

環境総局

「化学物質委員会」と「化学物質、農薬、バイオテクノロジーに関するワーキングパーティ」の合同会合

試験の実施と評価に関するOECDシリーズ No. 34

ハザードアセスメントのための新規又は改定試験法に対するバリデーションと国際的承認に関するガイダンス文書

Patric AMCOFF

Tel: +33 (0)1 45 24 16 19、Fax: +33 (0)1 44 30 61 80、Eメール : patric.amcoff@oecd.org

JT00188291

全文は、原書式のままOLISデータベースから入手が可能である。

環境・健康・安全関連OECD出版文書
試験の実施と評価に関するシリーズ
No. 34

ハザードアセスメントのための新規又は改定試験法に対する
バリデーションと国際的承認に関するガイダンス文書

環境総局

経済協力開発機構

パリ

2005年8月18日

Also published in the Series on Testing and Assessment:

- No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995)*
- No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*
- No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*
- No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*
- No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*
- No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*
- No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*
- No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*
- No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)*
- No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*
- No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*
- No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*
- No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries (1998)*
- No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*
- No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*
- No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

- No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*
- No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*
- No. 19, *Draft Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*
- No. 20, *Guidance Document for Neurotoxicity Testing (2004)*
- No. 21, *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals (2000)*
- No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*
- No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*
- No. 24, *Guidance Document on Acute Oral Toxicity Testing (2001)*
- No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*
- No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*
- No. 27, *Guidance Document On the Use of the Harmonised System for the Classification of Chemicals Which Are Hazardous For the Aquatic Environment (2001)*
- No. 28, *Guidance Document for the Conduct of Skin Absorption Studies (2004)*
- No. 29, *Draft Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*
- No. 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*
- No. 31, *Detailed Review Paper on Cell Transformation Assays for Detection of Chemical Carcinogens (draft)*

- No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies* (2000)
- No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures* (2001)
- No. 34, *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment* (2005)
- No. 35, *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies* (2002)
- No. 36, *Report of the OECD/UNEP Workshop on the use of Multimedia Models for estimating overall Environmental Persistence and long range Transport in the context of PBTS/POPS Assessment* (2002)
- No. 37, *Detailed Review Document on Classification Systems for Substances which Pose an Aspiration Hazard* (2002)
- No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disruptors Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay* (2003)
- No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (in preparation)*.
- No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which cause Respiratory Tract Irritation and Corrosion* (2003)
- No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in contact with Water release Toxic Gases* (2003)
- No. 42, *Guidance document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure* (2003)
- No. 43, *Guidance Document on Reproductive Toxicity Testing and Assessment (draft)*
- No. 44, *Descriptions of Selected Key Generic Terms used in Chemical Hazard/Risk Assessment* (2003)
- No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-Range Transport* (2004)

No. 46, *Detailed Review Paper on Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances (2004)*

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 49, *Report from the Expert group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs (2004)*

No. 50, *Report from the OECD/ICPS Workshop on Toxicogenomics, Kyoto, October 2004 (2005)*

© OECD 2005 全文か一部かを問わず、本資料の複製や翻訳の許可は、OECD出版局の局長まで申請のこと。OECD出版局局長2 rue André-Pascal, 75775 Paris, Cedex 16, France

OECDについて

経済協力開発機構（OECD）は、北米、欧州、太平洋岸、並びに欧州委員会に属する先進30ヶ国の代表が会合し、政策を調整かつ統一し、相互に関連のある懸案事項について協議するとともに、連携して国際問題に対処するための政府間機関である。OECDの事業の大部分は加盟国の代表で構成される200以上の専門委員会と補助グループによって行われており、多くのワークショップ等の会合は、OECDで特別な地位にあるいくつかの国や国際的な利害関連機関のオブザーバーの出席のもとで開催されている。委員会や補助グループの業務はフランスのパリに本部を置くOECD事務局が支援し、同事務局は部局及び部門に組織化されている。

環境・健康・安全部門は、「試験の実施と評価」、「優良試験所基準（GLP）と基準遵守のモニタリング」、「農薬及び殺生物剤」、「リスクマネジメント」、「バイオテクノロジーにおける規制監督の調和」、「新規の食品及び飼料の安全性」、「化学事故」、「環境汚染物質排出・移動登録制度」、「排出シナリオ文書」の9つの別冊シリーズによる無料文書を刊行している。環境、健康、安全（EHS）プログラムとEHS出版物の詳細は、OECDのWorld Wide Webのウェブサイトから入手が可能である（<http://www.oecd.org/ehs/>）。

本書は国際機関共同化学物質安全管理計画（IOMC）の枠組において作成されたものである。

国際機関共同化学物質安全管理計画（IOMC）は、化学物質の安全性領域における協力態勢を強化し、国際協調を高めることを目的とした国連環境開発会議の1992年勧告に従い、1995年に設立された。参加機関は国連食糧農業機関（FAO）、国際労働機関（ILO）、経済協力開発機構（OECD）、国連環境計画（UNEP）、国連工業開発機関（UNIDO）、国連訓練調査研究所（UNITAR）、世界保健機関（WHO）であり、世界銀行と国連開発計画（UNDP）がオブザーバーを務めている。IOMCの目的は、単独と共同のいずれで行うかに関わりなく、参加機関が遂行する政策と活動の協調を推進し、人間の健康と環境に関わる化学物質の安全な管理を実現することにある。

本出版物は、無料で電子提供している。

本書のほか、「環境・健康・安全」部門による他の多くの出版物は、OECDのWorld Wide Webのウェブサイト (www.oecd.org/ehs/) で照会するか、

以下に連絡先を挙げるOECD環境総局の環境・健康・安全部門まで問い合わせのこと。

**2 rue André-Pascal
75775 Paris Cedex 16
France**

**Fax番号 (33-1) 45 24 16 75
Eメール ehscont@oecd.org**

序 文

1. 本ガイダンス文書は、「代替毒性試験法のバリデーションと承認基準の統一化」に関する1996年のソルナワークショップの追跡的取り組みとして、1998年に作成が開始された。ソルナで合意し、ワークショップの報告書(1)にも記載された、新規及び改定試験法のバリデーションや行政上の承認に対する原則と規準は広く受け入れられたものの、未だに、これらの原則をさらに詳しく説明し、一般的なバリデーションについてさらに多くのガイダンスを提示する必要があるのではないかと考えられていた。また、ソルナワークショップ以降、バリデーション研究に対する実用的な経験が大幅に蓄積されたことから、この経験を適用したより具体的なガイダンスの提示が極めて有効であると判断されたのである。
2. 本ガイダンス文書の作成初期における重要なステップは、OECD事務局が試験法のバリデーション分野で国際的に認知されている特定数の専門家と協議したことである。これらの専門家による貢献は、実用的な助言や提言という意味でも、また、「ハザードアセスメントのための新規又は改定試験法に対するバリデーションと国際的承認」についてのガイダンス文書の最初の草案を最終的に確定するという点からも、極めて有益なものであった。ガイダンス文書の草案は、2001年後期に加盟国やその他の利害関係者に回覧され、OECDの事務局には、EC、国際動物保護委員会(ICAPO)、経済産業諮問委員会(BIAC)のほか、8つの加盟国からコメントが寄せられた。
3. 「ハザードアセスメントのための新規な、若しくは改定された試験方法に対するバリデーションと国際的承認」に関するOECDの会合は、2002年3月にスウェーデンのストックホルムで開催された。同会合の公式の目的は、行政上のハザードアセスメントを目的とした動物及び非動物による試験法のバリデーションと受け入れのための原則と過程について、実用的なガイダンスを構築し、これらに関するコンセンサスを確保することであった。この合意によるガイダンスを用いて、OECDガイダンス文書の草案を改定し、一層の向上を図るためである。ストックホルム会合では、草案ガイダンス文書をどのように改善し、拡張していったらよいかに関する40の具体的な勧告に合意した(2)。また、いくつかの全体的な提案のほか、草案ガイダンス文書の最初の回覧後に受領したコメントを考慮し、次回の改定に含めるべき編集上の訂正についても合意が得られた。
4. 会合の多くの参加者が草案文書の内容に適正に取り込まれていないと感じ、最も多くの議論を呼んだ領域のうちの二つが、データ解釈手順(DIP)とヒトデータの使用問題である。このことから、会合では、二つの個別のOECDワークショップを組織し、これらの問題をさらに詳しく調査するよう強く求める勧告を作成した。また、ガイダンス文書草案の完成をスピードアップするため、会合後、できるだけ早い時期に、修正案の作成に着手する小規模の作成グループを設置することを決定した。このグループは、EU、日本、北米の行政機関のほか、BIACとICAPOの代表で構成されるものである。2002年12月、事務局は正規の指名ラウンドを通じて作成グループを設置し、合わせて8人の専門家から支援の申し出があった。作成グループは2003年2月に一回目の電話会合を開き、修正の概略とアプローチについて協議するとともに、特に留意しなければならないいくつかの主題を特定した。これらのいわゆる「最重点問題」とは、(i) (定量的) 構造活性相関((Q)SAR)のバリデーション、(ii) データ解釈手順(DIP)、(iii) ヒト及び既存のデータの使用、(iv) バリデーション過程における異なる段階の区別、(v) 後向きバリデーションと慣例によるバリデーション、(vi) 妥当性検証済みプロトコルからのテストガイドライン(TG)の作成である。
5. 作成グループのメンバーは2003年3月にパリで会合し、ガイダンス文書草案の修正を開始した。作成グループは、ストックホルム会合や以前の草案に対する回覧によるコメントの成果を踏まえ、作業をどのように組織化したらよいかを協議し、合意した。修正した第二草案は2003年10月にTGプログラム国家調整官(ナショナルコーディネーター)会合(WNT)に提出され、再度、かなりの数のコメントが提出された。寄せられたコメントの対象や範囲を考慮し、事務局は第16回WNTに対して文書の修正をどのように継続すべきかの指針を提示するよう求めたところ、米国から、メリーランド州ベセスダで、これらの最後の論題を解決するための専門家会合を主催するという申し出があった。同会合は2004年の10月13日から15日にかけて開催され、受領したコメントに適切に対処するとともに、ガイダンス文書の草案に対し、かなり広い範囲にわたる多数の修

正を行った。

6. ストックホルム会合が提言したフォローアップ活動の一つはDIPに関するワークショップであったが、ドイツ（ドイツ動物実験代替試験法調査評価センター[ZEBET]）からの親切な申し出により、2004年7月1日から2日にかけてベルリンでワークショップが開催された。主たる目的は、バリデーシヨンの原則、並びに*in vivo*、*in vitro*、生態毒性試験方法のいずれにも適用できるDIPや予測モデル（PM）の使用について、コンセンサスの確立を試みることであった。ワークショップは用語を含むいくつかの項目について協議し、これらについて合意するとともに、ガイダンス文書第34号の草案の修正に関する詳しい指針を提示した。しかしながら、種類の異なる試験法へのDIPやPMの適用可能性については、コンセンサスに至ることができなかった。同会合については報告書が公表されており、入手が可能である（3）。
7. 1996年のソルナワークショップ以降、OECD会合に多くの専門家の参加があり、事務局では特に、次の専門家の方々が本書の展開と草案作成に積極的に関わったことに触れておきたい。
- Hans Ahr, Bayer AG, Wuppertal, Germany
 - Michael Balls, EC/ECVAM, Ispra, Italy
 - Robert Boethling, US-EPA, Washington, DC, USA
 - Dorothy Canter, US-EPA, Washington, DC, USA
 - Mark Chamberlain (BIAC) Unilever, UK
 - Alan Goldberg, CAAT, Baltimore, MD, USA
 - Petra Greiner, UBA, Germany
 - Kailash Gupta, CPSC, Bethesda, MD, USA
 - Karen Hamernik, US-EPA, Washington, DC, USA
 - David Hattan, FDA, Bethesda, MD, USA
 - Abigail Jacobs, FDA, Rockville, MD, USA
 - Manfred Liebsch, ZEBET, Berlin, Germany
 - Kimmo Louekari, Product Control Agency for Welfare and Health, Finland
 - Yasuo Ohno, NIHS, Tokyo, Japan

 - Willie Owens, (BIAC) Procter and Gamble, USA
 - Richard Phillips, Exxon/Mobile, East Millstone, NJ, USA
 - Amy Rispin, US-EPA, Washington, DC, USA
 - Andrew Rowan, Humane Society of the US, Washington, DC, USA
 - Len Schechtman, FDA, Rockville, MD, USA
 - Jerry Smrchek, US-EPA, Washington, DC, USA
 - Horst Spielmann, ZEBET, Berlin, Germany
 - Martin Stephens (ICAPO), USA
 - William Stokes, NIEHS, Research Triangle Park, NC, USA
 - Gary Timm, US-EPA, Washington, DC, USA
 - Leslie Touart, US-EPA, Washington, DC, USA
 - Neil Wilcox, formerly of FDA, Rockville, MD, USA
 - Marilyn Wind, CPSC, Bethesda, MD, USA
 - Andrew Worth, JRC-EC, Italy
 - Errol Zeiger, formerly of NIEHS, Research Triangle Park, NC, USA

OECD事務局は、これらの専門家を始め、本プロジェクトに専門的な見地からの支援と、この重要なガイダンス文書の完成になくてはならない助力をいただいた方々にお礼を申し上げる。

8. 本ガイダンス文書は、試験の実施と評価に関する同じガイダンス文書シリーズにおいて刊行された第1号ガイダンス文書（「化学物質の試験に対するOECDガイドラインの作成」（4））と連動するものと考えなければならない。

目次

序文	9
I. 序論	13
II. 試験法の定義	17
III. バリデーシヨンのためのアプローチ	20
前向きバリデーシヨン研究	20
バリデーシヨンの状況に対する後向き評価	21
バリデーシヨン状況を評価するためのモジュール方式	22
試験法の性能標準	24
追走的バリデーシヨンへの性能標準の適用	24
改定試験法への性能標準の適用	25
特許を取得している試験法のバリデーシヨン	25
試験法組合せ／試験戦略のバリデーシヨン	25
(定量的) 構造活性相関 ((Q)SARs) のバリデーシヨン	26
IV. バリデーシヨン研究の設計と実施	27
一般的考慮事項	27
バリデーシヨン研究の管理	30
統計学的専門知識	31
プレバリデーシヨンの実施	32
プロジェクト計画	34
参加施設	34
参照データ	35
被験化学物質の選択と化学物質の管理	36
被験サンプルのコード化と分配	38
試験法における判定基準とデータの解釈	39
参加施設の試験能力の監視	39
施設間における試験	39
データの収集	40
データの分析	41
報告	41
記録の保管とデータの配布	42
V. 独立したバリデーシヨン研究の評価 (第三者評価)	43
第三者評価のメカニズム	43
第三者評価委員の選定	44
第三者評価委員に課せられる責務	45
第三者評価過程	45
VI. バリデーシヨンされた試験法の国際的な行政上の承認	47
バリデーシヨン研究の成果	47
行政上の承認基準	47
プロトコルからTGまで	49
VII. 新規試験の提出—補助資料	50
提案する試験法の概要と根拠	50
試験法プロトコルの構成要素	50
提案する試験法のバリデーシヨンに使用する物質の特性付けと選択	51
提案する試験法の正確性評価に用いた <i>in vivo</i> 参照データ	52
試験法によるデータと結果	53
試験法の妥当性 (正確性)	53
試験法の信頼性 (反復性／再現性)	54
試験法におけるデータの質	54
その他の科学的報告書や精査	54
動物福祉に関わる考慮事項 (リファインメント (洗練)、削減、置換)	54
実用上の検討事項	55

参照文献	55
補足的資料	55
VIII. 参照文献	56
付属文書I 定義と用語	62
付属文書II 様々な試験法開発分野におけるバリデーション研究の例.....	69
<u>例1</u> OECDによるウキクサ (<i>Lemna</i>) 成長阻害試験、OECD TG草案の作成と共同比較試験の実施	72
<u>例2</u> オオミジンコ (<i>Daphnia magna</i>) 生殖試験の共同比較最終試験に関する報告書.....	75
<u>例3</u> 皮膚腐食性ポテンシャルの <i>in vitro</i> 試験のバリデーション、特別研究、 追走的バリデーションーヒト皮膚モデル試験	79
<u>例4</u> 局所リンパ節増殖試験 (LLNA)	86
<u>例5</u> 急性経口毒性に対する上げ下げ試験手順 (TG 425)	93
図表一覧	
表1 試験法のバリデーションにおける原則と基準.....	23
図1 試験法の開発からバリデーションまでー試験法最適化のエントリーポイント	28
図2 新規又は改定毒性試験法のバリデーションと行政上の承認における主要な要素.....	29
図3 バリデーション運営管理者及びバリデーション運営チームの役割と周辺組織関連性.....	33
表2 新規試験法及び改定試験法の行政上の承認における原則と基準.....	48
表3 試験対象化学物質の特性付け	52
表4 試験法の正確性評価	53

