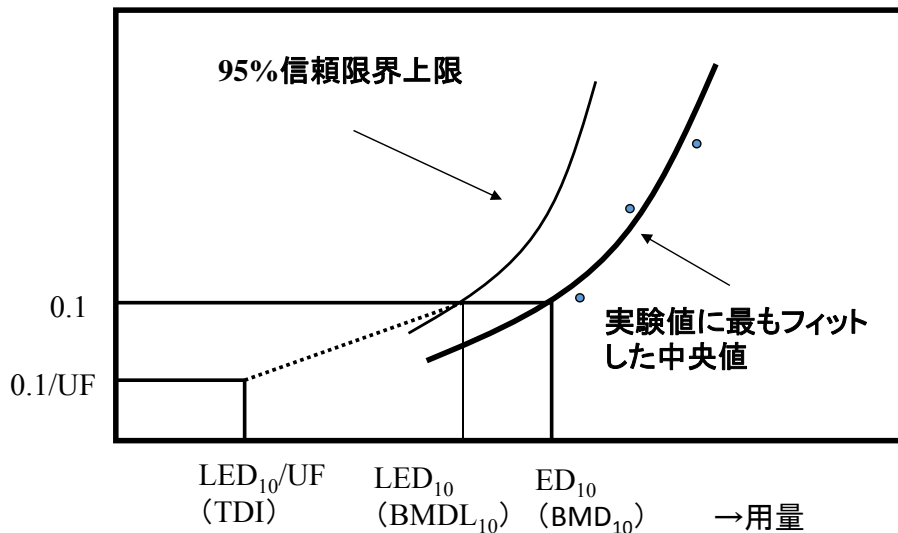
国立医薬品食品衛生研究所

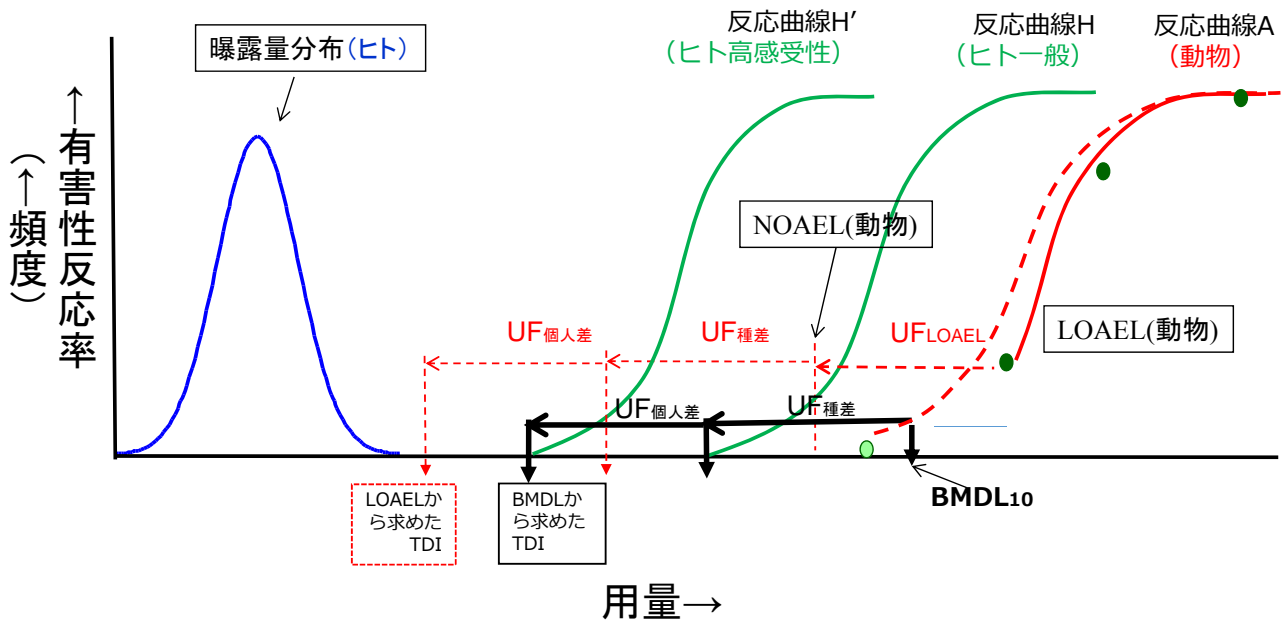
広瀬明彦

ベンチマークドース（Benchmark Dose : BMD）法

ベンチマークドース法は、毒性発現率（体重減少などの数的変化あるいは毒性発現頻度）と曝露量の相関性に数理モデルを適用し、統計学的に尤もフィットしたモデルにおいて、解析対象とする実験系で有意な影響を検出できる反応レベル（benchmark response, BMR）の用量に対する信頼限界の曝露量下限値をBMDL（benchmark dose lower confidence limit）として算出する手法である。

