

研究成果報告書（研究要旨）

研究課題名	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) (研究期間：平成22年度～平成24年度)
主任研究者名	研究者名：広瀬 明彦 所属機関：国立医薬品食品衛生研究所

リスクアセスメントの用量反応性評価において、近年その有用性が高まってきているベンチマークドース(BMD)法の適用に関して、我が国におけるガイダンス案の作成を目的として研究を行った。

手法としては、欧米や JECFA 等の国際評価機関で公表されているガイダンスや評価事例を調査、整理して問題点を検討すると共に、WHO や米国 EPA、我が国の既存化学物質安全性評価プログラム等の国内外の評価で採用された用量反応データを解析することにより、一般的な毒性データに BMD アプローチを適用する際の現実的な BMDL (BMD の信頼下限値) の算出基準を検討した。また、BMD アプローチの複合影響への適用性を検証するために、DEHP と DHP のラット複合暴露試験の実施とその結果の解析を行った。

その結果、BMDL の算定基準として、非連続データのベンチマークレスポンス (BMR) は 10%、連続データの BMR は 1SD (標準偏差) を提案した。モデル選択に関しては、BMD/BMDL 比は 10 未満、最低用量/BMDL は 100 未満という基準を提案した。疫学データに関してはハイブリッド法がより適切であると考えられた。一方、複合曝露試験を解析するためにはさらなる定量的アプローチの開発が必要であると考えられたが、今回の実験系は複合曝露影響を研究する系として有用であると考えられた。

最後に、本研究で得られた算出基準を基に BMD アプローチの適用に関するガイダンス案を作成し、解析に使用した BMDL の計算結果と共に Web ページ (<http://dra4.nihs.go.jp/bmd>) に公開した。

研究成果報告書（本体）

研究課題名	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 (研究期間：平成22年度～平成24年度)
主任研究者名	所属：広瀬 明彦 氏名：国立医薬品食品衛生研究所（研究課題番号：1007）

I 研究の期間及び研究目標等

1 研究期間

平成22年度～24年度（3年間）

2 研究目的

化学物質の健康影響評価におけるリスクアセスメントの中で、許容摂取量や耐容摂取量を設定するための用量反応性評価のステップで使用される無毒性量は、実験動物を用いた研究において、その数値が実験デザインに依存することとなり、実験条件によっては無毒性量が求められない場合や、各用量群の公比が大きいときは毒性発現の閾値を正確に反映していない値になる可能性がある。そのため、許容摂取量の設定に際しては、不確実係数の追加や調整が必要となる場合がある。近年これらの問題を回避するために、無毒性量に近い値を算出できるとされているベンチマークドースを無毒性量の代替値として使用し、より客観的な許容摂取量の算出を行うケースが増えてきている。一方、遺伝毒性発がん物質の用量反応性評価においても10%発がん率の信頼下限値（BMDL＝ベンチマークドース）を算出して定量的評価を行う傾向にある。ところが、この客観性が高いとされるベンチマークドースも実際に算出する際には、適用できるエンドポイントの選定や、計算に使用する反応率の設定、適合性の高い数理モデルの選択基準に関しては、ケースバイケースのアプローチが行われてきている。この点に関しては、IPCSやEFSAにおいて基本的なガイダンスがまとめられているところであるが、我が国ではベンチマークドース法の適用例も少なく、その使用に関するガイダンスも存在しない。そこで、本研究ではまず、IPCSやEFSAにおける用量反応性評価ガイダンスの中のベンチマークドースの適用に関する部分を抽出し整理すると同時に、そのガイダンスの中で推奨されているベンチマークドース計算ソフトウェアの特徴も並べて整理する。次に、これらの手法が、現実の様々な評価例でも有効且つ適正であるかどうかについて、これまでに国際機関等でTDI等が設定された物質や、我が国の既存化学物質点検事業で公表されている豊富な実験データを使用し、ベンチマークドース計算値の検証を行う。これらの研究を通して、それぞれのガイダンスや実データの計算過程における長所、短所を見極めたうえで、我が国におけるベンチマークドース値を算出するための基本的な指標を作成する。

上記の調査・計算を主体とした研究成果と分担研究による実証実験を用いた解析結果をもとに、我が国におけるベンチマークドース適用のためのガイダンス案を作成すると共に、得られた計算結果をWeb等で公表し、ベンチマークドース法の適用性に関しての透明性を高めることを目的とする。

3 研究体制

研究項目名	個別課題名	研究担当者名（所属機関名）
1. ベンチマークドース算出の為の指標に関する研究	ア. 国際的なベンチマークドース適用の現状調査研究	分担研究者④（平田睦子）、③（小野敦） （国立医薬品食品衛生研究所）
	イ. ベンチマークドースの算出結果の解析に関する研究	分担研究者③（小野敦）、④（平田睦子） （国立医薬品食品衛生研究所）
	ウ. 疫学データのベンチマークドース算出の為の指標に関する研究	主任研究者（国立医薬品食品衛生研究所）
	エ. ベンチマークドース算出の為の指標(案)の選定	主任研究者、分担研究者③（小野敦）、④（平田睦子） （国立医薬品食品衛生研究所）
2. ベンチマークドース法適用に関する実験的検証に関する研究	オ. 多用量群設定によるベンチマークドースの信頼性検証	分担研究者②（渋谷 淳）（東京農工大）、 ③（小野敦）（国立医薬品食品衛生研究所）
	カ. 基本骨格を同一にする物質の複合暴露影響評価	分担研究者②（渋谷 淳）（東京農工大）、 ③（小野敦）（国立医薬品食品衛生研究所）
	キ. 基本骨格を同一にするが毒性メカニズムが異なる物質の複合暴露影響評価	分担研究者②（渋谷 淳）（東京農工大）、 ③（小野敦）（国立医薬品食品衛生研究所）
3. ベンチマークドース適用のガイダンス案作成と透明性確保に関する研究	ク. ベンチマークドース適用のためのガイダンス案の作成	主任研究者（国立医薬品食品衛生研究所）、 分担研究者②（渋谷 淳）（東京農工大）、③（小野敦）、④（平田睦子） （国立医薬品食品衛生研究所）
	ケ. ベンチマークドース計算結果の公表	主任研究者、分担研究者③（小野敦） （国立医薬品食品衛生研究所）

4 倫理面への配慮について

研究項目1「ベンチマークドース算出の為の指標に関する研究」と3「ベンチマークドース適用のガイダンス案作成と透明性確保に関する研究」については、公表されたデータをもとに研究を行い、動物実験もヒト臨床および個人データも扱わないため、特に倫理面に対する配慮の必要はない。研究項目2「ベンチマークドース法適用に関する実験的検証

に関する研究」については、動物に対する投与実験は混餌ないしは飲水による経口投与が主体となる予定であり、動物の苦痛を最小限にとどめる。また、動物はすべてエーテルないしネンブタール深麻酔下で大動脈からの脱血により屠殺するため、動物に与える苦痛は最小限で済む。また、動物飼育、管理にあたっては、国立大学法人東京農工大学動物実験等に関する規定、米国国立保健研究所(NIH)が推奨している動物倫理に関するガイドラインに従う。

II 研究内容及び成果等

1 研究内容及び方法

(1) 研究項目名：1. ベンチマークドース算出の為の指標に関する研究

1) 個別課題名：ア. 国際的なベンチマークドース適用の現状調査研究

現状で国際的に使用されているベンチマークドースの算出例における問題点とその解決法を模索するために、現在IPCSやEFSA等で公表されている用量反応性評価ガイドランスに記載されているベンチマークドースの具体的な算出基準を抽出し、問題点を整理する。また、2012年6月に最終化された米国EPAの「Benchmark Dose Technical Guidance」を翻訳する（平田分担研究者）。それぞれのガイドランスで推奨されているベンチマークドースソフトウェアの特徴（長所、短所等）についての整理を行う。複合暴露評価に関してベンチマークドースを適用した例を調査する（小野分担研究者）。

2) 個別課題名：イ. ベンチマークドースの算出結果の解析に関する研究

実際の評価に使用されたNO(A)EL値とベンチマークドース値の比較を行いその妥当性を検証するために、これまでに国際的にTDI等が評価されてきた物質（WHOの飲料水ガイドラインや米国EPAのIRIS評価物質）を例にとり、ベンチマークドースの計算を試みる（平田分担研究者）。さらに検証範囲を広げる目的で、化審法における既存化学物質点検事業で反復投与試験が行われた物質について、NOELが設定されたエンドポイントについてのベンチマークドースの算出を試みる（小野分担研究者）。

3) 個別課題名：ウ. 疫学データのベンチマークドース算出の為の指標に関する研究

近年疫学データを用いたベンチマークドースの適用事例も少数例ながら増えてきており、それらの現状と問題点を整理すると共に、疫学データを用いた場合の今後の適正な使用基準（案）の作成を行う。（広瀬主任研究者）

4) 個別課題名：エ. ベンチマークドース算出の為の指標(案)の選定

上記3課題と研究項目2「ベンチマークドース法適用に関する実験的検証に関する研究」の研究成果を踏まえ、国際的な算出基準における問題点の克服や、TDI設定のためのより良い出発点（POD：Point of Departure）を設定するためのベンチマークドースの算出基準（案）を作成することを目的に、様々な算出/選定条件で求めたBMDL値とNOAEL及びLOAEL等との比較解析を行うことにより、モデルの選定基準やベンチマークレスポンス（10%値が良いかSD値が良いかなど）の選択基準の妥当性、複合暴露評価に

における適切な指標等を検証する（広瀬主任研究者、小野分担研究者、平田分担研究）。

（2）研究項目名：2. ベンチマークドース法適用に関する実験的検証に関する研究

1）個別課題名：オ. 多用量群設定によるベンチマークドースの信頼性検証

Diheptyl phthalate (DHP)について、公比2の割合で9用量を設定して、1群あたり15匹の雄性F344ラットに対して混餌投与による90日間の亜急性毒性試験を行う。次いで、肝臓のGST-P陽性細胞巢の数と面積の定量評価、肝臓のGST-P陰性細胞巢の α 2マクログロブリン(α 2M)やPTEN, TGF β シグナル、p38MAPK等の免疫染色による反応性の解析、DHPの肝臓でのPPAR α アゴニストとしての活性の用量反応性を検討するためのペルオキシゾーム酵素、 α 2Mの転写レベルの解析を行う。また、精巣の精上皮に対するPPAR α アゴニストとしての毒性評価（形態計測による定量評価、グレーディングによる評価）、精巣上皮への脱落精上皮細胞の出現頻度の評価（定量評価、グレーディングによる評価）と共に、細胞増殖活性、アポトーシスの定量解析を行う。次いで、数量データとカテゴリカルデータの用量反応性に関して、全用量群の詳細な実データを検証用データセットとして用いて、少ない用量群の組み合わせで算定されたベンチマークドースのバラツキの度合いを検討し、選択用量及びデータの種類による再現性の度合いにルールが存在するのかどうかと、信頼性のある用量設定を検証する。カテゴリ値は必要に応じて数値を割り振って数量値への変換を図り、その妥当性も検討する。

（渋谷分担研究者、小野分担研究者）

2）個別課題名：カ. 基本骨格を同一にする物質の複合暴露影響評価

基本骨格を同一にするフタル酸エステル類に属する物質の複合影響に関する実証実験として、PPAR α アゴニスト活性が明らかなdi (2-ethylhexyl) phthalate (DEHP)、di-*n*-butyl phthalate (DBP)と、22年度の検索でPPAR α アゴニスト活性による影響が明らかではなかったDHPを用いて、DEHPとDBP、DEHPとDHPの組み合わせで共投与による90日間のラット亜急性毒性試験を行う。次いで、肝臓毒性に関する毒性評価（病理組織学的変化に基づいたスコアリング、肝前がん病変指標としてGST-Pおよび α 2-macroglobulin (α 2M)の免疫組織化学染色によるGST-P陽性肝細胞巢ないしPPAR α アゴニスト投与で形成される α 2M陽性肝細胞巢の形態計測による定量評価、 α 2Mおよびペルオキシゾーム酵素(acyl-coenzyme A oxidaseなど)の転写レベルで評価)ならびに肝臓におけるDNA傷害の指標である8-OHdGレベルの測定、更には精巣毒性に関する毒性評価（病変の形態計測による定量評価、グレーディングによる評価）を行い、それぞれのエンドポイントでのベンチマークドース値に基づく相加作用の有無、あるいは干渉性の有無を検証する。（渋谷分担研究者、小野分担研究者）

3）個別課題名：キ. 基本骨格を同一にするが毒性メカニズムが異なる物質の複合暴露影響評価

基本骨格を同一にするが毒性メカニズムが異なる物質の複合影響として、23年度の検索で用いたDEHPとDHPについて、一方の物質の用量を固定した上でもう一方の物質

の用量について公比2の割合で4用量を設定し、雄性F344ラットに対して混餌にて共投与による90日間のラット亜急性毒性試験を行い、肝臓のGST-P陽性細胞巢や陰性細胞巢の形態計測による定量評価、精巢毒性に関する毒性評価（病変の形態計測による定量評価、グレーディングによる評価）を行い、それぞれのエンドポイントでのベンチマークドース値に基づく相加作用あるいは干渉性の有無を検証する。（渋谷分担研究者、小野分担研究者）

（3）研究項目名：3. ベンチマークドース適用のガイダンス案作成と透明性確保に関する研究

1）個別課題名：ク. ベンチマークドース適用のためのガイダンス案の作成

研究項目1「ベンチマークドース算出の為の指標に関する研究」および2「ベンチマークドース法適用に関する実験的検証に関する研究」の研究成果を受けて、用量反応性評価におけるベンチマークドースの適用基準（数理モデルおよびベンチマークドースレスポンスの選択、総合的算出方法、複合暴露評価に使用する際の基準等）に関するガイダンス案を作成する。（広瀬主任研究者、渋谷分担研究者、小野分担研究者、平田分担研究者）

2）個別課題名：ケ. ベンチマークドース計算結果の公表

個別課題イで行ったベンチマークドース値を上記ガイダンス案にしたがって再検討を行い、得られた結果を総合評価研究室のWebページから公表することにより、計算結果の透明性を確保する。（広瀬主任研究者、小野分担研究者）