I. 序論

9. OECDが作成したものを含め、テストガイドライン(TG)は、政府や産業界、独立分析機関等が化学物質のハザードや安全性を確定するのに使用している。妥当性検証済み試験法を土台とするTGを適用することで、ヒトや動物の健康、あるいは環境へのハザードアセスメントにおいて、信頼性の高いデータの作成が推進されるのである。これまで、試験法のバリデーションにおける一般的原則や重要な検討事項、説明例、潜在的な課題、得られる経験の結果等の概略を述べた国際社会向けガイダンスを作成する必要性が認識されてきた。本ガイダンス文書は、OECD加盟国や利害関係機関からの助言をもとに、加盟国の代表者が草案を作成したものである。
10. 本文書の目的は、新規作成又は改定試験法のバリデーションに関わる問題について、現在のアプローチと整合するガイダンスを提供することにある。本文書では、急速に変化し、進展を遂げている領域である試験法のバリデーションの現状の概要を提示する。OECD TG開発に関する、より全体的な観点に関するガイダンスは、化学物質の試験に関するOECD TG開発のためのOECDガイダンス文書(4)に提示されている。本書に詳述するバリデーションの原則は生物学ベースの試験法を対象に書かれたものであるが、これらは他の試験法にも適用できるものである。
11. 試験法のバリデーションとは、科学的に信頼できる原理(5)(6)にもとづき、ある試験やアプローチ、方法、過程の特定目的における信頼性や妥当性を実証する過程である。信頼性とは、標準化された同じプロトコルを用いて試験を実施した場合に、施設内や施設間において時間の経過を伴う当該試験から得られる結果の再現性の程度として定義される。一方、試験法の妥当性とは、限定された条件下において、試験と標的種における本質との関係や、当該試験法が定義した目的に対して意義と有用性を持ちうるかどうかを説明するものである。簡単に言えば、その試験法が、目的の(生物学的)影響を必要に応じて正確に測定し、予測できる範囲を指している。試験法における規制上の必要性や有用性、限界は、これらに関連を持つ重要な側面である。新規作成及び改定された試験法は、信頼性と妥当性の両方を併せ持つもの、すなわち、妥当性検証(バリデーション)済みのものでなければならないのである。
12. 「ソルナ原則」とも呼ばれるバリデーションのための基本原則は、1996年にスウェーデンのソルナにおけるOECDワークショップで作成されたもので、*in vivo*、*in vitro*、あるいはヒトの健康や環境に対する影響を測定する試験のいずれであるかにかかわらず、これらの原則は、新規又は改定ハザードアセスメント試験法のバリデーションに適用されるという点での合意が得られている(1)。これらの原則とは、次のようなものである。
 - 1) 試験法の理論的根拠が作成されていること。この中には、科学的必要性や規制上の目的が明確に記載されていなければならない。
 - 2) 当該試験法によって測定される指標と、*in vivo*での生物学的影響、並びに目的の毒性との関係について考察していること。例えば代謝能など、手法における限界についても説明されていなければならない。
 - 3) 正規の詳細なプロトコルが作成され、公共の場においてこれらを容易に入手できること。この場合のプロトコルは、ユーザーがこれらを確実に守れるような詳細なレベルで作成されていることとし、データ分析や判定基準についても内容に含まれていなければならない。これらの試験法と結果は、独立した査読済み刊行物において発表されていることが望ましい。また、試験の結果には、独立機関による科学的精査が実施されていることとする。
 - 4) 施設内、施設間において、当該試験法の試験内のばらつき、反復性、再現性がすでに実証されていること。施設内、施設間のばらつきのレベルと、時間の経過によるこれらの変動を詳述したデータが提示されていなければならない。
 - 5) バイアスを除外できるように、できればコード化した一連の参照物質を用いて、試験法の性能標準が実証されていること。

- 6) 関連性のある既存の毒性データや、関連する標的種の情報などに基づき、当該試験法の性能標準が評価されていること。
 - 7) バリデーション研究において収集した全データセットを含め、当該試験法の妥当性の評価結果を裏付けるすべてのデータが入手でき、精査を行うことができること。
 - 8) 通常、これらのデータは、OECDの優良試験所基準（GLP）原則に従って入手したものでなければならない。
13. 多様な目的で開発される試験法の数や種類が増大し続けていることを考えると、バリデーションのための過程は柔軟で、適応性の高いものでなければならない。これらのバリデーション原則をどの程度考察するかは、試験法の目的と性質、提案する用途によって異なる。*in vivo*試験と*in vitro*又は*ex vivo*試験では、バリデーション原則の適用において考慮しなければならない違いがある（7）。一般に、測定される影響と目的の毒性影響の関連性が深いほど、試験法の妥当性を実証するのは容易になる。また、当該試験法が懸念される影響と同一の影響を詳細に測定又は観察するものであるほど、その試験法によって懸念される標的種への影響を正確に特性付け、モデリングできる信頼性が高くなるのである（8）。
14. 毒物学や生態毒性学における試験法の開発は、時間の経過とともに著しく進歩してきた。毒性や生態毒性に関する多くの試験法は長年にわたって日常的に使用されており、化学物質の毒性に関連のある情報が継続的に得られている。行政当局や規制対象業界は、最新のバリデーション基準を用いてこれらの試験法を正規にバリデーションするという必要条件なく、これらの試験法をハザードやリスクの評価のために受容している。OECDのTGは、主にこれらの試験法に基づいている。
15. ヒトの健康被害の可能性を検討する多くの代替試験法には、国や超国家的組織又は機関（例えば米国代替試験法評価省庁間調整組織 [ICCVAM]や欧州代替試験法評価センター[ECVAM]）によるバリデーション原則の組織的調和に先立つ共同的なバリデーション研究が実施されてきた。国家的、国際的な試験機関や行政当局は、長期間にわたって使用され、有用性が実証されているという事実に基づいて、ハザードやリスクの評価においてこれらの試験法を承認している。これらは、行政当局の意思決定において、強力な科学的基盤となっているのである。
16. ヒトの健康影響に関する多くの試験法は、現在、新たに考案された代替試験法を評価する場合の基準として適用されている。また、生態毒性学の研究領域では、数十年にわたって試験法の信頼性を確定するための（「共同比較試験」とも呼ばれる）バリデーション研究が実施されてきた。これらの試験の重点は、プロトコルを標準化し、それぞれの技術移転性と再現性を実証することであった。
17. 新たに作成された試験法については、これらが信頼できる科学に基づいて考案され、規制上の必要性に適合しているかどうかを確認するためにバリデーションが実施される。これらの新たな試験法は、例えばヒト及び生態系の健全性に関わる懸念や、動物福祉上の検討事項、データの品質保証のほか、新技術や技術の向上を取り込むことを理由に作成されたものである。物質の毒性試験やその他の影響に対する試験動物の使用に関する懸念の高まりにより、代替試験法の開発における動物使用の3R原則[置換、削減、洗練（9）]が広く支持され、遵守されるようになった。行政当局は3R原則を支持しており、結果的に、動物の使用を非動物システムに置き換え、試験における動物の使用数を減らし、試験手順を精緻化して被験動物の苦痛やストレスを軽減することを目的に代替試験法が考案されるようになったのである（10）。
18. 新規の試験法が導入され、継続的な試験開発が重視され、新規又は改定試験法を国際的に承認するための取り組みの必要性が増大し、新規試験法のバリデーションに関わる一般原則や問題点、事例などについての国際的認識を推進していくための方法が必要であることが顕在化した。

19. 本書では、バリデーション研究の実施に関するガイダンスと、その試験法が適切にバリデーションされていることの実証の仕方を提示する。バリデーション研究は、小規模の活動から大規模な多国間プログラムまで多様である。バリデーションを裏付け、効力を生じさせるための試験は、個々のOECD加盟国、試験法提案者、あるいは複数の異なる機関によりOECDの援助の下で実施される。例えば、欧州連合レベルではECVAM、米国ではICCVAM、ドイツでは「ドイツ動物実験代替試験法調査評価センター」(ZEBET)などの、代替試験法のバリデーション実施専門機関、若しくは法律によって規定されたバリデーション機関である。(脚注1)
20. 試験がバリデーションされ、その性能標準(関連性と信頼性)が提案する用途に対して許容できるものであると判断された後、当該試験をOECDのTGとして採択するよう求める勧告が行われる場合がある。OECD TGとしての採択を提案する試験法提案者は、当該試験法が、例えばソルナワークショップ(1)で詳述されたもののような、国際的に認知された原則や基準に従って適切にバリデーションされたものであることを示す証拠を提示しなければならない。これらの基準は、新規及び改定試験法のいずれにも適用される。規制上、これらの試験法が受け入れられるかどうかは、上述の原則を考慮したバリデーションの結果によって決まることになる。試験法の行政上の承認を検討する場合、次の基準も極めて重要である(1)。
- 1) 当該試験法を適用した場合、(i) 新規の試験法と既存の試験法か、若しくは(ii) 新規の試験法と標的種における影響のいずれかにおける関連性を実証するという点において、対象の指標を十分に予測するデータを提示できること。
 - 2) 当該試験法によって、既存の方法を用いて得られるデータと少なくとも同等か、できればこれらよりも有用なリスク評価のためのデータが作成されること。これにより、従前と同等若しくは従前よりも高いレベルでヒトの健康や環境を保護することができる。
 - 3) 規制に関わるプログラムや機関によって管理される化学物質の種類(例えば農薬や化粧品など)を代表する化学物質及び製品に対し、十分な試験データであること。
 - 4) 当該試験が頑健で技術移転性の高いものであり、標準化が可能であること。高度な専用の機器や材料、専門的知識が必要な場合は、技術移転性を促進するための取り組みを模索しなければならない。これは、バリデーション研究の初期段階で検討しなければならない重要な基準である。(事務局によって追加された注釈—現在のOECDの方針に従えば、試験は、単独の供給源による機器や材料が必要なものであってはならない。しかしながら、この場合、特許を取得している方法は容認されないことになる。ソルナワークショップでは特許を取得した試験法の問題は協議せず、これをより高次のOECDの政策レベルに照会することとした。)
 - 5) 試験の費用対効果が高く、適用される可能性が高いものであること。
 - 6) 現在開発されている既存の方法を基準に、新規の試験法の(科学的、倫理的、経済的)正当性が提示されていること。この点において、3Rを含む動物福祉上の検討事項が適正に考察されていないと見なされる。

¹ これらの他にも、ジョンズ・ホプキンス大学動物実験代替試験法研究センター(CAAT)や医学動物実験代替試験法基金(FRAME)、欧州動物実験代替試験法研究グループ(ERGATT)、オランダの国立代替研究センター(NCA)、スウェーデン動物福祉局(SAWA)など、3R原則を推進する研究活動に資金を供給し、またこれらの活動を企画することによって、数の軽減、洗練、置換のための代替試験法の開発と使用に大きく貢献している組織が存在する。

21. 本書に詳述するように、試験のバリデーションの過程は個別の段階に分けられるが、これらの過程は実際には、バリデーションにおける多様な要素が連続するものであると言える。過程の個々の段階は、作業における別の段階の結果を適用し、これらの結果を土台に行われるからである。バリデーションのためのアプローチが前向き、後向き、モジュール式、あるいは他の形態のいずれであるかに関わりなく(11)、バリデーション過程の目標は、OECDのバリデーション原則の各々に適正に対処しているかどうかを確認することにある。すなわち、第VII章で概略を示して説明しているように、適正なバリデーション文書の提示が不可欠なのである。
22. 本書において使用している個々の単語や用語の定義は、補遺文書Iに提示した。この補遺文書では、バリデーション機関毎に使用している用語のばらつきの例についても簡単に言及している。*in vivo*、*in vitro*による健康影響及び生態毒性試験等、多様なバリデーション研究の例については付属文書のIIを参照のこと。
23. 本書では、バリデーション過程の実施前と実施中に考慮しなければならない重要な段階や側面についても考察している。これらは、例えば次のようなものである。
- (i) 試験方法とこれらに関わる問題の定義（例えば目的や判定基準、指標、限界等）
 - (ii) 試験法を最適化できるようなプレバリデーション研究の設計と実施
 - (iii) プレバリデーション研究の結果にもとづき、試験法の関連性と信頼性に関するデータの蓄積を目的とする正規の施設間バリデーション研究の設計と実施
 - (iv) 行政当局の要件を念頭に置き、新規又は改定試験手順の提出情報に対して行われる全体的なデータ評価と、評価によるバリデーション研究の結論

また、本書では、バリデーションされる試験法に対し、独立機関による評価や第三者評価の必要性はあるか、また、どの程度までこれらを実施したらよいかについても考察するとともに、試験法バリデーションの現状についての概要を提示するものとする。

II. 試験法の定義

24. 本ガイダンス文書、並びにOECD TGプログラム全体における試験法とは、化学物質の特性から物質の有害影響に至るまで、多様な情報を得るために適用できる実験システムのことである。「試験法」という用語は、生態毒性並びにヒトの健康に関する試験のための「アッセイ」と互換的に使用できる。
25. バリデーション研究の計画作成と実施は、極めて多様な試験法と手順に関わるものであり、これらを対象として行われることになる。この場合のバリデーション研究は、単独の試験として実施される場合もあれば、試験法組合せや階層的試験システムの構成要素として実施される場合もある。この場合、例えば化学的特性や (Q) SARの手順とモデル、*in vitro*試験、動物による*in vivo*試験、生態毒性や生態学に関わる試験法、これまでに得られているヒトでのデータと試験法などについて検討しなければならない。今後、ゲノミクスやプロテオミクス、あるいはその他の新規の手法などの新たな技術が試験法に取り込まれる可能性もある。また、これらの試験法は、指標が一つか複数かという違いがあり、指標の種類も異なっている（例えば細胞毒性や組織及び器官の重量、組織病理学的グレード、血清や臨床学的分析等）。バリデーション研究では、その試験が化学的特性やヒトの健康、環境影響のいずれを対象として行われるものかにかかわらず、新規の試験や置換試験、既存の試験における指標の変更、既存の試験法の改良などに対処しなければならない。また、これらのバリデーション研究で必要になる情報の量や試験対象となる化学物質の数、試験適用の時期と範囲、参加する施設数などは、これらの要素がバリデーション過程の結果に有害な影響を及ぼさない限りにおいて変動する場合が多い。これらの理由から、ある試験法に対するバリデーション研究の計画作成と実施は、ケースバイケースで行われるものとされる。
26. 新規試験法の開発や試験法の改定における最初のステップは、その試験法を定義することである。この中には、その試験法を化学と生化学、生物学のいずれを基礎にしたもので行うかの定義も含まれている。定義の内容には、測定対象の指標など、得られる結果の妥当性に関する根拠や、例えばデータ解釈手順 (DIP) や予測モデル (PM) (3) の適用などといった、結果の解釈や使用に関わる方法の根拠又は判定基準も含まれる。また、この場合の定義は、ヒトの健康や環境に対する影響評価への結果の使用のほか、適宜、試験法組合せや階層的試験戦略に占める位置など、他のデータ又は情報要件や試験法に対する当該試験法の位置関係を明らかにするものでなければならない。
27. 上記の事項をさらに詳しく述べるため、測定するよう設計されている影響のメカニズムや種類（例えば胚性幹細胞株の分化阻害や塩基対置換による突然変異誘発等）を用いて、試験がどのような基盤に依拠するものであるかの例を説明できる (12)。試験の役割や適用の例は、次の点について詳述することが可能である。
- 予測され得る生物学的影響（マウス局所リンパ節アッセイのメカニズムベースの生物学的反応にもとづくヒト過敏症の可能性など (13)）
 - 測定され得る生物学的影響（動物における発がん性、個々の魚種における毒性、ラットに有害健康影響を生じるような内分泌系の一つ又は複数の構成要素の変化など）
 - 生物学的試験における種の選択が、例えば健全な環境集落に認められる特殊な生態学的機能や、特殊な分類群、栄養レベル、微小生息環境などの代表標本となるよう設計されている。

その試験法がどのような科学的必要性や規制上の必要性によるものであるかの例としては、化学物質が肝臓に影響する可能性や、魚類の産卵数に影響を及ぼす可能性の特定などを挙げることができる。また、その試験法が階層的試験戦略で占める位置には、例えば長期的な生殖影響試験に先立つ*in vitro*又は*in vivo*試験などといった位置づけの例がある。その試験法の適用範囲とは、例えば、化学物質のクラス、作用メカニズム、試験の信頼性を担保して適用できる反応の範囲などを指している。適用範囲はまた、その試験法によって正確に特定又は測定できる物質のクラスの制限など、試験法における既知の適用限界を示している場合もある。さらに、試験法によっては、活性物質の特定には適しているが、不活性物質の特定には不向きであるというケースもある。また、その逆も存在する。

28. 次に挙げるのは、試験法及び試験戦略の5つのカテゴリーをさらに詳しく説明したものである。厳密に言うと、試験法のカテゴリーは最初の3つだけであり、これらはすべて置換試験法や試験法組合せの一部に相当するものである。

- スクリーニング試験法とは、物質を優先付けするか、若しくは潜在的な作用機構を一般的なカテゴリー（エストロゲン受容体への*in vitro*結合など）に分類するために行われる、迅速で、多くは簡単な試験のことである。スクリーニング試験の結果は、通常、暫定的な意思決定であり、より複雑な追加試験を実施する上での優先順位付けに用いられることが多い。スクリーニング試験の結果は、それだけではリスク評価には不十分であるが、これらの結果を階層的試験方式の他の試験結果と組み合わせることにより、ハザード並びにリスク評価に呈することができる場合がある。他のあらゆる試験の場合と同様に、この試験法においても試験の適用限界を考慮に入れなければならない。例えば、活性物質のみを特定することはできるが、不活性物質は特定できないといったスクリーニング試験もあるからである。また、（例えば*in vitro*でのエストロゲン受容体結合アッセイなど）アゴニストとアンタゴニスト活性を判別できない試験もある。
- 確定試験法は、必ず追加試験を要求する訳ではなく、物質における特定のハザードを特性付けるのに十分なデータを得る試験法である（例えば二世代の生殖毒性試験など）。理想的に言えば、用量反応性や有害影響について、リスク評価やリスクマネジメントに関わる機関がハザードやリスクを軽減又は最小化するための判断を下す上で十分な独立した情報を提供し得る試験法ということである。しかしながら、実際には、入手が可能であれば、リスク評価やマネジメントには他の試験法や実験のデータが組み入れられるのが一般的であり、確定試験法による結果は単独使用可能であるとは言えないことになる。
- 補完試験法では、データセットに追加したり、他の試験法の結果の解釈を補助したりすることにより、評価過程を支援するためのデータが作成される（異なる種における発がん所見の解釈を目的とした動物発がん物質のトキシコキネティクスなど）。
- 置換試験法は、スクリーニング試験か確定試験かにかかわらず、既存の試験法に代わることを目的に設計されるものである。その試験法が置換試験とみなされるためには、すでに受け入れられている試験法を超えるメリットを提示するものでなければならない。この場合のメリットとは、例えば動物数や苦痛軽減（すなわち3R原則の考慮）、ヒトの健康や環境上のハザードに対する同等若しくはさらに正確な予測能力、既存の試験よりも優れた感度や信頼性、決定を裏付けるのに有益な情報の増加、コスト削減、適用しやすさの向上などである。また、バリデーションは、新規又は改定試験手順が置換する試験法の能力を充足するか、若しくはこれに勝っていることを明白に実証できるよう、厳正に実行しなければならない。
- 試験法組合せはいくつかの試験法で構成され、通常は同時若しくは近接した順序で行われる（遺伝毒性試験法組合せ、又は藻類、ミジンコ、魚類のデータを含む水生生物ベースのセット試験等）。組合せにおけるそれぞれの試験法は、例えば異なる指標や遺伝毒性作用機構、異なる生態系システム要素の測定など、他の試験法を補えるよう設計されている。試験法組合せを構成する試験法はバリデーションでは個々の試験法として扱われるとともに、試験法の組合せによって信頼性や関連性の高い結果が得られ、個々の試験を実施するよりも効果的であることを実証しなければならない。また、通常、組合せを構成する構成要素置換の際には、より優れた試験性能が得られることが条件となる。