通常、BMRとしては一般毒性で10%、発生毒性で5%の毒性発現率を使用し、信頼限界が95%における曝露量をBMDLとして求める。BMDLは経験的にNOAELに相当する投与量を算出することが可能であると考えられており、実験から得られたNOAELよりも実際のNOAELが明らかに大きいと推定される場合、あるいはLOAELしか得られていない場合に、TDI等の算出の出発点（Point of Departure: POD）として用いることができる。遺伝毒性のある発がん性物質のリスク評価におけるPODとしても使われる。





2

ベンチマークドース法の特徴

- LOAELしか得られていない場合でも、適切にBMDLを求めることができれば、NOAELに相当する投与量を推定できる。
- 投与量依存性に基づいてBMDLを算出することから、同一の設定投与量で行った実験で同じNOAELが得られた場合でも毒性強度を比較できる。
- 信頼下限界を用いているので、データの質および統計学的考え方が含まれる。(動物数が少ない場合や、データのバラツキが大きい場合には信頼限界の幅が広くなり、BMDLはより低い値となる。)
- 過去の多くのデータ解析から、発生毒性では5%、一般毒性では10%の発現率(反応率)のBMDLがNOAELと同等であることが示されている。
- BMDLは実験投与量付近での計算値であるため、フィッティングが良好ないずれの計算式を用いてもBMDLの値に違いが少ない。
- 病理組織学所見データに対し適用する場合は、慎重な検討が必要である。BMDLは発生頻度データに基づき算出されることから、グレードが示されている試験を取り扱う場合や病理学的変化の進行により所見名が変わる場合等は注意が必要。
- 遺伝毒性発がん物質のVSDやMOEを求めるためのPOD(point of departure)を求める場合や、疫学データに対しても使用される。

3

病理組織学所見データの取扱い(例)

用量 (mg/kg/day)	0	30	100	300	1000
動物数	12	12	12	12	12
非連続データ					
肝細胞肥大 +	0	2	4	6	0
++	0	0	4	6	7
+++	0	0	0	0	5
計	0	2	8	12	12
連続データ					
相対肝重量 (g/100gBW)	3.20±0.089	3.30±0.165	3.52±0.099	3.96±0.132	4.49±0.161

4

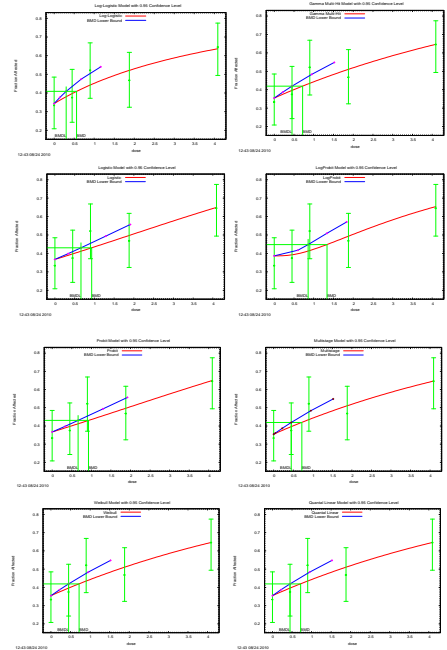
ベンチマークドースを算出するためのソフトウェア

- BMDS**
 米国EPAで開発され、インターネット上で無料公開されている。数理モデルとして、Gamma、Logistic、LogLogistic、Logprobit、Multistage、Probit、Weibul、Quantal-Linearなどの計算を行うことができる。インターフェイスの容易さから国際的によく使われている。
- PROAST**
 オランダのNational Institute for Public Health and the Environment in the Netherlands (RIVM)で開発している。統計解析ソフト(S-PlusまたはR)が必要)。数理モデルとしてBMDSのモデルに加えExponentialが使用できる。
- MADr-BMD**(Model Averaging for Dichotomous Response Benchmark Dose)
 Wheeler & Bailer(2007)により開発されたソフトウェアで、BMDSで使用される数理モデルから得られるBMDLをAIC (Akaike 1978),などの統計量に基づいてフィッティングの重み付けを行い、平均化する。
- EFSAが開発中のBMDモデル平均化の計算サイト**
 平均化手法はMADr-BMDと同じだが、BMDLの計算はPROASTを使う。
<https://shiny-efsa.openanalytics.eu/app/bmd>