2 研究成果、考察、今後の課題

（1）研究項目名：1. ベンチマークドース算出の為の指標に関する研究

1）個別課題名：ア. 国際的なベンチマークドース適用の現状調査研究（担当：平田分担研究者、小野分担研究者）

米国EPA、EFSA及びIPCSにより作成された用量反応性評価（ベンチマークドース法）に関するガイダンスの日本語訳を行い、Web公開した

(<http://dra4.nihs.go.jp/bmd/>)。

米国EPAは、2000年10月に“Benchmark Dose Technical Guidance Document (external review draft)”を公表した。このガイダンスドキュメントは長い間ドラフト版のままであったが、最近、最終化され、2012年6月に“Benchmark Dose Technical Guidance”が公表された。EPAのガイダンスでは、BMDアプローチを適用するために解決しなければならない多くの問題が取り上げられている。2009年に公表されたEFSAの“Guidance of the Scientific Committee on Use of the benchmark dose approach in risk assessment”では、BMDアプローチの適用法に関する簡潔な説明といくつかの勧告が示された。一方で、IPCSが、2009年に公表した、“Principles for Modelling Dose-Response for the Risk Assessment of Chemicals” (IPCS, 2009)は、EPAやEFSAのガイダンスとは異なり、その

対象をBMDアプローチに限定していない。用量反応性評価手法全体についての解釈や用量反応モデリングの長所及び短所が示された。

それぞれのガイダンスに記載されているBMD算出基準を比較した結果、主にBMR (Benchmark Dose Response)の選択法、モデル適合度の評価法及びモデル選択法について、下記の通り違いが認められた。

BMRの選択

非連続データに関しては、EPA及びEFSAのガイダンスでは、デフォルトBMRとして、過剰リスク10%が提案されている。一方、連続データについては、EFSAのガイダンスがデフォルトBMRを5%としたのに対し、EPAのガイダンスでは下記3つの方法が示された。① 生物学的に有意であると一般的に考えられる、最小の変化レベルがある場合は、その変化レベルをBMRとすることができる。② 個別データがあり、有害と考えられる変化レベルを決めることが可能な場合は、ハイブリッドモデルを用いるかカットオフ値に基づき、データを非連続データに変換することが可能である。③有害な反応レベルについてアイデアがない場合には、対照群の標準偏差 (SD)に相当する変化レベル (より重篤な影響については “0.5 x SD”などより低い値)をBMRとして用いるべきである。なお、IPCSのガイダンスでは、BMR選択に関わる技術的な問題やBMRを決定する際に考える必要のある影響について議論しているものの、特定の値については論じていない。

モデル適合度の評価：

EPA及びIPCSのガイダンスでは、包括的適合度を指標に、 $P = 0.1$ を適合度の臨界値として、評価を行うことが推奨されている。EFSAのガイドラインにおいても、適合度検定により評価を行うよう記載されていたが、推奨された適合度の臨界値は $P = 0.05$ とされた。EPAとIPCSのガイダンスには、さらにプロットの目視により、モデルが適切にデータを描写しているかチェックする必要があると述べられていた。

モデル選択及びBMDLの設定：

同じファミリーに属するモデル間の比較については、いずれのガイダンスにおいても、尤度比検定を用いて評価することができると記載されていた。異なるファミリーに属するモデル間の比較に関しては、EPAのガイダンスには、算出されたBMDL値が近い場合には、明確なモデル依存性はないと判断して、赤池情報量規準 (AIC)の最も低いモデルを選択することができるとの記載があった。BMDLが十分に近い値にならなかった場合には、モデル不確実性が大きすぎないか専門家による統計学的判断等が行われ、BMDL値の幅が妥当と判断された場合には、最も低いBMDLを選択することができる、とされた。EPAのガイダンスでは、追加モデルの検討やモデルのパラメータ値の調査、BMDの評価などさらなる解析や議論の可能性が指摘されており、モデル選択/BMDL設定の方法は常に議論されるべきであるとされた。IPCSのガイダンスではAICを用いることが推奨されているが、想定した確率分布が異なるモデル間の比較をする場合に

はさらなる検討の必要性が指摘された。一方、EFSAのガイダンスでは、より単純に最も低いBMDLを採用することが推奨された。

米国EPAのIntegrated Risk Information System (IRIS)で評価が行われ、Webから評価文書 (Toxicological Review)を入手することができた81物質の中から、BMDアプローチを用いてRfD/RfCが算出された46物質 (RfD: 21物質、RfC: 35物質)について、BMD解析手法を調査した。その結果、RfD/RfCの基となったデータは、連続データ21個及び非連続データ32個であった。連続データ21個のうち、12個はBMR=1SDとし、5個はBMR=10%、3個はBMR=5%として算出したBMDLを基にRfD/RfCが算出された。非連続データについては、32個中28個についてはBMR=10%、残りの4個はBMR=5%として算出したBMDLを基にRfD/RfCが算出された。NOAELが設定されたデータについて、BMDLとの比較を行った結果、図1のようになった。NOAEL及びLOAEL両方が設定されたデータ36個のうち、BMDLがNOAELとLOAELの範囲以内に収まったのは、12個であった。19個のBMDLはNOAELより低い値となり、5個のBMDLはLOAELを超える値となった。

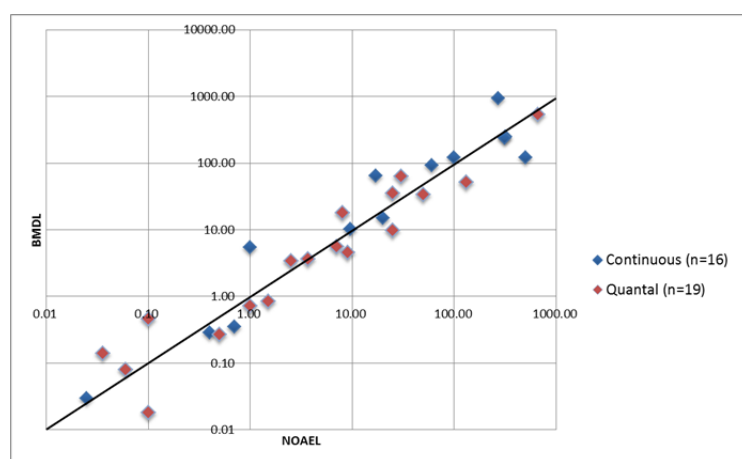


図1. BMDLとNOAELとの比較 [米国EPA IRIS評価物質]

2) 個別課題名：イ．ベンチマーク ドースの算出結果の解析に関する研究（担当：小野分担研究者、平田分担研究者）

実際の安全性評価に使用されたNO(A)ELやLO(A)ELとベンチマーク ドース値の比較を行い、ベンチマーク ドース手法の妥当性や有用性を検証することを目的として、これまでにTDI等が評価された物質について、TDI等の設定根拠となった用量反応データを収集し、解析を行った。

WHOで飲料水水質ガイドライン設定のために評価が行われた約160物質と米国EPAのIRISで評価が行われた約550物質について、評価文書を解析した。TDI、ADIやRfDの設定根拠となった毒性試験から、反復経口投与毒性試験を抽出し、原著を収集した

が、農薬関係の報告書のほとんどは非公開であり、原著を入手できたのはそれぞれ24物質 (WHO)及び104物質 (IRIS)のみであった。さらに、対照群を含めて4群以上のデータに絞り、その中から用量相関性を示すデータを抽出した結果、この基準に適合するデータは、WHOの飲料水ガイドライン評価物質では15エンドポイント (非連続データ: 8, 連続データ: 7)、米国EPAのIRIS評価物質では47エンドポイント (非連続データ: 23, 連続データ: 24)のみであった。次に、化審法のもとで行われている既存化学物質安全性点検の対象物質のうち、一群あたりの動物数が比較的多い、反復投与・生殖毒性併合試験の結果より、NOELの設定根拠となった一般毒性エンドポイントについてベンチマークドースの計算に使用可能なエンドポイントの抽出を試みた。その結果、232エンドポイント (非連続データ: 112, 連続データ: 120)を得ることができた。非連続データについては、さらに、化審法の既存化学物質点検事業で行われた28日間反復投与毒性試験からデータ抽出を行い、83エンドポイントを得ることができた。

米国EPAにより開発されたBenchmark Dose Software (BMDS, 2.1.2)とオランダの国立保健医療環境研究所 (RIVM)により開発されたPROAST(23.2及び28.1)を用いてBMDLの算出を試みた。非連続データに関しては、計226のデータを、BMDSでは8種のモデル (Gamma、Logistic、LogLogistic、LogProbit、Multistage、Probit、Weibull、Quantal-Linear)に、PROASTでは10種のモデル (one-stage、two-stage、log-logist、Weibull、log-prob、gamma、logistic、probit、E2、H2)に適合させた。ベンチマークドース算出のための指標 (案) の選定 (個別課題名エ)を行うために、様々な条件下でBMDLを算出した。BMRは5%もしくは10%とし、定数に制限を用いる (Constraint/Restriction) ことができる場合には、“ON”と“OFF”の両方の場合で計算を行った。さらに、EPAガイダンス (ドラフト版)で推奨された方法 (BMDLの範囲が3倍以内の場合は、AICの最も低いモデルから算出されたBMDLを採用。BMDLの範囲が3倍以上の場合はもっとも低いBMDLを採用; EPA方式)とEFSAのガイダンスで推奨されている方法 (最も低いBMDL値を採用; EFSA方式)両方で最終的なBMDL値の選定を試みた。

連続データについては、計151データについてBMDLを算出した。BMDSでは、5種のモデル (Exponential、Hill、Linear、Polynomial、Power)に、PROASTでは、2種のモデル (Hill及びExponentialモデル)にデータを適合させた。BMRは1SD (対照群の標準偏差に相当する変化)、5%もしくは10%とし、非連続データ同様に、Constraint/Restrictionが“ON”と“OFF”両方の条件下で算出したBMDLから、EPA方式及びEFSA方式により最終的な値を選択した。

3) 個別課題名: ウ. 疫学データのベンチマークドース算出の為の指標に関する研究
(担当: 広瀬主任研究者)

疫学データを用いた場合のベンチマークドースの適用に関する適正な使用基準（案）を作成するために、まず近年評価されている疫学データへの適用事例の情報収集と問題点の整理を行った。その結果、疫学データを用いたベンチマークドース手法に大きく分けて3つの手法が用いられていた。一つは、“全データに合致した量-影響曲線を選択し、その曝露0の健康影響指標の反応量が5%ないし10%変化する時の曝露量として数理統計学的に算出する”手法で、EFSAが食品中の鉛に関する科学的意見書で使用したものである。別の手法としては、原理的にBudtz-Jørgensenら（2001）のHybrid法に基づくものがある。この手法は、米国のNRC(National Research Council, 2000)で行われたメチル水銀の評価やEFSAのカドミウムの評価で使用された。さらに、有症率と相関する曝露指標値をランク分けして、そのランクの中央値や平均値などの代表値としたグループと有症率についてベンチマークドースを計算する手法がある。この手法は原理的には、動物実験データに於いて通常使用されるBIDSやPROASTをそのまま適用することが可能となり、2010年のJECFAにおけるヒ素や過塩素酸の評価に用いられた。

以下に、それぞれの手法に対しての適用状況の調査結果や問題点等について示す。

・EFSA法（EFSA Journal 2009; 7(10):1351）

EFSA が推奨するBMD 法では、非連続データの初期設定値としてベンチマークレスポンスとしてBMR=10%を推奨している。また、連続データの初期設定値はBMR=5%であるが、大数の標本を扱った疫学データの場合にはBMR=1%が用いられているとしている。EFSA の食品中の鉛に関する科学的意見書では、ベンチマーク算出の基となったエンドポイントである“収縮期血圧は集団全体の平均値または年間値が1%上昇するだけでも、心血管疾患の発症や冠動脈疾患死亡のリスク上昇につながるため、公衆の健康に関わる問題となると判断した。その結果、平均120 mmHg の収縮期血圧の1%に相当する1.2 mmHg の血圧の上昇をBMRとし、BMDL01（BMDの1%下限値）を算定した。EFSAでは成人血中鉛濃度に関する5つの選定した研究の結果に基づき、横断的研究データからは集団単位での全体平均値の上昇に関するBMD01のみしか算出できないのに対し、縦断データでは個人の収縮期血圧の年間平均上昇率1%のBMD01を算定できた。その結果、収縮期血圧に対し血中鉛レベル1.5から7.1 $\mu\text{g}/\text{dL}$ まで（縦断研究2.7～7.1 $\mu\text{g}/\text{dL}$ 、横断研究1.5～2.1 $\mu\text{g}/\text{dL}$ ）の4 つのBMDL01値を算定し、平均BMDL01を3.6 $\mu\text{g}/\text{dL}$ と算出した。

しかし、この算定手法については、連続データの初期設定値はBMR=5%に対してBMR=1% を設定した根拠、鉛曝露のない集団の平均収縮期血圧は120 mmHgや冠動脈疾患（または高血圧）に対する基準点の妥当性、飲酒などの血圧に影響を及ぼす交絡因子の調整について、疑問点のあると考えられた。因みに、EFSA が引用したNash らの論文では、4分割した集団の血中鉛濃度の増加に伴い血圧（4.0～31.1 $\mu\text{g}/\text{dL}$ 群のみ有意）と高血圧の発症Odds比は高くなるが、いずれのOdds比も有意でないことが報告されている。また、このEFSAが用いた対象集団の血中鉛の曝露範囲はIPCS による臨界濃度よりもかなり低いものであり、収縮期血圧の1%増加（1.2 mmHg の増加）は加齢によって

も引き起こされうる値である。したがって、この手法では、正常値や影響と見なすための基準点（またはBMR）の生物学的妥当性や交絡因子の調整の必要性について、別途検討する必要があることが示された。

・ハイブリッド法

この手法は、非曝露集団において設定された異常率 P_0 に対して、曝露集団における $\alpha\%$ （benchmark response、BMR）の異常増加をもたらす曝露濃度を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。

