

内分泌攪乱物質スクリーニングプログラム

エビデンスの重み付け：プログラム第1期の結果の評価
第2期試験の必要性を確認するためのスクリーニング



米国環境保護局 化学品安全および汚染防止部
Washington DC 20460

2011年9月14日

まえがき

EPAは2010年11月4日付 Federal Register Notice (75 FR 67963)に掲載のとおり、エビデンスの重み付け(Weight of Evidence, WoE)資料をパブリックレビューおよびコメントのため公開した。寄せられたコメントは集積し、共通点により分類して WoE 資料の改訂の際に容易かつ十分に参照できるようにした。内分泌攪乱物質スクリーニングプログラム (EDSP)において2009年10月29日に開始された第1期スクリーニングのための要請に対して提出されたデータを検討する EPAの担当者および管理者にとっては、ここに示す改訂 WoE アプローチが一般的な指針となるものと考えられる。またデータ提供者にとっても、第1期スクリーニングの結果の評価方法を知ることは興味のあるところであろう。本稿は一般的な指針を与えるものであって、EPAをも部外者をも拘束するものではなく、「である」「であろう」「し得る」「すべき」(will, is, may, can, should)などの表現も EPA・部外者いずれに対しても何らかの要求を示すものではない。したがって EPA は状況によって予告なく指針から逸脱することもあり得る。

WoE分析を用いることは総合的・解釈的プロセスであり、EPAは健康(USEPA 1991; 1996; 2002a; 2005)および生態学(USEPA, 1998)に関連する毒性の評価において、関連する科学的・技術的情報すべてを評価するために日常的に使用している。既存の EPA 文書に含まれる種々のエビデンスを重み付けし総合する際の原理や基準は、EDSP 第1期の試験データの評価に当たっても一般的に使用できるものと考えられる。

EPAは1998年にEDSPの開発・実施の作業を開始したが、それ以後の計算技術および分子関連技術の著しい進歩によって、毒性作用経路の評価に必要なマーカーの同定がますます迅速に行えるようになってきている。米国学術研究会議は2007年の報告書「21世紀の毒性試験ービジョンと戦略」(Toxicity Testing in the 21st Century: A Vision and a Strategy) (NRC, 2007)でこの進歩を認め、動物のみによる毒性試験の効率を高め、またそれへの依存度を下げ究極的には別法に置き換えることを目標に、EPAに対して最新の *in silico* モデルおよび分子技術による高効率スクリーニング試験を利用する戦略を開発することを勧告した。

現在 EPAの内外で、NRCの提案するこれらの最新の手法を利用するためのプロトタイプとして、内分泌攪乱物質のスクリーニングを利用することが試みられている。この作業の主目的はEDSPの迅速さ、信頼性、経済性、機構的特異性を改善することである。この進行中の研究を考慮して、EPAのリスク評価指針は「生きている文書」すなわち科学技術の進歩を反映して定期的に更新・改訂される文書の代表的な例であることを強調しておきたい。本稿で述べる WoE アプローチの利用に関する一般原則と基準はどのような種類の研究にも適用可能であるが、この方針も科学技術の新知見を取り込むため定期的に更新される可能性がある。

目次

略語.....	4
1 目的および適用範囲.....	5
2 内分泌攪乱物質スクリーニングプログラム(EDSP)の概要.....	5
2.1 EDSP 第1期スクリーニング試験.....	6
2.1.1 エストロゲン経路に対する物質の影響を検出する試験.....	8
2.1.2 アンドロゲン経路に対する物質の影響を検出する試験.....	9
2.1.3 ステロイド産生経路に対する物質の影響を検出する試験.....	11
2.1.4 HPG 軸に対する物質の作用を検出する試験.....	11
2.1.5 HPT 軸への物質の影響を検出する試験.....	12
2.2 EDSP 第2期試験.....	12
3 科学技術情報の出典.....	13
3.1 試験指針：第1期スクリーニング試験.....	13
3.2 学術的に有意義な情報.....	14
3.2.1 試験指針－健康および生態影響の研究.....	14
3.2.2 公刊または公表された査読済み論文.....	17
4 科学技術情報の品質.....	17
4.1 一般的評価因子 (GAF).....	18
4.1.1 健全性.....	19
4.1.2 適用可能性および有用性.....	19
4.1.3 明瞭性および完全性.....	19
4.1.4 不確実性および変動.....	19
4.1.5 評価およびレビュー.....	21
4.2 標準評価手順(SEP)およびデータ評価記録(DER).....	21
5 エビデンスの重み付けアプローチ.....	21
5.1 個別試験データの収集と評価.....	23
5.2 異種のエビデンスの総合.....	26
5.3 エビデンスの重み付けの叙述/特徴づけ.....	30
5.4 EDSP 第2期試験への勧告.....	31
6 要約.....	32
7 文献.....	34

略語

略語	原語・意味
A	Androgen アンドロゲン (ホルモン経路)
AR	Androgen Receptor アンドロゲン受容体
DER	Data Evaluation Record データ評価記録
EDSTAC	Endocrine Disruptor Screening and Testing Advisory Committee 内分泌攪乱物質スクリーニングおよび試験諮問委員会
EDSP	Endocrine Disruptor Screening Program 内分泌攪乱物質スクリーニングプログラム
E	Estrogen エストロゲン (ホルモン経路)
ER	Estrogen Receptor エストロゲン受容体
FIFRA	Federal Insecticide, Fungicide, Rodenticide Act 殺虫剤・殺菌剤・殺鼠剤法
FFDCA	Federal Food, Drug, and Cosmetic Act 食品・医薬・化粧品法
FQPA	Food Quality Protection Act 食品品質保護法
GAF	General Assessment Factors 一般的評価因子
GLP	Good Laboratory Practices 医薬品安全性試験実施基準
HPG	Hypothalamic-Pituitary-Gonadal Axis 視床下部-下垂体-性腺軸
HPT	Hypothalamic-Pituitary-Thyroidal Axis 視床下部-下垂体-甲状腺軸
MoA	Mode of Action 作用機序
OCSPP	Office of Chemical Safety Pollution and Prevention 化学物質安全・汚染防止局
OECD	Organization for Economic Co-Operation and Development 経済協力開発機構
PND	Postnatal Day 生後日数
SAB	Science Advisory Board 科学諮問委員会
SAP	Scientific Advisory Panel 科学諮問パネル
SEP	Standard Evaluation Procedure 標準評価手順
T	Thyroid (hormonal pathway) 甲状腺 (ホルモン経路)
WoE	Weight-of-Evidence エビデンスの重み付け

1 目的および適用範囲

この指針資料は、EDSP 第 1 期スクリーニングの評価と解釈のためにエビデンスの重み付け (WoE) アプローチを適用するための基本原則と基準を提供するものであり、第 1 期試験の結果と共に、第 2 期試験の対象候補物質を決定するための情報を対象とする。また試験に関連する情報、第 2 期試験に必要と考えられる情報を坑よして一般的な指針をも含めた。

EDSP 第 1 期のスクリーニング試験の目的は、エストロゲン(E)・アンドロゲン(A)・甲状腺(T)の各ホルモン経路と相互作用する可能性のある化学物質を同定することである。現在この試験には 11 項目が含まれ、それらは EPA のプログラムおよび研究関連部局が共同で開発・検証し、化学品安全性および汚染防止局より統一試験指針として発表されているものである (OCSPP 890 Guideline Series, Table 1)。EPA は第 1 期スクリーニング試験の結果を WoE アプローチによって評価することで、ある化学物質が E、A、T いずれかのホルモン経路と相互作用する可能性があるか否かを判定し、第 2 期試験の必要性の有無を評価することを意図している。第 2 期試験の目的は、第 1 期で認められた E、A、T 経路への影響を更に究明するため、考えられる有害作用に対する用量反応関係を *in vivo* 実験によって確認することである。

- EPA は WoE アプローチを「生物学的妥当性および一貫性を適切に考慮するため、有意な情報すべてを総合的に評価すること」としている (USEPA, 1999)。
- 内分泌攪乱物質スクリーニング・試験諮問委員会 (EDSTAC) は WoE アプローチを「熟練した専門家が一群の情報の利点と弱点を判断することにより、個々のデータの考慮のみでは明瞭になるとは限らない全体的な結論を得ること」としている (EDSTAC, 1998)。

第 1 期スクリーニングに対する WoE アプローチについては第 5 章で論ずるが、それに先立って序説および裏付け情報として、EDSP の 2 段階スクリーニングの簡潔な歴史的概観および試験パラダイム (第 2 章)、科学的・技術的情報源 (第 3 章)、科学的・技術的情報の品質および有意義性の判定に関する一般的指針 (第 4 章) を述べる。

2 内分泌攪乱物質スクリーニングプログラム(EDSP)の概要

プログラムの歴史的詳細は EDSP のウェブサイト (<http://epa.gov/endo/>)、および本稿に引用するその他の資料やウェブサイトで見ることができる。

連邦食品・医薬・化粧品法 (FFDCA) の 1996 年の改正によって、EPA に対して「ある種の物質がヒトに対して、天然のエストロゲンと同様な作用、または局長の指定する類似の内分泌作用を及ぼすか否かを決定するため、適切な検証済み試験システムおよび学術的に有意義な情報に基づいてスクリーニングプログラムを策定すること」が要求された [21 U.S.C. 346a(p)] (<http://www.epa.gov/pesticides/regulating/laws/fqpa/>)。

EPAは局長裁量権に基づき2段階のスクリーニングおよび試験戦略を採用し、EDSPを拡張して、下記のとおりアンドロゲンおよび甲状腺経路ならびに生態学的効果をも含めることとした。