5

第72回JECFA会議でのベンチマークドーズ法の適用

アクリルアミド	ラットの2年間の発がん性試験	BMDS
ヒ素	飲料水濃度と肺発がんの疫学データ	BMDS
水銀	無機水銀によるラットの腎臓重量増加	BMDS
フラン	雌ラットの肝臓がん	BMDS
過塩素酸	ボラティアに対するヨード取り込み阻害	PROAST
DON	ブタの嘔吐反応	PROAST



EPA BMDSVer2.1による計算結果の例

Model Name	Gamma	Logistic	LogLogistic	LogProbit	Multistage	Probit	Weibull	Quantal-Linear
AIC	321.59	321.786	321.489	322.546	321.59	321.779	321.59	321.59
Chi-square	1.94	2.14	1.84	2.91	1.94	2.13	1.94	1.94
P-value	0.5845	0.5437	0.6069	0.4059	0.5845	0.5451	0.5845	0.5845
Specified Effect	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Risk Type	Extra risk	Extra risk	Extra risk	Extra risk	Extra risk	Extra risk	Extra risk	Extra risk
Confidence Level	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
BMD	0.718974	0.930555	0.559789	1.33599	0.718974	0.9269	0.718962	0.718962
BMDL	0.449433	0.661157	0.301806	0.860224	0.449433	0.661213	0.449433	0.449433

食品安全委員会におけるベンチマークドーズ法の検討例

- ・メチル水銀(2004) ヒト(疫学) 運動機能、注意、視覚空間等
- ・1,1-ジクロロエチレン(2007) ラット 肝小葉中心性脂肪変性
- ・ブロモジクロロメタン(2009) ラット 肝脂肪変性
- ・トリクロロエチレン(2010) ラット 胎児心臓奇形
- ・無機ヒ素(2013) ヒト(疫学) 皮膚病変、胎児死亡・乳幼児死亡、発がん等
- ・オクラトキシン(2014) ラット 腎細胞腺腫及びがん組合せ頻度
- ・フタル酸ベンジルブチル(2015) ラット 胎児精巣の位置異常等
- ・アクリルアミド(2016) マウス ハーダー腺腫、ラット 坐骨神経軸索変性等
- ・フモニシン(8/31時点でパブリックコメント中) マウス 肝細胞の巨細胞化等

用量反応性評価ガイドンスの比較

- US EPA “Benchmark Dose Technical Guidance (June 2012)”
- EFSA “Guidance of the Scientific Committee on Use of the benchmark dose approach in risk assessment (2009)”
- WHO IPCS “Principles for Modeling Dose-Response for the Risk Assessment of Chemicals (2009)”

		US EPA	EFSA	IPCS
BMR	二値データ	10%	10%	記載なし
	連続値データ	①生物学的に有意な最も小さい変化レベル ②非連続データに変換 ③ 対照群の1SD	5%	記載なし
モデル適合度の評価		適合度の臨界値P = 0.1	P = 0.05	P = 0.1
モデル/BMDL 選択		モデル依存性がある (BMDLの幅が広い)場合は最も低いBMDL値 モデル依存性がない (BMDLの幅が狭い)場合は、AICが最も低いモデル	最も低いBMDL値	AIC推奨、ただし確率分布が異なるモデルに関しては検討が必要

AIC: 赤池情報規準

8

p-value

EPA BMDSでは、統計モデルに基づく用量反応曲線と実測データとの適合度をカイ二乗検定により評価している。p値が小さい統計モデルは、実測データから有意に乖離していると考えられ、EPA BMDSでは、 $p > 0.1$ となる(乖離しているとは言えない)モデルについて、フィッティングが適合していると判断している。

AIC 赤池情報量基準 Akaike Information Criterion

異なる統計モデルの良さを比較するための指標であり、モデルの複雑さと、測定データとの適合度とのバランスを表している。 $-2 \log(L) + 2p$ (モデルの対数尤度とモデルのパラメータ数)で求められる。AICが小さいモデルほど、バランスがよい統計モデルであるとされる。