米国のNRC(National Research Council, 2000)で行われたメチル水銀のベンチマークドース評価では、セイシェル、ニュージーランドおよびフェロー諸島での調査研究が評価の対象となり、Hybrid法が用いられBMRとしては、5%が採用された。しかし、その採用にあたっては、科学的情報に基づく政策的判断が含まれるとしている。たとえば、ニュージーランドおよびセイシェル調査の分析において0.1のBMR（すなわち、母集団内の10%がリスクを有する）を使用した場合、神経学的成績が異常領域に入る人のパーセンテージが3倍にもなり得る。NRCは、人々の健康をより確実に保護するため、 P_0 として0.05、BMRとして0.05（すなわち、母集団内の5%がリスクを有する）を使用した。しかし、これらの調査では暴露量が真にゼロの被験者がおらず、ゼロにおける平均反応率はフィットさせたモデルから外挿されたものとなっている。このため、リスク評価のベースとして量反応モデルを選択する際には、特に注意が必要であるとされている。最終的には、PCB暴露による偏りのおそれがあるもののフェロー諸島の調査研究が出発点を規定するための臨界用量を選択する上で、最も精度の高い評価が可能であるとされ、臍帯血中Hg濃度に基づく量反応データをKパワーモデル（ $K \geq 1$ ）でモデル化することが最良のモデルであると提案された。

EFSAでのカドミウムの評価（2009）においてHybrid法が用いられ、尿中カドミウムと尿中 β_2 -ミクログロブリンの間の量-反応関係を扱った研究を収集・選択してメタ分析した。この解析では、全集団と50歳以上の集団から得られたこれら2指標間にHillモデルを適用し、 β_2 -ミクログロブリンの尿中高値（ $300 \mu\text{g/g Cre}$ 以上）の有病率を5%上昇させる尿中カドミウムのBMDLはHybrid法を用いて $4 \mu\text{g/g Cre}$ であると推定しており、鉛の場合とは異なったアプローチを採用した。

・クラス分け法（JECFA 2010）

第72回JECFA会議（2010）においては、ヒ素と過塩素酸についてヒトの疫学データ等がベンチマークドースの算定に用いられた。無機ヒ素については、飲料水汚染による慢性中毒および発がん性に関する疫学研究が古くより行われており、JECFA会議（1983, 1989）においては、疫学研究における飲料水中濃度に基づき暫定週間耐容摂取量：PTWI：0.015 mg/kg体重/weekとして設定されていた。IARCの評価（1980）ではヒトに対して発がん性あり（皮膚、肺および膀胱癌）というグループ1に分類されている。第72回JECFA会議（2010）ではより定量性の高い評価を行うためにベン

チマークドースを用いた解析を行った。無機ヒ素は経口暴露による標準的な実験動物を用いた発がん性試験においては明確な催腫瘍性を示さないため、ベンチマークドース手法を用いた解析には、最近のTaiwanにおける研究として対象人数やフォローアップ期間の長いChenら（2010）の報告を用いて行われた。ここでは、表3に示す様に、飲料水中のヒ素濃度をランク分けして、それぞれランクの中央値と発症例数を用いてベンチマークドース手法を適用した。また、BMRとしては、用いた疫学研究の対象となった肺がんの背景値に比べて有意な増加を検出する最低の増加率を基に0.5%が選択された。さらに、飲水量から一日摂取量への変換には、飲水量として3 L/day、食品からの寄与は75µg/day、体重を55kgとして換算している。その結果、BMDL_{0.5}は、3.0 - 5.0 µg/kg体重/dayと見積もられた。また、実際には一日飲水量（2-4 L/day）や食品からの摂取量（50 - 200 µg/day）に範囲があることに基づく不確実性を考慮すると、算出された最低値のBMDL_{0.5}：3.0 µg/kg体重/dayは、2.0-7.0 µg/kg体重/dayの範囲を持つと推定された。この評価結果の値は、それまでのPTWIの値とオーバーラップしているため、結論としてPTWIは取り下げられることとなった。

表 1. ベンチマークドースに用いた ヒ素曝露による肺がん発生率 (Chen et al., 2010) の層別化データ

Stratified data for the BMD approach of lung cancer cases with arsenic exposure (Chen et al., 2010)							
Inorganic arsenic in water		Inorganic arsenic total		Cohort	RR	N	Adjusted
Category range	Central estimate	dietary exposure		incidence			cases
(µg/l)	(µg/l)	µg/day	µg/kg /day				
<10	5	90	1.6	0.021	1	2288	48
10-49.9	30	165	3.0	0.023	1.1	2093	48
50-99.9	75	300	5.5	0.021	0.99	907	19
100-299.9	200	675	12.3	0.032	1.54	909	29
≥ 300	450	1425	25.9	0.047	2.25	691	33

以上の結果より、疫学データに対するベンチマークドース手法を適用するためには、データの用量依存性を確認することに加え、以下の点が重要であると考えられた。

- BMRは、データの持つ生物学的、統計学的意義等に合わせた値に設定する。
- 交絡因子の存在の有無やその調整を行う。
- 用量（他の暴露源も含む）と結果（生涯リスクへの変換など）を適正に記述するためにデータ調整する。

上記の3つの手法のうち、EFSA法はすべてのデータを変換せずにそのまま扱うことが可能であり、生物学的反応率をBMRとして設定することが可能となり、解析結果の正確性が高いように思えるが、その生物学的意義あるいは毒性学的を決定するた

めの閾値を決定するための根拠を、モデル化した結果からは設定しにくいという欠点がある。また、クラス分け手法はベンチマークドースソフト等への適用性と言う観点からは、動物実験データで用いられるもの同様に行うことが可能なり簡便な解析手法であると考えられるが、クラス分けは生物学的意義よりも測定や解析のしやすさで恣意的に決定される欠点がある。一方、ハイブリッド法はBMRの設定はクラス分け手法と同様に生物学的反応率ではなく、生物学的に有意な反応を示した検体数の増加（あるいは減少）率の変化を基にしているが、生物学反応の有意性を解析可能である点や曝露量を恣意的にクラス分けする必要のないという特徴がある。これらの点を考慮すると、上記の3つの手法はそれぞれに特徴があるものの、最初に適用する手法としてはハイブリッド法が最も適切であると考えられた。

そこで、ハイブリッド手法の基本的な原理とその適用に関して検証すべき点を記載する。例えば、正常集団で影響指数の90%信頼区間を算出し、下限値（図2）では79）以下を異常値と定義すると、影響指数の異常値となる者は集団全体の5%いることになり、非曝露集団における異常率 P_0 は5%となる。曝露集団において、当該異常率 P_0 よりさらに $\alpha\%$ の異常増加（これをbenchmark response、BMRとする）をもたらし曝露濃度をBMDと定義する（ P_0 およびBMRの設定は影響指標により異なるが、通常 $P_0=0.05$ 、 $BMR=0.05$ がよく用いられる）。すなわち、図2の曝露濃度0の集団（非曝露集団）の正規分布を、曝露集団の有意な量－影響関係を示す関数に沿って右方移動し、正規分布の打ち切り（cutoff）値以下の範囲が P_0+BMR となるときの曝露濃度である。このBMDの95%信頼区間の下限値をBMDLと呼び、低濃度下の臨界濃度（閾値）を表すNOAELの代用値と考えることができる。

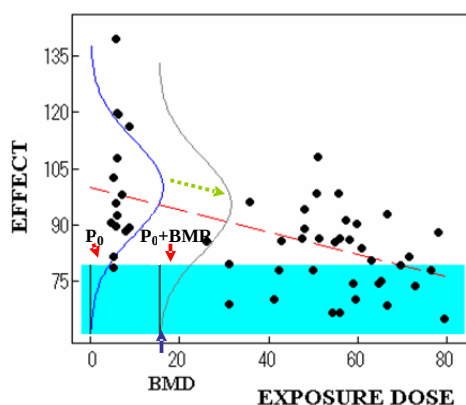


図2 Hybrid法の概念図

本法は原理的にBudtz-Jørgensenら（2001）のHybrid法を用いてBMDおよびBMDLを算出しており、曝露指標 d 、影響指標 $\mu(d)$ 、量－影響関係 $\mu(d) = b_0 + b_1 g(d)$ とすると、 $b_1 > 0$ ならば（曝露を受けると影響指標が増加する）、 $P_0 = 1 - \Phi([C - b_0] / \sigma)$ および $BMD = g^{-1} \{ [\Phi^{-1}(1 - P_0) - \Phi^{-1}(1 - P_0 - BMR)] \sigma$

$/b1\}$ から算出する。また $b1 < 0$ ならば (曝露を受けると影響指標が減少する)、
 $P0 = \Phi([C - b0] / \sigma)$ および $BMD = g^{-1}\{[\Phi^{-1}(P0) - \Phi^{-1}(P0 + BMR)] \sigma / b1\}$ で
 算出する (ベキ関数を $g(d) = d^K$ 、cutoff値を C と表現。また、 K は ≥ 1 である)。BMD
 算出に当たっては、以下の4点を念頭に入れておく必要がある。①図Xでは、説明を
 簡単にするために一次回帰式 $Y_i = b0 + b1X_i$ を当てはめ (Φ は逆正規累積分布関数、
 σ は非曝露集団の標準偏差)、 $BMD = [\Phi^{-1}(P0) - \Phi^{-1}(BMR + P0)] \sigma / b1$ の式で算
 出したが、実際にはベキ関数 (power function、 $b1 \cdot X^K$) が使用されている ($K=1$
 の時は上述の一次回帰式となる)。②非曝露集団の設定は事実上困難なことが多いの
 で、曝露量の極めて低い集団および高い集団を含む曝露集団全体で解析を行う (但
 し、曝露集団の平均曝露濃度が高いとBMDLが極端に低く算出され易くなる)。③性・
 年齢などの交絡因子の影響をBMDの算出過程において除外することも可能である。
 ④曝露集団において有意な量-影響関係が存在しない場合、この解析法は適用でき
 ない。また、BMDLはサンプル数が少ないと極端に低い値になりやすいので、解析に
 当たってサンプル数 (可能な限り100以上) は多いほうが望ましい。

4) 個別課題名：エ. ベンチマークドース算出の為の指標(案)の選定

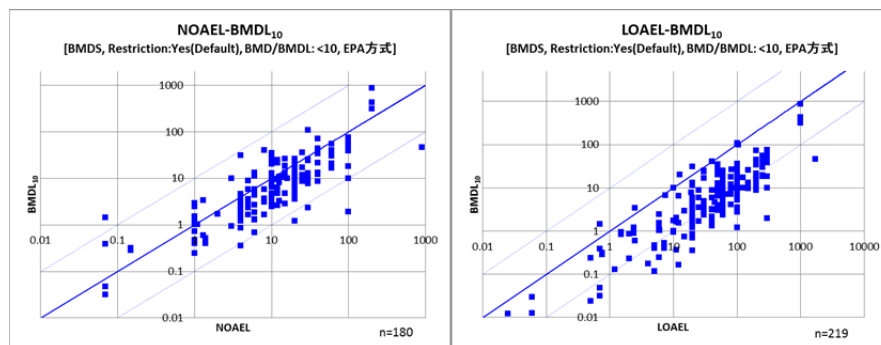
(担当：広瀬主任研究者、小野分担研究者、平田分担研究者)

個別課題イで得られた計算結果を比較解析することにより、各種計算モデルの選定基準やBMRの妥当性を検討した。

非連続データ

最初に、BMR (5%もしくは10%)に関する検討を行った。BMDsを用いて算出し [Restriction: Yes (Default)、BMD/BMDL: <10]、EPA方式で選択したBMDL₁₀及びBMDL₀₅それぞれをNOAEL及びLOAELと比較した結果を図3に示す。対象とした非連続値データ226個のうち、219個についてBMDL₁₀及びBMDL₀₅を算出することができた。NO(A)ELが求められているデータ180個のうち、BMDL₁₀及びBMDL₀₅がNO(A)ELとLO(A)ELの範囲内に収まったものは、それぞれ43個 (23.9%)及び15個 (8.3%)であった。

A. BMR=10%



B. BMR=5%

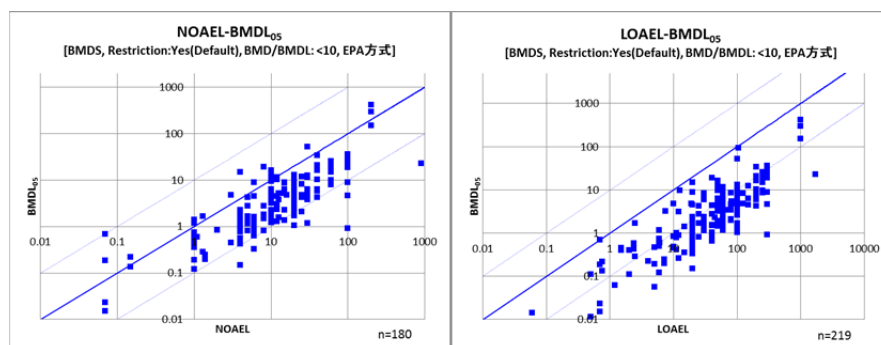
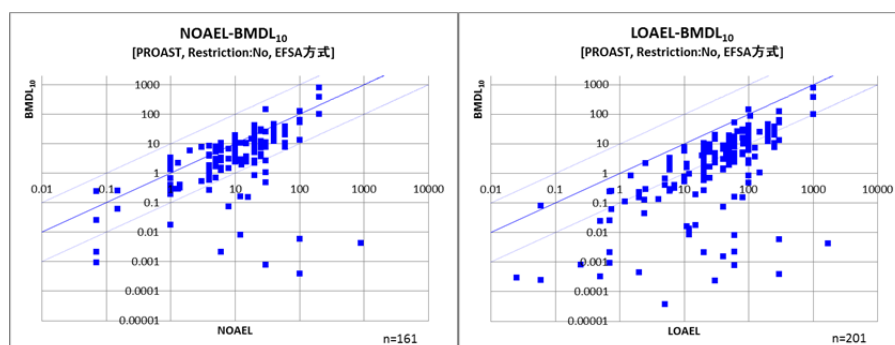


図3. BMDsを用いて算出したBMDL₁₀及びBMDL₀₅とLOAELおよびNOAELの比較

一方、PROASTを用いた結果、201個のデータについてBMDL₁₀及びBMDL₀₅を算出することができた (Restriction No, EFSA方式)。42/161個 (26.1%)のBMDL₁₀及び13/161

個（8.1%）の $BMDL_{05}$ がNOELとLOELの範囲内に収まった（図4）。これらの結果から、非連続データのBMDLを算出する際のBMRとしては10%がより適切と考えられた。