29. 高度化や改善によって一層優れた試験法を作成するため、現在の試験法に対して定期的な改定が行われる場合がある。高度化の例には、試験法における感度の向上や既存の指標の拡充を目的とした新たな指標の追加などが含まれる。また、例えば実験動物数の削減や試験実施時間の短縮を目的とし、新たな方法や指標が設計される可能性もある。新規の試験法の場合と同様に、これまでとは別の作用機構や毒性を対象とした新たな指標についても、科学的に、及び規制面での高い信頼性を持って結果を解釈できるよう、バリデーションを実施しなければならない。既存の複雑な試験法に対し、新しい測定や指標を追加したことによって、論理及び解釈の両面で複雑さが増大するような場合には、バリデーションの際にこれらを適切に考慮する必要がある。

III. バリデーションのためのアプローチ

30. 特定目的に対する試験法の科学的妥当性を実証するには、試験法をバリデーションするための基準や科学的原則を充足する情報を入手し、これらの原則をどの程度まで充足しているかを評価しなければならない。これらの情報は、既存の情報源の照会（バリデーション状況の後向き評価）や新たな実験データの作成（前向きバリデーション研究）、並びに前向き及び後向きバリデーションアプローチの組合せを採用することにより入手できる。
31. バリデーション原則はすべての試験法に適用される。バリデーションの適用範囲や試験法の種類、またその方法が新規作成版又は改訂版か、あるいは従来のものかに関わりなく、科学的な厳正さは常に求められる。しかしながら、個々の目的に対して必要であると考えられる保証のレベルはまちまちであり、ケースバイケースに特定しなければならない。
32. 必要な情報の量と種類、並びに新規の試験法に適用される基準は、いくつかの要素によって決まる。これらの要素とは、例えば次のようなものである。
- 当該試験法の適用に対する規制上、科学上の根拠
 - 評価対象となる試験法の種類（既存の試験、新規の試験等）
 - 提唱されている試験方法の用途（メカニズムの補完試験、スクリーニング試験、確定試験、動物又は非動物による試験法の全面的又は部分的置換等）
 - 提唱されている試験法の適用範囲（限定的な化学物質クラスの試験、複数の化学物質クラスに対する試験等）
 - 被験種と影響が懸念される種の関係
 - 試験のメカニズム的基本理論と懸念される影響との関係
 - 該当があれば、科学界及び規制領域における当該試験法の適用履歴
33. 新規の試験法がこれまでに考察されたことのない指標を対象としている場合や、実証されている参照データが全く若しくはほとんど存在しない場合には、当該試験の予測能力を厳正に評価することは不可能であろう。しかし、入手可能なあらゆる情報を用いて、妥当性を評価しなければならない。一方、信頼性は実証されているが、その妥当性が関連する参照データとの比較によって部分的にしか評価されていないような試験（例：TG 414（14）やTG 424（15））の場合、入手可能なあらゆる関連データを正当に検討し、バリデーションの状況を定期的に見直すことが望ましい。また、既存の複雑な試験設計に新規の指標を追加することにより、試験の高度化を試みているケースでは、ある程度のバリデーションを実施して、これらの追加が論理上可能であり、既存の手順や測定の効果を下げることなく、また過剰でもないことを実証する必要がある。

前向きバリデーション研究

34. 試験法の妥当性を適正に評価するのに必要であると考えられる情報の一部又は全部が得られていない場合、前向きなアプローチによるバリデーション研究を実施して新たな実験に取り組みなければならない。作成されたデータを用いて、当該試験法の試験性能特性（妥当性や施設内及び施設間の再現性等）、メリット、並びに限界を確定する。前向きバリデーション研究が完了したら、得られた情報やデータのすべてを評価することにより、具体的な用途案に対する試験法の妥当性を判定しなければならない。
35. 前向きバリデーション研究（16）に資する前に、できればある種の情報が得られていることが望ましい。これらの情報とは、例えば次のようなものである。

- 試験法の科学的根拠と規制上の目的を明確に示す記述
- 試験法が必要とされている理由の説明。新規の試験法を開発する場合に最も明白な理由は、まだ試験が存在しない毒性学的懸念領域に対処するというものである。また、迅速性や実施のしやすさ、コスト削減、使用される動物数の削減、洗練（すなわち、動物におけるストレスや痛みの軽減）、感度や正確性の向上、個々の物質クラスの同定における向上、*in vitro*でのスクリーニング試験や*in vivo*試験の置換試験としての適用など、新規の試験法や新規の被験種に既存の試験法よりも大きなメリットがあるという点も別の理由として考えられる。
- 標準化された明確で包括的な試験法プロトコル。できれば、GLP原則に準拠し、適宜、標準操作手順書（SOP）とともに提示されていることが望ましい。このプロトコルは、試験システムの詳細、曝露条件、用量選択手順、評価対象の指標、実施する測定、必要になると考えられる専用の装置や補給品、ばらつきの尺度、結果の計算及び提示方法、陽性及び陰性対照の使用並びにその他の試験性能チェックなどを内容に含むものでなければならない。
- 試験結果の分析と解釈に適用する根拠及び基準の詳細（判定基準、若しくは該当する場合にはPM又はDIP等）
- 選択した試験法に対して予想若しくは実証される有用性と限界を裏付ける背景情報。当該試験法に関する出版物や未発表の報告の一覧を、これらの背景情報とともに提示しなければならない。

バリデーシヨンの状況に対する後向き評価

36. 試験法のバリデーシヨンの状況に対する後向き評価は、すでに実施された以前のバリデーシヨンの研究の結果など、試験法の妥当性を立証する、若しくはこれに疑問を呈する入手可能な全ての情報やデータの精査によって実施することができる。このようなケースでは、新たな実験実施は不要であるが、追加のデータ分析や評価作業が必要になる場合がある。追加バリデーシヨンを実施するかどうかは、本文書に概略を述べる原則に整合した上で、ケースバイケースで判断しなければならない。
37. 発表データが得られている試験法であれば、最初に行うべき試みは、当該試験法の試験性能に関する情報について第三者評価済みの科学文献や、信頼性の高い他の関連報告や出版物を詳細に調査することである。このような作業の実施により、文献において、前向きバリデーシヨンの過程全体若しくはその一部の要素を置換するだけの十分な情報が得られているかどうかを判断しなければならない。出版された情報の中には、刊行物のスペース上の限界やコストによって、試験のバリデーシヨンに関わる情報の記述や提示が制限されているケースもある。また、バリデーシヨンデータが、例えば技術報告書や論文の草稿、あるいは極めて入手の難しいその他の多様な未発表文書などの「灰色文献」でしか入手できない場合もある。さらに、これらの文書は翻訳を要する言語で書かれていることも多い。いずれにしても、完全な前向きバリデーシヨンの研究は行われていないが、代わりに科学文献において使用の履歴や発表データが得られているような試験法の場合、表1に明記した原則が完全に充足されていなければならない。規制への採択が提唱されている試験法の試験性能やメリット、限界を判断するには、これらの原則が必要なのである。これらの試験法のバリデーシヨンの状況に対する後向き評価を行うには、試験プロトコルを入手し、生データを独自に分析することによって、これらの重要なバリデーシヨンの特性を適正に評価できるようにしておかなければならない。このようなケースでは、追加の分析試験を実施しなくても、上述のバリデーシヨンの原則に従って収集したデータを評価し、試験の信頼性と妥当性、メリットと限界を判断できる独立した専門家委員会を組織すると有効である。ただし、提唱されている個々の用途に対し、十分なバリデーシヨンを実施するための過程の一環として、データの不足を解消しなければならないような状況が生じる可能性がある。このようなケースでは、新規の動物実験の割合を極力制限するよう努力しなければならない。また、後向きバリデーシヨンの評価には、独立した第三者評価（第V章「独立したバリデーシヨンの研究の評価－第三者評価」を参照のこと）を実施する必要がある。

38. 試験法のバリデーション状況に対し、後向き評価を実施する場合には、いくつかの特殊な問題に対処しなければならない。例えば、ある試験法プロトコルから派生した多様な形態を用いて得られた実験データを、一つの分析のためにプールしておくことができるか、あるいは、生データがない状態で試験法の妥当性を評価できるかなどといった問題を明らかにする必要がある。

バリデーション状況を評価するためのモジュール方式

39. バリデーション状況の後向き評価で十分であるか、あるいは前向きバリデーション研究が必要であるかどうかの判断は、その試験法の妥当性を実証できるような、信頼性の高い情報が得られているかどうかによって決まる。実際に、十分な情報やデータを入手するためには、後向きアプローチと前向きアプローチの組合せを適用しなければならない場合が多い。
40. 試験法のバリデーション状況を評価するのに必要な試験を記録するための一般的な概念的枠組、バリデーションのためのいわゆるモジュール方式が提唱されている（11）。このアプローチでは、ある試験法の妥当性を実証するのに必要な情報が、次に挙げる情報を提示するモジュールに組織化される。
- (i) 試験の定義（目的や必要性、科学的基礎を含む。）
 - (ii) 施設内での反復性と再現性
 - (iii) 施設間における技術移転性
 - (iv) 施設間における再現性
 - (v) 予測能力（正確性）
 - (vi) 適用範囲
 - (vii) 性能標準

この提案によれば、提案されている個々の用途において試験法のバリデーション状況を評価するには、これらの情報「モジュール」がすべて入手可能である必要がある。ただし、これらを線形や連続の形態で完成させる必要はない。

表1 試験法のバリデーションにおける原則と基準

- a) その試験法の根拠が明確に記載されていること。
 この中には、科学的基礎や規制上の目的、試験の必要性などが明確に記述されていなければならない。
- b) その試験法の指標と目的の（生物学的）現象の関係が説明されていること。
 ここでは、その試験法によって測定される影響の科学的妥当性について、メカニズム的（生物学的）若しくは観察的（相関的）関係という点から、目的の個々の影響若しくは毒性とどのような関連があるか言及していなければならない。これらの関連はメカニズム的なものである場合もあれば、相関的なものである場合もあるが、評価対象の影響や毒性に生物学的な関連性を持つ試験法が望ましい。
- c) 試験法の詳細なプロトコルが作成されていること。
 この場合のプロトコルは、十分に詳細であるとともに、（該当する場合には）試験に使用されると考えられる具体的な細胞の種類や構成、動物種など、必要材料の詳細のほか、測定の対象と方法の説明、データ分析方法の説明、データの評価における判定基準、試験性能の承認基準などが内容に含まれていなければならない。
- d) その試験法における施設内、及び施設間の再現性が実証されていること。
 施設内及び施設間における経時的な再現性とばらつきのレベルを示すデータが入手可能でなければならない。また、生物学的ばらつきがその試験法の再現性に及ぼす程度についても考察する必要がある。
- e) 試験法を適用する物質の種類を代表する参照物質の試験結果に基づいて、その試験法の試験性能が実証されていること。
 ここでは、バイアスを除外するためにコード化を適用した上で、十分な数の参照物質の試験が実施されていない（「被験サンプルのコード化と分配」に関する項を参照のこと）。
- f) 影響が懸念される種から得られている関連情報や、既存の関連毒性試験データとの関わりにおいて、その試験法の性能を評価していること。
 代替試験法の場合、提案されている代替試験法の性能を確実に分析し、これらによって置換する試験との比較を行えるだけの十分なデータが得られていなければならない。
- g) できれば、その試験法の妥当性を支持するすべてのデータが、GLPの原則に従って得られたものであることが望ましい。
 GLPに準拠せずにデータ収集を実施した点を明確に特定し、これらが試験法のバリデーション状況にどのような影響を及ぼしうるかを示さなければならない。
- h) 専門家が精査できるよう、その試験法の妥当性が評価されたことを裏付けるすべてのデータが入手可能であること。
 試験法の詳細なプロトコルは、速やかに入手できるように公開しておかなければならない。また、必要に応じて独立して精査できるよう、その試験法の妥当性を立証できるデータを系統的に整理し、すぐに使用できる状態にしておく必要がある。試験法は、独立した施設が手順を再現し、同等のデータを作成できるように詳細なレベルで記載されていなければならない。また、ベンチマークを確保し、独立した施設自身がプロトコルに正しく従っていることを確認できるよう配慮する必要がある。

試験法の性能標準

41. 性能標準の目的は、営利目的（著作権による保護、商標登録、登録等）及び非営利目的の両方における新規試験法が、具体的な試験目的に対し、十分な正確性と信頼性を持つものかどうかを確定できる根拠を伝達することにある。すでに承認されているバリデーション済みの試験法をもとにしたこれらの性能標準を適用することにより、同様の科学的原理にもとづく、あるいは同じ生物学的影響や毒性影響を測定又は予測する他の類似試験法（口語的にいわゆる「類似試験法」）の正確性や信頼性を評価することができる（17）。ただし、この場合も、第VII章に概略を述べる文書類を作成し、バリデーションや承認基準にどの程度適合しているのかを適正に説明しなければならない。また、次項に提示するような性能標準を、必要に応じて、新規試験法に対して発行されるTG中に提示する必要がある。

性能標準における3つの要素について述べる。

試験法における必須の構成要素 これらは、提案されているメカニズム的、機能的に類似の試験法のプロトコル中に含まれるバリデーション済み試験法の構造上、機能上、手順上の必須要素から構成される。これらの要素とは、例えば試験法に独自の特性、重要な手順上の細部、品質管理手段などである。試験法の必須構成要素を遵守することで、提案されている試験法が、これらに対応するバリデーション済みの試験法と同じ概念に基づくものであることの保証が容易になる。

最小限の参照物質リスト これらは、提案されているメカニズム的、機能的類似試験法の正確性と信頼性を評価するのに適用される。これらの化学物質は、すでにバリデーションされている試験法の信頼性と正確性の実証に用いられる化学物質を代表するサブセットである。可能な限りにおいて、これらの参照物質は、次の性質を持つものでなければならない。

- － バリデーション済みの試験法が測定又は予測できる反応や影響の範囲を代表的に示すもの。
- － バリデーション済みの試験法や参照試験法、あるいは目的種への試験において、ばらつきのない結果が得られているもの。
- － バリデーション済みの試験法の正確性を示すことができるもの。
- － 化学的構造や純度が明確に定義されているもの。
- － 容易に入手できるもの（すなわち、市販品として入手できるもの）。
- － 過剰な危険性や法外に高い廃棄処分コストを伴わないもの。

提案されているメカニズム的、機能的に類似の試験法の性能を評価するために、最低限、これらの参照物質は使用されるべきである。推奨する化学物質が入手できない場合には、十分な参照データが得られている他の化学物質を代わりに使用してもよい。この場合の代用化学物質は、可能な限り、元の化学物質と同じ化学物質クラスや活性を持つものであることが望ましい。十分な参照データが得られていれば、希望に応じ、他の化学物質クラスや製品クラスに相当する追加の化学物質を使用し、提案されている試験法を総合的に評価することもできる。ただし、これら追加の化学物質に、提案されている試験法の開発に使用された物質が含まれてはならないものとする。

正確性と信頼性の値 これらは、最小限の参照物質リストを用いて評価を実施した場合に、提案されている試験法が達成しなければならない同等の性能の必須要件である。

追走的バリデーションへの性能標準の適用

42. 提案されている特定の用途に試験法が有効であると判断したら、この試験法に基づいて確立した性能標準を適用することにより、構造的、機能的に類似した他の試験の妥当性を評価することが可能である。これらの類似法を評価するための標準物質を特定し、使用することで、いわゆる追走的バリデーション（18）（19）を簡易化することができる。このように迅速化したアプローチが、皮膚腐食性に関わるヒトの皮膚モデルのバリデーションに適用され、成果を上げている（20）。試験法のバリデーション状況の評価に用いられるアプローチがどのようなものであるにしろ、全ての新規又は改定試験法には、正規の第三者評価を実施しなければならない。それぞれ

の試験法はそれぞれのメリットに基づいて検討される必要があり、試験法のバリデーションに対する一般原則に適合するものでなければならない。

改定試験法への性能標準の適用

43. 新規の試験法が適用されるにつれ、試験から得られる結果の数が増大することになる。試験法の有用性が変わっていないかどうかを確認するために、これらの結果を定期的に見直すことが望ましい。また、時々、ハザードやリスク評価のための既定の試験法におけるバリデーション状況を見直し、再評価することが適切であると考えられる。これらの見直しは、（個々に再評価の必要性を示唆する新たな科学的所見に応じた）「事象主導型」の場合もあれば、（前回の見直し以降に蓄積された新たな収集情報が再評価に値するかどうかを確定するための）定期的性質を持つものである場合もある。また、試験システムの元々の特性や反応性が時間経過により変化する可能性や、試験法の結果に影響を及ぼす可能性のある手順の修正を認識しておくことは重要である。陽性、陰性、及び基準対照物質の監視により、変化の発生や影響の判断に役立てることができる。
44. 承認済み試験法に対して提案されている改善案については、評価を実施することが適切であるかも知れない。これらの変更を採用してしまう前に評価を実施することによって、提案する変更が試験の性能にどのような影響を及ぼすか、また、これらの変更がバリデーション過程の他の構成要素に対して得られている情報にどの程度影響するかを確定しなければならない。提案されている変更の数と性質、これらの変更において作成されるデータと補足書類などに応じて、これらには新規の試験に対して詳述したものと同一バリデーション過程か、若しくは、該当する場合には、すでに確立されている性能標準を用いた限定的な信頼性と妥当性の評価を実施する必要がある。この場合、適切であると考えられるバリデーション研究の範囲は、ケースバイケースに決定しなければならない。規模の大きな変更では新たに独立した第三者評価を実施するものとし、このような第三者評価の際には、試験がすでに採用されており、提案された個々の変更や強化（又は改善）は、試験性能の向上を目的としていることを考慮しなければならない。また、使用される実験動物の数や提案されている改定案の実現の可能性、費用及び時間対効果、施設的能力などについても考察する必要がある。

特許を取得している試験法のバリデーション

45. バリデーションの目的においては、特許取得済み若しくは営利目的試験法も他の方法と同じように扱うものとし、これらの方法も、提案されているそれぞれの用途に対し、科学的に有効でなければならない。現在、OECDは、特許を所有している独自の装置や過程を使用しなければならないようなTGを作成する予定はない。その一つの理由は、試験法はユーザーとなる可能性のある全ての人々が容易に活用できるものでなければならないからであり、一つの民間企業によるOECD試験法の独占を避けたいというのがもう一つの理由でもある。特許を取得している試験法をOECD TGとして採用できるようにするための一つの選択肢は、その試験法の詳細かつ一般的な記述を提示するとともに、すでにバリデーションされている同試験法の特許取得版の適切な参考文献を一連の性能標準とともに提示することである（17）。特許を取得しているかどうかにかかわらず、試験法の他の版は、すべてこれらの性能標準を満たすものでなければならない。