- 1998年、EDSTACの勧告(EDSTAC, 1998)に基づき、EDSP案を12月28日付 Federal Register Notice (63 FR 71542)に公開、審査のため局内の科学諮問会議(SAB)およびFIFRA 科学諮問パネル(SAP)に提出。この合同審査の結果は公開されている (SAB/SAP, 1999)。
- 2008年、各々の試験の審査を含む広範囲な検証(USEPA, 2007)を経て、EDSP第1期スクリーニング試験の案を1月24日付 Federal Register Notice (73 FR 4216)に公開、審査のためFIFRA SAPに提出。この審査の結果は公開されている (SAP, 2008)。
- 2009年、現在のEDSP第1期スクリーニング試験(表1)、および各試験につき利用可能な統一試験指針(OCSP 890 Guideline Series)を10月21日付 Federal Register Notice (74 FR 54416)に公開。
- 2009年、パブリックレビューおよびコメントを経て、最終的に67物質のリストおよび第1期スクリーニングにおける試験発注計画を10月21日付 Federal Register Notice (74 FR 54422)に公開。

2.1 EDSP 第1期スクリーニング試験

現在のEDSP第1期試験には11項目が含まれている。これらはFIFRA SAPの勧告(SAP, 2008)に従ったもので、表1に示すように、多様ではあるが互いに補完的な*in vitro*試験と*in vivo*試験から成っている。この組み合わせは全体として、E、A、T各ホルモン経路との相互作用の可能性に対して最大の感度を持つように設計されている(EDSTAC, 1998)。第1期スクリーニングは一般に下記を評価し得る試験法から、種々の要因を勘案して選定した。

- 種々の分類群におけるE、A、Tのいずれかのホルモン経路
- 受容体との結合によるエストロゲンまたはアンドロゲン媒介性効果(作動性および拮抗性)
- エストロゲン媒介性トランス活性化
- 生殖腺ステロイド産生経路の関係する酵素阻害
- 視床下部-下垂体-生殖腺(HPG)軸の関係するフィードバック機構を変質させる可能性のある生殖腺エストロゲンおよびアンドロゲン産生との相互作用
- ある試験内において他の試験と相補的な、アンドロゲンおよびエストロゲンに影響されるエンドポイント
- 甲状腺ホルモン産生または機能との相互作用、およびそれに関連する視床下部-下垂体-甲状腺(HPT)軸の関係するフィードバック関係の変質

第1期試験のロバスト性は、各試験の長所と、試験の組み合わせ相互間の相補的エンドポイ

ントに基づいている。したがって「個々の試験の結果は、他の試験の結果から切り離して考察することはできない。なぜなら一つの試験の限界を他の試験の長所によって補完するように全体が組み立てられているからである」(EDSTAC, 1998)。

表 1：EDSP 第 1 期スクリーニング試験は化学物質がエストロゲン(E)、アンドロゲン(A)、甲状腺(T)の各ホルモン経路と相互作用する可能性のある単一の作用機序(受容体結合)の、および内分泌経路に沿って(ステロイド産生、視床下部-下垂体-生殖腺および甲状腺軸)の主要エンドポイントを包括する*

スクリーニング試験	試験指針	受容体結合				ステロイド 産生		HPG 軸	HPT 軸
		E	抗 E	A	抗 A	E	A		
in vitro									
ER 結合 (ラット膈サイトゾル)	OCSPP 890.1250	■	■						
ERα 転写活性化 (ヒト細胞株 HeLa-9903)	OCSPP 890.1300 OECD 455	■							
AR 結合 (ラット前立腺サイトゾル)	OCSPP 890.1150			■	■				
ステロイド産生 (ヒト細胞株 H295R)	OCSPP 890.1550					■	■		
アロマターゼ (ヒト標的細胞または細胞株 ミクロソーム)	OCSPP 890.1200					■			
in vivo									
子宮肥大 (ラット)	OCSPP 890.1600 OECD 440	■							
Hershberger (ラット)	OCSPP 890.1400 OECD 441			■	■	■ ¹			
思春期雄性 (ラット)	OCSPP 890.1500			■	■		■	■	■
思春期雌性 (ラット)	OCSPP 890.1450	■	■			■		■	■
魚類短期間繁殖	OCSPP	■	■	■	■	■	■	■	

	890.1350 OECD 229								
両生類変態 (カエル)	OCSPP 890.1110 OECD 231								■

* 試験項目間で相補的なエンドポイントを各欄に■で示す。

1 5 α -レダクターゼ阻害のみ

試験指針のほか、第1期スクリーニング試験の各々に関する詳細な記述（開発、検証、長所および限界）も EDSP のウェブサイト

(<http://epa.gov/endo/pubs/assayvalidation/index.htm>)にある各実験項目についての EPA Integrated Summary Report および OECD Final Report で見ることができるが、本稿の目的のため、以下に各試験の特徴および相補的なエンドポイントについて概観することとする。

2.1.1 エストロゲン経路に対する物質の影響を検出する試験

第1期試験のうち下記の5種はエストロゲン活性および抗エストロゲン活性を持つ化学物質を検出することができる（表1）。

- 1 ER 結合（ラット子宮サイトゾル）
- 2 ER α 転写活性化（ヒト細胞株 HeLa-9903）
- 3 子宮肥大（ラット）
- 4 思春期雌性（ラット）
- 5 魚類短期間繁殖

in vitro の ER 結合試験は、ラットから切除した子宮組織のサイトゾルから分離したエストロゲン受容体（ER α または ER β ）に、ある化学物質が結合する可能性を検討するものである。しかし結合のみではその物質がエストロゲンアゴニストであるかアンタゴニストであるかを識別することができない。in vitro の ER 転写活性化試験は、ヒト子宮頸部腫瘍から得られた細胞において、エストロゲンアゴニストの可能性のある物質が ER(ER α)に媒介される遺伝子転写を活性化する可能性を検討するものである。

エストロゲン経路の評価に用いる in vivo 試験は、化学物質に対する種々の曝露経路、すなわち皮下注射（子宮肥大）、経口摂取（思春期）、水（魚類）に関係している。子宮肥大試験は卵巣を切除した成熟した雌、または性的に未成熟な無傷雌ラットを用いて行う。下記に基づいてエストロゲンアゴニスト活性を検出することができる。

- 子宮重量の増加
- 場合によっては子宮および膣の組織観察

思春期雌性試験は離乳後のラットについて行う。下記の複数のエンドポイントによりエストロゲンアゴニストおよびアンタゴニスト活性を検出することができる。

- 性的発達の特徴（膣開口日齢、発情周期、発情周期を示す個体の比率）
- 生殖器の重量と組織（卵巣および卵管、子宮および子宮液、下垂体）

魚類短期間繁殖試験は成熟した雌雄のファットヘッドミノーを用いて行う。下記の複数のエンドポイントにより、エストロゲンアゴニストおよびアンタゴニスト活性を検出することができる。

- 行動
- 生殖能力
- 受精成功率
- 二次性徴（生殖結節の数と大きさ、雄の臀鰭乳頭状突起）
- 生存率
- 体重・体長
- 生殖腺の大きさ（生殖腺指数）および組織病理学
- 雌雄の血漿中ビテロゲニン濃度
- 場合によっては雌血漿中のエストラジオール濃度

2.1.2 アンドロゲン経路に対する物質の影響を検出する試験

第1期試験のうち4項目が化学物質のアンドロゲンまたは抗アンドロゲン効果を検出することができる（表1）。

- 1 AR 結合（ラット前立腺サイトゾル）
- 2 Hershberger（ラット）
- 3 思春期雄性（ラット）
- 4 魚類短期間繁殖

in vitro の AR 結合試験は、ラットから切除した前立腺組織のサイトゾルから分離したアンドロゲン受容体に、ある化学物質が結合する可能性を検討するものである。しかし結合のみではその物質がアンドロゲンアゴニストであるかアンタゴニストであるかを識別することができない。

エストロゲン経路の評価に用いる *in vivo* 試験は、化学物質に対する種々の曝露経路、すなわち皮下注射または経口摂取(Hershberger)、経口摂取(思春期)、水(魚類)に関係している。Hershberger 試験は思春期前後の去勢ラットを用いて行う。下記のアンドロゲン依存器官または組織の重量変化(増減)を含む複数のエンドポイントに基づいてアンドロゲンアゴニストおよびアンタゴニスト活性ならびに 5α -レダクターゼ阻害剤(外因性テストステロンのジヒドロ

テストステロンへの転換の阻害)を検出することができる。

- 腹側前立腺
- 精嚢、凝固腺および分泌液
- 肛門挙筋・球海綿体筋
- カウパー腺
- 陰茎亀頭

思春期雄性試験は離乳後のラットについて行う。下記の複数のエンドポイントによりアンドロゲンアゴニストおよびアンタゴニスト活性を検出することができる。

- 性的発達の特徴（包皮分離の日齢）
- 生殖器官および組織の重量と組織学（睾丸、副睾丸、腹側および背側前立腺、精嚢・凝固腺および分泌液、肛門挙筋・球海綿体筋、下垂体）
- 血清中全テストステロン濃度の測定

魚類短期間繁殖試験は成熟した雌雄のファットヘッドミノーを用いて行う。2.1.1 項に示した複数のエンドポイント、および場合によっては雄の血漿中テストステロン濃度の測定により、アンドロゲンアゴニストおよびアンタゴニスト活性を検出することができる。

2.1.3 ステロイド産生経路に対する物質の影響を検出する試験

第1期試験のうち6項目がステロイド産生経路の攪乱を検出することができる(表1)。

- 1 ステロイド産生(ヒト細胞株 H295R)
- 2 アロマターゼ(ヒト標的細胞または細胞株マイクロソーム)
- 3 Hershberger(ラット)
- 4 思春期雄性(ラット)
- 5 思春期雌性(ラット)
- 6 魚類短期間繁殖

in vitro のステロイド産生試験ではヒト細胞株(H295R)を用いて、化学物質とステロイド産生経路との相互作用を、テストステロンおよびエストラジオール産生の変化(増減)により検出する。*in vivo* のアロマターゼ試験では各種標的組織または細胞株のマイクロソームを用い、アロマターゼ活性の阻害およびアンドロゲンのエストロゲンへの転換を検出する。この試験はアロマターゼ活性の誘導を検出するには不適當である。