統計モデルが測定データによく適合していれば、 $-2 \log(L)$ の値は小さくなる。パラメータの数を増やせば、測定データとの適合度を高めることができるが、ノイズなどの偶発的な変動にも無理にあわせてしまう(オーバーフィッティング)問題が生じるとされる。そこで、パラメータ数が増えると $2p$ の値が増加してAICは大きくなることとなり、この問題を回避する。

9

BMDS 2.6.0.1に収載されている 非連続データの用量反応モデル

モデル	数式	パラメータの数	Restriction
Logistic	$\frac{1}{1 + e^{-(\alpha + \beta \text{dose})}}$	2	-
Probit	$\phi(\alpha + \beta \text{dose})$	2	原点付近の傾きが無限大になることを防ぐための制限
Log-logistic	$\gamma + \frac{1 - \gamma}{1 + e^{-(\alpha + \beta \ln(\text{dose}))}}$	3	$\beta (\text{slope}) \geq 1$
Log-probit	$\gamma + (1 - \gamma)\phi(\alpha + \beta \ln(\text{dose}))$	3	$\beta (\text{slope}) \geq 1$
Weibull	$\gamma + (1 - \gamma) \cdot (1 - e^{-\beta \text{dose}^\alpha})$	3	$\alpha (\text{power}) \geq 1$
Gamma	$\gamma + (1 - \gamma) \cdot \frac{1}{\Gamma(\alpha)} \int_0^{\beta \text{dose}} t^{\alpha-1} e^{-t} dt$	3	$\alpha (\text{power}) \geq 1$
Quantal linear	$\gamma + (1 - \gamma) \cdot (1 - e^{-\beta \text{dose}})$	2	単調増加でない曲線になることを防ぐための制限
Multistage two-stage (n=2) three-stage (n=3)	$\gamma + (1 - \gamma) \cdot (1 - e^{-\sum_{j=1}^n \beta_j \text{dose}^j})$ n = degree of polynomial	3 4	Betas ≥ 0 Betas ≥ 0

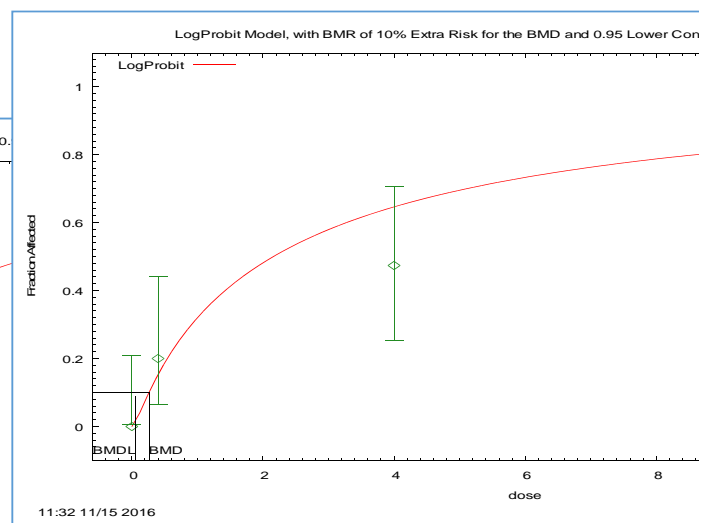
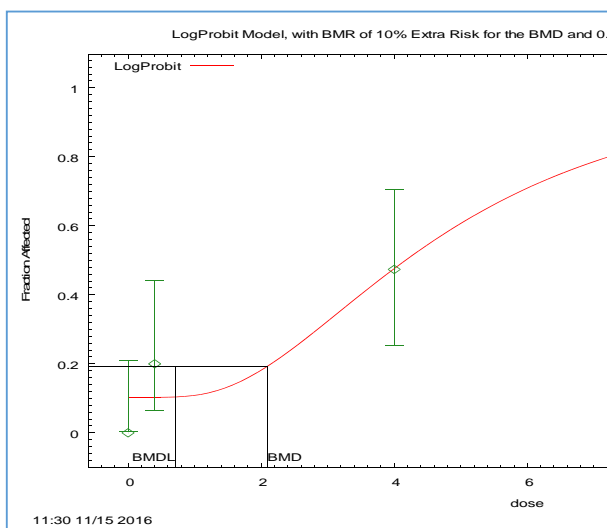
- ✓ EPA BMDSでは、実測データをモデルにフィッティングさせる段階で、パラメータに制限 (Restriction) を設けるオプションを選択することが可能。
- ✓ 生物学的に説明できない用量反応曲線にならないように、Restriction onとoffの両方でフィッティングすることを推奨。