A. BMR=10%



B. BMR=5%

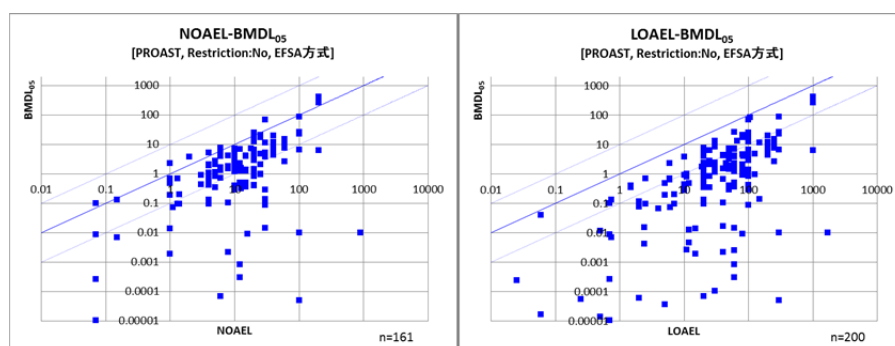


図4. PROAST (Restriction No) を用いて算出した
 $BMDL_{10}$ 及び $BMDL_{05}$ とLOELおよびNOELの比較

次に、モデルへのフィッティングパラメータとして係数のRestrictionの有無について解析を行った。BMDs（2.1.2）では、適用可能な8種のモデルのうち、Gamma、LogLogistic、LogProbit、Multistage及びWeibullモデルについてはdefaultでrestrictionがonとなっており、restrictionをoffにした状態でもBMDLを算出することができる。この5種のモデルについてrestriction offの状態では $BMDL_{10}$ を算出し、それらの値を含めた場合と含めない場合で、最終的な $BMDL_{10}$ 値がどのように変化するか、比較解析した。BMDLの選択はEPA方式で行った。その結果、restriction offで計算した値を含めると、219個中74個（34%）の値が低い値となり、そのうち23個（11%）の $BMDL_{10}$ はdefaultで計算した値から選択した $BMDL_{10}$ の2分の1未満であった。Restriction Noで計算した値を含めて選択した $BMDL_{10}$ とNOEL及びLOELを比較した結果を図5に示す。Restriction Yes（default）で計算した値から選択した結果（図3-A）と比較したところ、RestrictionをNoとして算出した $BMDL_{10}$ を含めない方が、

BMDL₁₀がNOEL/LOELから大きく外れるケースが減り、値が収束していることが明らかになった。

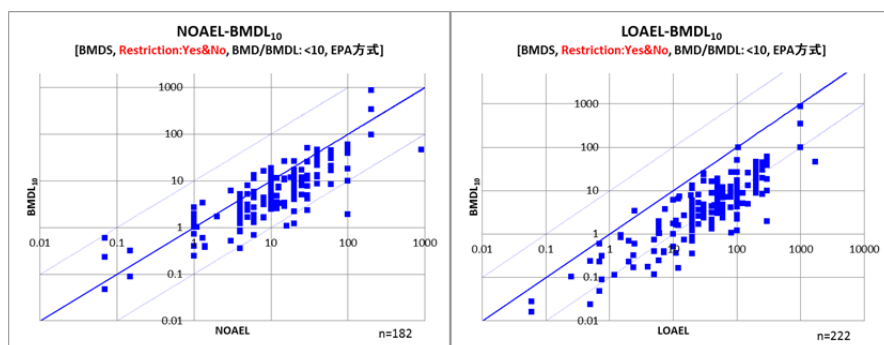


図5. BMDsを用いてRestriction Noで算出したBMDL₁₀とNOEL及びLOELの比較

PROASTを用いてrestriction Yesの状態ではBMDL₁₀を算出し、EFSA方式で最終値を選択した結果、およそ半数（54%）のデータについてはRestriction Noで計算した値と同じ値となった。25%のデータについてはRestrictionをYesとすることでBMDL₁₀が1～2倍に増加し、10%は2～10倍、6%は10倍以上の高い値となった。Restriction Yesの状態では求めたBMDL₁₀とNOEL及びLOELを比較した結果を図6に示す。Restriction Noで計算した結果（図4-A）と比較したところ、BMDsを用いて検討した結果と同様に、RestrictionをYesとした場合の方が、BMDL₁₀がNOEL/LOELから大きく外れるケースが減ることが明らかになった。しかし、Conservativeな意味では、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、両条件で計算を行い、最終値を選択することがより適切と考えられる。

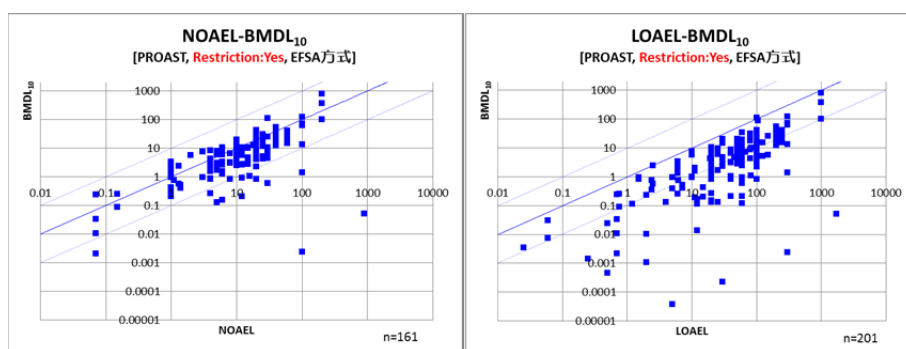


図6. PROASTを用いてRestriction Yesで算出したBMDL₁₀とNOELおよびLOELの比較

BMDsを用い、BMDL値の選択方法について検討を行った。Restriction Yes (default)、BMD/BMDL<10の条件で算出したBMDL₁₀の中から、EPA方式及びEFSA方式で最終値の選択を行った。しかし、EPA方式で最終値を選択しても、ほとんどのデータ

(219個中173個) については、算出されたBMDL値の範囲が3倍を超えていたため、結果として、EFSA方式と同様に最も低いBMDL値が選択された。両方式で異なる値が選択されたのは、216個中31個のみであり、値の変化は3倍未満であった。異なるファミリーのモデルの適合度を比較する際の指標としてAICを用いることは適切ではないとする見解もあり、Conservativeな意味では、EFSAガイダンスで推奨されている通り、最低値を選択することがより適切と考えられる。

算出されたBMDLの不確実性を示す指標として、BMD/BMDL値、最低用量/BMDL値、BMDL値の幅などがある。これらが大きい場合はBMDL値の信頼性が低いと考えられており、EFSAのガイダンスでは、これらの新たな基準を追加する必要性が指摘されている。PROASTを用い、BMR=10%、RestrictionをNoとして、“BMD/BMDLが10未満”となったモデルからEFSA方式で選択したBMDL₁₀とNOAEL及びLOAELを比較した結果を図7に示す。この条件を加えずに計算した結果は図4-Aに示す通りである。新たな条件“BMD/BMDLが10未満”を追加することで、BMDL₁₀がNOAEL/LOAELから大きく外れるケースを顕著に減らすことができた。BMR=10%、RestrictionをNoとして算出したBMDL₁₀の幅とBMD/BMDL値を比較したところ、明確な相関関係がみられたことから（図8）、“BMD/BMDLは10未満”の条件を追加することで、BMDL₁₀の幅の広いデータも概ね排除することができると考えられる。

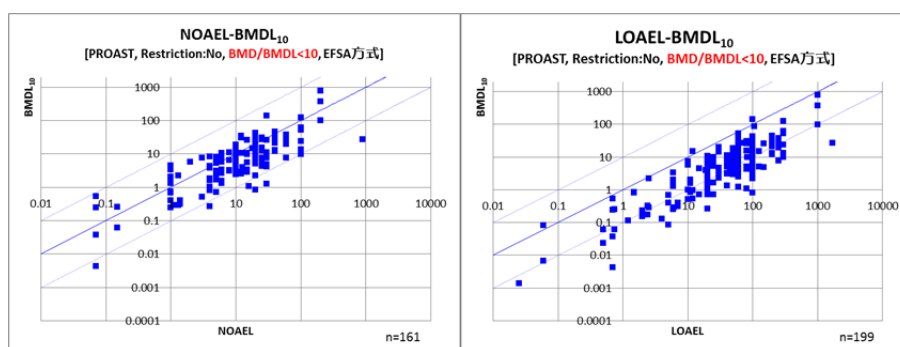


図7. “BMD/BMDLが10未満”となったモデルから選択したBMDL₁₀（PROAST, Restriction No EFSA方式）とNOAEL及びLOAELの比較

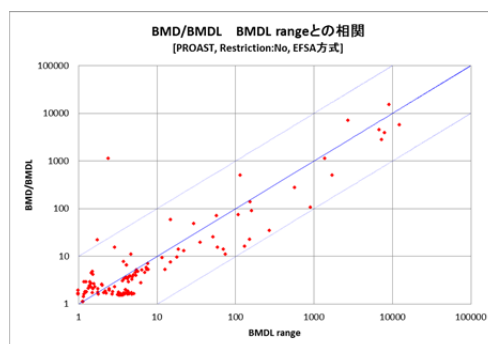


図8. BMDL₁₀の幅とBMD/BMDL値の比較（PROAST、Restriction No、EFSA方式）

BMDs及びPROASTを用い、同じ条件 [Restriction Yes & No, BMD/BMDL<10、最低用量/BMDL<100、EFSA方式]で算出/選択したBMDL₁₀を比較したところ、概ね近い値が算出/選択された (図9)。

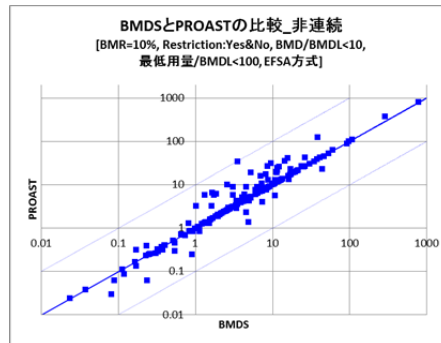
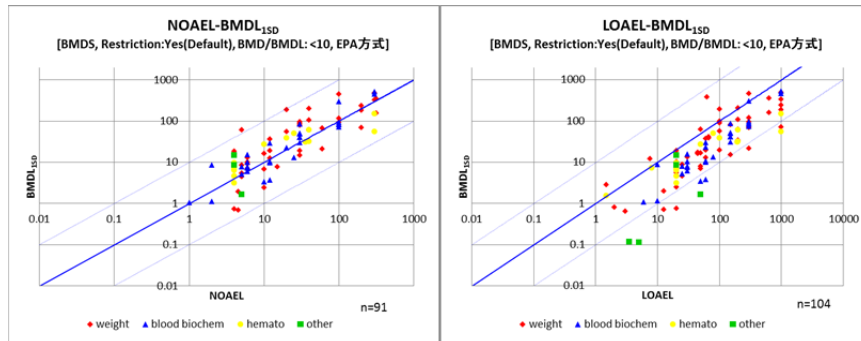


図9. BMDs及びPROASTを用いて算出したBMDL₁₀の比較
(Restriction Yes&No、BMD/BMDL<10、EFSA方式)

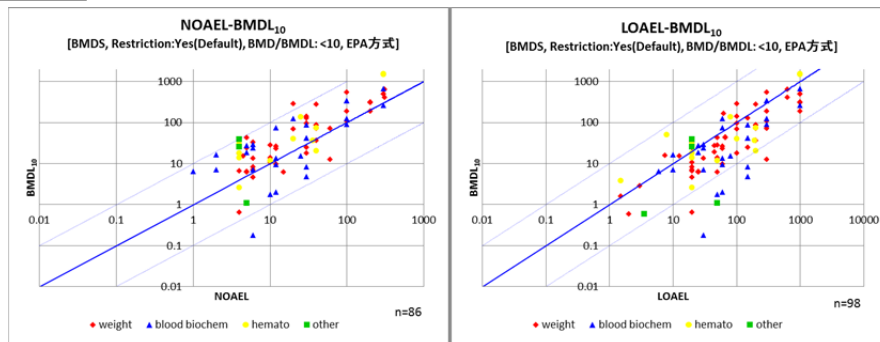
連続データ

BMRに関する検討では、まず、BMDsを用いてBMDL_{1SD}、BMDL₁₀及びBMDL₀₅を算出し [Restriction Yes (default)]、EPA方式でBMD/BMDL<10を満たす最終値を選択した。その結果、対象とした連続値データ151個のうち、それぞれ104個、98個及び102個のデータについてBMDL値を求めることができた。NO(A)ELが求められているデータのうち、BMDL_{1SD}、BMDL₁₀及びBMDL₀₅がNOAELとLOAELの範囲内に収まったものは、それぞれ、47個 (52%)、43個 (50%) 及び27個 (30%)であった。PROASTを用いて、BMDL_{1SD}、BMDL₁₀ 及びBMDL₀₅を算出/選択した結果 (Restriction No, EFSA方式)、それぞれ104個、98 個及び102個のデータについて値を求めることができた。NOAELとLOAELの範囲内に収まったものは、それぞれ41個 (37%)、35個 (32%) 及び39個 (34%)であった。図10と 図11にBMDs及びPROASTを用いて計算した各BMDLとLOAELおよびNOAELとの比較を示す。解析は、連続データの種類 (臓器重量、血清生化学値、血球系測定値、その他) ごとに分類して行ったが、各パラメータに特徴的な傾向はみられず、全体的な分布 としては、NOAEL値とBMDL_{1SD}との相関性が高い傾向があるように感じられた。