前項に述べた *in vivo* 試験とそのエンドポイントに関して、ステロイド産生(生殖腺でのエストロゲンまたはアンドロゲンの産生)に対する見かけ上の影響が、*in vitro* 試験ではそれを裏付けるステロイド産生効果が認められないにも関わらず観察されることがある。化学物質がゴナドトロピンの合成または分泌に直接作用するか、または HPG 軸によるフィードバック反応を惹き起こすことにより、ステロイド産生を間接的に攪乱することも考えられる。たとえばアンドロゲンまたはエストロゲン活性を持つ物質に曝露されたとき、HPG 軸を通じてのネガティブフィードバックの結果として内因性アンドロゲンまたはエストロゲンの濃度が減少するかもしれない。視床下部または下垂体レベルの効果の意味については以下で更に検討する。

2.1.4 HPG 軸に対する物質の作用を検出する試験

一般に HPG 軸の調節には、ポジティブおよびネガティブフィードバックの複雑な組み合わせが関係している。生殖腺エストロゲンまたはアンドロゲンは、視床下部または下垂体で対応する受容体に結合することで、それぞれゴナドトロピン放出ホルモンまたは性腺刺激ホルモンを調節し、これらが卵巣または睾丸におけるステロイド産生を調節する。現在の第1期スクリーニング試験には、視床下部または下垂体による生殖腺エストロゲンまたはアンドロゲンホルモン産生の調節に影響する化学物質を特異的に検出する *in vitro* 試験は含まれていないが、下記3種の *in vivo* 試験は HPG 軸に対するこのような作用を検出することができる(表1)。

- 1 思春期雌性(ラット)
- 2 思春期雄性(ラット)
- 3 魚類短期間繁殖

第1期スクリーニングの枠内で *in vivo* の思春期試験および魚類短期間繁殖試験の結果を解釈し、当該物質が生殖軸の視床下部-下垂体による調節に影響するかどうかを判定することができる。簡単な例として、思春期雌性または雄性試験で陰嚢開口または包皮分離が遅れても、*in vitro* の ER または AR 結合試験、ステロイド産生試験、アロマターゼ試験の結果と組み合わせれば陰性の判定になることがある。そのような結果はステロイド産生への化学物質の直接的な影響を示すものとは解されないにせよ、視床下部または下垂体レベルでの効果を通じて間接的な影響を及ぼす可能性を示唆するものである。

2.1.5 HPT 軸への物質の影響を検出する試験

HPT 軸の調節は一般的には HPG 軸と同様であるが、フィードバック関係に甲状腺ホルモン（チロキシンなど）が関与している点が異なる。甲状腺ホルモンの視床下部および下垂体へのフィードバックによって、それぞれチロトロピン放出ホルモンおよび甲状腺刺激ホルモン (TSH) が調節され、それらが甲状腺によるホルモン産生を調節する。現在の第1期スクリーニング試験には、視床下部または下垂体による甲状腺ホルモン産生の調節に影響する化学物質を特異的に検出する *in vitro* 試験は含まれていないが、下記3種の *in vivo* 試験は HPT 軸に対するこのような作用を検出することができる（表1）。

- 1 思春期雌性（ラット）
- 2 思春期雄性（ラット）
- 3 両生類変態（カエル）

思春期試験については前述したが、甲状腺ホルモン経路に影響する可能性のある化学物質の検出に関連して、下記のエンドポイントを利用する。

- 甲状腺の重量および組織
- 血清中の下垂体 TSH および甲状腺チロキシン(T4)濃度の分析

両生類変態試験はオタマジャクシを用いて行う。主として7日～21日齢における、変態に関連する下記の発達エンドポイントの変化および甲状腺組織に基づいて、水を介しての化学物質への曝露の影響を評価する。

- 発達段階
- 甲状腺組織
- 下肢の長さ
- 全身の長さ（鼻から排出口まで）

2.2 EDSP 第2期試験

第2期試験では、分類群（哺乳類、鳥類、両生類、魚類、無脊椎動物など）にまたがる一層包括的な研究によって、内分泌系に限らず神経系、免疫系、肝臓、腎臓なども含めた毒性その他の有害性を考慮して、広くヒトの健康や生態系に対するリスクの評価を行うことになる。第

2 期試験の詳細な説明は EDSTAC 報告書(EDSTAC, 1998)、1998 年 12 月 28 日付 Federal Register Notice (63 FR 71542)、および EDSP のウェブサイト (<http://epa.gov/endo/>)で見ることができる。

3 科学技術情報の出典

EDSP 第 1 期スクリーニング試験は、化学物質が E、A、T のいずれかのホルモン経路と相互作用する可能性を評価するために設計されており、したがって第 1 期スクリーニングの総合的な結果は WoE 評価において考慮される科学技術情報の主要な源泉となることが期待される。しかし評価に際しては他の情報源、たとえば第 1 期スクリーニングに関連あるものとして提供された情報も考慮される可能性がある。そのような情報の出处としては、農薬登録者や化学企業の行った研究、公刊または公開された査読済み論文をはじめ多様なものがある。

3.1 試験指針：第 1 期スクリーニング試験

EDSP の第 1 期スクリーニング試験の発注に基づいて得られるデータは、表 1 に示されているとおり OCSP 試験指針シリーズ 890 に規定されている。これらのスクリーニング試験は広範囲にわたる検証、個別のピアレビュー、および FIFRA SAP による独立のレビュー(SAP, 2008)を経たものである。またこの試験指針は利用者に対する明確な指針となり得るような詳細さで記述されており、各試験の実施および結果の解釈に関する推奨事項を含んでいる。更に各試験について結果の評価を助けるため標準評価手順(SEP)が作成されている(4.2 項)。試験指針の公開に続き、EPA は外部から表明された疑問に答える形で、各スクリーニング試験の性能基準の適用に際しての技術的な操作および結果予想に関する追加的な指針となる資料を作成した(USEPA, 2011a)。更に EPA は EDSP ウェブサイトを開設し、SEP やデータ評価レコード(DER)テンプレートなどの情報、試験発注に対する応答状況、現状の追跡などを提供している (<http://epa.gov/endo/pubs/toresources/index.htm>)。

このように EPA は、第 1 期スクリーニング試験と結果の評価に職員も外部も利用できる情報を複数のレベルで提供している。第 5 章に説明する WoE アプローチで強調されているように、どの試験あるいはエンドポイントもそれだけを独立に解釈すべきものではない。第 1 期スクリーニングは全体として実行したとき相互に補完的な情報が得られるように設計されており、更に WoE 評価では第 2 期試験の必要性を判定する際に、学術的に意味のある他の情報も加味することがある。

3.2 学術的に有意義な情報

任意に提供された情報で EPA の措置を支持または明確化するものは一般に「その他の」学術的に有意義な情報と呼ばれる。そのような科学的・技術的情報の出处としては、EPA 試験指針または OECD の同等の試験指針に関する研究結果、公刊または公開された査読済み論文などがある。情報は出处の如何を問わず、EPA の情報の品質に関する指針(USEPA, 2002b)による品質および意義の評価を経た上で利用される。

EPA は EDSP 第 1 期スクリーニング試験の発注を 2009 年秋から開始した。当初の応答として、受注者から受注の意思表示と共に、第 1 期試験に代わるものとして考慮すべき既存の学術情報がしばしば提供された(USEPA, 2009a)。EPA は提供された情報が第 1 期試験の発注条件に適合するかどうかを評価し、結果を受注者に通知した。

試験発注への適合のためには、受注者は EDSP 第 1 期スクリーニングの結果を提出しなければならない。この際にその他の学術的に有意義な情報を含めてもよい。その情報には当初受注時に提出したものも、新たな追加情報も含めることができる。EPA は追加的に提出された情報の品質と意義を考慮して、第 1 期スクリーニング試験結果と共に、化学物質の E, A, T ホルモン経路との相互作用の可能性の有無を決定するための WoE 分析において考慮する。

3.2.1 試験指針 — 健康および生態影響の研究

40 CFR Part 158, subpart F, G に示されている EPR の規制には、健康および生態系への影響が定義されている(それぞれ 870 および 850 Guideline Series)。FIFRA の要求に対応して整備されたデータは、上述のピアレビュー済み試験指針または対応する OECD 指針によるものである。

これらの試験は GLP 規制(40 CFR Part 160, Part 792)に従って行われ、データの一貫性・再現性・完全性が保証されている。従来毒性研究では、内分泌系もそれ以外も含めた複数の作用機序(MoA)の影響を受ける最終的エンドポイントを利用していることが少なくない。このため EPA 試験指針または同等の OECD 指針には、化学物質の E, A, T 各ホルモン経路との相互作用の可能性を定める WoE 分析に際して、第 1 期スクリーニング結果と並んで考慮すべき科学技術情報を示したものがあ

る。EDSTAC の EPA への報告書(EDSTAC, 1998)において、特に哺乳類の二世世代生殖毒性試験について次のように言及されている。

「…ホルモン効果の可能性は、行動の変化、妊娠能力、妊娠期間、出産の困難または時間延長の徴候、仔の見かけの性比(肛門性器間距離で確認した)、仔の雌性化または雄性化、産仔数、死産数、膈・子宮・卵巣・睾丸・副睾丸・精囊・前立腺その他の標的器官の組織病理によって検出することができる」

哺乳類の二世世代生殖毒性試験、および新しい OECD の拡張一世代生殖毒性試験指針が EDSP の第 2 期試験として提案されている。拡張一世代生殖毒性試験は生殖試験から従来型の

情報を得るように設計されたものであるが、それを拡張して、40 CFR Part 158 指針の他の試験では必ずしもカバーできない、雌雄齧歯類の成獣および新生期・離乳期・思春期の仔の内分泌系・神経系・免疫系に関連する生殖および発達エンドポイントの評価が可能になっている。