試験法組合せ／試験戦略のバリデーション

46. 試験法組合せや基本セット、階層的試験アプローチは、通常、同時若しくは近接した順序で行われる一連の試験である。組合せにおける個々の試験は、例えば多因子影響における他の構成要素の特定や、別の試験の確認など、他の試験を補完できるよう考慮して選択される。試験の組合せは、階層的（段階的）な試験アプローチにアレンジすることができる。階層的アプローチでは試験は連続的に適用され、連続する個々のレベルでどの試験を選択するかは、前階層の試験の結果によって決定される。これらのケースでは、それぞれの階層毎に、試験を終了するに十分な情報が得られているか、さらに追加の試験が必要かの判断がなされ、現段階では、これが、例えば生態毒性試験などにおける一般的なアプローチになっている。
47. 試験法組合せや試験戦略のバリデーションに関する包括的なガイダンスは、まだ作成されていない。

い。ただし、ここでは、いくつかの一般的な留意点を提示している。

48. 試験法組合せ（又は試験戦略）における個々の試験は、これらの試験法組合せや試験戦略における制限的な役割を考慮し、本書に詳述するバリデーション原則を用いてバリデーションしなければならない。試験法組合せを受け入れるには、主に意図する目的に対する全体的な性能をもとに、これを正当化する必要がある。試験を階層的アプローチで適用する場合、各階層における特定の試験が確認目的で使用されるか、「セーフティネット」としての機能を持つものでない限り、試験の全体的な結果は、各階層の個々の試験の強度によって決まることになる。
49. 階層的試験戦略の試験性能、すなわち、予測能力や、動物の使用の置換、削減、洗練などに発揮する能力は、皮膚腐食性（21）（22）（23）や眼刺激性（24）（25）に関わる階層的試験戦略の評価が例証しているように、既存のデータを使用する戦略から予想される結果をシミュレーションすることによって評価することができる。

（定量的）構造活性相関（(Q)SARs）のバリデーション

50. まとめて（定量的）構造活性相関（(Q)SARs）と呼ばれる構造活性相関、定量的構造活性相関は、化学構造を物理化学的特性や環境動態に関わるパラメータや生物学的活性と関連付ける理論上のモデルである。このようなモデルにより、これらの指標に対する有益な予測結果を提示し、ハザード評価のための実験データを入手する必要性を軽減できる可能性がある。
51. (Q)SARsのバリデーションに関わる個々の問題については、本書とは別に、(Q)SAR領域における近年の出版物（26）（27）（28）（29）（30）（31）（32）を考慮してOECDが現在作成している「(Q)SARのバリデーションに関するガイダンス文書」において考察するものとする。

IV. バリデーション研究の設計と実施

一般的考慮事項

52. バリデーション過程の目的は、規制に関わる状況での使用を検討している試験法の試験性能や有用性、限界を明らかにし、その試験から得られる結果をどの程度までハザード同定に適用できるかを確定し、リスク評価やその他の健康上、安全上の判断を裏付けることにある。すなわち、バリデーションは、試験法の開発とは異なる過程なのである。
53. バリデーションのための原則と基準は、すべての試験法に適用できるものである（第17項を参照）。バリデーションの適用範囲や試験法の種類に関わりなく、また、その方法が新規か改定された試験法であるかに関係なく、科学的な厳正さは常に求められることになる。しかしながら、個々の目的や試験の種類に対して適切であると考えられる保証のレベルはまちまちであり、ケースバイケースに評価しなければならない。この場合、特に、特定の種類の試験に対し、いくつかの一般的ガイダンスを提示することが可能である。
54. 図の1及び2に、バリデーション過程の鍵となる要素とバリデーションの流れ図を示す。バリデーション研究の結果を提出する前に、これらの試験結果を分析し、要約するよう、行政当局が特別要件を課している場合もある。
55. 状況によっては、研究ツールとして、あるいは製品開発の中で適用されてきた既存の試験に対し、規制下での使用が提案される場合がある。これらの試験はこれまで広範囲にわたって使用されてきたかもしれないが、その信頼性や妥当性を正規に評価できるような標準化されたプロトコルを用いてデータを組織的に蓄積していないケースがある。これらの試験のバリデーションは、得られているデータベースに応じ、二つの異なる経路で実施することができる。
 - (i) 試験の信頼性や妥当性の後向き分析を実施できるよう、得られているすべてのデータを精査する。十分かつ適切なバリデーションデータが得られていない場合も、この方法が行われる。
 - (ii) 追加の前向きバリデーション研究を実施することにより、関連性のある化学物質のデータセットを作成し、これまでに得られている既存のデータと合わせて検討する。
56. ケースによっては、両方のアプローチを組み合わせることにより、試験の妥当性の評価に必要な関連情報を提示した方がよい場合もある。
57. 試験法が規制上受け入れられるかどうかは、科学的なバリデーションの結果によって決まる。行政当局がバリデーション研究の計画作成や設計に早い段階から関与し、各当局にとって関心や懸念のある領域に対処することで、承認過程を大幅に簡易化できる。
58. 試験法に対して施設間での正規のバリデーション研究を実施する前に、提案する試験法を標準化し、初期評価を行うことが望ましい。この初期段階は多くの場合において「プレバリデーション」と呼ばれ、いくつかの理由から実施される（33）。最も重要な三つの理由は次のようなものである。
 - (a) 試験プロトコル、並びに該当する場合には標準操作手順書（SOP）を精緻化し、最適化し、標準化することにより、他の施設がこれらを容易に適用できるようにすること。
 - (b) 試験法の信頼性や妥当性に関する予備的データを作成すること。
 - (c) 試験性能の疑わしい試験法の多施設試験に費やされる不必要な資源の消費を回避すること（図1）。

図1 試験法の開発からバリデーションまで—試験法最適化のエントリーポイント

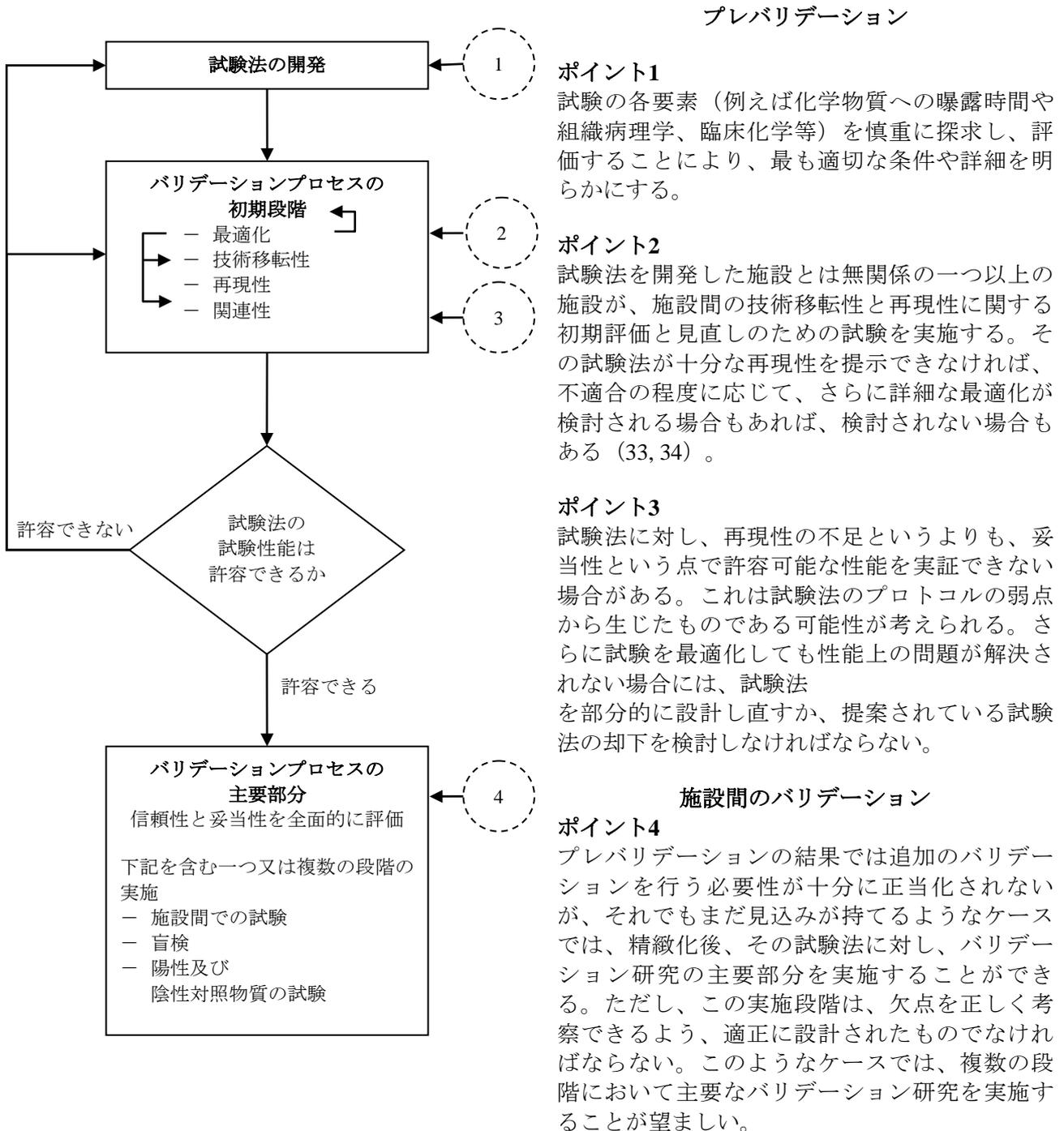
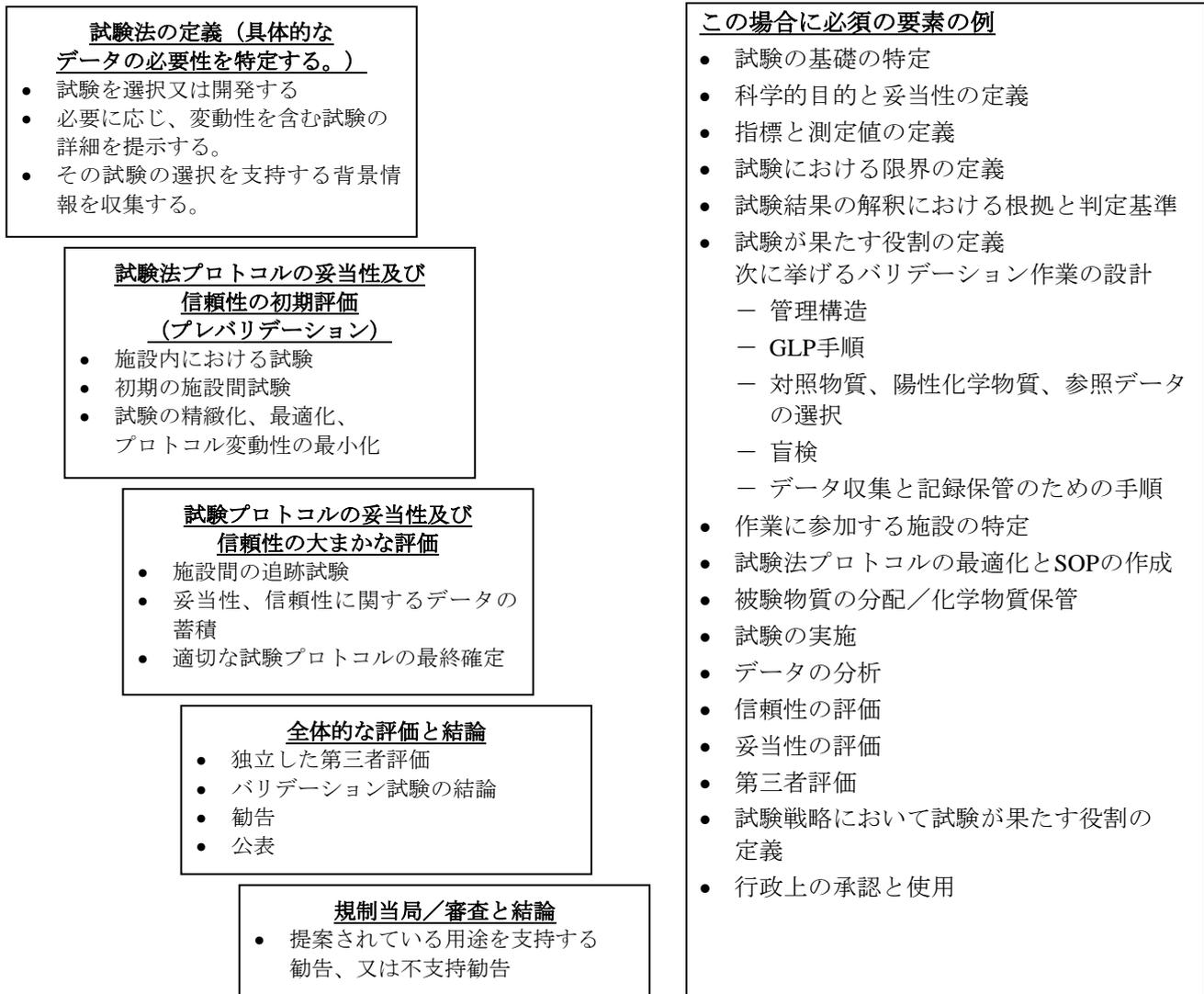


図2 新規又は改定毒性試験法のバリデーションと行政上の承認における主要な要素



59. また、これらの目的のため、プレバリデーション過程は、陽性及び陰性の対照物質における適切な許容反応範囲を確立してデータの記録と評価の基準を構築し、予備的な判定基準を確認又は精緻化するとともに、試験の技術移転性、再現性、信頼性の予備的な評価結果を確保し、試験の限界を特定するために適用される。ここでは、最後に挙げた項目に特に留意しなければならない。これまでの経験から、開発した施設や、限られた範囲の化学物質においては十分に機能していた試験が、事前に特定や定義が行われていなかった試験上の限界により、バリデーション研究では同様の成果を示さない場合があるからである。また、プレバリデーションは、いくつかの派生プロトコルの評価を行うことにより、正式な施設間のバリデーション研究に最も効果的なプロトコルを選択することや、得られている情報では明確になっていない可能性のある試験プロトコル上の側面を明確化することを理由に実施される場合もある。
60. 初期評価が完了したら、正式な施設間のバリデーション研究の計画を作成し、開始する前に、試験の結論がデータによって立証されているか、提案されているプロトコルが十分かつ適切なものであるか、などの点に関する評価を含め、得られた結果を全面的に評価しなければならない。このような精査の結果をもとに、次の点に関する判断が行われることになる。
- 再現性や正確性を最大限に高めることを目的とし、選択した試験法プロトコルやSOP（作成されている場合）、必要に応じたデータ評価若しくは判定基準を用いて、正式な施設間バリデーション研究を実施する。
 - 試験法を改善又は改定するため、さらに詳しい研究や開発を提言する。
 - 正式な施設間のバリデーション研究が行われていなくても、試験法が規制施行において許容されるよう、試験法の独立した評価を実施する。この選択肢は、プレバリデーション研究や既存データから得られた結果、また、該当する場合には試験法の作用メカニズムに関する知識が、それ以上詳しいバリデーション作業が不要であるという勧告を理論的に支持しているような状況において、選択されることになる。
 - さらにバリデーション研究に値しないものとして、その試験法を放棄する。
61. プレバリデーションが適正に完了した後、正式な施設間のバリデーション過程が開始される場合がある（脚柱 共同比較試験は、(a) 当該試験法の技術的側面を開発することを目的としてプレバリデーションにおいて実施される場合もあれば、(b) バリデーションに適用し、優先度の高い方法の性能特性を評価する場合もある）。プレバリデーションの結果を用いて施設間による正規のバリデーション研究が設計されることから、プレバリデーションが十分な計画と設計のもとに行われていれば、その後のバリデーション研究の多施設による試験段階において、コストや作業の軽減を実現できるものと考えられる。
62. バリデーション過程の結論に続き、国や地域の行政当局による承認プロセスが進む場合もある。これら両方の過程セットには、それぞれの試験が充足しなければならない原則や基準が存在する（図2）。表1に、検証済み試験法と認められるための原則と基準を示す。これらの原則や基準は多様な出典（1）（5）（6）（8）（16）（33）（35）（36）から採択したもので、試験の開発やバリデーション過程の多様な時点で必要になる情報の全体的な説明を提示している。試験法の定義からバリデーションの様々な段階、さらには最終報告書の作成に至る過程のそれぞれの段階で提示される情報は、次の段階に進むという判断を理論的に立証できるものでなければならない。
63. 試験法のバリデーションから得られるデータの管理と追跡に提唱されている一つの概念的アプローチは、**第43項**（11）で詳述したように、7つの別々の「モジュール」においてデータや情報を収集するというものである。どのようなアプローチを適用するかに関わりなく、バリデーション状況の精査を求めて提出される試験法には、**第VII章**（新規試験法の提出—補助資料類）に概略を述べる文書類を添付しなければならない。

バリデーション研究の管理

64. バリデーション研究は通常、国際団体や政府機関、代替試験法のためのバリデーション組織（ECVAM、ICCVAM、ZEBET等）、国家機関、その他の独立機関、あるいは商業上の後援組織等、スポンサーの後援のもとで行われる。スポンサーは、バリデーション運営管理者に十分に定義された役割や責任を指定することで、試験の調整を行うのが普通である（図3）。OECDの後

援のもとでバリデーション研究が実施される場合は、スポンサーはOECDの専門家グループやタスクフォース、ワーキンググループ、ワーキングパーティであることが多く、これらの組織の委員はそれぞれの加盟国政府が指名している。加盟国のほか、国家機関や国際機関のような利害関係者、産業連合会、非政府組織などがスポンサーになる場合もある。