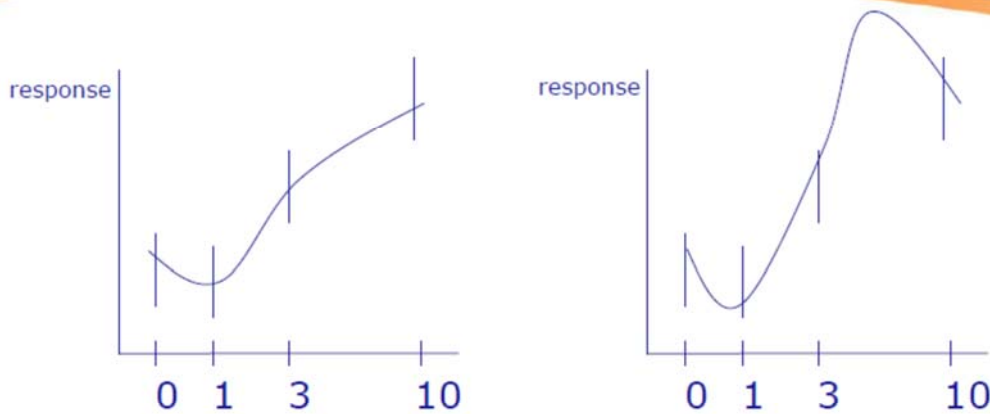
(EPAテクニカルガイダンス2.3.3.3より)

10

Log-Probit: Restriction off

Log-Probit: Restriction on





accepted by the data
but biologically implausible

Restrictions on the models are needed

8

EFSA (2011) TECHNICAL REPORT: Use of BMDS and PROAST software packages by EFSA Scientific Panels and Units for applying the Benchmark Dose (BMD) approach in risk assessment

12

「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」のガイダンス（案）の要点

平成22-24年度 食品安全委員会食品健康影響評価技術研究

BMDSまたはPROASTを使用する場合

- BMRの設定

二値データの場合: 10%の過剰リスクを推奨

連続データの場合: 対照群の1SD (1標準偏差)を推奨

- モデルへのフィッティングパラメータとして係数のRestrictionの有無

選択したモデルの生物学的意義付けによるRestrictionの有無の選択に科学的理由が無い限り、Restrictionがonとoffの両方でのフィッティングを試行することを推奨。

- 適合モデルの選択

原則的には、各々のソフトウェアの適合判定で適合とされたモデルから算出されたBMDLの中から、安全サイドの観点に立って、最も低いBMDLを選定する。

BMD/BMDL比が10以上、もしくは最低用量/BMDL比が100以上となる場合は、BMDL値の信頼度が低いことから、除外することを推奨する。

最終的には、作成された各々のモデルのグラフを目視して、モデルが適合しているかどうか判定する。

13

アクリルアミドの評価におけるBMDL算定方法（食品安全委員会2015）

① BMR の設定

統計学的に検出可能なレベルを考慮し、BMR として10%を採用した。

②BMD 関連指標の算出

EPA のBMDS ver. 2.5 を使用し、Gamma、Logistic、Log-Logistic、Multistage、Probit、Log-Probit、Quantal-Linear及びWeibull の各モデルを用いて解析した。明らかに生物学的な用量反応関係に適合しないモデル曲線(用量0 で無限大の傾きをとる曲線)を除くため、選択ができるモデルについては「Restrict ON」とした。(尚、確認のため「Restrict OFF」条件でも計算を行い、モデル間でBMD 値に実質的な差がない(10 倍未満)ことを確認した。)