エストロゲン経路については、健康影響に関する 40 CFR Part 158 試験指針 Subpart F（または同等の OECD 指針）の試験が学術的に有意義な情報源となり得る。エストロゲンに影響されるエンドポイントは、二世代および拡張一世代生殖毒性試験、発達神経毒性試験、亜慢性・慢性および癌バイオアッセイに存在する。

すべてのエンドポイントおよび測定の詳細については、これらの試験に対する指針を参照する必要があるが、エストロゲンに影響されると考えられるエンドポイントのいくつかの例を挙げれば下記のとおりである。

- 膻開口時の日齢
- 発情周期
- 生殖器官の重量および組織病理
- 生殖能力

アンドロゲン経路については、上述の試験が学術的に有意義な情報源となり得る。すべてのエンドポイントおよび測定の詳細については、これらの試験に対する指針を参照する必要があるが、アンドロゲンに影響されると考えられるエンドポイントのいくつかの例を挙げれば下記のとおりである。

- 肛門生殖器間距離
- 包皮分離の日齢
- 齧歯類雄幼獣の尿道下裂、陰茎裂、乳輪/乳頭遺残
- 生殖器官の重量および組織病理、精子形成
- 生殖能力

甲状腺ホルモン経路については、40 CFR Part 158 試験指針 Subpart F（または同等の OECD 指針）の他の試験（齧歯類およびイヌの 90 日試験、イヌの 1 年慢性試験、マウスおよびラットの慢性試験）が学術的に有意義な情報源となり得る。測定の詳細については、これらの試験に対する指針を参照する必要があるが、甲状腺固有のエンドポイントには下記のようなものがある。

- イヌの 90 日試験における甲状腺重量および組織病理
- マウスおよびラットの 90 日試験における甲状腺重量および組織病理
- イヌ慢性毒性試験、ラットの慢性毒性/発癌性試験、マウスの発癌性試験における甲状腺の組織病理

- 甲状腺に影響することが知られている、またはその可能性のある化学物質については、甲状腺ホルモン(T3, T4, TSH)、甲状腺の重量および組織病理の試験を行ってもよい。

甲状腺ホルモン関連のエンドポイントは従来の哺乳類二世代生殖毒性試験指針には含まれていないが、拡張一世代生殖毒性試験指針では推奨されている。また下垂体ホルモン(TSH)や甲状腺ホルモン(T3, T4)の測定値を含む研究も存在する。

40 CFR Part 158 の試験指針のうち生態系への影響に関する **Subpart G** も学術的に有意義な情報源となり得る。これらの試験は、化学物質と内分泌系の相互作用の可能性を反映し得るエンドポイントを含んではいるが、予測性を持つものとは解されない。すべてのエンドポイントおよび測定の詳細については、鳥類生殖試験、魚類全ライフサイクル試験、魚類幼若期試験の指針を参照する必要があるが、内分泌系との影響に関する情報を与える可能性のあるエンドポイントとして下記のものがある。

- 生殖能力
- 生殖の成功率
- 卵の発達
- 胚/幼生の生存率および成長

特に、生殖器を含む体内諸器官の総合的な形態が病的と見られる場合には組織分析も行われる。

化学物質と内分泌系との相互作用については、無脊椎動物からも知見が得られることがある。無脊椎動物の内分泌系は、脱皮を調節する一群のステロイド様ホルモン（エクジソンなど）に依存していると思われる。無脊椎動物には哺乳類のような機能的なエストロゲンまたはアンドロゲン受容体は知られていないが、エストロゲン受容体を含む保存核内受容体スーパーファミリーの一部をなす核内受容体にエクジソンが結合する。現在のところ、EDSP 第1期スクリーニング試験で評価されるようなエストロゲン結合やトランス活性化とエクジソンとの関係はよく知られていない。

しかし全ライフサイクル慢性試験指針による研究は、典型的にはミジンコ(*Daphnia magna*)と河口性または海洋性のアミ(*Americamysis bahia*)を用いて行われる。すべてのエンドポイントおよび測定の詳細については、無脊椎動物に対する各々の試験の指針を参照する必要があるが、内分泌系との影響に関する情報を与える可能性のあるエンドポイントとして下記のものがある。

- 成長
- 生殖
- 生存率

要約すれば、40 CFR Part 158, Subpart F および G の EPA 規制は、それぞれ健康および生態系への影響に関する毒性データの要件を定めてはいるが、特に化学物質の E, A, T ホルモン経路との相互作用の可能性を試験するためのものではない。しかし EPA ないし同等の OECD 指針による試験の中には、内分泌系に関する効果に関する有用な科学技術情報を提供するものもある。より広範囲な毒性の枠内での、内分泌系以外への化学物質の直接的影響（容量応答関係、有害作用など）が、内分泌系への間接的影響の解明に役立つこともある。このため、EPA の Part 158 試験指針および OECD の対応指針は、化学物質の E, A, T ホルモン経路との相互作用の可能性の有無を決定するための WoE 分析において、第 1 期スクリーニング試験結果と並んで考慮すべき科学技術情報を提供する可能性がある。

3.2.2 公刊または公表された査読済み論文

EPA はリスクアセスメントにおいて、健康および生態系に対する化学物質の悪影響の可能性に関する知識を明示するために、公刊または公表された査読済み論文を用いている。これに対応して、公刊または公表された査読済み論文が入手できるならば、化学物質の E, A, T ホルモン経路との相互作用の可能性の有無を決定するための WoE 分析において、第 1 期スクリーニング試験結果と並んで考慮すべき科学技術情報を提供する可能性がある。

公刊または公表された査読済み論文の研究は一般に、仮説を設定し観察・実験によってこれを検証し確認するという、科学の標準的な方法論に従っている。しかし指針による研究と異なり、公開文献に発表されている（指針によらない）研究の多くは GLP に従ってはいない。したがって WoE 評価において第 1 期スクリーニング試験の結果と並んで考慮すべき研究は、指針によると否とを問わず、EPA 情報品質指針に従って品質と意義の評価が行われる。

4 科学技術情報の品質

試験指針によるか否かを問わず、論文に含まれる科学技術情報の評価は、規制に関する意思決定を支援すべき WoE 分析において基本的に重要である。行政管理予算局(OMB, 2002)への対応として「EPA の発信する情報の品質・客観性・有用性・完全性の保証および最大化のための指針」(Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency) (USEPA, 2002b)には情報品質の保証と最大化のための方針と行動指針が含まれている。これを含め、科学技術情報の評価に関する各種指針が EPA ウェブサイト (<http://www.epa.gov/quality/informationguidelines/>)に掲載されている。

毒性および生態毒性データの質の評価のための系統的アプローチを述べた或る重要な論文に基づくとと思われるデータ品質評価手順を述べた報告書の存在も把握している (Klimisch et al., 1997)。

- 大生産量の化学品の研究に関する OECD マニュアル(OECD Manual for Investigation of High Production Volume (HPV) Chemicals) (OECD, 2005)
- オーストラリア生態毒性データベースの品質評価手順(Australian Ecotoxicity Database Quality Assessment Scheme) (Hobbs et al., 2005)
- 毒性データの信頼性評価ツール(Toxicological data Reliability assessment Tool (ToxRTool)) (Schneider et al., 2009)

これらのアプローチはすべて原理的には同様であり、主として「適切性(adequacy)」「信頼性(reliability)」「有意義性(relevance)」に基づいている。中でも信頼性(すなわち妥当性または健全性、完全性)が中核的な評価基準として強調されている。

EPA 内部では、情報の質に関する指針(USEPA, 2002b)が、情報の品質に関する多くの既存のシステム・活動・指針に基づいて定められている。この章ではこれらの指針の基礎を要約して一般的評価因子(USEPA, 2003)を示し、科学技術情報の品質評価について考察する。

4.1 一般的評価因子 (GAF)

行政管理予算局(OMB, 2002)への対応として EPA は情報の品質に関する指針 (USEPA, 2002b)を作成・公表した。これは部分的には EPA 全体の方針あるいは他のプログラム固有の方針から発展したもので、出処の如何を問わず情報の質を保証し最大化するための EPA の方針と手順指針を定めている。これによって任意に提供された情報、EPA が収集または自ら作成した科学技術情報の品質評価への一般的なアプローチの透明性が高められる。EPA の科学政策委員会 (Science Policy Council, SPC) は下記 5 項目の一般的評価因子 (General Assessment Factor, GAF)の使用を推奨している(USEPA, 2003)。

- 1 健全性(soundness)
- 2 適用可能性および有用性(applicability and utility)
- 3 明瞭性および完全性(clarity and completeness)
- 4 不確実性および変動(uncertainty and variability)
- 5 評価およびレビュー

これらは EPA の活動を支援する科学技術情報の品質および意義の評価を記述する、既存の EPA 品質システム、慣行、指針から引き出されており(USEPA, 2002b)、品質に関する新しい考察や新規な情報評価方法を提示するものではない。

GAF は、EPA Exposure Handbook (USEPA, 2009b)に収録する試験を選択した場合のように個別の情報に適用することも、あるいは一群のエビデンスに適用して WoE アプローチによる全体的評価を行うことも可能である。WoE アプローチは既に定義されているように(USEPA 1999; EDSTAC, 1998)、学術的に有意義な情報すべてを考慮して総合的に分析することによる解釈の手順であって、入手可能な各種エビデンス、エビデンスの質と量、各々の種

類のエビデンスの長所と短所を考慮し、各種エビデンスが全体としてどのようにして結論を支持するかを説明している。

以下では SPC (USEPA, 2003) に従って 5 項目の GAF の概要を説明し、学術的に有意義な情報として提出された科学技術情報の品質評価のガイダンスとして、項目ごとに具体的な考察を加える。これらは必ずしも網羅的なものではなく、また項目間で重複する場合もある。

4.1.1 健全性

「情報を得るために使用した科学的・技術的手順、手段、方法あるいはモデルが所期の目的に対して合理的であり適合していること」

考察：1) 目的とする効果の検出に対して試験方法が適切であること、2) 試験が仮説設定、観察・実験による検証、確認の科学的方法に従って行われていること、3) 試験の目的に従って、特定の結果と一般的な結果とを識別できること、4) 結果の解釈と結論が統計的に有意であり、生物学的に妥当性があり、データと矛盾しないこと。

