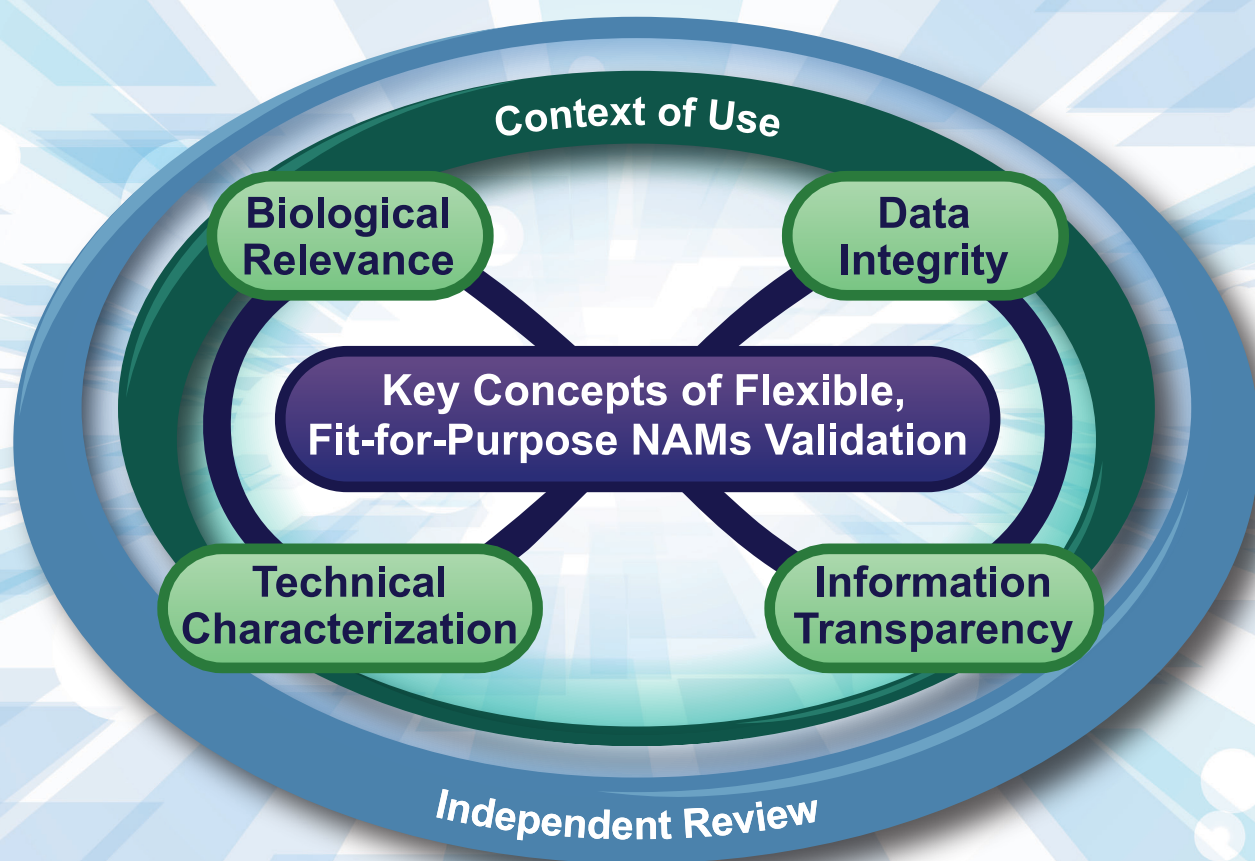


New Approach Methodology の妥当性検証、適格性、 および規制承認 2024年3月



目次

表の一覧.....	v
図の一覧.....	vi
略語と頭字語.....	vii
ICCVAM 妥当性検証作業グループ.....	ix
ICCVAM 機関代表者.....	xi
免責事項.....	1
エグゼクティブサマリー.....	1
1.0 はじめに.....	2
2.0 柔軟で目的に適合した NAM の妥当性検証における主要概念	4
3.0 NAM の信頼性を構築するための主要概念の適用	6
3.1 使用方法と目的 (COU).....	6
3.2 生物学的妥当性.....	8
3.2.1 メカニズムの理解.....	8
3.2.2 参照物質.....	12
3.2.3 既存の実験動物法との比較.....	15
3.3 技術的特性評価.....	17
3.3.1 選択された品質ツールの組み込み.....	17
3.3.2 品質管理のベストプラクティス.....	19
3.3.3 文書化.....	24
3.4 データの完全性.....	28
3.5 情報の透明性.....	28