65. バリデーション運営管理者やバリデーション運営チームは、試験の趣旨や目的が明確に定義され、プロジェクト計画書が作成され、バリデーション研究の慎重な監督過程が構築されていることを確認しなければならない。バリデーションの管理においては、その試験法や基礎となっている科学と科学的設計、バリデーション研究の管理と評価に関わる専門的知識が全体として確保されていないなければならない。
66. 試験法のバリデーション研究を実施する場合、スポンサーは、同試験の実施において遵守しなければならない利益相反防止措置を構築するよう求められるのが普通である。このような利益相反防止措置が構築されていない場合、あるいはこれらが利益相反のすべての側面に対処していない場合、その試験やバリデーション研究の結果に個人的、経済的な利益相反を持たない人が試験を監督しなければならない。また、バリデーション研究に責任を負う個人は、その試験法や基礎となっている科学、科学的設計、バリデーション研究の管理と評価に関わる専門的知識を有する者でなければならない。
67. バリデーションの管理に責任を負う個人やチームはバリデーションを監督し、参加施設による試験の調整を実施し、これらの施設からデータを受領する場合の窓口として中心的な役割を果たさなければならない。
68. バリデーション運営管理者やバリデーション運営チームは、統計学的分析や被験化学物質の選択、適用すべきGLP基準遵守の範囲に関わる決定を含めた、参加する施設の選定など、基本的な業務をタスクグループや個人に委託することができる。
69. バリデーション運営管理者は、試験に関して得られているすべての情報を精査し、これらが完全であることを確認しなければならない。図3は、バリデーション過程におけるバリデーション運営管理者やバリデーション運営チームの役割を提示するとともに、バリデーション研究の管理における本質的な業務や相互作用を要約したものである。バリデーション研究の組織や管理構造のための多様なアプローチを詳述した体系については、他の文献(5)(35)(37)(38)(39)(40)(41)で詳述されている。スポンサーは、試験法のバリデーションの際に発生したすべての情報やデータが保管されており、独立した精査に利用できるように確保しなければならない。

統計学的専門知識

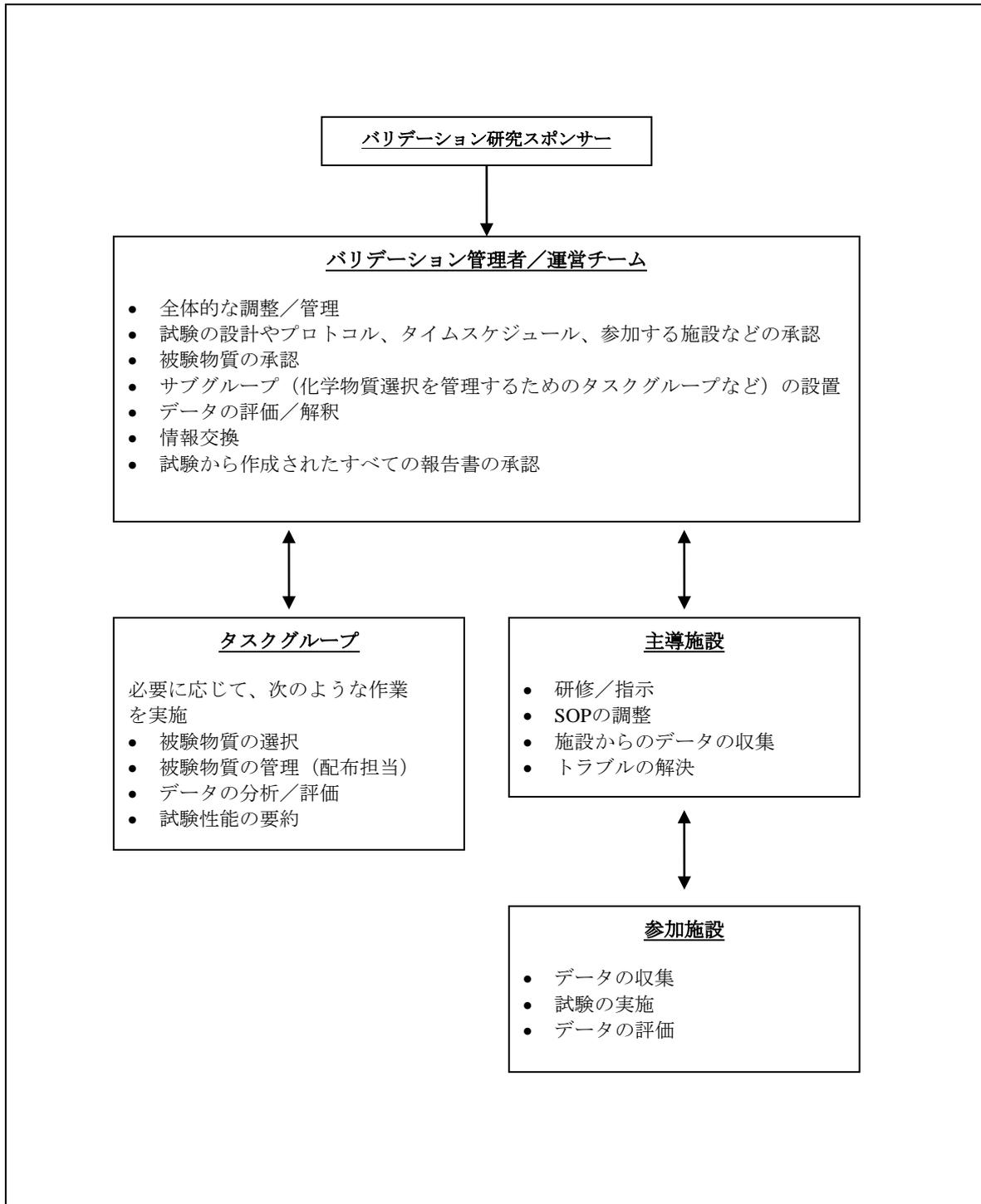
70. バリデーション研究を適切に設計し、得られたデータを評価できるだけの統計学的専門知識を確保しておかなければならない。統計学のアドバイザー(生物統計学者)をバリデーション運営管理者若しくはバリデーション運営チームの一員か顧問にし、新規又は改定された試験のすべてのバリデーション段階に関与させる必要がある。この場合における統計学アドバイザーは、提案されている試験や参照試験法の生物学的基盤や実用上の限界、あるいは最終的な対象種における指標に精通していなければならない。このような知識は、該当する統計学的手順の選択や適切な判定規準の構築を助長し、試験結果の伝達に大いに役立つものと考えられる。
71. 生物統計学者は、全バリデーション段階の設計のほか、記録の保管やデータ収集、データの提出における適切な手順の確立に関与しなければならない。また、この場合に適用される統計学的手法は、試験の開始前に定義し、提案されている試験の設計の立証を目的に適用されるものでなければならない。ただし、データの詳細が常に予想されるとは限らないため、試験データの受領後に統計学的手順を修正したり、新たな手順やモデルを特定又は開発したりしなければならない場合もある。統計学的评价は、施設内又は施設間におけるばらつきを評価することにより、試験における再現性を伝達できるものでなければならない。また、ばらつきの程度や、有効な試験として許容できると判断される合意について、試験の開始時に承認基準を提示しておく必要がある。

72. バリデーション過程の様々な段階において、判定基準を用いることにより、新規若しくは改定試験から得られた結果がどの程度参照データと一致しているかに関する統計学的評価を実施しなければならない。また、試験条件を統計学的に評価し、プロトコルの効率を改善するための勧告を提示する必要がある。このような分析は、例えば、再現性やばらつき、並びに測定対象の指標の反応性などに関わる試験の効力をグループあたりの動物数の関数として示すことにより、動物の使用を最適化し、軽減するために役立つものと考えられる。

プレバリデーションの実施

73. プレバリデーションは、その新規又は改定試験法に類似した試験法の実施に経験のある、一つ又は複数の施設によって行われることが多い。試験法の正確な設計や参加する施設の数、施設に求められる適格性、被験対象の物質数、使用される動物数などは、適宜、個々の試験法や提案されている用途、生物測定学上の考慮事項などによって異なる。
74. プレバリデーションには、通常、試験の開発に責任を負っていた人々のほかに、一つ又は複数の施設や専門家の参加が必要になる。この段階における活動の程度は、個々の試験や、その性能に関して既に得られている情報の量によって大きく異なり、それぞれ以前の作業段階を土台とする、単独若しくは複合的な実験段階を伴うことが多い。
75. プレバリデーション段階におけるその他の活動には、その試験法や必要な手法に経験のない施設への試験プロトコル技術移転性の評価のほか、この段階で生じる疑問や不整合の解決などを挙げることができる。このような過程から、最適な試験プロトコルや、必要に応じてこれらに随伴するSOPが作成されるとともに、その試験の評価や判定基準の精緻化に適用される補助的なデータが得られるものと考えられる。
76. 試験の中には、参加する施設の人員に対する研修のほか、スコア化の基準や科学的専門用語の調整を試験の必須構成要素として実施しなければならない試験もある。

図3 バリデーション管理者及びバリデーション運営チームの役割と周辺組織との関連性



プロジェクト計画

77. バリデーション研究の目的を具体的に定義し、明確かつ明白で実現可能な目標になりうるような形で提示しなければならない。プロジェクトのために設置されるバリデーション運営チームは、その責任において、バリデーションの趣旨と目的を定義する必要がある。また、プロジェクト計画は、関与するすべての当事者が実施されるバリデーション作業を明確に理解できるような形態で作成しなければならない。プロジェクト計画は、スポンサー並びにバリデーション運営グループ、主導及び参加施設の間における合意の基盤となりうるものである。プロジェクト及び適用されるプロトコルの目的や詳細を試験のすべての参加者が入手できるよう図るとともに、参加者は、事前にプロジェクト計画や試験プロトコルに綿密に従うことに合意しなければならない。
78. 本文の前の箇所に提示した、有効な試験法と認められるための原則と基準に加え、正式な施設間バリデーション研究への試験の組み入れが許容される前に、次に挙げる準備上の具体的な指示事項が明確に作成されていなければならない。
- 試験プログラムにおける当該試験の位置、そのメリットと限界、試験が可能若しくは不可能な試験物質のクラスや種類などを含む試験の定義
 - データ評価の範囲、あるいはプレバリデーションの結果にもとづく判定基準。これらは、試験によって作成されるデータのほか、予測やモデリング、測定（該当する場合に応じてPMやDIPなど）の対象となる影響に適したものでなければならない。
 - SOP（作成される場合）や実施する測定、試験から得られる具体的なデータの種類など、バリデーションの施設間試験の段階に提案されている具体的で詳細な単独又は複数のプロトコル。
 - 施設内での試験の再現性を示す証拠、並びに施設間の技術移転性に関する初期評価（プレバリデーション）の結果。
79. プロジェクト計画は、その内容に参加する施設の必須情報を含むものとし、これらの機関の役目と責任を説明するものでなければならない。また、計画作成が進行するにつれて、必要に応じた改定を行わなければならない。予想される試験の規模と複雑さに応じて、被験化学物質の選択や参照データベースの構築、主導施設の選定、他の参加施設の選定並びにこれらの機関に関する最低基準の作成、データ分析手順及び統計学的分析の性能の選択、などといった必須業務はサブグループや個々のメンバーに委託できるものとし、これらの人々はバリデーション管理者に報告を行わなければならない（図3）。

参加施設

80. バリデーション研究に参加する施設は、然るべき能力を有するスタッフの確保や施設、安全性、動物の福祉（該当する場合）、品質保証手順などに関する最低限の基準に適合しなければならない。できれば、これらの参加機関は、一般的な試験法や類似試験法に対する実施能力をすでに実証しており、少なくとも、例えば優良試験所基準（GLP）への適合性など、データ文書化手順が適正であることを示す情報を提示することが望ましい。また、経験のない施設を選択した場合、バリデーションの結果に影響を及ぼす可能性があることを考慮することが重要である。バリデーション研究は、GLP（42）（43）（44）（45）（46）（47）に従って実施し、データを記録し、保管するのが理想的である。少なくともこれらの試験は、必ずしもこれらに限定するわけではないが、SOPの使用や適正なデータの記録、記録保管などといったGLPの原則に常に従うことによって実施しなければならない。この点では、標準化されたソフトウェアプログラムを用いてデータを収集すると有効である。

81. バリデーション研究に参加するすべての施設が具体的な責務や責任を明確に理解していることが重要である。これらの責務や責任の中には、試験法プロトコルの厳正な遵守要件やデータ提出に関わる規定の計画予定表、結果の所有権の明確化、発表手順（例えば秘密情報の使用許可、共同発表又は単独での発表等の決定）などが含まれる。試験開始前に、正規の合意や契約においてこれらを明確に解決しておくこと明らかに有効である。また、国際的な試験の場合、多くの言語や法律システムのほか、異なる経済習慣や契約システムが関与する場合がある。事前に合意を得ておかないと、これらの問題はいずれもバリデーション管理を複雑にする可能性がある。
82. 試験法の妥当性を包括的に評価するのに必要な施設の最小数又は最大数は、試験の種類や考察しなければならない疑問、各機関に求められる総体的な試験の量、試験実施のコスト、動物の消費などによって異なる。多くの場合、一つの試験法毎に3又は4というのが、施設間再現性を評価するのに適正な機関数であると考えられている。ただし、十分なデータが得られていれば、もっと少ない数の施設において、予測能力（試験法の性能）の評価を実施することが可能である。参加機関数が多くなりすぎると、コストや物流管理の複雑さ、動物数が増大するだけでなく、必ずしも結果の科学的な質が向上するとは限らない。代わりに、参加する施設の数が少なすぎると、得られるデータが不足し、施設間のばらつきを正確に評価することができなくなる。バリデーション研究毎に望ましいと考えられる施設数は、生物統計学的考察と、本段落に列挙した統計学的な意味で適正な事項に基づいて決定しなければならない。
83. 主導施設は指名によって決定するものとし、この施設は当該試験法の属する学問領域の専門的知識と経験を事前に実証していなければならない。また、評価対象の試験法の実施経験に関する定評があることが条件である。バリデーション運営管理者は初期プロトコルの開発と精緻化を監督するとともに、すべての参加施設が同じ手順と基準に従って試験を実施し、データを評価できるよう、参加施設の人員に必要な応じた適切な研修を提供しなければならない（図3）。また、各施設の試験監督者は、バリデーション運営チームとの明確な意思疎通を確保する必要がある。

参照データ

84. あらゆる種類の試験に対し、妥当性の判断を行わなければならない。置換試験法の場合、もとの試験法において作成され、新規の（若しくは改定された）試験法と比較するための適正な参照データベースを確保する必要がある。この参照データベースは、予想される反応の範囲（強い、適度、微弱、陰性）や試験が適用される化学物質の範囲（適用範囲）すべてをカバーするものでなければならない。また、この種の参照データベースは、多様な範囲の生物学的、化学的活動に対する試験性能の尺度を提示できるものである必要がある。高次の*in vivo*試験では、指標の性質や毒性の直接的尺度（組織病理学的変化や器官重量の変化など）により、それほど広範囲に及ぶデータセットを確保しなくても、妥当性を評価できる。このようなデータセットは通常、反応の範囲をカバーするのに必要な最小限の数の化学物質で構成される。これらの試験の施設間バリデーション研究が必要になった場合、これらの試験は、動物の使用などといった実用上の限界を考慮に入れた上で、試験法の妥当性と信頼性の実証に必須のデータを作成できるよう設計されたものでなければならない。
85. 毒性試験の最終的な目標は、ヒトや生態系の健康と安全性を保護することにある。既存の試験と置き換えるために設計されている試験の場合、既存の試験や、可能であれば最終的な目的種から参照データを得るためのあらゆる取り組みを実施しなければならない。生態毒性試験において目的種の直接的な試験の実施は理論的には可能であるが、一般には標準的なサロゲートが使用される。これらはすでに研究室で飼育できるよう適応がなされているだけでなく、野生個体群から十分な数を確保できたとしても、対象となるすべての魚類や鳥類の試験は不可能だからである。

86. ヒトの健康に対する影響を評価する場合、ヒトから得られるデータが最も妥当性が高い。ヒトのデータは、次のようないくつかの情報源（50）から入手することができる。
- 疫学研究
 - 職業曝露
 - 事故や中毒の事例
 - 臨床試験
 - 倫理的に認可されたヒトボランティア被験者による試験

ただし、ヒトデータの使用では、対処しなければならないいくつかの問題が生じる。実際に、質の高いヒトデータの入手は難しいのが普通である。

- 後向きデータの比較に関わる特殊な課題がある。
- 目的とする影響を示さない化学物質について利用可能なヒトデータが得られていないため、その試験の偽陽性率が確定できない。
- ヒトデータの使用と入手に対処する国家的な倫理観の相違など、ヒトデータの構築に関する倫理的な課題がある。
- 化学物質への曝露に関する正確な情報が潜在的に欠落している。

結果的に、十分なヒトデータが得られていないということは、動物データが既定の参照データであることを意味している。例えば、試験法の目標がヒトへの影響の代替となる動物試験の置換であり、ヒトデータが不足しているような場合、動物種を用いて得られるデータが最も妥当性の高いものということになり、これを使用しなければならない。

87. 標準種における化学活性を分類する場合、詳細な情報によってこれを補足しなければならない。この場合、これらの基準に適合する参照物質の数が十分でない可能性が考えられる。これらのケースでは、追加の参照データ作成の必要性について、動物福祉の考察を含めた慎重な判断を行わなければならない。
88. 生態毒性領域では現在、置換試験法の問題や過程が他の領域といくらか異なる形態で扱われている。環境中には多くの多様な植物種や動物種が存在するため、試験が可能なのは各グループにおける少数の種のみであり、各生物群に対して最小限の試験を構築するための多くの研究が行われてきた。これらの初期試験をいったん構築したら、各グループの（生態学的、経済的意味における）「重要性」に応じて、同じグループのその他の種に対する調査の実施により、追加の被験種として構築できる。すなわち、（例えば魚類などの）重要な生物グループには、試験対象となりうる許容種のリストが存在することが多い。どの種を試験の対象とし、ハザード特定や評価過程のどの時点でこの試験を実施したらよいかのガイダンスは行政当局の裁量によって提示される。置換試験法は、例えば既存の甲殻類試験に置き換える場合の甲殻類など、同じ種類の生物を基盤にしたものでなければならないが、生態毒性ではこのような直接の置換はめったに行われない。

被験化学物質の選択と化学物質管理

89. バリデーション運営チームは、使用する化学物質選択のとりまとめに全体的な責任を負っている。一般には、この目的のため、化学物質選択のためのサブグループが設置されることが多い（図3）。行政当局はバリデーション対象の新規試験法の将来のユーザーになるため、これらの行政当局の専門家を被験化学物質の選択に関与させることが望ましい。生物統計学的見地からの助言を要請することにより、試験の目的を達成するのに必要な最小限の被験化学物質の数を求めなければならない。関連する化学物質グループの代表標本となりえない場合には、この数が補正されることがある。また、この数は、予測能力（試験法の性能）並びに、施設内及び施設間再現性に関して、同時評価と別個評価のいずれを優先させるかによって異なるものと考えられる。被験化学物質や安全上の関連情報を購入し、コード化し、参加施設に供給するための化学物質配布担当の設置が検討される場合もある。