4.1.2 適用可能性および有用性

「情報が EPA の意図する利用法に対して有意義であること」

考察：1) 試験目的に関連する根拠、目的、仮説に照らして試験資材および方法、試験計画、エンドポイントが適切であること、2) データの収集・解析・表現・解釈および結論において十分な能力を示すものであること、3) 新旧の方法論による情報に信頼性があること。

4.1.3 明瞭性および完全性

「データ、仮定、方法、品質保証、後援機関、および情報を発生するために用いた解析方法が明記されていること」

考察：1) 著者、共著者、寄稿者、およびそれぞれの所属機関または団体、後援機関が明確であること、2) 背景または根拠、試験目的、検定しようとする仮説、実験計画（対照群、試験目的に関連する測定またはグループ数など）、3) 反復性、正確度、精度を保証する標準化の程度、または科学的に妥当な方法論、4) 生データの入手可能性、5) 統計的解析のアプローチ、6) 統計的優位性の解釈、生物学的結果の妥当性、科学的に健全な結論。

4.1.4 不確実性および変動

「情報または手順・方法・モデルの不確実性および変動（定性的および定量的）の評価と特徴づけが行われていること」

考察：1) 試験方法またはモデル、実験計画、あるいはエンドポイントの選択性および感度に関する参考資料の引用、2) 試験方法の再現性・反復性のエビデンス、3) 性能基準および品質管理または保証対策（履歴、参照管理情報、変動係数、GLP への適合、独立のピアレビューなど）、4) 動物数または測定/グループ数、グループ間の差異の検出に十分かつ適切な統計解

析のアプローチ。

4.1.5 評価およびレビュー

「情報、または手順・尺度・方法またはモデルが独立に確認・検証されピアレビューが行われていること」

考察：1) 試験方法および特異的かつ鋭敏な測定単位としてのエンドポイントの有意義性および信頼性を評価するための確認または検証手順の説明または引用、2) 査読済み論文において方法が一般的に受容されていること、3) 検証結果が入手できること、4) 性能または評価の基準が入手できること。

4.2 標準評価手順(SEP)およびデータ評価記録(DER)

EPA は一般に、試験指針による各個の試験のために作成された SEP を、各試験の実施方法および結果の解釈の評価に用いる。評価に続いて各試験に対する DER が作成される。また EPA は、追加的情報として提供された指針によらない試験を WoE 分析に利用する前にレビューを行う。DER は EPA によるレビューの正式な記録であり、試験の実施方法の適切性、指針への適合性の評価の要約と、データに支持される解釈および結論を含んでいる。EDSP 第 1 期スクリーニング試験の各々に対して SEP とそれに対応する DER テンプレートが作成されており、それぞれ結果を評価・解釈・要約したものが WoE 分析に使用される。EDSP 第 1 期スクリーニング試験のための SEP および DER テンプレートは EDSP ウェブサイト (<http://epa.gov/endo/pubs/toresources/index.htm>) に掲載されている。

5 エビデンスの重み付けアプローチ

この章では、第 2 期試験に関する意思決定として、物質が内分泌系に媒介されるプロセス（すなわち E, A, T 経路）と相互作用する可能性を WoE 分析で決定する際の原理、基準、アプローチについて説明する。

WoE は一般的に、得られたデータが、ある物質が特定の効果を持つとの仮説をどの程度支持するかを判定する手順と定義される (USEPA, 1999; 2002a; 2005)。この手順には有意義なデータの収集、データの質と意義の評価、物質の性質に関する結論を支持するための各種エビデンスの総合など多くの段階がある。WoE は単に陽性・陰性それぞれの結果を示す試験の数を比較するだけではなく (USEPA, 2002a)、専門的な判断を必要とする。したがって WoE 分析においては透明性が重要である。WoE による評価では、利用できるデータの種類、その選択と評価の方法を述べ、異なった種類のエビデンスが相互に適合して結論が引き出されることを示す。またデータの利点と欠点や不確定性など、特に考慮すべき主な論点を示し、解釈における重要な点を強調する。

第 2 章で説明したように、第 1 期試験は受容器結合（エストロゲンおよびアンドロゲンアゴニストおよびアンタゴニスト）への影響、ステロイド産生への影響、視床下部－下垂体－生殖

腺（HPG）および甲状腺(HPT)軸に対するその他の影響など生物学的事象を評価するために設計されている。したがってこの場合のWoEアプローチは、化学物質がE, A, T経路と相互作用する可能性の有無を決定するような結論への到達および結論の支持に関してEDSP第1期試験からのデータを検討することになる。3.2項に述べたように、他の情報源も適切と見なされることがある。

前述のようにこのWoE分析は、当該化学物質をEDSP第2期試験において更に試験する根拠の有無の決定を支援することが目的である。このためWoEには下記のような考察が含まれることになる。

- 存在するデータは、当該物質がE, A, T経路の正常な機能と相互作用する可能性があるとのエビデンス（たとえば個々の試験結果内部および異なる試験間での一致）として意義・ロバストネス・一貫性のあるものを提供するか。
- データが上記内分泌経路と相互作用する可能性を示す場合、どのホルモン経路（E, A, Tのいずれか）が影響されるか、また第2期試験としてどのようなものが適当か。

使用するWoEアプローチは仮説に基づくアプローチとして特徴づけることができる(USEPA, 2005; Boobis et al., 2006, 2008; Rhomberg, 2010)。特にMoAデータの評価に構造・厳密さ・透明性を与えるため、国際化学品安全プログラム(IPCS)におけるEPAの作業(2005)に関連してWoEの枠組みを構築した(Boobis et al., 2006; 2008)。ここで使用した基準は、生物学的妥当性、エビデンスの総体としての整合性、強さ、一貫性を含めたEDSPのWoE評価にも適用できる。内分泌系に媒介される過程の複雑さを反映する異種のエビデンスもこのWoEの枠内で評価され、化学物質がE, A, Tホルモン経路と相互作用するかどうかに関する仮説ないし問題を検討する。この問題へのアプローチでは一般に第1期試験を用いて、生体組織の種々のレベルでの効果を考慮する。生体組織の種々のレベルにおけるエビデンスに支持された仮説に基づくアプローチの適用例を以下に示す。

- 化学物質と分子的ターゲットとの相互作用、たとえばエストロゲン受容体に対する拮抗作用（*in vitro*のER結合試験で測定される）
- これにより細胞の機能的応答が変化し、たとえば雌のビテロゲン産生が減少する（魚類短期間生殖試験で測定される）。
- このことは器官ないし組織レベルでの構造的応答の変化、たとえば生殖腺指数(GSI)の減少、卵母細胞の変質、卵胞の発達の変化（魚類短期間生殖試験における雌において評価される）によって裏付けられる。
- 最終的に、エストロゲン受容体拮抗作用の結果として全身レベルの有害作用、たとえば生殖能力の減退が現れる。

この例では、最初の3項目に例示されているように、第1期スクリーニング試験によって化学物質と内分泌系に媒介される過程との相互作用が見出され、最後の項目に示すように、第2

期の *in vivo* 試験によって全身レベルの有害作用とその容量応答関係の情報が得られる。

第2章に述べたとおり、EDSP 第1期試験の各項目は相補的になるように設計されているので、個別の試験を組み合わせることで解析することにより E, A, T 経路との相互作用について一層詳細な知見が得られる。WoE 評価においては異種のエビデンスを総合的に評価するのであり、単一の試験またはエンドポイントのいずれもそれ自体で第2期試験の要否の判定を支持するのに十分なロバストネスを持つと期待されていないことが本稿の基本的な眼目である。

この WoE 分析は個別的に行われ、まず個々のエビデンス（すなわち個別の試験項目、5.1 項参照）を集めて評価し、ついですべてのエビデンス（すなわちすべての試験項目、5.2 項参照）を総合的に評価する。

次項に述べるように、WoE 分析で考慮するデータはすべて記録されたものであること、科学的に妥当なものであることが必要である。

5.1 個別試験データの収集と評価

WoE 分析は典型的には個別の研究を注意深く評価することから始まる。1つの種類のエビデンスの評価プロセスには、データの収集、現行の受容および品質基準によるそのデータの評価、各試験結果に関する結論の記述が含まれる。試験のレビューに際しては、どの研究を対象としたか（あるいは対象から除いたか）、試験の質をどのように判断したかを明確にしなければならない。5.2 項で述べるように、個別の試験に対する結果を試験の種類およびエンドポイントに従って表とすることにより、結果が構造的かつ透明性をもって示され、WoE による決定が容易になる。

データを集めて評価するに際して、考慮する情報は科学的に健全なだけでなく、第2期に追加試験が必要か否かの判断に対して意義を持つものでなければならない。第2章で述べたように、EDSP 第1期試験はある物質が E, A, T ホルモン経路と相互作用するか否かを決定するために計画され、学術的ピアレビューを経た試験プロトコルに従って実施される。第1期試験が正しく実施されれば、データの質は一般に健全であり、化合物が E, A, T 経路と相互作用するか否かを判定するのに適切であることが期待される。したがって評価では各試験の、相互作用の可能性を検出する上での方法論的な長所・短所が重要になる。いくつかの第1期試験の長所および考えられる短所の例が、個別の試験に関する EPA Integrated Summary Reports および OECD Final Reports (EDSP ウェブサイトに掲載されている)、2008 FIFRA SAP report (SAP, 2008)、およびその他の総合報告 (Eldridge & Laws, 2010; Bogert et al., 2011 など) に見られる。EPA は各々の試験（指針に従っているか否かを問わず）について、用いられた方法および試験の実施条件を確認して科学的な質を、したがって WoE への寄与としての信頼性のレベルを評価する。第4章で一般的に述べた科学技術情報の品質と有意義性の評価に加えて、WoE の文脈では分析の助けとして下記の質問を用いることで試験の評価が容易になる。ただし必ずしもすべての質問がすべての試験に適用できるわけではない。また個々の評価担当者が適切と判断すれば別の質問を用いることもあり得る。