③モデルの適合性の評価

適合モデルの適合性を判定するために、以下の基準を適用した。

- 1) 適合度検定P 値 > 0.1
- 2) BMDL/BMD > 0.1
- 3) BMDL/各試験の最低用量 > 0.1

④モデル選択の基準

BMD10 値が最も低いエンドポイントを選択し、サンプリングに伴う不確実性を考慮して信頼区間下限値であるBMDL₁₀ 値を基準点とした。

14

雄マウスのハーダー腺腫/腺癌（NTP 2012）

Model Name	Restrict	P 値	BMD ₁₀ [mg/kg体重/日]	BMDL ₁₀ [mg/kg体重/日]	BMDL ₁₀ /BMD ₁₀	BMDL ₁₀ /最低用量 (1.04 mg/kg 体重/日)
Gamma	ON	0.03	0.38	0.31	0.8	0.3
Logistic	- *	0.00	0.93	0.78	0.8	0.7
Log-Logistic	ON	0.30	0.37	0.17	0.5	0.2
Multistage2	ON	0.03	0.38	0.31	0.8	0.3
Multistage3	ON	0.03	0.38	0.31	0.8	0.3
Probit	- *	0.00	0.97	0.83	0.9	0.8
Log-Probit	ON	0.08	0.62	0.51	0.8	0.5
Quantal-Linear	- *	0.03	0.38	0.31	0.8	0.3
Weibull	ON	0.03	0.38	0.31	0.8	0.3
Gamma	OFF	0.08	0.14	0.02	0.1	0.0
Log-Logistic	OFF	0.30	0.37	0.15	0.4	0.1
Multistage2	OFF	0.40	0.26	0.20	0.8	0.2
Multistage3	OFF	0.68	0.40	0.22	0.5	0.2
Log-Probit	OFF	0.25	0.39	0.16	0.4	0.2
Weibull	OFF	0.11	0.17	0.05	0.3	0.0

*Logistic、Probit、Quantal-Linear のモデルは、Restrict の有無が設定されていないモデルである。

15

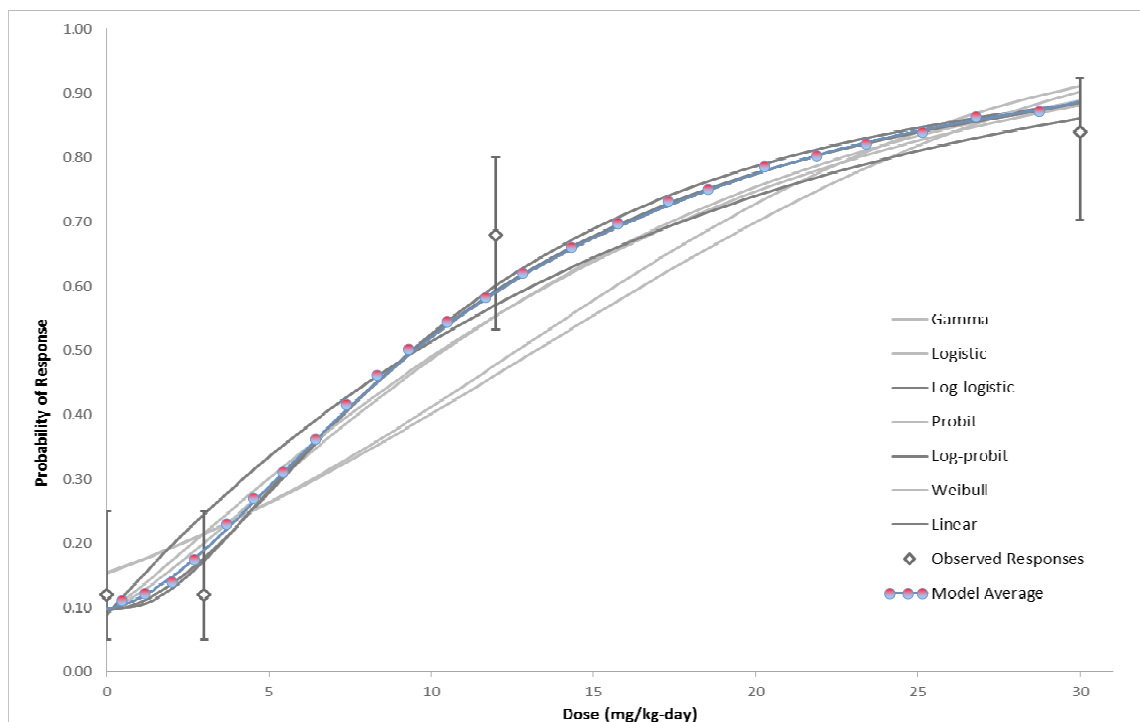
Update: use of the benchmark dose approach in risk assessment