90. 化学物質は、試験法の目的や、バリデーション対象となる試験法の種類に基づいて選択しなければならない(49)。すなわち、被験対象の化学物質の選択において、定義された試験目的との整合性を確保しなければならない。これは、懸念される有害影響や作用機構に妥当性を持つだけでなく、懸念対象種についても妥当性のある化学物質を選択しなければならないことを意味している。例えば、pH値が極端に高く、皮膚に腐食作用を及ぼす化学物質など、(すでに許容されている物理化学的特性や化学的構造特性などの)生物学的試験以外の尺度に基づいて毒物学的ハザードを評価できる化学物質は選択してはならない(若しくは最小限に抑制する必要がある)。化学物質は、妥当性と信頼性の高い参照データが得られるかどうかに基づいて、できれば、参照試験や懸念対象生物においてその活性の明確な評価結果を提示できるものを選択することが望ましい。バリデーション対象の試験法の性能は、これらの化学物質に対する試験の反応に基づいて判断されることになる。すなわち、試験に最終的に組み入れられる化学物質や化学物質クラスのスペクトルは、バリデーション済みの試験法が使用される化学物質と関連性を持つものでなければならないのである。これらの選択を行う場合、参照データから十分な数の陽性及び陰性対照を組み入れることが重要である。いずれの場合においても、被験化学物質の選択の基準及び理由付けは、透明性があり、正しく公表され、必要に応じて発表文献の内容に含めなければならない。
91. 被験対象の化学物質は、できる限り、入手できる最高純度のものか、組成が明らかなものを使用しなければならない。バリデーション研究のすべての参加施設は、可能な限りにおいて、同じバッチの被験化学物質を使用する必要がある。できれば、すべての被験化学物質は、参照データの作成に採用した物質と化学的組成や物理化学的特性が極力近いものであることが望ましい。ただし、参照試験法のデータを作成した際に使用したものと同一バッチや正確な配合を適用できない場合も多い。化学物質は、予想外の損失をカバーし、限定的な度合いではあるが後でフォローアップ調査が必要になった場合に対処するためのバックアップサンプルを含め、試験の際に(できれば市販の供給源から)すべての参加施設が使用できるだけの十分な量を容易に入手しなければならない。また、これらの化学物質は、少なくともバリデーション研究中、規定の保管条件の下で安定性を維持することが可能でなければならない。保管や使用、処分上の推奨と同様、既知の範囲で、安定性と溶解性、揮発性などに関する物理化学的特性を確保する必要がある。
92. 試験法の正確性や信頼性を特性づけるために適用しなければならない化学物質の数は、試験法の目的や性質によって異なる。施設間の再現性を評価する場合、被験物質のサブセットを用いて正確性を評価するのが適切であると考えられる。ただし、この場合のサブセットは、試験法が適用可能であるとされる適切な反応範囲や物理化学的特性を十分に代表できるものでなければならない。発がん性試験や神経毒性試験のように、目的の指標が直接観察される一部の*in vivo*試験のバリデーションの場合、バリデーション研究に使用される化学物質は予想される反応の範囲(例えば強い、中等度、微弱、陰性など)をカバーできるものとし、使用する物質の数や特性に対する根拠が提示されていなければならない。また、時間や資源を浪費しないよう、試験の開発やプレバリデーションの際に得られた経験を考慮に入れて被験化学物質を選択する必要がある。被験化学物質の数や種類は、次のような多くの要素によって異なる。
- バリデーション対象となる試験の種類
 - 参加施設の数
 - 予想される毒性のハザード分類
 - 調査対象の毒性影響や強度の範囲

- 対象の化学物質クラスの多様性と化学物質間における構造的関係
- 考察すべき物理的状态の範囲
- *in vivo*試験の複雑さや期間に関する実用上の限界
- 当該化学物質の妥当性や信頼性に関する参照データの有無
- 統計学的検討事項

被験サンプルのコード化と分配

93. バリデーション研究の場合、科学や技術、安全などに関わる考慮事項によって適用除外とならない限り、使用される化学物質はコード化するのが通常である。試験を実施する施設は、危険な化学物質の輸送や保管に関わる該当国の規則に準拠しているかどうかを確認する必要がある。使用される物質は独自のコードを用いてコード化され、それぞれを識別できないような方法で包装されていなければならない(51)。安全を考慮するため、また、被験物質の希釈剤や媒体の選択、取り扱い、廃棄手順を補助することを目的とし、参加施設や施設の安全担当者に各被験化学物質の物理化学的特性に関する十分な情報を提示する必要がある。施設に発送する前に、被験物質をコード化しておくことが望ましい。あまり望ましい方法ではないが、指名した試験機関の安全担当者か、バリデーション研究に関与していないその他の個人に化学物質のコードを割り当てさせるという選択肢もある。いずれのケースにおいても、参加する全施設において登録された安全担当者に、研究対象のすべての化学物質の製品安全データシート(SDS)を保管させておくことが望ましい。加盟国によっては、SDSの規定に関し、特別な規制を設けているところもある。
94. コード化した被験物質は、できれば一つの配布担当から適正な容器で輸送するものとし、該当する輸送規制に従ってラベル表示しなければならない(52)。施設には適宜、それぞれの被験化学物質についての次のような情報を提示しておく必要がある。
- a) 視覚的外観
 - b) 物理的状态
 - c) 発送する被験化学物質サンプルの重量又は体積
 - d) pHや揮発性、安定性、溶解性、化学的反応性などの物理化学的データ
 - e) 保管に関わる指示事項
95. 個々の化学物質が付加的な防護措置や管理を必要とする場合、化学物質が同定されない範囲で、物理化学的情報シートの内容に該当する情報を含めなければならない。このようなケースでは、安全な使用や取り扱い、廃棄処分を確実にするための代替手段を確立しておく必要がある。参加施設には、被験化学物質の健康被害に関する必要情報を同封した密封パッケージを提供するものとする。この中には、事故発生時の対処措置に関する明確な指示事項のほか、汚染された動物や施設備品、余剰化学物質及び作業溶液の除染又は廃棄処分に必要な情報が含まれていなければならない。特定の化学物質についての出荷の事前通知、保管要件に関する情報及び安全関連の情報が含まれていなければならない。
96. コード化した危険有害化学物質を使用するバリデーション研究では、コード化されたすべての化学物質をそのグループで最も危険かつ有害な化学物質と同じ毒性を持つものとして扱うことが望ましい。このような場合、この化学物質に関して施設に提示する健康上、安全上の情報は、そのグループのすべての化学物質に適用されるものと考えなければならない。状況によっては、このようなアプローチが適用できない場合もある(例えば、一つ又は複数の化学物質に対する緊急対応措置が、水への溶解性の著しい違いによって、最も危険かつ有害な化学物質と異なるような場合である)。このような状況における一つのアプローチは、コード化した化学物質毎に、それぞれの健康上、安全上のデータを同封し、その化学物質のコード番号を表示した十分な数の密封封筒を別個に用意するというものである。この密封封筒は、緊急事態の発生時に確保できるよう、その化学物質を取り扱うすべての作業領域に用意しておかなければならない。ただし、施設のスタッフには、このような極端な状況以外では提供されている封筒を開けたり、内容を読んだりしないよう指示しておく必要がある。参加施設は試験の終了時に(密封されたままか開封したかに関わらず)すべての封筒を返却し、開封したすべての封筒については書面での説明を提示しなければならない。ケースによっては、試験を実施する施設が健康や安全に関わる具体的な計画を作

成しななければならない場合もある。ここでは、試験の実施スタッフではなく、健康や安全を担当するスタッフが、被験物質の特徴を把握しておくことになる。いずれのケースにおいても、参加施設に提示される情報は、国や地域の法的要件に従うものでなければならない。

試験法における判定基準とデータの解釈

97. プレバリデーション研究の結果によって、試験法が十分に構築され、施設間のバリデーション研究に移行できるかどうか判断されることになる。試験結果の解釈や化学物質の分類（陽性、陰性、未確定など）、毒性値やカテゴリーの確定などに関わる根拠や判定基準を、バリデーション研究に使用するために標準化した最適化プロトコル中で明確に定義しなければならない。ここで留意しなければならないのは、通常はバリデーションに不可欠な要素である試験結果の解釈のための判定基準が、透明かつ明確でなければならないという点である。試験や、適用される試験の選別によって得られるデータの性質に応じ、異なるアプローチを用いて根拠や判定基準が定義されることが多い。*in vivo*や*ex vivo*システムに多く見られるように、対象とする毒性指標をモニタリング又は測定するようなケースでは、例えば陽性や陰性、未確定などの結果に対し、明確な判定基準を（DIPなどにおいて）定義しておかななければならない。
98. 対象とする毒性の指標を直接測定又はモニタリングできないような場合、アルゴリズムを用いることによって、試験法の適用から得られた結果を対象とする毒性影響の測定値や予測値に変換できる。アルゴリズムは、*in vitro*試験のように、値や値の範囲として表わされる結果が、対象とする毒性影響と相関関係にあるような試験に適用されることが多い。ただし、これらが生態毒性の*in vivo*試験に使用されることはめったにない。3T3 NRU光毒性試験（OECD TG432）のバリデーションでは、適正に構築された確率ベースのPMによって、処置と無処置の対照の識別が実証されている（53）（54）（55）（56）。

参加施設の試験能力の監視

99. 数多くの被験物質を使用する試験の場合、段階的な形態で試験を進行させるのが最も効果的であると考えられる。このようなアプローチを用いて、まず物質のサブセットのみを試験し、データデータセット全体の精査を含め、各施設の基準遵守のチェックを実施する。この精査が完了した後、それぞれの施設が満足のいく実績を挙げたかどうか判断されることになる。この最初のステップで規定のプロトコルやSOP（該当する場合）に準拠する手順が実行できなかつたり、適切な記録保管手順を維持できなかつたりする施設は、それ以上の参加を検討してはならない。この場合、そのような施設から得られたデータがその後のバリデーション過程に使用できるだけの頑健なものかどうかを判断しなければならない。
100. 試験実施過程の中間的時期に、バリデーション運営チームは各施設が合意された試験プロトコルや記録保管手順に従っているかどうかを判断しなければならない。このような判断は、施設が作成する基準遵守供述書によって実現できることが多い。必要であるとみなされれば、基準遵守のチェックが実施される場合もある。このようなチェックには、限られた数のサンプルに対する施設の記録の精査が含まれるものと考えられる。また、施設の作業の質（サンプルの汚染レベルや実施における過誤、データポイントの再現性の高さ）が評価される場合もある。試験プロトコルや記録保管手順からの逸脱、あるいは施設としての不適正な実施の証拠がないかどうかを特定し、これらを文書化しなければならない。問題を是正して（不適正に実施された試験サンプルに対して再試験を実施することにより）試験を継続するか、バリデーション研究から不適正なデータを削除するかを判断する必要がある。若しくはその代わりに、規定のプロトコルを用いて得られたもの以外のデータを特定はするが、その後の分析には使用せず、排除の理由を文書化してもよい。一般的な原則としては、分析や、主導施設又はバリデーション運営チームの作成する報告書には、参加する全施設から得られたすべてのデータが使用される。

施設間における試験

101. 通常、試験法のプロトコルが明確で疑問の余地がなく、試験がすべての参加機関によって適正に実施（共同比較試験については第64項を参照）されることをバリデーション運営管理者が納得す

るまで、この段階の試験に移行することはない。また、バリデーション運営チームは、データの収集や分析、報告に用いられる手順が適正であるかどうかを確認しなければならない。被験物質を規定の期限内で全施設に供給し、これらの機関がこれを受領するとともに、すべての施設が規定の期限内でそれぞれに任命された試験を完了できるよう、最終的なスケジュールを確立しておく必要がある。試験の終了時やあらかじめ合意した中間時点において、プロジェクト計画書に詳述した適切な書式を使用することにより、試験を完了したサンプルの完全なデータ記録をバリデーション運営チームやデータ分析者に提出しなければならない。

102. 試験法を開発又はバリデーションするすべての施設が、GLPに精通し、厳正なGLPへの基準遵守のもとで試験を実施できるよう組織されているとは限らない。GLPの構成要素のうち、あらゆるケースにおいて遵守しなければならないのは、すべてのプロトコルや実験に関するメモ、データの記載事項が完全かつ詳細で正確なものであり、作業を実施し、データを入力する個人の氏名や作業日時の注釈が付けられているかどうかという点と、これらに關与するすべてのスタッフに対し、適正な訓練が行われているかどうかという点である。

データの収集

103. バリデーション研究に参加する施設数が多い場合、あるいは試験データが膨大になると予想される場合、試験における施設間の調整や研修、データ収集、管理、分析などに関わる側面の実施責任を個別のタスクサブグループに委託してもよい。場合によっては、データの受領や処理、分析に独立した契約業者が指定され、透明性に関する要件に従って、主導施設としての役割を果たす場合がある（図3）。
104. バリデーション研究の計画を立案し、データ収集や評価手順を設計する場合、次に挙げる要素に対処しなければならない。
- 収集すべきデータやデータ記録手順、並びにデータ品質チェックを実施し、データ分析する手順を明確に特定しなければならない。
 - 試験法の手順毎に、適宜、各施設が実施する実験の反復回数と、一回の実験における指標の測定回数（N数）を明記しなければならない。
 - 全ての参加施設が使用するデータ記載のための標準書式を指定しなければならない。これらの形式や書式において不明瞭なところがなく、データや他の試験関連情報の報告方法について明確な指示を提示できるよう注意する必要がある。
 - もとの試験データを入手し、これらから値を導き出す際の品質保証や品質管理に関わる手順を特定しなければならない。
105. データの収集や分析、報告を実施するために、一定の具体的なコンピュータベースのシステムが用いられる。しかし、これらのシステムの使用は、試験の設計段階で、プロジェクトの一部とするよう合意を得ておかななければならない。バリデーション運営チームによって承認される試験計画は、提出すべきデータを明確に指定するとともに、要請に応じて、あらゆる個々のデータを確保できるものでなければならない。試験計画書に特に記載がない限り、データはすべて提出する必要がある。要約や変換データを提出する場合には、もとのデータを確保しておくことにより、分析が実施できるように配慮しなければならない。

データの分析

106. 適用する統計学的分析手順は、分析対象のデータの種類やデータによって明らかにしようとする疑問に妥当性を持つものでなければならない。また、バリデーション運営チームによって承認される試験計画書には、参加施設が（事前に合意した手順によって）自分達の試験データの分析を実施するのか、独立した統計学的分析が別途行われるのかを明記する必要がある。ここでは後者のアプローチが望ましいと考えられる。この場合、生物統計学者は、適切な分析を実行し、データにおける異常を認識するために、試験手順のほか、バリデーション研究の設計について十分な知識を持っていなければならない。必要に応じ、バリデーション運営チームの生物統計学者が個々の施設に直接連絡を取り、データ報告に関わる点を明らかにできるような規定を設けておくべきである。
107. 一般に、バリデーション研究からは二つの種類の情報が求められる。すなわち、試験の信頼性や再現性に関する尺度と、試験法の妥当性に関する尺度の二つである。個々の施設内、並びに異なる施設間でのばらつきの測定を含め、参加する施設間における試験の定性的、定量的再現性の評価が行われる。また、必要に応じて、再現性に影響を及ぼす要因や、ばらつきを生じる原因になるプロトコルの多様な部分に対し、評価が実施される場合もある。
108. 個々の施設のデータの分析により、試験結果の妥当性が判断される。該当する場合には、参照試験法から得られる結果の予測に、提案されている試験法がどの程度の正確性を示すかの評価が行われることになる。試験の際に実施した判定基準（PMなど）の調整は完全に文書化し、科学的に正当化しておかなければならない。バリデーション研究が完了した時点で、試験法の本来の目的である予測能力を向上させるために、判定基準を精緻化する必要があることが、データから明確に示される状況が生じる場合がある（57）。この場合における重要な考慮事項は、判定基準におけるいかなる変更も適正に文書化し、変更の影響や結果を明確に説明しておかなければならないという点である。生態毒性試験に関わる試験法の場合、判定基準はヒトの健康に関わる試験とは異なる形態で適用されることが多い。生態毒性においては、判定基準がバリデーションを構成する要素の一部に含まれず、試験結果の解釈を必要としない場合があるのである。
109. データの分析には、いくつかの異なる統計学的方法が使用されることが多い。最も適切な統計学的方法のアプローチの選択は、部分的にはあるが、データの性質やバリデーション研究のデザインに依存している。定性データや分類データは、従来、多くの場合において、試験結果の感度や特異度といった、クーパーの統計学を用いて説明されてきた（58）（59）。一方、定量データの評価には、適用が可能な多くの手順が考案されている。この場合に適用される個々の手順は、分析対象のデータの種類や、参照値若しくはその他のデータを対象にして行われる比較の結果、試験の目的に妥当性を持つものでなければならない。
110. バリデーションの施設間試験段階やデータの集計と分析が完了したら、バリデーション運営チームは試験から得られた結果に対して批評的な評価をしなければならない。参加したそれぞれの施設から得られたすべてのデータを評価し、すべての要素を考慮に入れることによって、試験の当初の目的が充足されているかどうかを判断する必要がある。この評価は、バリデーション研究のすべての側面を組織的に精査するものでなければならない。

報告

111. 研究が完了したら、すべてのデータや施設の記録を長期保管しなければならない。この場合の目標は、正当な通知があった場合に、容易に検索できる形態で長期保管されたデータや他の施設の記録を入手し、独自の評価を行えるようにしておくことである。

112. バリデーション研究の結果や最終的な結論については、研究におけるすべての参加者と協議を行わなければならない。バリデーション運営チームは、試験報告書を作成する前に、バリデーション研究の結果や結論に関わる合意を得られるよう努力する必要がある。試験報告書の草案は主導施設並びにすべての参加施設に回覧し、それぞれの精査とコメント提出が完了してから最終決定するものとする。バリデーション運営チームは受領したすべてのコメントを精査し、グループが適切であると認めた場合、草案の修正を行わなければならない。
113. バリデーション研究や得られた結果、バリデーション運営チームによる勧告について説明し、考察した試験の最終報告書は、スポンサーに提出しなければならない。最終報告書は、独立した第三者評価委員の委員会による評価が行われ（第V章「独立したバリデーション研究の評価—第三者評価」を参照）、その後行政当局に提出されて規制上の許容に関わる検討を受けることになる。独立した第三者評価の報告とバリデーション研究報告書は一般に公表し、できれば、報告書をもとに、科学文献として提出する論文を作成しておくことが望ましい。