3.6	独立審査.....	29
4.0	米国連邦機関による NAM の承認	30
4.1	規制のニーズと意思決定の背景の理解.....	30
4.2	使用方法と目的 (COU) に関する考慮事項.....	31
4.3	獲得した経験に基づく信頼性の変化.....	31
5.0	米国と国際調和.....	32
5.1	米国の調和: ICCVAM と NICEATM の役割.....	32
5.2	米国の調和: 3Rs を推進する機関間連携の強化.....	33
5.3	国際調和.....	34
6.0	NAM の使用を促進するためのコミュニケーションとトレーニング	35
7.0	結論と実装.....	36
	参考文献.....	37
	付録 A: 用語集.....	48
	付録 B: 品質ツール.....	54
1.0	フローチャート.....	54
2.0	因果関係分析.....	55
3.0	管理図.....	55
4.0	チェックシート.....	58
5.0	散布図.....	58
	付録 C: 分析方法評価.....	60
1.0	検出と定量化の限界.....	60
2.0	干渉物質の同定.....	60

3.0	分析精度の評価.....	60
4.0	材料の安定性.....	60
5.0	頑健性試験.....	61
6.0	回収の分析.....	62
7.0	適用範囲の技術分析.....	62
8.0	陽性対照.....	63
9.0	機器校正の参照標準物質.....	63
10.0	規格の設定.....	63

表の一覧

表 1. NAM の妥当性検証、適格性、規制上の使用に関する既存の米国および国際文書の例.....	3
表 2. 特定のトピックに対する政府機関の試験ニーズの詳細について ICCVAM 作業グループが作成した文書.....	7
表 3. NAM の生物学的妥当性とメカニズム上の妥当性が実証されているエンドポイントの例*.....	9
表 4. 参照物質の主な用途	12
表 5. NAM の技術的特性評価によく使用される適格性評価ツール*	19
表 6. 分析方法評価*	22

図の一覧

図 1. 柔軟性と目的適合性に優れた NAM バリデーション戦略の開発および実施中に考慮すべき重要な概念.....	5
図 2. 頑健な NAM を開発するためのフレームワーク	18
補足図 1. 修正 MTS を説明するフローチャート	54
補足図 2. 因果関係図	55
補足図 3. 対照表データ	57
補足図 4. 硫酸カドミウム EC ₅₀ 値の相関関係	59

略語と頭字語

3Rs: 動物を用いない方法への置き換え・動物の使用数の削減・動物使用に伴う苦痛の低減

AOP: 有害性発現経路

C&E: 因果関係

COU: 使用方法と目的

CPSC: 米国消費者製品安全委員会

DA: ディファインド アプローチ（定義済み総合判定方式）

EPA: 米国環境保護庁

EURL ECVAM: 欧州動物実験代替法評価センター

FDA: 米国食品医薬品局

GD: ガイダンス文書

GHS: 化学品の分類および表示に関する世界調和システム

GIVIMP: 優良インビトロ試験基準に関するガイダンス

GLP: 優良試験所基準

IATA: 試験の実施と評価のための統合的アプローチ

ICATM: 代替試験法国際協力

ICCVAM: 米国の代替法に関する省庁間連絡会議

ICH: 医薬品規制調和国際会議

IQ/OQ/PQ: 設備据付時適格性評価/運転時適格性評価/性能適格性評価

JaCVAM: 日本動物実験代替法評価センター

MTS: (3- (4,5-ジメチルチアゾール-2-イル) -5- (3-カルボキシメトキシフェニル) -2- (4-スルホフェニル) -2H-テトラゾリウム)

NAM: 新しいアプローチや方法論

NCATS: 国立先進トランスレーショナル科学センター

NICEATM: 代替法評価に関する毒性学プログラム省庁間センター

NIEHS: 国立環境衛生学研究所

OD: 吸光度

OECD: 経済協力開発機構

SOPs: 標準操作手順書

TG: 試験法ガイドライン

Tox21: 21 世紀の毒性試験（研究プロジェクト）

TSAR: 規制当局の承認に向けた代替法の追跡システム

ICCVAM 妥当性検証作業グループ

有害物質疾病登録庁

Moiz Mumtaz, PhD

米国消費者製品安全委員会

John Gordon, PhD (Co-Chair)

米国防衛省

Donald Cronce, PhD

Natalia Garcia-Reyero Vinas, PhD

Matthew Johnson, DVM, DACLAM (2022 年 6 月まで)

Emily N. Reinke, PhD (2022 年 4 月まで)

米国退役軍人省

George Lathrop, Jr., DVM, MS, DACLAM

米国環境保護庁

農薬プログラム部

Anna Lowit, PhD

Scott Lynn, PhD

Monique Perron, ScD

研究開発局

Kelly Carstens, PhD

Alison Harrill, PhD

Nisha Sipes, PhD

米国食品医薬品局

医療機器・放射線保健センター

Jennifer Goode, BS

医薬品評価研究センター

Paul C. Brown, PhD

食品安全・応用栄養センター

Suzanne Fitzpatrick, PhD (Co-Chair)

Anneliese Striz, PhD

タバコ製品センター

Jueichuan (Connie) Kang, PhD

チーフサイエンティスト部

Tracy Chen, PhD

国立環境衛生学研究所

Warren Casey, PhD

代替法評価に関する毒性学プログラム省庁間センター (NICEATM)

Helena Hogberg, PhD

Nicole Kleinstreuer, PhD

国立衛生研究所

国立先進トランスレーショナル科学センター

Menghang Xia, PhD

国立標準技術研究所

Elijah Petersen, PhD (Co-Chair)