EFSA Scientific Committee,
 Anthony Hardy, Diane Benford, Thorhallur Halldorsson, Michael John Jeger,
 Katrine Helle Knutsen, Simon More, Alicja Mortensen, Hanspeter Naegeli, Hubert Noteborn,
 Colin Ockleford, Antonia Ricci, Guido Rychen, Vittorio Silano, Roland Solecki, Dominique Turck,
 Marc Aerts, Laurent Bodin, Allen Davis, Lutz Edler, Ursula Gundert-Remy, Salomon Sand,
 Wout Slob, Bernard Bottex, Jose Cortiñas Abrahantes, Daniele Court Marques,
 George Kass and Josef R. Schlatter

・ガイダンス(2009)を改訂

モデルの平均化

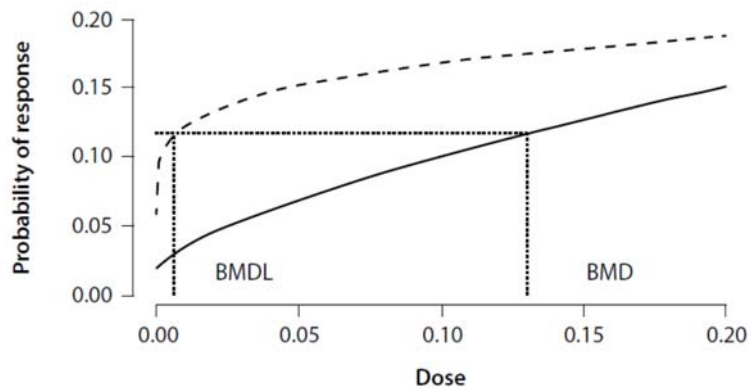
$$w_j = \frac{\exp(-0.5 \cdot IC_j)}{\sum_{i=1}^K \exp(-0.5 \cdot IC_j)}$$



JECFAの動向

第83回JECFA(2016.11)

「統計学的に妥当であるが生物学的に受け入れられない」モデルの取扱いについて議論



対応方針

- ① Restrict ONでモデルを補正
- ② そのまま
- ③ モデルの平均化

実線: Dose-Response 破線: 95%CI Upper bound

- ・第83回JECFAでは、Restrict ONのみBMDLの検討に利用(上記①)。
- ・Restrict OFFのみで検討したBMDL(上記②)及び平均化により算出したBMDL(上記③)は、①で導出したBMDLとの比較目的で利用。

→WHO IPCS(2009)をより適した内容に更新するため、専門家によるワーキンググループで議論すべきと結論

18

3-クロロプロパン-1,2-ジオール (3-MCPD) エステル類の評価

◆ JECFAの評価

67回会合でラットがん原性試験での腎尿細管過形成を根拠
LOEL 1.1 mg/kg/day PMTDI: 2 μ g/kg/day(安全係数500)

83回会合でBMDL10を計算: 0.87 mg/kg/day

PMTDI: 4 μ g/kg/day

(不確実係数200(追加の2は生殖毒性試験の不適切さに対応))

◆ EFSAの評価

JECFAと同じ試験、同じエンドポイント

BMDL10: 0.077 mg/kg/day

TDI: 0.8 μ g/kg/day

(不確実係数: 100)

Renal Tubule hyperplasia

Dose (ppm)	N	Obs
0	50	1
1.97	50	11
8.27	50	21
29.5	50	36

19

Model	Restricted Model	BMD	BMDL	p-value
Logistic	Yes	5.62	4.60	0.00
Log-Logistic	Yes	1.21	0.87	0.61
Probit	Yes	5.36	4.47	0.00
Log-Probit	Yes	3.98	2.96	0.01
Weibull	Yes	2.14	1.66	0.07
Multistage 2°	Yes	2.14	1.66	0.07
Quantal-Linear	Yes	2.14	1.66	0.07
Gamma	Yes	2.14	1.66	0.07
Gamma	No	0.526	0.07	0.92
Log-Logistic	No	0.831	0.22	0.57
Log-Probit	No	0.917	0.27	0.54
Weibull	No	0.63	0.13	0.81

20

一貫性のあるBMD法適用に向けた課題

以下について、統計学的な共通理解が必要

○BMRの設定

○各数理モデルの統計学的背景の理解

○「生物学的に受け入れられないモデル」の取扱い
Restriction ON and/or OFF

○「モデルの平均化」の統計学的意義

○各機関の選択方法の統計学的理解

21