A. BMR=1SD



B. BMR=10%



C. BMR=5%

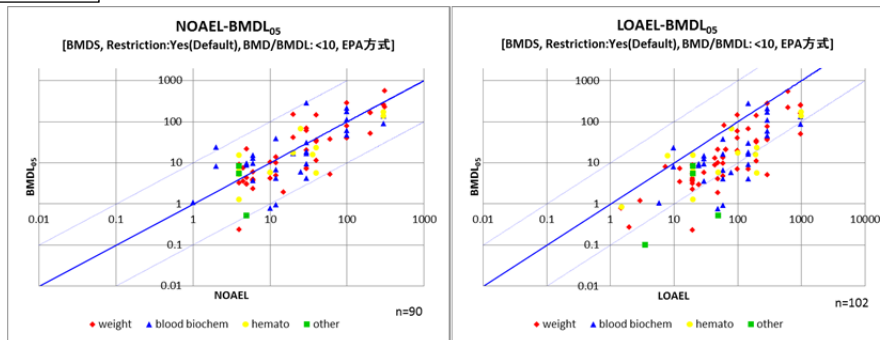
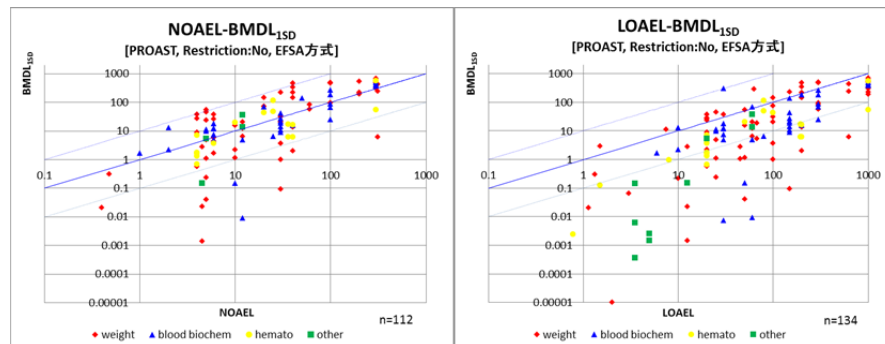
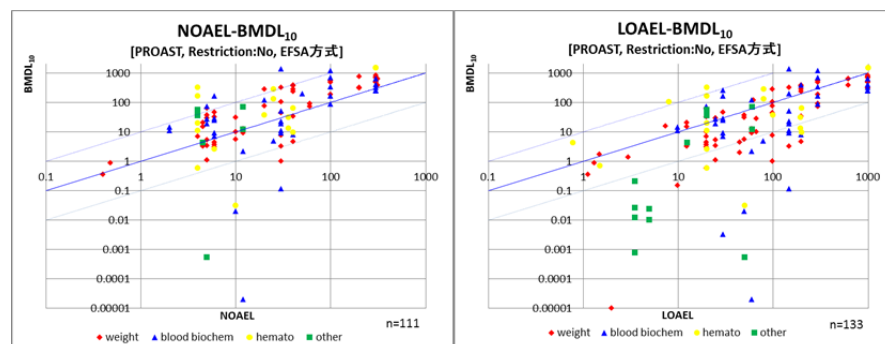


図10. BMDsを用いて算出した各BMDL値とLOAELおよびNOAELの比較

A. BMR=1SD



B. BMR=10%



C. BMR=5%

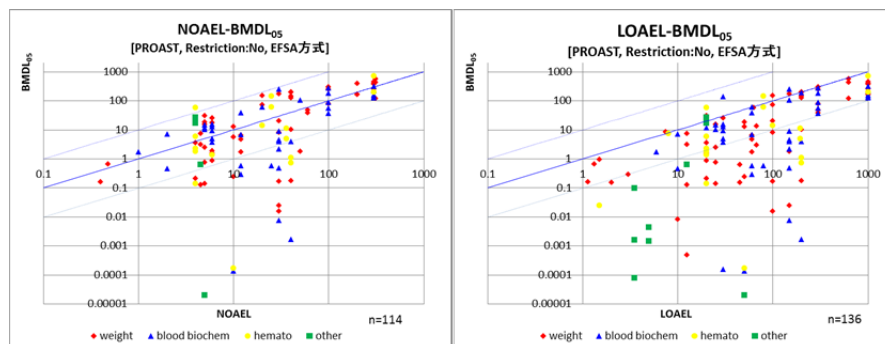


図11. PROASTを用いて算出した各BMDL値とLOAELおよびNOAELの比較

BMDSを用い、モデルへのフィッティングパラメータとして係数のRestrictionの有無について解析を行った。EPA BMDSでは、適用可能な5種のモデルのうち、Exponentialモデル、Hillモデル及びPowerモデルについてはdefaultでrestrictionがonとなっており、このうちHillモデルとPowerモデルについては、restrictionをoffにした状態でもBMDLを算出することができる。この二種のモデルについてrestriction offの状態でもBMDL_{1SD}を算出し、それらの値を含めた場合と含めない場合で、最終的なBMDL_{1SD}値がどのように変化するか、比較解析した。なお、BMDLの選択はEPA方式で行った。

その結果、restriction offで計算した値を含めると、104個中39個（38%）の値が低い値となり、そのうち27個（26%）の $BMDL_{1SD}$ はdefaultで計算した値から選択した $BMDL_{1SD}$ の2分の1未満であった。Restriction Noで計算した値を含めて選択した $BMDL_{1SD}$ とNOAEL及びLOAELを比較した結果を図12に示す。Restriction Yes（default）で計算した値から選択した結果（図10-A）と比較したところ、非連続データと同様に、RestrictionをNoとして算出した $BMDL_{1SD}$ を含めない方が、 $BMDL_{10}$ がNOAEL/LOAELから大きく外れるケースが減り、値が収束していることが明らかになったが、Conservativeな意味では、両条件で計算を行い、最終値を選択することがより適切と考えられた。

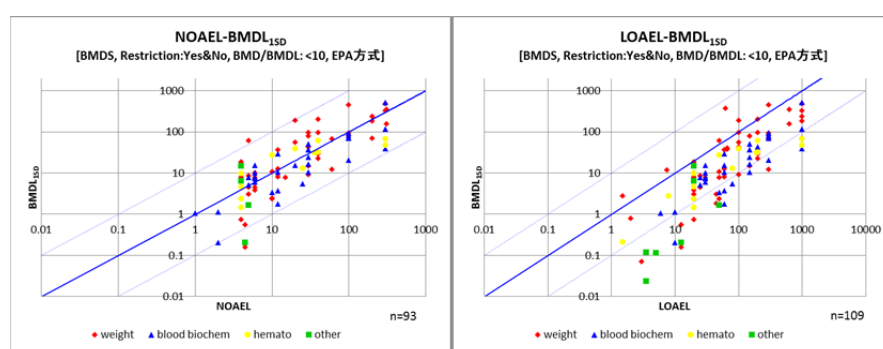


図12. BMDSを用いてRestriction Yes&Noで算出した $BMDL_{1SD}$ とNOAEL及びLOAELの比較

BMDSを用い、Restriction Yes（default）及び $BMD/BMDL < 10$ の条件で $BMDL_{1SD}$ を算出し、最終値の選択方法（EPA方式かEFSA方式か）について検討を行った。非連続データとは異なり、算出された $BMDL$ 値の幅はほとんど（80/104個）が3倍未満であったが、EPA方式とEFSA方式で異なる値が選択されたのは、104個中36個のみであり、値の変化は3倍未満であった。非連続データと同様に、Conservativeな意味では、EFSAガイダンスで推奨されている通り、最低値を選択することがより適切と考えられる。

PROASTを用い、 $BMR=1SD$ 、RestrictionをNoとして、“ $BMD/BMDL$ が10未満”となったモデルからEFSA方式で選択した $BMDL_{1SD}$ とNOAEL及びLOAELを比較した結果を図13に示す。この条件を加えずに計算した結果は図11-Aに示す通りである。新たな条件

“ $BMD/BMDL$ が10未満”を追加することで、 $BMDL_{1SD}$ がNOAEL/LOAELから大きく外れるケースを減らせることができた。しかし、“ $BMD/BMDL$ が10未満”という条件を追加してもいくつかの $BMDL_{1SD}$ については、NOAEL/LOAELから大きく外れた値を示したことから、さらなる条件の追加を検討した。図13において、 $BMDL_{1SD}$ がNOAEL/LOAELから大きく外れたケースでは、 $BMDL_{1SD}$ が最低用量よりもかなり低い値となっていたことから、最低用量/ $BMDL < 100$ という条件を新たに追加した結果、図14に示すように、 $BMDL_{1SD}$ がNOAEL/LOAELから大きく外れるケースはなくなった。

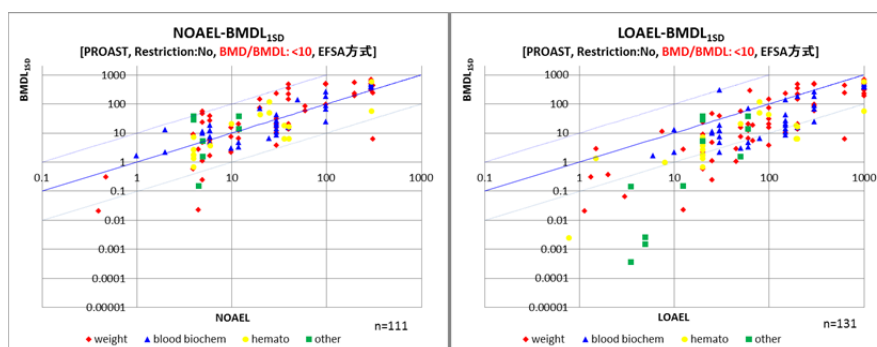


図13. “BMD/BMDLが10未満” となったモデルから選択した
BMDL_{1SD}とNOAEL及びLOAELの比較

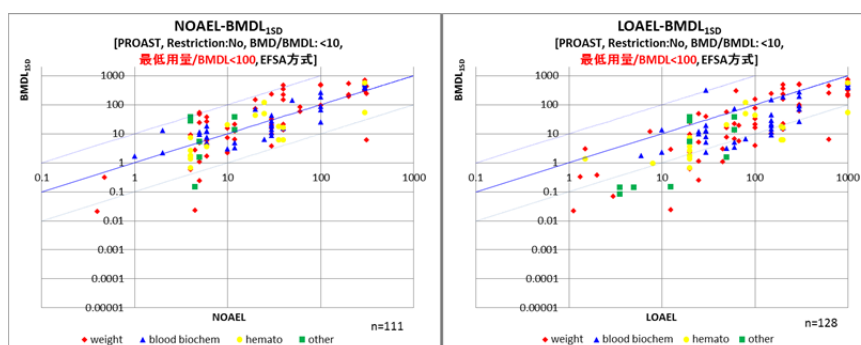


図14. “BMD/BMDLが10未満” 及び “最低用量/BMDL” となったモデルから選択し
たBMDL_{1SD}とNOAEL及びLOAELの比較

BMDs及びPROASTを用い、同じ条件 [Restriction Yes & No, BMD/BMDL<10、最低用量/BMDL<100、EFSA方式] で算出/選択したBMDL_{1SD}を比較したところ、概ね近い値が算出/選択された (図15)。

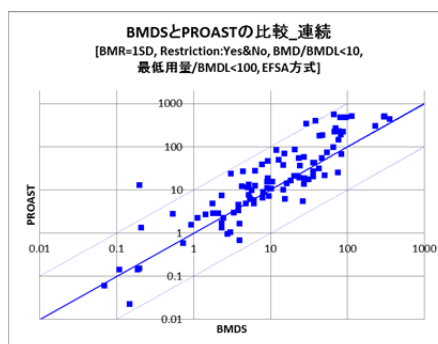


図15. BMDs及びPROASTを用いて算出したBMDL_{1SD}の比較

(2) 研究項目名：2. ベンチマークドース法適用に関する実験的検証に関する研究

1) 個別課題名：オ. 多用量群設定によるベンチマークドースの信頼性検証

(担当：渋谷分担研究者、小野分担研究者)

ベンチマークドース適用性に関する問題点として、グレード化された病理組織所見には適用しにくいことが挙げられる。そこで本個別課題では、ベンチマークドースの病理組織所見への適用と用量群設定が信頼性に与える影響を検討するため、通常的安全性試験よりも多くの用量群を設定した動物実験を実施して、病理組織学的変化に関して、形態計測、あるいは発生頻度により表現される数量値と、病変の程度のグレーディングにより表現されるカテゴリー値に対して、計算に用いる用量群数を変えてベンチマークドースを適用し、ベンチマークドースの信頼性をたどるために必要な用量群数の比較・検証を行うことを目的とした。

Diheptyl phthalate (DHP)について、最高用量を20,000 ppmとして、混餌投与にて公比2の割合で9用量を設定して、1群あたり15匹の雄性F344ラットに90日間の亜急性反復投与を行った。その結果、10,000 ppm以上の高用量群のみで体重増加抑制が見られ、臓器重量では肝臓において10,000 ppm以上の群で相対重量の増加が見られ、精巣の絶対重量は20,000 ppmでわずかに減少した。組織学的には肝臓では5,000 ppm以上で肝細胞の腫大、空胞変性が用量依存的に重篤化したが、精巣ではいずれの群においても特筆すべき変化は認められなかった。免疫染色では1,250 ppmからGST-P陽性単細胞が増加し、200 μ m径を超えるGST-P陽性巣の数および面積は5,000 ppmより増加した。また、 α 2マクログロブリン(α 2M)陽性巣ならびにGST-P陽性巣との一致率及びKi-67陽性細胞率も5,000 ppm以上で用量依存性に増加した。肝臓におけるDNA傷害の指標である8-OHdGレベルも5,000 ppm以上で有意な増加が認められた。Real-time RT-PCRでは、*A2m*の発現が10,000 ppm以上で増加したが、ペルオキシソーム酵素(*Acox*、*Ehhadh*、*Acaa1*、*Hadhb*)発現は投与により変化しなかった。また、10,000 ppm以上で*Gadd45g*の発現低下と*Bcl2*の発現増加もみられた。

以上の結果は以下の様にまとめることができる。

成果1. 体重ならびに肝臓・精巣重量：10,000 ppm以上の高用量群のみで体重増加抑制が見られた。肝臓は10,000 ppm以上の群で相対重量の増加が見られたが、精巣の絶対重量は20,000 ppmでわずかに減少した。

成果2. 肝臓の組織学的変化：5,000 ppm投与群から肝細胞の腫大、空胞化などの変性が認められ、この用量より変性の程度および範囲が用量依存性に増大した。

成果3. 精巣の組織学的変化：高用量群を含むいずれの群においても特筆すべき組織学的変化は認めなかった。

成果 4. 肝臓の免疫染色結果：GST-P陽性単細胞は1,250 ppmから、200 μ m径を超えるGST-P陽性巣の数および面積、 α 2M陽性巣ならびにGST-P陽性巣との一致率及びKi-67陽性細胞率は5,000 ppm 以上で用量依存性に増加した。