すべてのデータ源に対する考慮：

方法の質/妥当性

- EDSP 第 1 期試験において、個々の試験が試験指針をどの程度遵守したか。不一致があったか、あった場合明確に記述されているか。不一致は試験結果あるいはその解釈可能性に影響するか。
- 指針に準拠しない試験において、実験手順・方法・モデルが科学的に健全であるか、十分に記録されているか、E, A, T 活性の評価に適しているか。
- 指針に準拠しない試験の場合、使用資材、装置の条件、測定手順、対照群、試験の長所・短所の記載は、独立の評価が可能な程度に詳細であるか。結果の記載は、独立の評価が可能な程度に詳細であるか。
- EDSP 第 1 期試験のプロトコルに準拠しない試験の場合、他の品質基準に適合しているか。（注：EPA 指針および同等の OECD 指針による試験は通常 GLP (40 CFR Part 160 および Part 792) に従って実施される。）

結果の信頼性

- EDSP 第 1 期試験指針は検証されているが、指針および SEP に記述されている試験性能または基準に基づいて第 1 期試験を実施する実験室の能力が十分に示されているか。
- EDSP 第 1 期試験プロトコルに準拠しない試験の場合、エンドポイントの測定の信頼性評価がなされているか。
- 実験計画は適切か（実験資材の純度と安定性、ビヒクルまたは溶媒の適切性、投与方法、実験動物の数・種・系統の適切性など）。
- 実験計画とその実施（該当する場合）を評価するための適切な陽性および陰性対照群が含まれているか。
- 動物または *in vitro* の複製の数は試験指針の推奨に従っているか。投与量の選定の根拠が明示されているか、その投与量は適切か。
- 統計解析は適切に説明されているか。適切な解析法が選択されているか、また正しく実施されているか。

観察された影響の本質

- 試験物質の影響が明確に記述されているか。
- 陽性および陰性対照（使用している場合）に観察された応答は適切か。
- 報告されている応答はどのような条件でのものであるか。たとえば試験系の環境的・生理的条件はどのようなものであるか。曝露経路が応答に影響している可能性はないか。
- 応答の程度はどのくらいか。複数の処理投与量が評価されている場合、用量応答はど

のようなものであるか。

- *in vitro* 試験において、濃度応答の形状はどうか。細胞毒性のエビデンスはあるか。どのような細胞毒性試験を行ったか。試験は適切な濃度範囲にわたって行われているか。観察された影響は細胞毒性濃度のみにおいて現れたものか。
- 試験物質の溶解度に問題があったか (*in vitro*, *in vivo* いずれも)。試験物質の溶解度の限界が明示されているか。その限界値はどのようにして決定され、また試験においてどのように処理されているか。
- *in vitro* 試験において、動物の臨床的徴候、体重変化など目標以外の変化としてはどのようなものが記録されているか。
- 試験の種類と測定する影響に応じて、その影響は強弱いずれか、持続性・可逆性・一過性を評価した場合その程度はどうか。
- 応答に顕著な変動があったか。応答は当該の試験、種または系統において正常な範囲内であるか。

各試験において報告されたエンドポイントの間の整合性と関係

- 複数のエンドポイントを評価した試験（一般に *in vivo* 試験）において、測定されたエンドポイントの間に、E, A, T 経路との相互作用の可能性を示すような影響の整合的なパターンが認められるか。

測定されたエンドポイントの有意義性、特異性、感度

- 第1期指針によらない試験の場合、化学物質と E, A, T 経路との相互作用の可能性を評価するのに有用な情報の得られるエンドポイントを測定しているか。
- 試験は鋭敏なモデルを用いて、感受期に実施されているか(たとえば 22~42 日齢に試験物質に曝露した思春期前後の雌について 22 日齢以降の膣開口および体重を確認)。
- 報告された影響が内分泌系以外の事象(たとえば全身一般毒性)から来ている可能性はないか。

4.2項に述べたように、試験の品質および試験結果からの結論の記述は通常 DER に記入する。DER は EDSP 第1期試験の各項目ごとに作成し、試験発注時の要求条件を満たしたか否かを明記する。WoE 評価で考慮した公刊または公開の査読済み論文についてもレビューを作成する。DER は 40 CFR Part 158 のデータ要件への適合のための標準毒性試験指針にも利用できる。

試験の信頼度の決定を助けるため、試験の長所および付随する短所と不確実性を一般的に評価・説明・報告する。E, A, T 経路への影響のような複雑な問題の評価においては、EPA が試験の方法およびデータに関する詳細な情報を持つことが極めて重要である。

この詳細情報は試験方法の全体的な適切性・信頼性(たとえば実験群の大きさは十分か、対

照群は適切か、用量は適切か、など) の判定に用いられる。質の高い試験とは一般的に、科学的に許容される方法論に適合し、方法とデータの両方を十分に記録した研究である。一般に GLP に適合し、ピアレビューを経た試験指針に従って実施された試験、またはその他の品質保証制度あるいは規格に適合する試験であれば、得られた情報に高い信頼が置かれる。したがって推奨される方法との不一致あるいはそれからの逸脱は WoE 評価において考慮しなければならない事柄である。試験指針または SEP は試験の信頼性を判定するための有用な手段となる。方法が十分記録されていない、または許容できない試験、あるいは計画・実施あるいは知見の報告に許容できない欠陥がある研究は質的に許容されず、有用ないし信頼性ある情報を提供するものとは見なされない。

E, A, T 経路に関する情報の得られるエンドポイントを測定した試験、あるいは複数のエンドポイントを測定して、E, A, T 経路の影響を受ける相互に関連したエンドポイントの間に一貫した応答を認めた試験は全体的な WoE 分析(5.2 項)において主要なエビデンスを提供する。相互に関連するエンドポイントの間に、特に理由なく一貫したパターンの見られない試験や、変動あるいは著しい細胞毒性などの要因が交絡している試験は一般に、化学物質と E, A, T 経路との相互作用の可能性について有用かつ信頼性あるエビデンスを提供するものとは見なされない。第 1 期試験に欠陥が見出されたときは、WoE による決定に寄与する可能性のある試験またはエンドポイントのデータを含め、関連情報すべてを考慮した上で、試験を再度行うか、あるいは欠陥を回復するための特別な試験を行うよう要請する場合がある。

応答が不十分なものであるか明らかに陽性であるかの考慮および特徴づけも、E, A, T 経路との相互作用、したがって第 2 期試験の対象である健康および生態系への悪影響を及ぼす可能性の大きい化合物と小さい化合物との区別に役立つ点で有意義な情報を提供する。

5.2 項に述べるように、個別の化学物質が第 2 期試験を要するかどうかの決定は、エビデンスの総体 (EDSP 第 1 期試験における分析および他の学術的に有意義な公刊または公開されている査読済み論文) に基づいて行う。

5.2 異種のエビデンスの総合

データの総合的分析とは、十分な品質と信頼性を持つ公刊・公開の査読済み論文すべての結果を、各論文およびエンドポイントを通じて総合的に評価することである。ある化学物質が E, A, T 経路と相互作用する可能性があり第 2 期試験の対象候補となるか否かの決定は第一義的には第 1 期スクリーニング試験の結果に基づくことになろうが、他の学術的に有意義な情報も考慮されることがあり得る。

一般に WoE 分析では下記を考慮して異種のエビデンスを検討する。

- 1 つの試験内の、および複数の試験にまたがる、影響性質 (数、種類、程度など)
- 影響の生じた条件 (投与量、投与経路、期間など)
- 単独および複数の試験、種、系統、性別における影響のパターン、範囲、相互関係
- in vitro および in vivo 情報の長所と短所

- E, A, T 経路との相互作用の可能性の生物学的妥当性

この WoE 評価においては、単独の試験内で、および複数の試験間で、測定されたエンドポイントの一貫性、整合性、相互関係 (E, A, T 経路との特定の相互作用に関して大多数の試験が予想されるとおりの同様な傾向を示しているか、その MoA に関する生物学的知識に基づいて予想されるパターンであるか) を検討し、異種のエビデンスから得られる知見の生物学的妥当性を考慮することが重要である。内分泌経由の MoA には第 1 期試験での評価対象に含まれない各種の表現型的結果を生ずるものがあるため、毒性データベースにおける当該物質の情報、たとえば標準的な毒性試験における生殖系への影響や腫瘍応答などホルモンの影響に関する可能性のあるデータも WoE 評価に寄与することがある。したがって WoE 分析で取り上げるべき問題は、主張または仮定されている内分泌系との相互作用に対して毒性データベースが内部的に整合しているか否かである。

第 1 期試験は相互補完的なエンドポイントのデータを得るように計画されているので、表 2 に示すようにデータを表形式で示すのが便利である。これによって整合性のある影響と孤立的ないし矛盾した応答とを識別し WoE による判定を容易にすることができる。

表 2：各種エビデンスを組織化するための表形式の一例。この表は例示のためであって網羅的ではない。ここに含める試験やその組織化の方法はデータの量および試験の質に依存する。この表はある化合物とエストロゲン経路との相互作用の可能性を示す各種エビデンスを組織化するために考えられる一つの方法を示したものである。[凡例：P = 陽性応答、N = 陰性応答、E = 曖昧な応答。矢印（↑ ↓）は応答の方向を示す。ダッシュ（--）は評価を受けなかったパラメータを示す。]