記録の保管とデータの配布

114. 常に厳正に義務付けられるわけではないが、バリデーション研究はOECDの優良試験所基準（GLP）（43）に従って実施され、報告されることが望ましい。記録には、GLPに対する基準遵守の記述を含め、コード化した化学物質を使用したか、使用したとすればどの時点かを明記しなければならない。GLPに準拠して実施しなかったデータ収集については詳しく説明するとともに、GLPを適用しなかったことによる潜在的な影響を詳述する必要がある。
115. バリデーション対象の試験を立証するデータや、データの作成において依拠した詳細なプロトコル、並びにその他の補足文書類は、バリデーション運営チームが承認した試験計画書に詳述する要件に従い、いつでも精査できるように確保しておかなければならない。数学的に変換したデータや要約による結論が提示されている場合には、これらを元のデータと関連づけておくことが必要である。
116. できれば、試験の障害にならない限りにおいて、バリデーション研究の全体にわたり、継続的に利害関係団体に概要情報を配布しておくことが望ましい。これらの情報とは、例えば次のようなものである。
- (i) 試験実施の告示。適切であると判断された場合、試験の目的も含めるものとする。
 - (ii) 進捗状況の報告と試験による成果
 - (iii) レビューシステムを有する科学誌における試験報告書の発表

バリデーション研究と第三者評価の完了後は、試験の概要報告書や参加施設の報告書の出版を積極的に検討することが望ましい。また、スポンサーは、バリデーション研究の透明性をどのように確保するかについて検討する必要がある。いくつかのバリデーション過程のデータについてホームページ等に登録して公開し、利害関係団体がこれを精査できるようにしておくことができる。

V. 独立したバリデーション研究の評価（第三者評価）

117. 提案されている試験法が提案上の個々の用途に対して、定められているバリデーションや承認基準にどこまで適正に対処したかを判定するとともに、これらの試験法に専門家の助言を提供するため、独立した科学的評価（第三者評価）が行われる。このような独立した評価は通常、試験法の開発やバリデーションには関与しておらず、精査の結果、経済的、職業的にメリットを得る立場にない、適格な科学者の集団によって実施される。第三者評価は、提案されている用途に対する試験法の有用性と限界を批評的に評価し、行政当局による試験法の許容判断を支援するものであり、この過程は新規又は改定毒性試験法のバリデーション後に行われる重要なステップである。すべての第三者評価のすべての側面は完全に透明であるものとし、利害関係者からのコメントを考慮できるものでなければならない。

第三者評価のメカニズム

118. OECD TGの作成では、独立した試験法第三者評価に対し、次のようないくつかの選択肢がある。
- (i) 試験法の評価のため、法律によって設置された組織、若しくはこのような評価を法律によって義務付けられている組織が調整する第三者評価
 - (ii) 加盟国の政府機関が調整する第三者評価
 - (iii) OECDが調整する第三者評価
119. スポンサーや他の適切な当事者は、例えばECVAMやICCVAMなど、試験法の評価を専門にする組織に連絡を取り、独立した第三者評価を手配することができる。これらの組織は、試験法のバリデーション状態を評価するためのリソースや他の組織のためのガイダンス情報源を提供するとともに、バリデーション状況を指揮し評価できる。
120. 試験法のスポンサーにとって、バリデーションの精査過程を支援する上で入手可能なデータを綿密に文書化することが重要である。この際、表1に詳述した原則や基準を遵守し、これらの情報を第三者評価委員に提示しなければならない。すなわち、バリデーション状況を記録した提出データ一式には、次の内容が含まれる。
- 提案する試験法の科学的根拠、並びに規制上の根拠
 - 試験法プロトコルにおける本質的要素の根拠
 - 提案する試験法のバリデーションに使用した物質の完全な詳細と、これらを選択した根拠
 - 試験法の正確性評価に用いた被験物質の*in vivo*若しくはその他の適切な参照値（該当する場合）
 - 提案されている試験法と参照試験法から得られたすべてのデータと結果
 - 試験法の性能（正確性）
 - 試験法の信頼性（反復性／再現性）
 - 試験法のデータの質を評価した記述
 - 関連のあるその他の科学的報告やレビュー
 - 試験法における洗練、削減、置換の評価結果

- 試験法における有用性と限界の評価結果
 - 参照文献
 - 補助資料
121. 試験法の提出書類の取りまとめに関する実用的なガイダンスがICCVAM（60）などのバリデーション機関によって効果的に活用されている。このガイダンスは、第三者評価のための委員会が開発された試験法のバリデーション状況の評価するのに十分な情報を確保する上で役立つことが実証されている。これを他の当事者に対して拘束力のある規準と解釈してはならないが、第三者評価に対し、適正な文書類を提示するための一つの方法と考えることができる（第145～146項の「行政上の承認基準」も合わせて参照のこと）。
122. TG作成の土台であるOECD TGプログラムに試験法を提出する当事者は、第三者評価がすでに実施されたことを明確に記載するか、若しくは評価に用いられる第三者評価アプローチを提案しなければならない。第三者評価がすでに完了している場合には、結論や推奨の詳細な根拠を含め、試験法に関して提示されたすべてのコメントとともに、独立した第三者評価委員会による詳細かつ完全な報告書を提示する必要がある。独立した第三者評価がまだ行われていない試験法に基づいてOECDにTG案が提出された場合には、OECDが必要な独立した第三者評価の企画に責任を負う適切な団体について協議し、これを決定しなければならない。
123. 第三者評価過程では柔軟性が極めて重要であると考えられる。ただし、第三者評価が完全に公然かつ透明で、均衡のとれた専門的な過程を維持できるよう、十分に留意しなければならない。

第三者評価委員の選定

124. バリデーション研究の第三者評価委員を務める専門家は、技術的資格と他者からの信頼性、能力を有し、自らに割り当てられた評価や書類の作成作業を進んで履行する人物でなければならない。また、利益相反に関わる基準についても明確に詳述する必要がある。例えば、いくつかのバリデーション組織では、本人若しくはその近親者がその試験や第三者評価の成果に経済的若しくはその他の利益相反がある場合、第三者評価委員にしないよう定めている。第三者評価過程の完全性を確保するため、この過程を監督する組織は、精査の候補者となる専門家を評価するための具体的な基準を構築し、これを適用しなければならない。ここでは、次に挙げるような具体的な選定基準が考えられる。
- 専門家が関連する一つ又は複数の科学的学問分野における専門的知識を実証していること。第三者評価のための委員会には、特定されている学問分野から少なくとも一人以上の専門家を確保しなければならない。特定の必須分野については、精査のための委員会に複数の専門家を参加させると有効である。
 - また、複数の第三者評価委員において、新規及び改定毒性試験法のバリデーション研究のデザイン、実施、評価に集合的な専門的知識が得られることが重要である。また、該当する場合、動物の福祉に関わる問題や動物福祉の3R原則を理解している第三者評価委員を一人又は複数確保しておかなければならない。
 - さらに、プロセスへの事前の関わりの点でも、あるいは精査の結果に対する経済的利益などの物質的利益の点でも、第三者評価委員が完全に独立しており、これらの利益相反による影響を受けないことが極めて重要である。若しくは、結果における特定の利害は第三者評価過程の開始時に明確に宣言するものとし、委員会は全体的に均衡と独立性を確保した構成でなければならない。
125. 第三者評価委員会は、ICCVAM（60）やECVAM、米国科学アカデミーが実施しているように、OECD内でバリデーションプロジェクト毎にケースバイケースで設置される場合もあれば、既に設置されている独立専門家委員会に適切な専門的知識を持つ人を補うこともできる。第三者評価委員を指名又は選定するための機構は、精査の主権に責任を負う組織によって異なる。どのような機構を適用するかは、ケースバイケースに決定される。いずれの場合も、指名する適格な専門

家は、科学や利害関係のある該当領域から広く募らなければならない。委員会の全体的な構成が均衡のとれたものであり、第三者評価過程が公然かつ透明であることが極めて重要である。バリデーションに関与した利害関係者や専門家が第三者評価委員会の一員であるような場合には、この点が特に重要になる。バリデーションされた試験法を広く国際的に許容するためには、バリデーション過程に対する独立した評価の手配に従事する国や組織が、第三者評価委員会を国際的な構成にするよう考慮しなければならない。

第三者評価委員に課せられる責務

126. 第三者評価過程を簡易化するため、第三者評価委員会には評価対象の試験法の性質に相応しい、科学的方向を指向した質問の一覧を提示しておくことができる。ただし、このリストによって第三者評価の適用範囲が制限されるようなことがあってはならない。これらの質問は第三者評価委員会による評価のテンプレートの役割を果たすものと考えられる。これらの質問は、例えば試験法のデータはGLPのための適切な国際基準に準拠するよう設計され、実施された試験によって収集したものか、などといったように、標準的な性質のものであることが多い。その他に、提案されている試験法の目的の達成という点から考えて、プロトコルの構成要素が適切であるかどうかを評価するため、より具体的な性質の質問が行われる場合もある。第三者評価過程を監督する組織は、第三者評価を行う人々に対し、委員会がすべての基準を評価できるような質問を作成する責任を負うものとする。
127. 第三者評価委員に課せられる責務は、提案されている個々の用途に関し、試験法の重要な側面における有用性や限界についての新たな見解を引き出すような、探求的で洞察に満ちた質問を提起するものでなければならない。評価委員は、すべてのバリデーション基準をどの程度考察したかに関する包括的な精査を実施するとともに、提案された標準化プロトコルや判定基準に関する見解を提示する必要がある。委員会は試験法が適用される可能性のある化学物質や製品のほか、試験法が適正に評価されなかった、若しくは有用であることが実証されなかった化学物質や製品の物理化学的特性を特定しなければならない。最後に、委員会は、本来の用途における試験法の性能（すなわち、正確性と信頼性）の適格性について見解を提示する責務を負うものとする。精査は完全かつ客観的であり、信頼性の高いものでなければならない。

第三者評価過程

128. 前述のとおり、第三者評価の最初のステップは、完全な試験法の提出文書パッケージを第三者評価委員に提示することである。このパッケージには、責務やバリデーション研究の報告書、第三者評価委員によって評価されるその他の文書のほか、責務に応じて第三者評価委員が検討すべき他の関連文書が含まれるものと考えられる。
129. 可能な限りにおいて、第三者評価委員会が精査するすべての資料は、該当する利害関係者に公表しなければならない。この際、利害関係者によるコメントは、委員が検討できるよう、第三者評価委員会に提示しなければならない。加盟国によっては、このコメント募集過程に一般の人々を組み入れるよう定めた要件を課しているところもある。理想的には、第三者評価過程は公開で透明なものであることが望ましい。一部のバリデーション機関は、第三者評価委員会を公開の会合で招集することで、この問題に対処している。特に国際的な精査に関しては、現実的な対応により第三者評価委員会による最も適切な精査プロセスが決定されるはずである。第三者評価過程の終了時には、同委員会の最終報告書に試験法の有用性と限界の全体的な評価結果を含める必要がある。この最終的な評価は、委員会に提起された質問に基づいて実施されることになる。
130. 第三者評価では、必ずしも、すべての問題に関する完全な合意が得られなくてもよい。ただし、異なる見解の詳細とそれぞれを裏付ける根拠を提示しなければならない。委員会によって試験法のバリデーション状況に関する最終勧告が作成されたのち、第三者評価委員会による討議の結果は書面において公表されるものとする。第三者評価報告書は、提案する試験法の有用性と限界のみを考察したものではなく、同試験法に関する今後の活動についての勧告を提示するものでなければならない。

131. 試験法のバリデーションスポンサーは、第三者評価委員会による勧告の承認や、試験法が本来の用途に対してバリデーションされ、行政上の承認を推奨すべきかどうかを判断しなければならない。また、スポンサーは、例えば第三者評価委員会の個々の委員による見解にどのように対処するか、実施状況が不適正であった施設の結果をどのように扱うか、などといった第三者評価時における重大問題への対処に責任を負うものとする。
132. できれば、第三者評価報告書は一般に広く配布できるような形で公表し、報告書や報告書の概要についてレビューシステムを有する科学刊行誌に発表することが望ましい。最終的には、それぞれの行政当局が、バリデーションされた試験法を目的に適っているものとして許容するかどうかを判断しなければならない。
133. 中には、以前に精査の行われた試験法が詳細な改定や開発過程を経て提出され、新たな、若しくはこれらの過程に続くバリデーション状況の精査を受ける事例もある。このような追加のバリデーション審査に取り組むレベルは、試験法におけるプロトコルの構成要素に加えられた変更の程度や重要性に釣り合ったものでなければならない。

VI. バリデーションされた試験法の国際的な行政上の承認

134. 資源や時間を節減するため、試験法は、バリデーション研究や第三者評価を実施してからOECDのTGとして考察することができる。ただし、これは、今のところOECDのTG作成過程の要件にはなっていない。
135. 研究開発に関わる状況においては、通常の研究過程において毒性試験手順が実施されるため、手順を実施し、データを使用する科学者は、データに基づいて判断を行うための手順に十分な知識と経験を持っていることになる。バリデーションされた試験法は、試験結果が意思決定過程を指示するために使用される規制状況において、毒性影響を示す信頼性の高い指標である。
136. これまで、行政上の承認過程はケースバイケースに行われるのが一般的であり、今日、正規のバリデーションであると考えられている過程が適用されていない試験法（例えば他の試験から得られた結果を補強し、説明するのに役立つメカニズム的研究に用いられる方法など）を用いて得られた結果を承認するかどうかは、行政当局の裁量で決定されている。しかしながら、個々の行政当局が試験法を承認したからと言って、必ずしも他の行政当局によって普遍的に受け入れられるとは限らない。承認方針は国によって異なるものであり、時には同じ国であっても行政当局によって相違がある場合がある。
137. 本ガイダンス文書に詳述するバリデーションは、提案されている試験法の国際的な承認に大いに貢献するだけでなく、国際規模における「データの相互受け入れ（MAD）」（61）を推奨し、支援するものである。ただし、行政当局がすでに実証された信頼性や妥当性以上に、その試験法に対して疑問を持ち、これが行政上の承認に影響を及ぼす場合がある。このような場合、本文書に提示したガイダンスを考慮することにより、適正にバリデーションされた試験法の国際的な規制への受け入れを調和できる。ここでは、一般的なバリデーション過程が実施されていない試験の行政上の承認は控えらる。バリデーションの必要性がない、若しくは実施が適切でないと判断されるケースでは、これらを書面で正当化しなければならない。

バリデーション研究の成果

138. 適正なバリデーション研究によって試験法の有用性や適用可能性が明示されると、行政当局による採択の検討を行うことができる。バリデーションされた試験法がOECDに提出され、OECDのTGとして正規に採択される場合もある。

行政上の承認基準

139. 試験法に正規のバリデーションが実施され、提案されている特定の用途に対して承認できると判断された場合、OECDのTGとしての採択に関する勧告が行われる。本書の前の箇所に述べたように、バリデーション過程の極力早い段階で、バリデーションされた試験法から得られた試験結果の提出先の行政当局が関与した場合、行政上の承認は大いに迅速化されるものと考えられる。新規試験法の行政上の承認について検討する場合には、表2に詳述した原則や基準が極めて重要である。

表2 新規試験法及び改定試験法の行政上の承認における原則と基準

- a) 提出される試験法や、これを裏付けるバリデーショndataには、独立した透明性のある第三者評価を実施しなければならない。
第三者評価は、試験の考案やバリデーションに実質的に関与せず、試験法やデータ評価に十分な知識を持ち、精査の結果に物質的、経済的な影響を受けない、その分野の専門家が実施しなければならない。第三者評価委員は、雇用者か組織の所属者かに関わりなく、科学的適格性や専門的知識、経験に基づいて選定する必要がある。
- b) 試験法によって作成されるデータは、対象となる指標を適正に測定若しくは予測できるものでなければならない。代替試験法の場合、これらのデータによって、提案されている試験法と既存の試験法、あるいは提案されている試験法と標的種若しくはモデル種における影響、若しくはこれらの両方の関係性が明らかにならなければならない。
- c) 試験法は、ハザード/リスク評価のために有益なデータを作成するものでなければならない。
新規の試験法は、認識されているデータの不足を補うものであることが多い。試験法は単独で有用である場合もあれば、試験法組合せや階層的試験アプローチの一部として有効性を発揮する場合もある。代替試験法の場合、少なくとも既存の試験法から得られるものと同等か、できればこれらよりも優れたデータを得ることが可能でなければならない。
- d) 提出される試験法や、これらを裏付けるバリデーショndataは、試験法を提案する規制上のプログラムや行政当局によって管理される代表的な化学物質や製品のスペクトルを適正にカバーできるものでなければならない。また、試験法の適用可能性と限界を明確に詳述する必要がある。
- e) 試験法は十分に頑健であるとともに（プロトコルにおける軽微な変更によって、比較的差が生じにくい）、十分に訓練されたスタッフを擁し、適切な装置を備えた施設間で移転可能なものでなければならない。
試験法は、できれば標準化が可能なものであることが望ましい。専門性の高い装置や材料、専門的知識が必要な場合には、試験法の技術移転性を促進するための手段を確保しなければならない。
- f) 試験法は時間やコストに対する効果が高く、規制に関わる状況下で使用される可能性が高いものでなければならない。
- g) 新規及び改定試験法には、既存の試験法に照らした、（科学的、倫理的、経済的）正当性を提示しなければならない。
該当する場合、置換、削減、洗練という3R原則を含む動物福祉上の問題に対処しなければならない。

140. いくつかのOECD加盟国では、次に挙げる3つの基準を充足していれば、新規又は改定試験法の使用を採択することができる。
- 適正にバリデーションされている。
 - 少なくとも既存の試験と同等の科学的信憑性を確保することができるとともに、動物を全く使用しないか、使用する数を減らせる。若しくは、動物の苦痛を軽減することができる(7)。
 - 人や環境に関わる安全性の評価を向上できる。