成果 5. 肝病変形成への酸化ストレスの関与：8-OHdGは5,000 ppm以上で有意な増加が認められた。

成果 6. 肝臓におけるmRNA発現変化：ペルオキシソーム酵素の発現には変化が見られなかったが、10,000 ppm以上で*Gadd45g*の発現低下と*A2m*及び*Bcl2*の発現増加がみられた。

これらの結果より、DHPによる肝発がん性は高用量域においてのみ見られ、その機序にはPPAR α アゴニスト活性を介した作用が関与しないと推察された。また、精巣に対しては少なくとも20,000 ppmまでは病変を惹起しないことが示された。

精巣や副生殖器官には組織学的に明らかな変化を認められなかったものの、肝臓では低～中用量域から胆管増生や肝細胞の変性変化を認め、肝細胞変化は投与用量に応じた一連の毒性変化と判断された。このような病理組織学的所見をそれぞれスコアリングし、その合算した数値がベンチマークドース法に適用しうるかについて検討を行ったが、高用量でスコアが急激に増加したため、異なる病変のグレードを合算して得られる数値の妥当性は乏しいと判断された。また、組織学的所見および（中間評価以後の研究による）免疫染色結果や転写レベルの変化等のカギとなる所見（あるいはその組み合わせ）を抽出して解析した結果においてもそれらが高用量域に留まっていたため、広い用量域をカバーした反応性が得られず、ベンチマークドース法のバリデーションには向かない反応性であった。

以上のことは、用量設定が少ない動物実験に対するベンチマークドース法の妥当性検証のためには、用量範囲を広くするのではなく、対象とする影響に特化して狭い範囲で用量を多く設定する必要があることが示された。しかし、用量反応性が大きく変化する用量域は、妥当性を評価する対象となる側性影響毎に異なることが想定され、一つの実験では検証されるエンドポイントは限定されるか、もしくは様々なエンドポイントを一度の実験で検証するためには、さらに大規模な用量設定が必要となると考えられ、現実的なアプローチではないことが明らかとなった。それ故、以下の「個別課題名：カ」以降は、複合影響評価へのベンチマークドース法の適用に焦点を絞ることとした。

2) 個別課題名：カ. 基本骨格を同一にする物質の複合暴露影響評価

（担当：渋谷分担研究者、小野分担研究者）

実験的検証に関する研究においては、グループTDIの設定根拠を与える手法として、主要な毒性メカニズムを同一にする複数の物質群の複合暴露影響評価に関して、ベンチマークドース法の適用が有効であると考えられるが、実際的な評価事例で実証はさ

れていない。そこで、同じ化学物質類に属する物質間での相加性（あるいは干渉性）影響の検討を行うことは、複合暴露影響評価の実証性を与える上で重要な要件となる。そこで本課題では、毒性影響の相加性を示す物質群、あるいは干渉性を示す可能性のある物質群の複合暴露評価に関する実証実験として、基本骨格が同じで毒性影響も同じ物質の組み合わせと共に、基本骨格が同じでも毒性影響が異なる可能性のある物質の組み合わせで行い、毒性影響のベンチマークドース値に基づく相加性、あるいは干渉性の有無を明らかにすることを目的とした。

Di (2-ethylhexyl) phthalate (DEHP; 12,000 ppm)、di-*n*-butyl phthalate (DBP; 12,000 ppm)、DHP (10,000 ppm)を単独あるいはDEHPとDBP、DEHPとDHPの組み合わせで1群あたり15匹の6週齢、雄性F344ラットに対して混餌による90日間の亜急性反復投与を行った。その結果、体重増加抑制がDEHP+DBP、DEHP+DHPの両併用投与群で開始1週目から、各単独投与群でも3週目から見られた。肝臓の相対重量の増加がDHP単独投与以外の群で見られ、DEHP+DBP併用投与群では相加作用が見られた。一方、生殖器系では、精巣の絶対重量の減少がDEHP+DBP、DEHP+DHPの両併用投与群においてのみ見られ、前立腺および精嚢の絶対重量の減少はDEHP単独投与以外の群で見られ、DEHP+DBP、DEHP+DHPの両併用投与群において相加的な影響を示す傾向が見られた。血液生化学検査では、AST、ALTの増加がDHP単独投与群のみに見られ、総コレステロールはDHP単独投与以外の群で減少した。トリグリセリドの減少は全ての投与群で見られ、DHP単独投与群で最も顕著に減少していた。この結果はPPAR α 活性の有無によるフタル酸エステルの脂質代謝の違いを示唆した。組織学的変化としては、肝臓では、ペルオキシゾーム増生を反映した好酸性肝細胞肥大がDEHP単独投与群で見られ、DBPとの併用により増強された。一方、DHP単独投与群では肝細胞の腫大・空胞変性が見られたが、DEHPとの併用により消失した。生殖器系においては、各単独群では有意な変化が見られなかったが、精巣の精細胞の脱落を主としたびまん性の精細管萎縮がDEHP+DBP、DEHP+DHPの両併用投与群で見られた。肝臓の免疫染色では、肝前がん病変マーカーであるGST-PはDHP単独投与群においてのみ陽性細胞巣を形成したが、DEHPとの併用により陽性巣は消失した。PPAR α アゴニスト誘発肝前がん病変の指標として報告のある α 2Mは、DEHPとDBP及びそれらの併用群ではびまん性に弱陽性を示す肝細胞が見られ、DHP単独投与群においてのみGST-Pと同様な陽性巣を形成した。これらの陽性巣もDEHPとの併用で消失した。また、肝臓におけるDNA傷害の指標である8-OHdGレベルを検索したところ、全ての投与群で対照群との比較において有意な増加が認められたが、併用投与による共力作用は見られなかった。一方、ミクロソーム分画の活性酸素種(ROS)産生レベルならびに脂質過酸化の指標であるTBARSレベルも検討したが、投与各群で変化が認められなかった。発現アレイではペルオキシゾーム酵素などのPPAR α 関連因子、薬物代謝、抗酸化酵素、アポトーシス、細胞周期、DNA修復およびNOX関連因子においてDHP単独投与群で他の群とは異なる発現動態を示す因子が認められた。この結果に基

づき、Real-time RT-PCRによる遺伝子発現の検証を行なったところ、PPAR α 関連因子 (*Acox*, *Ehhadh*, *Fgf21*, *Me1*) 及び抗酸化酵素のうち *Aldh1a1* と *Aldh1a7* はDHP単独投与群に比べてその他の投与群で有意な発現増加が認められた一方、細胞周期関連因子 (*AurkB*, *Cdc20*, *Cdc2a*, *Ccnb1*)、NOX関連因子 (*Cybb*, *Cyba*, *Ncf1*) 及び抗酸化酵素の *Nqo1* はDHP単独投与群のみで有意な発現増加が認められた。また、細胞増殖マーカーであるPCNAならびにNOX2の免疫染色では、DHP単独投与群のみで有意な陽性細胞比率の増加が見られた。

以上の結果は以下の様にまとめることができる。

成果 1. 体重ならびに肝臓・精巣重量：体重増加抑制はDEHP+DBP, DEHP+DHPの両併用投与群は開始 1 週目から、各単独投与群でも3週目から見られた。肝臓の相対重量増加はDHP単独投与以外の群で見られ、DEHP+DBP併用投与群では相加作用が見られた。精巣の絶対重量はDEHP+DBP, DEHP+DHPの併用投与群においてのみ減少しており、相加作用を示す傾向が見られた。

成果 2. 血液生化学検査：DHP単独投与群のみににおいてAST、ALTが上昇し、コレステロール値はDHP単独投与以外の群で減少した。トリグリセリドの減少は全ての投与群で見られ、DHP単独投与群で最も顕著な減少を示した。

成果 3. 肝臓の組織学的変化：好酸性肝細胞肥大はDEHP単独群で見られ、DBPとの併用により増強された。一方、DHP単独投与群では肝細胞の腫大・空胞変性が見られ、DEHPとの併用により消失した。

成果 4. 精巣の組織学的変化：両併用投与群でのみ精巣の精細胞の脱落を主としたびまん性の精細管萎縮が見られた。

成果 5. 肝臓の免疫染色結果：GST-PはDHP単独投与群においてのみ陽性細胞巣を形成したが、DEHPとの併用により陽性巣は消失した。 α 2MはDEHPとDBP及びそれらの併用群ではびまん性に弱陽性細胞が見られ、GST-P同様にDHP単独投与群においてのみ陽性巣を形成したが、 α 2M陽性巣もDEHPとの併用で消失した。PCNA及びNOX2の免疫染色ではDHP単独投与群のみで有意な陽性細胞比率の増加が見られた。

成果 6. 肝病変形成への酸化ストレスの関与：DNA傷害の指標である8-OHdGは全ての投与群で対照群と比較して発現レベルの増加が見られたが、いずれの併用投与群でもその増加は相加的ではなかった。

成果 7. 肝臓におけるmRNA発現変化：発現アレイではPPAR α 関連因子、薬物代謝、抗酸化酵素、アポトーシス、細胞周期、DNA修復およびNOX関連因子においてDHP単独投与群で他の群とは異なる発現動態を示す因子が認められた。また、Real-time RT-PCRによる遺伝子発現の検証においてもPPAR α 関連因子及びいくつかの抗酸化酵素はDHP単独投与群に比べてその他の投与群で有意な発現増加が認められた一方、細胞周期関連因子、NOX関連因子及び抗酸化

酵素の*Nqo1*はDHP単独投与群のみで有意な発現増加が認められた。

これらの結果より、肝臓の病変形成に際してはDEHPによるPPAR α 活性に対するDBPの増強作用と、DHPのPPAR α 非依存的な発がん作用に対するDEHPによる拮抗メカニズムの存在が推察された。一方、生殖器系では精細管萎縮に関して組み合わせによって相加的、あるいは一方の作用を増強することが推定された。

以上、基本骨格が同じでも作用強度や機序の異なる物質を90日間併用投与した結果、標的臓器により併用による影響、特にDEHPとDHPの併用による影響が異なっていた。しかし、今回は各物質につき1用量のみの設定であったため、肝臓ではDHPによる組織学的変化をDEHPが完全に抑制していた一方で、精巣による変化は併用投与によってのみ見られたために片方の物質の作用を増強したものか相加作用であるかが明確にできなかった。従って、ベンチマークドース法の適用に適う毒性評価を実施できなかったため、臓器ごとに相加性および干渉性を示す組み合わせであるDEHPとDHPについて、それぞれ複数用量を設定した併用投与実験を実施することで、これらの物質の用量反応関係を再検討する必要がある。

3) 個別課題名：キ．基本骨格を同一にするが毒性メカニズムが異なる物質の複合暴露影響評価（担当：渋谷分担研究者、小野分担研究者）

基本骨格を同一にするが毒性メカニズムが異なる物質の複合影響として「個別課題：カ」の検索で用いたDEHPとDHP について、一方の物質の用量を固定した上でもう一方の物質の用量について公比2の割合で4用量を設定し、雄性F344ラットに対して混餌にて共投与による90日間のラット亜急性毒性試験を行い、肝臓の影響や精巣毒性に関する毒性評価を行い、それぞれのエンドポイントでのベンチマークドース値に基づく相加作用あるいは干渉性の有無を検証することを目的とした。

Di (2-ethylhexyl) phthalate (DEHP)のみ (12,000 ppm、24,000 ppm)、Diheptyl phthalate (DHP)のみ (10,000 ppm、20,000 ppm)、DEHP + DHP (12,000 ppm + 10,000 ppm、12,000 ppm + 5,000 ppm、12,000 ppm + 2,500 ppm、12,000 ppm + 1,250 ppm、6,000 ppm + 10,000 ppm、3,000 ppm + 10,000 ppmあるいは1,500 ppm + 10,000 ppm)を1群あたり12匹の6週齢、雄性F344ラットに対して混餌により90日間の亜急性反復投与を行い、DHP誘発肝病変に対する DEHPによる干渉作用と精巣におけるDEHPとDHPによる相加あるいは増強作用の用量反応性の検討を行なった。

その結果、体重増加抑制は無処置対照群に比べ全ての投与群で見られたが、DEHP 12,000 ppm 単独群での体重増加抑制が最も遅く、8週目から観察された。肝臓の相対重量はDEHP単独群で、無処置対照群に比べ用量に応じて増加したが、DHP単独群では増加が見られなかった。併用群ではDEHP 12,000 ppm + DHP 10,000 ppm 併用群で、DEHP 12,000 ppm 単独群よりは有意に低値、DHP 10,000 ppm単独群よりは有意に高値を示し、DEHP 12,000 ppm + DHP 1,250～5,000 ppm 併用群ではいずれもDEHP 12,000 ppm 単独

群と同程度の増加を示したが、DEHP 1,500～6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の低下に伴い、DEHP 12,000 ppm 単独群と比較して重量増加率の減少が見られた。一方、生殖器系では、無処置対照群に比べ精巣及び精巣上体の絶対重量の減少がDEHP 24,000 ppm 単独群とDEHP 12,000 ppm + DHP 10,000 ppm 併用群においてのみ見られ、前立腺および精嚢の絶対重量の減少はDHP 24,000 ppm 単独群のみで見られた。

血液生化学検査では、ALTがDHP 10,000 ppm以上を投与した全ての群で無処置対照群と比べて有意に増加し、DEHP 12,000 ppmとの併用では増加抑制が見られた。しかし、DHP 5,000 ppm以下とDEHP 12,000 ppmとの併用群では変化は見られなかった。トリグリセリドはDEHP単独群、DHP単独群共に用量に応じて、無処置対照群と比べて有意に減少した。DEHPとDHPの併用群においても相加的に減少し、それぞれの併用量の低下に伴ってDHP 10,000 ppm + DEHP 12,000 ppm併用群と比較して有意な減少率の低下が認められた。総コレステロールはDEHP 10,000 ppm以上を投与した群で無処置対照群と比べて有意に減少した。DHP 10,000 ppm + DEHP 12,000 ppm併用群では相加的に減少し、DEHP併用量の低下に伴う減少率の低下は見られたが、DHP併用量の低下に伴う減少率の低下は見られなかった。