エストロゲン経路との相互作用（エストロゲン拮抗作用）の可能性を示すエビデンス												
試験/ 引用 文献	ER 結合	ERT トランス 活性化	性 ステロイド	子宮 重量	卵巣 重量	卵巣の 組織病理	下垂体 重量	下垂体の 組織病理	発情周期（周期を示す 動物の日齢、体調、 %）	生殖能力	重 陰 開口時の日齢と体	ビテロゲン
試験 1	--	--	↓	↓	↓	--	N	--	--	--	--	--
試験 2	P	N	--	--	--	--	--	--	--	--	--	--
試験 3	--	--	--	--	--	--	--	--	--	--	--	--
試験 4	--	--	E	--	--	N	--	--	--	--	--	↓
試験 5	--	--	--	--	↓	P	N	N	↓	--	↓	--
試験 6	--	--	--	--	--	--	--	--	↓	--	--	--
試験 7	--	--	--	--	↓	--	N	N	E	↓	↓	--
試験 8	--	--	--	--	↓	--	N	N	--	--	--	--
試験 9	--	--	--	--	E	N	N	N	--	--	↓	--

現在の EDSP 第 1 期試験の枠内では、*in vitro* 試験の結果のみで第 2 期試験の要否を決定する十分な論拠が得られるとは期待されていない。*in vitro* 試験固有の限界を考慮すれば、異種のエビデンスの重み付けを行い、陽性と陰性の結果のバランスを検討する際には *in vivo* 試験のエビデンスが WoE 評価により大きな影響を持つと予想される。すなわち *in vitro* 試験は、MoA に関する洞察を与えることもあるが、化合物の正常な代謝的活性化、クリアランス、生理的条件(内分泌系の変化を補償する能力など)を考慮することができないなどの難点がある。異種のエビデンスにおける測定エンドポイントの相対的な感度や特異性も考慮する必要がある。与えられたデータに関して、観察された結果に対する他の説明(たとえば内分泌系以外の MoA の二次的結果、あるいは一般毒性など)を考察した上で排除することも重要である。このように WoE 評価における重要かつ有意義な考察は、異なる仮説の可能性を考慮し、それを支持するエビデンスの有無(データは真に化学物質と E, A, T 経路との相互作用を反映しているか)を検討することである。

第 1 期スクリーニング試験とその他の学術的に有意義な情報から得られる結果は、*in vitro* および *in vivo* 試験で見出される分子、細胞から組織、器官に至る生物体の様々なレベルでの事象を含んでいる。生体の異なるレベルにおけるエンドポイントの相互関係や、それらが正常な E, A, T 経路に媒介される過程に及ぼす影響は、化学物質とそれらホルモン経路の相互作用の有無を決定する重要な要因である。

E, A, T 経路との同一の相互作用を反映する、相互に関連した複数のエンドポイントの間で一貫して一致した影響を示す化合物は内分泌系機能に鉛供する可能性が高いと考えられる。反対に影響が孤立的な、あるいは矛盾するものであれば、E, A, T 経路への影響の可能性は低くなる。例えば、ある化学物質が受容体拮抗作用を介してアンドロゲン経路と相互作用することは下記のようなエビデンスによって支持される。

- *in vitro* でのアンドロゲン受容体との結合およびその裏付けとなる Hershberger 試験での知見
- その裏付けとなる、哺乳類の *in vivo* 試験におけるアンドロゲン感受性組織の測定値(思春期到達の遅れ、睾丸の組織病理、副睾丸重量の減少、テストステロンの変化など)
- 魚類の *in vivo* 試験における雄の二次性徴の弱化、睾丸の退化、雄性器重量および生殖腺指数の低下

雌の性的成熟・膣開口期の日齢および体重の対照群に比べての変化は、多くの要因の複雑な絡みによって生じ、内分泌系によってもそれ以外の MoA によっても影響されている可能性がある。しかし思春期雌性試験の種々の知見が整合的なパターンを示し(膣開口の促進、子宮重量の増加、膣における持続的な発情発現と無排卵、血中エストラジオール濃度の増加など)、

化学物質のエストロゲン受容体との相互作用（*in vitro* の ER 結合試験およびトランス活性化試験で測定される）が認められればエストロゲン性を支持するエビデンスとなる。魚類の雄のビテロゲニン産生量の増加や二次性徴の変化（魚類短期間生殖試験で測定される）のように、異なる分類群にまたがるエビデンスが得られればエストロゲン性の存在が更に強く支持されることになる。

ある化合物がステロイド産生に影響するか否かの決定に対して *in vitro* および *in vivo* の両試験のデータが必要な情報を提供することがある。*in vivo* 試験においてのみホルモン産生が、予想される組織の応答および生殖・発達過程への影響と共に見られ、*in vitro* での確認ができない場合は、化学物質が視床下部一下垂体機能を阻害するためゴナドトロピンの合成または分泌に変化をもたらすことが一つの可能性であるが、報告された知見に対して可能な別の説明を考慮した上で排除する必要がある（たとえば内分泌系以外の MoA の二次的影響または一般毒性）。以上の例は、ある物質が単独の試験および異なった種類の試験および相互作用を示すエンドポイントを通じて一貫したパターンを示す場合、WoE 分析によってその物質が第 2 期試験を要することが示される例である。

第 2 期試験の要否判定には、EDSP 第 1 期試験の結果に加えて、他の学術的に有意義な情報（たとえば 40 CFR Part 158 の試験指針による試験、公刊または公開された査読済み論文）の考慮が望ましい場合があり、もし試験内または試験間で曖昧ないし微弱な、あるいは一貫性のない関係が認められるならば特に重要性が大きくなることもある。このようなデータは、報告された知見の他の説明（たとえば内分泌系以外の MoA の二次的影響または一般毒性）を考察する際にも有用である。正常な内分泌機能への影響から予想される効果（ある種の生殖または腫瘍関連の知見）の有無も更なる洞察を与えるエビデンスとなり得る。また公刊・公開されている査読済み論文には第 1 期試験と同様のデータが含まれていれば、第 1 期試験の結果の解釈に役立つことがある。関連化合物に関する情報、コンピュータモデル（QSAR など）による予測も第 1 期の結果の解釈に有効な場合がある。代謝、トキシコキネティクス、攪乱を受けるターゲット（ER、AR、ステロイド産生酵素など）の分子的保存に関するデータも、生理学的差異および環境曝露条件（曝露経路など）の想定を考慮するならば、特定の試験されていない種（野生生物、ヒトなど）が同様に応答するか否かの評価を助けるものである。

5.3 エビデンスの重み付けの叙述/特徴づけ

個別の試験の詳細な分析や異種エビデンスの総合的分析には WoE の概略の叙述または特徴づけを併記するのが普通である。WoE 評価および判定には通常叙述が付随する(USEPA, 2005)。この叙述/特徴づけは透明性を旨とし、読者に第 2 期試験の要否の結論の背後にある推論を明確に理解させることを目的とする。物質が E, A, T 経路と相互作用するか否かの判定は、分析全体を考慮した利用可能なデータによって支持されなければならない。

叙述は一般に、主要なエビデンスとして用いた試験または効果の選択、および試験対象物質が E, A, T 経路と相互作用するか否かの結論に関わる結論を説明するものであり、本質的な内

容として最も重要で関係の深い情報と主要な結論を含む。一般に WoE 分析の要約的記述は、なかならず以下の項目を取り上げることとなる。

- 試験対象物質の E, A, T 経路への影響（知られているならば各性および問題になる種について）の主要なエビデンス（試験範囲、および見出された内分泌系に媒介される過程との相互作用に対する選択性または特異性を含む）
- 不確実性、およびそれが結論に及ぼす程度
- 結論に対して決定的とされた試験の説明、なぜ決定的であるかの理由
- 一貫性のない、または矛盾するデータの説明、そのような結果に対する説明（たとえば種、代謝、投与量、投与経路などの差）の可能性の有無
- 第 2 期試験が必要と結論したときは、利用可能なデータに基づいて別の説明（たとえば内分泌系以外の MoA の二次的効果、または一般毒性）が可能でないかどうかの議論
- 追加データ（第 1 期、第 2 期、または特別の研究）が必要であるか否か、必要ならばその理由

結論を支持する（すなわち試験対象物質が E, A, T 経路と相互作用する可能性の有無を示す）、WoE 評価によるエビデンスの全体的な強さも記述を必要とする。

5.4 EDSP 第 2 期試験への勧告

EDSP 第 2 期試験には多大の時間、労力、実験動物を含む資源を必要とすることから、第 2 期への移行の決定は、当該物質が E, A, T 経路と相互作用する可能性の有無についての、科学的に健全でロバストな WoE 評価に従わなくてはならない。

- EDSTAC は次のように述べている。「(…) スクリーニングおよび試験の戦略は、健全な決定を下すためのスクリーニングないし試験を必要最小限とすることで、決定に必要な時間を短縮しなければならない (…)」 (EDSTAC, 1998)。

EPA は一般的に、最終的エンドポイントの評価を進めて用量応答関係を確立する必要性の判断は、当該物質に関する既存のすべての知識（第 1 期試験の結果とその他の曝露・有害性等の情報）を考慮した仮説に基づくアプローチ（すなわち化学・生物学両者の理解）により、最も可能性の高い有害性とリスクを特徴づけるために *in vivo* 試験を追加する必要性の有無を決定することで行うこととしている。第 1 期試験に追加して行う試験の必要性と種類の決定に関連する一般的な検討事項のいくつかを以下に述べる。

全体的な WoE 評価において、異種のエビデンスに基づいて E, A, T 経路と相互作用する可能性がないことが示された化合物は以後試験の必要がない。反対に異種のエビデンスを通じて一貫したパターンを示して陽性と判断された化合物は第 2 期試験の対象として考慮されることになる。この考慮とは、その物質の内分泌活性から生ずるリスクを評価するために追加のデータ

が必要だとすれば、それはどのようなデータかを検討することである。化学物質が E, A, T 経路と相互作用するか否かの決定と同じく、追加データの要否に関しても複数のデータセットを考慮し WoE 分析によって決定する。第 1 期スクリーニングを超える試験の要否を考える際に取り上げるべき問題として次のようなものが考えられる。