このようなケースでは、バリデーションと規制上の承認の両方に対して提示した原則と基準が充足されていなければならない。

プロトコルからTGまで

141. 行政当局は、過去の経験をもとに、規制目的に適う正確性及び柔軟性の両方を確保できる試験手順を模索するのが慣例である。このことから、具体的な試験法のプロトコルというよりも、むしろ一般的な国際的テストガイドライン (TG) を使用することになる。
142. その試験法に専門的知識を持つ専門家グループによるTG草案の作成が提唱されている。専門家グループは、有効な試験結果の作成に重要であり、ゆえにTGに必要である個々の要素のほか、専門家の判断に委ねられる可能性のある、すなわち自由裁量の項目を個々に識別できなければならない。そのパラメータが必要か、若しくは自由裁量の要素であるかについて専門家グループの委員の同意が確保できない場合には、試験の実施により、これらのパラメータにおけるばらつきが試験結果に及ぼす影響を確定するのが適切な方法であると考えられる。試験プロトコルを最適化し、標準化し、バリデーションした後で、専門家グループがTGを作成するのが望ましい。

VII. 新規試験の提出－補助資料

143. 新規試験法をOECDのTGとして提案する準備が整ったら、次のいずれかの過程によってOECD事務局に提出することができる。
- i. 加盟国が国家調整官（ナショナルコーディネーター）を通じて提出
 - ii. EUがECを通じて提出
 - iii. 産業連合体がOECD経済産業諮問委員会（BIAC）を通じて提出
 - iv. 招聘された専門家が国家調整官（ナショナルコーディネーター）を通じて提出
144. 試験法をOECDのTGとして採択するための過程の詳細は、別のガイダンス文書に提示する（4）。
145. 行政上の承認の検討を提案する試験法に対し、その有効性を立証するのに必要な情報をどのように構成したらよいかを示す実用的な例として、ICCVAMが作成した「新規又は改定試験法及び代替試験法の推挙と提出」のためのガイドラインの資料が引用されている（60）。このガイドラインは、ICCVAM（重点は代替試験法のバリデーションに置かれている）並びに同団体に加盟する15の連邦機関に適用されるものであり、これを他の団体への拘束力を有する規格とみなすことはできない。しかしながら、同ガイドラインは、OECDの加盟国において必要な情報やデータを取りまとめ、構成化するとともに、OECD若しくはその他の既定のバリデーション基準に準拠した試験法のバリデーション状況の評価を簡易化できる、効果的で効率的な方法の実例としての役割を果たすものである。この提出文書を構成する要素を以下に説明する。

提案する試験法の概要と根拠

146. 新規試験の提出文書では、現在若しくは今後予想される規制への適用状況において、提案する試験法に関するどのような（例えば階層的試験戦略中のスクリーニング試験として、メカニズム的情報を確保するための補完試験として、あるいは既存の試験法に代わる代替又は置換試験として）適用の可能性が考えられるか詳細を提示できるものでなければならない。提案されている試験法のメカニズム的基礎や、これらの試験法が被験材料又は物質の毒物学的活性の測定又は予測に用いられる状況、並びに対象種（ヒトに関連する健康毒性試験におけるヒト等）と比較した場合に、これらの試験法の試験システムにおける作用様式及び作用機構にどのような類似点や相違点があるかに関わる知識について考察する必要がある。また、すでにバリデーションされ、承認されているメカニズム的、作用的に類似の試験法がある場合、提案する試験法がこれらの試験法の性能標準をどの程度充足しているかという問題に対処しなければならない。最後に、スポンサーは、提案する試験法が、安全性評価又はハザードアセスメント過程のどの時点にどのように組み入れられるかを明示しなければならない。特に、提案する試験法が階層的方式や組合せによる方式の一部を構成する場合、その新規の試験法に階層的方式や組合せ方式の中の他の試験法と比べてどの程度の重要性が置かれるのかについても考察するものとする。

試験法プロトコルの構成要素

147. スポンサーは、この章を用いて、試験法プロトコルにおける機能上、構造上、手順上の重要な構成要素に関する決定の根拠を説明し、詳述しなければならない（提案する試験法の完全かつ詳細なプロトコルは、提出文書の付属文書において提示するよう義務付けられている）。この中には、提案する試験法のプロトコルが、すでにバリデーションされ、性能標準が定められている、メカニズム上、機能上の類似試験法のプロトコルと、どの程度類似しているかの説明が含まれるものと思われる。また、提案する試験法のバリデーションの際に行われたプロトコル変更の根拠とその影響についても考察しなければならない。提案する試験法の技術的パラメータ（媒体や曝露時間等）や評価対象の反応の性質、提案される併用対照物質の根拠なども詳しく説明する必要がある。併用対照（陰性、溶媒、陽性）は、必要に応じて試験法が試験条件下において有効であることを示唆し、実験間の比較の基礎になるものであることから、所定の実験に対する許容基準の一部に含まれることが多い。許容可能な範囲を実証するのに用いた対照物質の反応の許容範囲と過去のデータを内容に組み入れなければならない。

148. 収集するデータの性質やデータ収集に用いた方法、データを保存する媒体の種類、ばらつき基準、（用量－反応関係の探索に用いた方法を含め）データの分析や評価に適用した統計学的、非統計学的方法、また、該当する場合、反応を陽性又は陰性に分類するのに用いた判定基準（並びにその根拠）について説明しなければならない。また、用量選択基準のほか、用量選択や実際の試験に動物を使用した場合にはその必要数を明記する必要がある。データの評価に適用した統計学的方法と非統計学的方法の両方について、適正な正当化を実施しなければならない。提案する試験法に付随する機密情報があれば、この点を明確に示すこととする。ただし、機密情報の内容への記載は勧められない。
149. 試験を適正に実施するのに要した試料のN数及び／又は反復実験数を提示し、試験の設計における根拠を詳述しなければならない。複数サンプル（レプリケート）の適用や反復実験が提案する試験のプロトコルの一部に含まれていない場合には、その根拠を提示することとする。
150. 提案する試験法システムを選択した根拠を明確に提示しなければならない。動物モデルを使用している場合には、動物モデル使用の根拠とともに、種や系統又は株、雌雄の別、許容週齢範囲、飼料、投与の頻度、投与回数、並びにプロトコルにおけるその他の該当要素を選択した根拠を内容に記載する。
151. 試験法において所有権のある構成要素がある場合、ロット間並びに経時的な（信頼性や正確性という点での）完全性を保証するために適用した手順を詳述しなければならない。また、これら所有権のある構成要素の完全性を検証するのにユーザーが適用できる手順を明確に説明する必要がある。

提案する試験法のバリデーションに使用する物質の特性付けと選択

152. 試験法は、あらゆる種類の化学物質や製品ではなく、特定のクラスの化学物質の評価にのみ有効であると判断される場合があることから、バリデーション研究に使用した化学物質やクラスについての情報の提示が重要である。この際、バリデーション過程で試験を行った物質の数や種類に対する根拠を説明しなければならない。化学物質や製剤の具体的な名称と、試験した物質に該当する化学物質及び製品のクラスを明記する必要がある。また、（例えば水への溶解度など）化学物質の物性におけるすべてのデータセットが均質であるとは限らないという事実に留意しなければならない。このようなケースでは、物理化学的特性や化学物質クラスをもとに、データセットをより小さく均一なサブセットに分け、データ分析を行うと有益かも知れない。可能な限り、それぞれの被験物質に対し、次の情報を提示することが望ましい。
- Chemical Abstracts Serviceの登録番号（CASRN）
 - 物理化学に関わる特性
 - 溶解度（オクタノール／水分配係数[Kow]等）
 - 試験を実施した濃度
 - 純度
 - 供給源
 - 試験媒体における試験物質の安定性

153. 試験法の正確性及び／又は信頼性に直接的な影響を及ぼす可能性のある被験化学物質の特性について詳述しなければならない。この場合、バリデーション過程においてコード化した物質や盲検を適用したかどうかに関する情報も内容に含める必要がある。混合物では、可能な限り、成分やその相対的な濃度を記述するものとする。すでにバリデーションが行われ、性能標準が確立されている試験法とメカニズム的、機能的に類似した試験法を提案する場合、提案する試験法において、性能標準において推奨される参照物質を試験した範囲について考察するとともに、このリストからの逸脱に対する正当な理由を提示しなければならない。リスト化されている参照物質が入手できなかった場合には、代替化学物質の選択に適用した基準を詳述する必要がある。代替化学物質は、もとの参照物質と比較した場合に、可能な限り、同じ化学物質や製品のクラスから選択し、*in vivo*の参照試験法において類似した影響を発現するものでなければならない。また、すでにバリデーションされている同等の試験法があれば、代替化学物質にはこれらの試験法による試験を実施していなければならない。使用した参照化学物質がリストされているものよりも少ない場合には、これらの根拠を提示する。試験を実施した物質の詳細を書面で記述するほかに、表3に示すように、表形式で情報を提示することが望ましい。できれば、これらの情報を印字と電子媒体の両方の形式で作成するとよい。

表3 試験対象化学物質の特性付け

化学物質又は製品の名称	CAS登録番号	化学物質クラス	製品クラス	試験濃度	純度	物質のサプライヤ若しくは供給源	物理化学的特性

提案する試験法の正確性評価に用いた*in vivo*参照データ

154. 提案する試験法が*in vivo*による既存の参照試験法の置換や代替を目的とするものであるような場合、提案する試験法と*in vivo*による参照試験法のデータの比較が必要である。対象とする生物学的指標や毒性指標に対し、標的種から得られたデータも合わせて提示しなければならない。この場合、提出書類には次の内容が含まれることになる。
- *in vivo*の参照試験法により得られた比較データのほか、入手の可能性や適用の可能性に応じて、ヒトなどの標的種に対する試験から作成された比較データ。このような場合、可能な限り個々の動物やヒトに関するデータを提示しなければならない。
 - *in vivo*の参照試験法（又はヒト）によるデータの選択に適用した基準。
 - *in vivo*の参照試験法によるデータの供給源（例えば、すでに発表されている情報の文献での引用、施設の試験責任者、未発表データの作成年等）。
 - *in vivo*の参照試験法やヒトによるデータの作成に採用したプロトコルの詳細。*in vivo*の参照試験法プロトコルを修正した場合には、データセット毎にこれを明記するとともに、これらの修正が提案されている試験法の正確性評価に潜在的に及ぼす影響について考察しなければならない。
 - GLPへの基準遵守の程度（45）（64）（65）（66）（67）（68）やコード化した被験化学物質の使用など、*in vivo*の参照試験法データの質に関する詳細。
 - *in vivo*の参照試験法によるオリジナル試験データ、若しくは、これらの有効性に関わる参考文献。

- 対象種から得られる他の関連毒性情報の入手の可能性や使用に関する概要（例えば、ヒトでの試験や、ヒトの健康に関わる毒性試験法に関する偶発的曝露、市販後調査の結果等）。

試験法データと結果

155. この項では、提案する試験法のプロトコルを用いた物質の試験によって得られるデータを報告する。開発過程でのプロトコルの修正と、このような修正による影響は、データセット毎に明確に記述しなければならない。もとのデータと推論によって導き出されたデータの両方を含むすべてのデータに加え、それぞれの試験の成果に関する各施設の判断の概要を提出する必要がある。提出書類には、うまくいったものもそうでなかったものも含め、すべての試験から得られたデータ（及び説明）を内容に組み入れなければならない。また、データの評価に用いた統計学的アプローチについて詳述し、これらに対する正当な理由を提示する。
156. また、被験化学物質のロット間での均質性や、多様な試験の実施期間、試験を実施した施設についての詳述が重要である。試験法の信頼性や正確性の施設内評価に従事した各施設は、コードで指定してもよい。もとのデータのうち、提出されていないものは、要請に応じて精査に提出できるようにしておかなければならない。
157. 参照試験法と提案する試験法のそれぞれから得られた結果を比較しやすいよう、結果は表の形式で提示することが望ましい。表4は結果を提示するための表形式の一案であり、これを用いて正確性評価に適用した情報を提示することができる。これらの情報は、印字と電子媒体の両形態で作成しておかなければならない。

表4 試験法の正確性評価

化学物質 又は 製品の 名称	CAS 登録 番号	化学 物質ク ラス	製品 クラス	提案する 試験法を 用いて得られた 結果 (定量的結果)	提案する 試験法を用い て得られた 結果 (+/-)	参照試験法を 用いて得られた 結果 (定量的結果) *	参照試験法 を用いて 得られた 結果 (+/-)	参照文献 又は データの 供給源	補足 説明

* 可能な限り、*in vivo*での参照試験法から作成されたデータは、情報の得られている種毎に一つの段組にまとめて分類しなければならない。ヒトによるデータは、必ず、ヒト以外のデータとは分けて提示する。すでにバリデーションされ、性能標準が確立されているメカニズム的、機能的類似試験法を用いて得られた同様のデータがあれば、これらも提示しなければならない。

試験法の妥当性（正確性）

158. この項には提案する試験法の正確性（感度や特異度、陽性及び陰性予測度、偽陽性率や偽陰性率等）について詳述するとともに、これを、現在行政当局によって受け入れられている参照試験法の正確性と比較し、さらには対象種（例えばヒトの健康に関わる毒性試験におけるヒト等）から得られたデータや既知の毒性情報とも比較を行わなければならない。提案する試験法がまだ存在していない試験法指標の測定や予測を行うものであるような場合には、対象種から得られた関連情報との比較によって、反応の正確性を判断する必要がある。また、提案する試験法が、すでにバリデーションされ、性能標準が確立されている試験法とメカニズム的、機能的に類似しているようなケースでは、両方の試験法における正確性を比較してみなければならない。提案する試験法を用いて得られた結果がバリデーション済みの同等の試験法の結果と一致しない場合は、それぞれの試験法において正確に結果を予測できる頻度を対象種から得られている毒性情報と比較し、その結果を提示する。次の比較において結果に不一致が認められた場合、その根拠について考察しなければならない。

- 提案する試験法と、現時点で受け入れられている標準的な試験法の比較
- すでにバリデーションされ、性能標準が確立されている同等の試験法が存在する場合、提案する試験法とこれらの試験法の比較
- 対象種によるデータが得られている場合、提案する試験法と、受け入れられている参照試験法のこれらの種の反応予測における比較

159. 提出書類には提案する試験法の長所と限界に関する考察結果を記載するものとし、データの解釈における顕著な問題について説明しなければならない。

試験法の信頼性（反復性／再現性）

160. 試験法の信頼性（反復性及び再現性）に対する評価結果を提示しなければならない。この場合の評価結果には、施設内及び施設間の再現性評価に適用した物質の選択における根拠や、これらの物質により示すことの可能な試験結果の範囲についての考察を含める。また、外れ値を特定し、これらについて考察しなければならない。米国試験材料協会（ASTM）の規格番号E691-92（62）や変動係数分析で説明されるもののような、施設内、施設間でのばらつきの程度に関する定量的統計分析を内容に取り込む必要がある。ここでは、過去の類似対照データ（該当する場合、陰性、溶媒、陽性等）において、代表値やばらつきの尺度をまとめておかなければならない。提案する試験法が、すでにバリデーションされ、性能標準が確立されている試験とメカニズム的、機能的に類似しているようなケースでは、二つの試験法の信頼性を比較し、これらの相違によって生じる潜在的な影響について考察する必要がある。

試験法データの質

161. この項には、提出書類中で提示するデータの国家的、国際的GLPガイドライン（46）（63）（64）（65）（66）（67）の遵守レベルや、データの質に関わる監査の結果を記載しなければならない。GLPガイドラインからの逸脱や、監査において特定された不適合が及ぼす影響について詳述する必要がある。また、施設のラボノートやスポンサーの保管している他のデータが外部監査時に入手できていたかどうかについての情報を明記することとする。すべてのデータは、施設のラボノートにおける関連文書によって立証されなければならない。

その他の科学的報告書や精査

162. この項では、提出書類により、発表されているか未発表かを問わず、提案されている試験法を用いて行われた試験のすべてのデータについて考察する。提案する試験法から得られた結果や、これらの試験の試験性能に関する独立した精査から提起された結論、あるいは提案する試験法の他の科学的精査による結論についての見解を提示しなければならない。これらの科学的報告書や精査の結論を、提出書類において到達した結論と比較する必要がある。また、提案する試験法に対して進行中若しくは計画済みの評価があれば、これらについても詳述しなければならない。提案する試験法が、既にバリデーションされ、性能標準が確立されている試験法とメカニズム的、機能的に類似しているようなケースでは、評価後に実施された試験の結果を内容に記載するとともに、これらが提案する試験法の信頼性や正確性に及ぼす影響についても考察することとする。

動物福祉に関わる考慮事項（洗練、削減、置換）

163. 提案する試験法が、対象の指標に対してその時点で採用されている方法に比べ、どのように動物の苦痛を軽減し、数を削減し、これらに代わることができるかを説明する。提案する試験法において動物を使用しなければならない場合は、その根拠を提示しなければならない。対象の指標に対し、動物の苦痛を軽減し、数を削減し、動物使用に代わる代替試験法が確立できる可能性があるという判断に適用した情報源を提示する必要がある（68）（69）。これらの説明には、少なくとも、調査したデータベースや調査戦略、調査日、データベース調査の結果のほか、入手可能な代替試験法を適用しなかった根拠についても内容に含めなければならない。また、提案する試験法における適切な動物数を確定した根拠について説明する必要がある。その試験が動物の痛みや苦痛を潜在的に伴うものである場合、これらの痛みや苦痛の発生を最小限に抑制し、可能な限りにおいて解消するために取り込まれている手順やアプローチについて考察すべきである（10）。

実用上の検討事項

164. この項では、提案する試験法を用いて試験を実施する場合に伴うコストや時間を明記し、参照試験法、並びに該当があれば、既にバリデーションされ、性能標準が確立されている、メカニズム的、機能的に類似している試験法と比較しなければならない。また、この項には次の点を内容に盛り込む。

- 試験法の実施に必要な設備や主要な据付機器に関する考察
- 他の必要な機器や補給品の一般的な確保の可能性
- 試験の実施者に対して必要になる訓練のレベルや専門的知識、習熟度の実証

参照文献

165. 提出書類において引用しているすべての出版物のリストを提示しなければならない。

補足的資料

166. 付属文書は次の内容を含むものとする。

- 提案する試験法の詳細なプロトコル
- 関連のあるすべての出版物の写し。この中には、提案する試験法や*in vivo*での参照試験法、また、もし該当する場合は、バリデーション済みの同等の試験法から得られたデータを掲載した出版物も含まれるものとする。
- これまでに作成され、提案する試験法の妥当性を評価するのに使用した、すべてのオリジナルデータ
- 提案する試験法に対する性能標準が作成されていない場合、検討すべき性能標準案

本翻訳は、厚生労働科学研究費補助金（医薬安全総合研究事業）「国際協調を重視した化粧品・医薬部外品における安全性試験法の再評価に関する研究」の支援を受けて実施された。