組織学的変化としては、肝臓ではDEHP単独群でペルオキシゾーム増生を反映した好酸性肝細胞肥大が、DHP単独群で肝細胞の腫大・空胞変性が見られ、いずれも用量に応じて増強した。併用群ではDEHP 12,000 ppm + DHP 1,250～10,000 ppm 併用群で肝細胞の腫大・空胞変性が消失した一方、DEHP 1,500～6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の減少に伴い、肝細胞の腫大・空胞変性が頻度及び程度ともDEHP 12,000 ppm + DHP 10,000 ppm併用群と比較して有意に増加した。また、好酸性肝細胞肥大については、DEHP 3,000～1,500 ppm + DHP 10,000 ppm併用群においてDEHP 12,000 ppm + DHP 10,000 ppm併用群と比較して頻度及び程度とも有意に減少した。生殖器系においては、精巣の精細胞の脱落を主とした精細管萎縮がDEHP 24,000 ppm 単独群とDEHP 12,000 ppm + DHP 10,000 ppm 併用群でびまん性に見られ、DEHP 12,000 ppm + DHP 1,250～5,000 ppm 併用群ではDEHP 12,000 ppm + DHP 10,000 ppm併用群と比較して頻度及び程度とも有意に減少した。肝臓の免疫染色では、肝前がん病変マーカーであるGST-P陽性巣の数と面積、ならびに細胞増殖指標であるKi-67陽性細胞比率がDHP単独群で用量に応じて増加し、DEHP 12,000 ppm + DHP 1,250～10,000 ppm 併用群ではGST-P陽性巣はほぼ完全に抑制され、Ki-67陽性細胞比率はDHP併用量の減少に伴い用量に応じて減少した。DEHP 1,500～6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の減少に伴い、GST-P陽性巣の数と面積及びKi-67陽性細胞比率の増加が用量に応じて回復した。Real-time RT-PCRによる各種分子の転写量を検討したところ、ペルオキシゾーム酵素 (Acyl-CoA oxidase 1 palmitoyl, Enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase, Acetyl-CoA acyltransferase, hydroxyacyl-CoA

dehydrogenase/3-ketoacyl-CoA thiolase/ enoyl-CoA hydratase) はDEHP単独群及びDEHP 12,000 ppm + DHP1,250~10,000 ppm 併用群で有意な発現増加が見られ、DEHP 1,500~6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の減少に伴い、用量に応じて発現増加抑制が見られた。細胞増殖関連因子(Aurora kinase B, Cdc20, Cdc2a) 及びNOX関連因子(Cybb, Cyba, Ncf1)はDHP単独群及びDEHP 1,500~12,000 ppm + DHP 10,000 ppm 併用群で有意な発現増加が見られた。

以上の結果は以下の様にまとめることができる。

成果1．体重ならびに肝臓・精巣重量：体重増加抑制は全ての投与群で見られた。

肝臓の相対重量はDEHP単独群で用量に応じた増加を示したが、DHP単独群では増加が見られなかった。併用群ではDEHP 12,000 ppm + DHP 1,250~5,000 ppm 併用群ではいずれもDEHP 12,000 ppm 単独群と同程度の増加を示したが、DEHP 1,500~6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の低下に伴い、重量増加率の減少が見られた。精巣の絶対重量の減少はDEHP 24,000 ppm 単独群とDEHP 12,000 ppm + DHP 10,000 ppm 併用群において見られた。

成果2．血液生化学検査：ALTはDHP10,000ppm以上の投与により有意に増加し、DEHP12,000 ppm併用では増加抑制が見られたが、他の併用群では併用量の低下に伴う変化は見られなかった。一方、総コレステロールはDEHP投与により有意に減少し、DHP との併用においてはDEHP投与量の低下に伴う減少率の低下が見られた。トリグリセリドは全ての群で無処置対照群と比べて有意に減少し、併用群においても相加的に減少し、それぞれの併用量の低下に伴って有意な減少率の低下が認められた。

成果3．肝臓の組織学的変化：DEHP単独群で好酸性肝細胞肥大が、DHP単独群で肝細胞の腫大・空胞変性が見られ、いずれも用量に応じて増強した。DEHP 12,000 ppm との併用群ではいずれも肝細胞の腫大・空胞変性が消失した。一方、DEHP 1,500~6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の減少に伴い、消失していた肝細胞の腫大・空胞変性が用量に応じて回復した。

成果4．精巣の組織学的変化：精巣の精細胞の脱落を主とした精細管萎縮がDEHP 24,000 ppm 単独群とDEHP 12,000 ppm + DHP 10,000 ppm 併用群でびまん性に見られ、DEHP 12,000 ppm + DHP 1,250~5,000 ppm 併用群では精巣辺縁の一部でごく軽度の精細管萎縮のわずかな用量に応じた増加が見られた。

成果5．肝臓の免疫染色結果：GST-P陽性巣の数と面積ならびにKi-67陽性細胞比率がDHP単独群で用量に応じて増加し、DEHP 12,000 ppm + DHP 1,250~10,000 ppm 併用群ではGST-P陽性巣はほぼ完全に抑制され、Ki-67陽性細胞比率はDHP併用量の減少に伴い用量に応じて減少した。DEHP 1,500~6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の減少に伴い、GST-P陽性巣の

数と面積及びKi-67陽性細胞比率の増加が用量に応じて回復した。

成果 6. 肝臓におけるmRNA発現変化：ペルオキシゾーム酵素はDEHP単独群及びDEHP 12,000 ppm + DHP1,250～10,000 ppm 併用群で有意な発現増加が見られ、DEHP 1,500～6,000 ppm + DHP 10,000 ppm 併用群ではDEHP併用量の減少に伴い、用量に応じた発現増加抑制が見られた。細胞増殖関連因子及びNOX関連因子はDHP単独群及びDEHP 1,500～12,000 ppm + DHP 10,000 ppm 併用群で有意な発現増加が見られた。

これらの結果により、肝臓ではPPAR α 非依存的なDHP誘発性肝前がん病変形成へのDEHPによる干渉作用は低用量域から用量反応性を示す一方で、DHPによるDEHPへの干渉作用はDEHP併用量が低用量の場合のみ認められ、精巣における精細管萎縮は高用量域においてのみDEHPへのDHPの増強作用が認められることが推測された。

以上の結果を基に、複合影響の評価にベンチマークドース法の適用可能性について検証した。複合影響が比較的はっきりとした影響として検出でき、且つBMDの算出が可能であった以下のエンドポイントについて、化合物単剤での用量反応と複合曝露による用量反応についての、それぞれのBMDを算出して比較した。

表2. DEHP投与による影響指標のBMDがDHP投与により受ける影響

エンドポイント	DEHP (0→24,000ppm) 単独投与 BMD (mol/kg/day)	DHP (10,000ppm) +DEHP (0→ 12,000ppm) BMD (mol/kg/day)	
絶対精巣重量減少	2.032	0.830*	DHPの投与がDEHPの作用を増強
精巣萎縮	1.146	0.545	DHPの投与がDEHPの作用を増強
肝細胞肥大	0.096	0.110	DHPの投与はDEHPの作用に影響しない
相対肝重量増加	0.170	0.203	DHPの投与はDEHPの作用に影響しない
血中総Cho減少	0.046*	0.156	DHPの投与がDEHPの作用を減弱
血中TG減少	0.108*	0.357	DHPの投与がDEHPの用量相関性を減弱
血中ALP増加	0.058	0.357	DHPの投与がDEHPの作用を減弱

*: Varianceが適合しない (test3: $p < 0.1$)

表3. DHP投与による影響指標のBMDがDEHP投与により受ける影響

エンドポイント	DHP (0→20,000ppm) 単独投与 BMD (mol/kg/day)	DEHP (12,000ppm) +DHP (0→10,000ppm) BMD (mol/kg/day)	
絶対精巣重量増加	2.907*	0.591*	DEHPの投与がDHPの作用を増強
血中総Cho減少	2.043	1.379	DEHPの投与がDHPの作用を僅かに増強
血中TG減少	0.060	0.328*, **	DEHPの投与がDHPの作用を減弱
Ki-67陽性細胞比率の増加	0.018	0.108*	DEHPの投与がDHPの作用を減弱

*: Varianceが適合しない (test3: $p < 0.1$), **: 適合しない (Test4: $P < 0.1$)

上記の結果は、定性的には片方の化合物の単独投与に比べて他方の化合物を複合投与した場合のBMD値がほとんど変わらない場合は、ほぼ相加的に影響が現れていることを示し、BMD値が減少（作用が増強）している場合は相乗的な作用、BMD値が増加（作用が減弱）している場合は、拮抗作用が認められる可能性を示唆している。しかし、複合影響の強さを定量的に評価する場合には、相加的に作用した場合においてのみBMD値を基準として使用することが可能となるはずで、相乗及び拮抗作用を定量解析するためには別の指標の開発が必要であると考えられる。しかし、DEHPとDHPの複合曝露系は、一つの実験系で、相乗、相加および拮抗作用のすべてを解析することが今回の研究で明らかとなり、将来的な指標開発共に複合曝露影響を定量的に解析する系としては、有用な実験系であること示された。

(3) 研究項目名：3. ベンチマークドース適用のガイダンス案作成と透明性確保に関する研究

1) 個別課題名：ク. ベンチマークドース適用のためのガイダンス案の作成

(担当：広瀬主任研究者、渋谷分担研究者、小野分担研究者、平田分担研究者)

研究項目1の研究結果を基に、動物実験及び疫学データについて、ベンチマークドースを算出する際の基本的なガイダンスを以下のように作成した。

原則的にベンチマークドースを計算するソフトウェアは、動物実験に基づくデータに関しては米国EPAにより開発されたBenchmark Dose Software (BMDs)またはオランダの国立保健医療環境研究所 (RIVM)により開発されたPROASTを使用し、疫学データに関してはSPBS統計パッケージを用いたハイブリッド法を使用する。

以下に各ソフトウェアを用いた場合の適用ガイダンスを示す。

<BMDS>

BMDSは、米国EPAのホームページ (<http://www.epa.gov/ncea/bmds/>)より、無償でダウンロードすることができる。

最初に、データ入力フォームにデータを入力する。病理所見の発現頻度等の非連続データの場合は、モデルタイプとしてDichotomousを選択し、体重、血液/血液生化学検査値や器官重量等の連続データの場合は、モデルタイプとしてContinuousを選択する。非連続データ及び連続データそれぞれに対する適用方法を以下に示す。

非連続データ

1. Benchmark dose response (BMR)の設定

10%の過剰リスクを選択することが推奨される。

2. 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

Gamma、LogLogistic、LogProbit、Multistage及びWeibullモデルについては、フィッティングパラメータとして係数のRestrictionの有無を選択できる。Defaultでは、restrictionがonとなっているが、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思考することを推奨する。

3. 適合モデルの選択

選択できるすべてのモデルのフィッティングを試みる（Multistageモデルについては、デフォルトであるPoly=2を使用するが、用量群が多い場合は3以上も検討する）。適合モデルの判定基準として $P > 0.1$ となるモデルをフィッティングが適合したモデルとする。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。

4. BMDLの選定

原則的には、適合し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。しかし、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。

連続データ

1. BMRの設定

デフォルトとして対照群の1SD（1標準偏差）の変化分に対応するリスクを選択することが推奨される（例数が少ない場合等で、対照群の分散が極端に小さくなる場合には、10%の変化分に対応するリスクを選択することも検討する）。

2. 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

HillモデルとPowerモデルについて、Restrictionの有無を選択できる。非連続データと同様に、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思考することを推奨する。

3. 適合モデルの選択

選択できるすべてのモデルのフィッティングを試みる。下記a, c, dを満たすモデルを適合モデルと判断する。

- a. Test 1 ; 用量間で反応/分散に差がある $\Rightarrow <0.05$
- b. Test 2 ; 分散均一モデルが適切である $\Rightarrow \geq 0.1$
※不適切だった場合、分散非均一モデルを用い再計算する。
- c. Test 3 ; モデル化された分散が適切である $\Rightarrow \geq 0.1$
- d. Test 4, 5a, 6a, 7a ; モデル平均がデータにフィットしている $\Rightarrow \geq 0.1$

なお、適合したモデルとFullモデルでパラメータ数が等しいため尤度比検定ができない場合があるが、その場合は、目視で判断する。Exponential familyのModel2～5については、下記の尤度比検定の結果から最適モデルを選択する。

- a. Test 5b<0.05 $\Rightarrow$ Model 2より3が適切。
- b. Test 6b<0.05 $\Rightarrow$ Model 3より4が適切。
- c. Test 7b<0.05 $\Rightarrow$ Model 3より5が適切。
- d. Test 7c<0.05 $\Rightarrow$ Model 4より5が適切。

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。

4. BMDLの選定

原則的には、適合し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。しかし、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。

<PROAST>

PROASTは、RIVMのホームページ(<http://www.rivm.nl/en/Library/Scientific/Models/PROAST>)より無償でダウンロードすることができる。S言語で書かれているため、使用にあたっては、統計ソフトS plusもしくはRをインストールする必要がある。

最初に、マニュアルに従って、データファイルを作成する。病理所見の発現頻度等の非連続データの場合は、モデルタイプとしてQuantalを選択し、体重、血液/血液生化学検査値や器官重量等の連続データの場合は、モデルタイプとしてContinuousを選択する。非連続データ及び連続データそれぞれに対する適用方法を以下に示す。

非連続データ

1. BMRの設定

10%の過剰リスクを選択することが推奨される。

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

フィッティングパラメータとして係数のConstraintの有無を選択できる。選択したモデルの生物学的意義付けによるConstraintの有無の選択に科学的理由が無い限り、ConstraintがYesとNoの両方でのフィッティングを思考することを推奨する。

3. 適合モデルの選択

選択できるすべてのモデルのフィッティングを試みる。適合モデルの判定結果がyesとなったモデルのみ採用する。

4. BMDLの選定

原則的には、適合し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。しかし、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。

連続データ

1. BMRの設定

デフォルトとして対照群の1SD（1標準偏差）の変化分に対応するリスクを選択することが推奨される（例数が少ない場合等で、対照群の分散が極端に小さくなる場合には、10%の変化分に対応するリスクを選択することも検討する）。