- 第 1 期スクリーニングで見出された特定の内分泌系相互作用について、標準的な毒性試験指針による試験（生殖試験を含む）では、異なる分類群に共通な最終的エンドポイントとしてどのようなものが予想されるか。そのようなエンドポイントは適切に測定されているか。分類群による差は認められるか、あるいは予想されるか。最終的エンドポイントに関する既存のデータと第 1 期の *in vivo* 試験のデータとの組み合わせは、有害性やリスクの評価のために最終的影響を特徴づけるに十分な基礎となるか。
- 対象とする特定の内分泌系相互作用およびエンドポイントを最も有効に特徴づけるにはどうすればよいか。たとえば思春期試験において、ある物質が哺乳類の甲状腺ホルモン経路のみに影響するとのエビデンスが得られているならば、第 2 期で哺乳類二世世代生殖試験を行うことは科学的に最も有効または効率的とは必ずしも言えない。より有効かつ効率的なアプローチとしては、たとえば周産期の曝露と甲状腺軸の攪乱の容量応答関係の評価を含む特別の短期試験が考えられる。
- 対象化合物についてのリスク評価が既に利用できる場合、どの毒性を用いてリスク評価の出発点あるいはその他の局面を選択しているか。第 1 期スクリーニングの結果およびその他の情報は当該化学物質の現行のリスク評価と適合しているか。適合しているならば、追加の試験によって規制に関係するエンドポイントが影響される可能性はどの程度か。

EPA は関連するすべての危険性および曝露の情報を考慮した上で、追加の試験は不要と判定し、あるいは反対に第 2 期試験が当該物質のリスク評価に寄与すると判断することもある。当該物質の情報を全体的に見た結果、問題になる最終的エンドポイントの特徴づけや用量応答関係の確立のためには、標準あるいは指針による第 2 期試験は必ずしも最も有効な方法ではないと判断され、そのため特定の規制上の必要を満たす特別の試験を EPA が勧告することもある。

第 2 期試験が必要とされた場合 EPA は、要約的叙述および特徴づけにより、E, A, T 経路との相互作用の可能性に関する結論の基礎が説明されるだけでなく、リスク評価のために追加的データが必要である理由の考察も提供されることを期待する。第 2 期試験への最終的勧告には一般に、適切と考えられる試験、および特定の試験が必要な理由が記述される。

当該物質および特定のエンドポイントに対して個別に検証された第 2 期試験プロトコルに変更が求められる場合は、評価が可能なように説明を加えなければならない。

6 要約

要約して述べれば、EDSP 第 1 期スクリーニング試験の解釈と第 2 期試験の要否の判定は科学的エビデンスの総体に基づいて行われる。多くの場合第 1 期スクリーニング試験の結果が主要なデータ源となるであろうが、当該物質に関する他の学術的に有意義な情報（たとえば 40 CFR Part 158 指針、公刊または公開された査読済み論文）も併せて考慮される。WoE 分析は専門家の判断を尊重する。ある物質が E, A, T 経路と相互作用する可能性に関する結論の基礎は、分析結果の利用者のために、明確に記述・表現しなければならない。不確実性や不整合にも注意を促す必要がある。どのようなデータにも何らかの不確実性や不整合は存在する可能性はあるので、それらが結論にどの程度影響するかの説明を考慮することも必要であり、またエビデンスが内分泌系との相互作用またはそれが存在しないことを示すのに十分でありデータベースの不備を補うものであるか否かの説明も同様である。提出されたデータの不備が顕著であるため、第 1 期試験の繰り返し、あるいは何らかの変更を施した第 1 期試験ないしは目的を限定した実験計画が要請されることもあり得る。WoE 分析の結果、得られたすべてのエビデンスによって第 2 期試験の実施に十分な理由があると認められたならば、EPA は第 2 期試験が必要とされる根拠およびそのリスクアセスメントにおける重要性、ならびに修正された、または目的を限定されたアプローチの科学的基礎の説明を公表する。

7 文献

Bogert, C.J., Mihaich, E.M., Quill, T.F., Marty, M.S., Levine, S.L., Becker, R.A. (2011b) Evaluation of EPA's Tier 1 endocrine screening battery and recommendations for improving the interpretation of screening results. *Regulatory Toxicology and Pharmacology* 59:397-411.

Boobis, A.R., Cohen, S.M., Dellarco, V., McGregor, D., Meek, M.E., Vickers, C., Willcocks, D., Farland, W. (2006) IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Critical Reviews in Toxicology*. 36(10):781-92.

Boobis, A.R., Doe, J.E., Heinrich-Hirsch, B., Meek, M.E., Munn, S., Ruchirawat, M., Schlatter, J., Seed, J., Vickers, C. (2008) IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Critical Reviews in Toxicology*. 38(2):87-96.

EDSTAC (1998) Endocrine Disruptor Screening and Testing Advisory Committee, Final Report, Volume I-II.

<http://www.epa.gov/scipoly/oscpendo/pubs/edspoverview/finalrpt.htm>

Eldridge, J.C., Laws, S.C. (2010) The U.S. EPA's Tier I screening for endocrine disruptor compounds. In: *Endocrine Toxicology III Edition*. J.C. Eldridge and J.T. Stevens eds. Informa Health Care, N.Y. 1-26.

Hobbs, D.A., Warne, M.St.J., Markich, S.J. (2005) Evaluation of criteria used to assess the quality of aquatic toxicity data. *Integrated Environmental Assessment and Management* 1:174-180.

Klimisch, H.J., Andreae, M., Tillmann, U. (1997) A systematic approach for evaluating the quality of experimental and toxicological and ecotoxicological data. *Regulatory Toxicology and Pharmacology* 25:1-5.

National Research Council (NRC). (2007) *Toxicity Testing in the 21st Century: A Vision and a Strategy*. National Academy of Sciences, Washington, D.C.

http://www.nap.edu/catalog.php?record_id=11970#toc

OECD (2005) Data Evaluation In: Manual for Investigation of HPV Chemicals, Chapter 3.
<http://www.oecd.org/dataoecd/13/15/36045203.pdf>

OMB (2002) Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by federal agencies. (67 FR 8452).
<http://www.whitehouse.gov/omb/fedreg/reproducible2.pdf>

Rhomberg L. (2010) Hypothesis-based weight of evidence: a tool for evaluating and communicating uncertainties and inconsistencies in the large body of evidence in proposing a carcinogenic mode of action--naphthalene as an example. *Critical Reviews in Toxicology* 40:671-696.

SAB/SAP (1999) Review of the Endocrine Disruptor Screening Program by a Joint Subcommittee of the Science Advisory Board and Scientific Advisory Panel. EPA-SAB-0EC-99-013.
[http://yosemite.epa.gov/sab/sabproduct.nsf/C8ABD410E357DBCF85257193004C42C4/\\$File/ec13.pdf](http://yosemite.epa.gov/sab/sabproduct.nsf/C8ABD410E357DBCF85257193004C42C4/$File/ec13.pdf)

SAP (2008) Meeting Minutes of the FIFRA Scientific Advisory Panel to review and consider the Endocrine Disruptor Screening Program (EDSP) Proposed Tier 1 Screening Battery. SAP Minutes No. 2008-03.
<http://www.epa.gov/scipoly/sap/meetings/2008/march/minutes2008-03-25.pdf>

Schneider, K., Schwarz, M., Burkholder, I., Kopp-Schneider, A., Edler, L., Kinsner-Ovaskainen, A., Hartung, T., Hoffmann, S. (2009) ToxRTool, a new tool to assess the reliability of toxicological data. *Toxicology Letters* 189:138-144.

USEPA (1991) Guidelines for Developmental Toxicity Risk Assessment. *Federal Register* 56(234):63798-63826.
<http://www.epa.gov/raf/publications/pdfs/DEVTOX.PDF>

USEPA (1996) Guidelines for Reproductive Toxicity Risk Assessment. *Federal Register* 61(212):56274-56322.
<http://www.epa.gov/raf/publications/guidelines-reproductive-tox-risk-assessment.htm>

USEPA (1998) Guidelines for Ecological Risk Assessment, PA/630/R095/002F, Risk

Assessment Forum. <http://www.epa.gov/raf/publications/pdfs/ECOTXTBX.PDF>

USEPA (1999) Guidelines for Carcinogen Risk Assessment, Review Draft, CEA-F-0644, Office of Research and Development. <http://cfpub.epa.gov/ncea/raf/cancer.cfm>

USEPA (2002a) A Review of the Reference Dose and Reference Concentration Processes. Risk Assessment Forum, Washington DC EPA/630/P-02/002F December 2002 Final Report. <http://www.epa.gov/raf/publications/pdfs/rfd-final.pdf>

USEPA (2002b) Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency, EPA 260R-02-008, Office of Environmental Information. http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf

USEPA (2003) A summary of general assessment factors for evaluating the quality of scientific and technical information. Science Policy Council, Washington DC. EPA/100/B-03/001. <http://www.epa.gov/osa/spc/pdfs/assess2.pdf>

USEPA (2005) Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum, Washington, DC. EPA/630/P-03/001F. http://www.epa.gov/raf/publications/pdfs/CANCER_GUIDELINES_FINAL_3-25-05.PDF

USEPA (2007) Validation of Screening and Testing Assays Proposed for the EDSP. Version 5.4, July, 2007. http://www.epa.gov/scipoly/oscpendo/pubs/edsp_validation_paper_v%205.4.pdf

USEPA (2009a) EPA's Approach for Considering Other Scientifically Relevant Information (OSRI) Under the Endocrine Disruptor Screening Program. Accessed at Federal Docket Management System, EPA-HQ-OPPT-2007-1080-0032

USEPA (2009b) Exposure Factors Handbook: 2009 Update. External Review Draft. EPA/600/R-09/052A. <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=209866#Download>

USEPA (2011a) Corrections and Clarifications on Technical Aspects of the Test Guidelines for the Endocrine Disruptor Screening Program Tier 1 Assays (OCSPP Test Guideline Series 890).

<http://www.epa.gov/scipoly/oscpdo/pubs/assayvalidation/clarificationdoc.pdf>