2. 適合モデル及びBMDLの選定

選択できるすべてのモデルのフィッティングを試みる。原則的には、結果が表示された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。しかし、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。

<SPBS統計パッケージ>

BMDの計算は、秋田大学大学院医学系研究科環境保健学講座の村田勝敬先生により開発され、公開されているSPBS統計パッケージ（[http:// www.med.akita-u.ac.jp/~eisei/link.html](http://www.med.akita-u.ac.jp/~eisei/link.html)）を使ってベンチマークドースの算出を簡便に行うことが可能となっている。データの入力等は、プログラムと共にダウンロードされるヘルプファイルに使用法が記載されている。

非曝露集団の異常率の設定には $P_0=0.05$ と $P_0=0.10$ のいずれかを選択する。また、ベキ関数によるフィッティングを行うには、「曝露量の $\log(X+1)$ 変換値」を選択する必要がある。BMRの値は0.02、0.05、0.10の3種類の値が自動的に算出される。このモデルでは、曝露指標に直接関係しなくとも影響指標に関係すれば、その影響を共変量（交絡因子）として調整してBMDを算出する。さらに、当該 P_0 の非曝露群における影響指標のcutoff値も算出される（この値が臨床上の基準値に近いかな否かでモデルあるいは対象集団の妥当性を判断する必要がある）。なお、影響指標が正規分布しないで対数正規分布する時には、影響指標を対数（底10）変換してから解析する。

さらに得られた結果に対しては、以下の観点を考慮して使用の判断を行い必要に応じて再試行することが必要である。

■ モデルがフィットしているかどうかは目視にもよるが、モデルの尤度比（ $-2 \log L$ ）検定の結果（P値）が有意でない場合は原則的にモデルを使用すべきでない。

■ 共変量については、集めたデータに共変量を含んでいない場合には計算に加えることができない。他の変数がある場合、過去の論文を確認し、共変量として調整した方が良いならば使用する。また未確認の変数の場合、一度加えてみて、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し、有意でなければ削除する（この判断は、試行錯誤的に行う）。一方、統計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。

■ 曝露指標は $\log$ 変換しない。一方、影響指標は最初無変換で行い目視にて対数変換するか否か検討する。 $\log(X+1)$ 変換モデルは経験でフィットする例はあまりなく、一次回帰モデルかベキ関数モデルを優先的に用いる。いずれにしても目視で図を見ながら、フィットするモデルを探索的に試行する必要がある。

2) 個別課題名：ケ．ベンチマークドース計算結果の公表

（広瀬主任研究者、小野分担研究者）

上記ガイダンス案に従って算出/選択したベンチマークドース値を国立医薬品食品衛生研究所の総合評価研究室のホームページからリンクされる公開Webページ

（<http://dra4.nihs.go.jp/bmd>）を作成した。Webページには、下図に示すようにベンチマークドース算出結果データベースに加えて、以下の項目に関する情報提供も行うように設計した。

- ・ベンチマークドース（BMD）法について
- ・用量反応評価手法/BMD法に関する国際的ガイダンス文書（和訳）
- ・BMD法の適用に関するガイダンス（提案）

ベンチマークドース法の適用に関するガイダンス

HOME

ベンチマークドース法の適用に関するガイダンス

本サイトは、化学物質の健康影響リスク評価における用量反応性評価手法として、耐容一日摂取量（TDI）や曝露マージン（MOE）を算出する際に適用されるベンチマークドース（BMDL）を算出する際のガイダンスを提供することを目的としている。

このページは、食品安全委員会食品健康影響評価技術研究における「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」の研究成果の一部として公開するものである。

- [ベンチマークドース（BMD）法について](#)
- 用量反応評価手法/BMD法に関する国際的ガイダンス文書（和訳）
 - [International Programme on Chemical Safety \(IPCS\)](#)（[原文](#)）
 - [European Food Safety Authority \(EFSA\)](#)（[原文](#)）
 - [U.S. Environmental Protection Agency \(EPA\)](#)（[原文](#)）
- BMD法の適用に関するガイダンス（提案）
 - [BMDs](#)（[米国EPA](#)）
 - [PROAST](#)（[オランダRIVM](#)）
 - [SPBS](#)（[秋田大学](#)）
- [BMDLシミュレーション結果データベース](#)
- 研究報告書（食品安全委員会の公開ページへのLink）

個人情報保護方針

Copyright © 2013 National Institute of Health Sciences, Division of Risk Assessment. All Right reserved.

（４）研究全体の成果、考察及び結論

化学物質のリスクアセスメントにおける用量反応性評価において、重要な手法の一つであるベンチマークドース手法の適用について、IPCSやEFSA、EPA等により公表されている資料を参考に、より具体的な適用ガイダンスを作成することを目的として研究を進めた。IPCSのガイダンスでは用量反応性評価全体の概念について解説が示されているが、EFSAとEPAのガイダンスでは、より具体的な適用手法の解説が述べられている。調査の結果、EFSAとEPAのガイダンスでは、主に、BMRの選択、モデル適合度の評価、モデル及びBMDLの最終選定方法の3点について、相違が存在していることが明らかとなった。そこで、WHOやEPA（IRIS）あるいは化審法の既存化学物質評価等に於いて実際の安全性評価に使用されたNO(A)EL値

等のデータを検証用データセットとし、EPAやRIVMで開発されたベンチマークドースソフトであるBMDsとPROASTを用いて、様々な条件でBMDを算出することで、BMRの選択、モデル選択及びBMDLの最終選定方法の妥当性の検証を行った。その結果、BMD算出のためのBMRとしては、非連続データに対しては10%、連続データについては1 SDが最初の選択として適切であると考えられた。また、BMDLの選択基準として、適合したモデルの中から、BMD/BMDLの比が10未満、最低用量/BMDLの比が100未満であるモデルを選択し、それらのモデルから算出された最も低いBMDL値を採用することが妥当であると考えられた。

一方、疫学データにおいてもベンチマークドース手法の適用は有用であると考えられており、IPSCやEFSA等では、小数例ながら鉛や水銀曝露による疫学データを用いた用量反応性評価が行われている。EFSA法、Hybrid法、クラス分け法など大きく分けて3種類の手法が適用されていたが、BMRの設定や交絡因子の調節に関してHybrid法（SPBS統計パッケージ）が第一選択として適切であると考えられた。以上の解析結果を基に、BMDs, PROAST, SPBSを用いた場合のベンチマークドース適用ガイダンスを作成し、IPSCやEFSA、EPAのガイダンスの和訳及び解析に使用した用量反応データに関するBMD/BMDL計算結果と共に公開Webページを作成した。

ベンチマークドース手法に関しては、さらに、データの型に応じたlog変換などのデータ調整や入れ子になった場合のデータの扱い、順位データを扱うための手法も開発されているが、今回の解析では対応していない。しかし、それらの対応には、より統計学的専門性を考慮に入れたケースバイケースの対応が必要であり、一般的な適用基準を作成するのは困難であると思われる。今回のガイダンスは、実際に評価に使用され、公開された実験データをそのまま使用して妥当性検証を行うことにより作成したものであることから、少なくとも汎用性の高いデータに適用できるガイダンスの作成には成功したと考えている。

また、ベンチマークドース手法の多用量実験や複合影響への適用性を検証するために、9用量群を用いた実験や、相乗・相加作用等を示す化合物の複合曝露実験を行った。その結果、多用量実験については対象とする影響に特化して狭い範囲で用量を多く設定する必要があることが示された。しかし、反应用量域が異なる複数のエンドポイントを解析するには用量範囲を広くする必要があり、単一の実験で解析するのは困難であったことから、複合影響への適用性に焦点を当てた解析を行った。相互作用があることが示されたDEHPとDHPの複合曝露実験を行い、各々の影響についてBMD解析を行った。その結果、相乗、相加および拮抗作用の判定についてBMDを指標としてある程度定性的に解析可能であることが示されたが、より定量的な解析手法としてのBMDの有用性は、現在のところ単純な相加作用を示した場合のみで、更なる指標の開発が必要であると考えられた。しかし、DEHPとDHPの複合曝露系は、一つの実験系で、相乗、相加および拮抗作用のすべてを解析することが可能であり、将来的な指標開発と共に複合曝露影響を定量的に解析する系としては、有用な実験系であることが示された。

結論として、リスクアセスメントにおける用量反応性評価において、一般的な動物実験

および疫学研究で得られる用量反応データに対して、ベンチマークドース計算ソフトウェア（BMDs、PROAST、SPBS）を適用してベンチマークドースを算出する際のガイダンス案の作成に成功した。

Ⅲ 本研究を基に発表した論文等

1 本研究を基に発表した論文と掲載された雑誌名のリスト

広瀬明彦、加藤日奈、平田睦子、小野 敦（2011）JECFAにおける汚染化学物質のBMD手法をもちいたMOE評価について、イルシー（ILSI）No. 106：36-45.

村田勝敬、荻田香苗、堀口兵剛、岩田豊人、広瀬明彦（2011）ベンチマークドース法の臨床的基準をもつ健康影響指標への適用、産業衛生学雑誌 53(3)，67-77.

2 本研究を基にした学会発表の実績

広瀬明彦（2010）ベンチマークドーズ法によるヒ素、水銀、アクリルアミド及びフランのリスク評価、第8回食品安全フォーラム、2010年11月、東京

鈴木和彦、谷合枝里子、嶋本敬介、小野 敦、林 仁美、Wang Liyun、大石 巧、三森国敏、渋谷 淳：フタル酸ジヘプチルのラット90日間混餌投与による肝前がん病変形成機序、第38回日本トキシコロジー学会学術集会、横浜、第38回日本トキシコロジー学会学術集会講演要旨集：p. 127（P-9）、7月11～13日、2011

鈴木和彦、谷合枝里子、嶋本敬介、林 仁美、三森国敏、渋谷 淳：フタル酸ジヘプチル反復投与によるPPAR α 非依存的ラット肝前がん病変形成機序、第70回日本癌学会、名古屋、第70回日本癌学会学術総会演題要旨集：p. 474（P-3077）、10月3-5日、2011

鈴木和彦、谷合枝里子、小野 敦、八舟宏典、大石 巧、剣持 明、三森国敏、渋谷 淳：異なるフタル酸エステル類のラットへの90日間併用混餌投与による肝臓及び雄性生殖器へのインパクト、第28回日本毒性病理学会学術集会、東京、第28回日本毒性病理学会学術集会講演要旨集：P-022（p. 78）、2月2日～3日、2012

鈴木和彦、谷合枝里子、小野 敦、石井雄二、盛田怜子、八舟宏典、三森国敏、渋谷 淳：異なるフタル酸エステル類の90日間併用混餌投与によるラットの肝臓と雄性生殖器への影響の相互作用、第39回日本毒性学会、仙台、第39回日本毒性学会学術年会プログラム：p. S239（P-0166）、7月17日-19日、2012

鈴木和彦、谷合枝里子、盛田怜子、八舟宏典、石井雄二、小野 敦、三森国敏、渋谷 淳：異なるフタル酸エステル類の併用投与によるラット肝臓及び雄性生殖器への影響、第154回日本獣医学会、岩手、第154回日本獣医学会学術集会演題要旨集：p. 196（B-45）、9月14-16日、2012

鈴木和彦、谷合枝里子、盛田怜子、八舟宏典、三森国敏、渋谷 淳：異なるフタル酸エステル類のラットへの90日間併用混餌投与による肝臓と雄性生殖器への影響の

用量反応性. 第29回日本毒性病理学会学術集会, 筑波, 第29回日本毒性病理学会学術集会講演要旨集: p. 75 (P-26)、1月31日-2月1日, 2013

3 特許及び特許出願の数と概要

該当なし

4 その他 (各種受賞、プレスリリース、開発ソフト・データベースの構築等)

「European Food Safety Authority (EFSA): Use of the benchmark dose approach in risk assessment. The EFSA Journal 1150:1-46, 2009」の日本語訳“リスク評価におけるベンチマークドーズ法の利用”をWeb公開
(<http://www.med.akita-u.ac.jp/~eisei/EFSA2009BMD.pdf>)

IV 主任研究者による研究全体の自己評価

項目	評価結果	評価コメント
1 研究の妥当性	4	研究体制としては、専門性に従って研究課題が割り当てられていた。研究の計画、方法については、情報収集では複数の情報源を活用し、様々計算条件で検証を行い、妥当性の高い手法であった。 研究の実施機関における遂行可能性について、研究項目1と3では、翻訳等の外部委託、計算のための人的リソース、研究協力者との協力により達成可能である。研究項目2も多用量の複合曝露実験を行うのに十分な研究機関であった。 費用対効果については、23、24年度と研究費の総額が減少したことにより、実験的規模の縮小を余儀なくされたが、最小限の目的を得るための研究体制は維持ができた。データ解析や翻訳作業等でも効果的な資金配分と費用対効果の高い運営が可能であった。
2 研究目標の達成度	4	研究項目1では、予定通りの情報収集に加えて、研究期間内に公表されたEPAのガイダンスの翻訳にも対応することができた。BMDL算出のモデル選定基準に基づく解析や疫学データの調査から、有用な基準やプログラムの選定が可能となった。 研究項目2では、多用量の実験条件は、BMD解析に向かないことが明らかとなり、複合曝露影響の解析を目標として実験を行えた。予算縮小のため十分な実験はできなかったが、最終年度には、最小限の解析を行うためのデータを得ることができた他、複合曝露影響の解析に有用な系を見つけることができた。 研究項目3では、予定通りガイダンス案を作成し、計算結果や国際機関のガイダンス和訳などを公開することができた。
3 研究成果の有用性	5	食品健康影響評価への貢献等の可能性について、近年ベンチマークドースを使用した評価事例が増加しつつあり、食品安全委員会でもヒ素や遺伝毒性不純物質等の評価において、その適用例が増えてきている。様々な評価の場面でBMDの算定手法のガイダンスとして本研究成果の貢献の可能性は高いと考えられる。
合 計	13	
総合コメント 全体的にみて、研究項目の2において実験的検証の方向性を変更せざるを得なくなったが、最終年度にはBMD解析に適合するデータを得ると共に、複合曝露影響の解析に有用な系を見つけることができた。それ以外は、研究の妥当性および研究成果の達成度、有用性の観点において、当所の研究計画と目標を達成することができた。		