職業安全衛生局

Janet Carter, MS

NICEATM サポート契約スタッフ (Inotiv)

David Allen, PhD (2023 年 12 月まで)

Michaela Blaylock

Amber Daniel, MTTox

Agnes Karmaus, PhD (to April 2023)

ICCVAM 機関代表者

本報告書の公開時点での ICCVAM 委員会メンバー

*主要機関代表者;+代理主要機関代表者

有害物質疾病登録庁

- * Moiz Mumtaz, PhD
Patricia Ruiz, PhD

米国消費者製品安全委員会

- * John Gordon, PhD
- + Kristina Hatlelid, PhD, MPH
- + Eric Hooker, MS
- + Joanna Matheson, PhD

米国農務省

- * Jessie Carder
- + Patrice Klein, MS, VMD, DACPV, DACVPM
Erika Edwards
Ben Green, PhD
Katherine Horak, PhD

米国防衛省

- * Shannon Marko, DVM, DACLAM
- + Natalia Garcia-Reyero Vinas, PhD (Co-Chair)
- + Saber Hussain, PhD, ATS Fellow, AFRL Fellow
- + Elaine Merrill, PhD
Matthew Grogg, PhD
Donald Cronic, PhD

米国エネルギー省

- R. Todd Anderson, PhD

米国内務省

- * Barnett A. Rattner, PhD
- + Jessica K. Leet, PhD

米国運輸省

- * Steve Hwang, PhD
- + Rebecca Rothhaas
- + Ryan Vierling, PhD

米国退役軍人省研究開発局

- * Holly Krull, PhD
- + George Lathrop, Jr., DVM, MS, DACLAM

米国環境保護庁

化学物質安全・汚染防止局

Charles Kovatch

農薬プログラム部

Monique Perron, PhD

William (Bill) Eckel, PhD

Cecilia Tan, PhD

汚染防止および毒性局

- * Anna Lowit, PhD
- Louis (Gino) Scarano, PhD

研究開発局

- + Alison Harrill, PhD
- Kelly Carstens, PhD
- Grace Patlewicz, PhD

米国食品医薬品局

生物製剤評価研究センター

Leslie Wagner

Allen Wensky, PhD

医療機器・放射線保健センター

- + Jennifer Goode, BS
- Simona Bancos, PhD
- Rakhi M. Dalal-Panguluri, PhD

医薬品評価研究センター

Paul C. Brown, PhD

Nakissa Sadrieh, PhD

食品安全・応用栄養センター

- * Suzanne Fitzpatrick, PhD (Co-Chair)
- Omari J. Bandele, PhD
- Patrick Crittenden, PhD
- Shruti Kabadi, PhD

タバコ製品センター

Jueichuan (Connie) Kang, PhD

動物用医薬品センター

M. Cecilia Aguila, DVM

Li You, PhD

国立毒性研究センター

Mugimane (Manju) Manjanatha, PhD

チーフサイエンティスト部

Tracy Chen, PhD

Chad P. Nelson, PhD, MSPH

国立がん研究所

- * Brian Cholewa, PhD
- + Ron Johnson, PhD
- Gary Robinson, PhD

国立労働安全衛生研究所

- * Stephen Leonard, PhD

国立環境衛生学研究所

- * Warren Casey, PhD
- + Stephen Ferguson, PhD
- + David M. Reif, PhD
- Nicole Kleinstreuer, PhD

国立標準技術研究所

- * John Elliott, PhD
- + Elijah Petersen, PhD

国立衛生研究所

- * Nicolette Petervary, VMD, MS, DACAW
- Michael Eichner, DVM, DACLAM

国立医学図書館

* Dina N. Paltoo, PhD, MPH, CPI

職業安全衛生局

* Deana Holmes, MT

+ Janet Carter, MS

免責事項

本報告書は、動物実験において、動物を用いない方法への置き換え、動物使用数の削減、動物使用に伴う苦痛の低減への新しいアプローチの信頼性確立を求める米国連邦機関および関係者向けのリソースとして作成された。本報告書に記載されている原則は、17 の連邦機関の職員、複数の機関内作業グループ、一般市民、および代替毒性法に関する科学諮問委員会からの意見を取り入れて策定された。そのため、本報告書は必ずしも特定の機関または作業グループの意見または方針を反映するものではない。特定の個人または団体に権利をもたらすものではなく、特定の連邦機関による取り組みとして見なすべきではない。本報告書は、法的強制力のある責任を定めるものではない。代わりに、本報告書は、あるトピックに関する米国の代替法に関する省庁間連絡会議(ICCVAM)の現在の考えを説明するものであり、特定の規制または法的要件が引用されていない限り、推奨のみとして見なすべきである。「すべき」という言葉の使用は、何かが提案されたり推奨されたりしても、必須ではないことを意味する。

エグゼクティブサマリー

New approach methodology (NAM) が開発される頻度はますますたかくなっており、化学物質や製品が人の健康と環境に及ぼす潜在的な毒性効果の規制および非規制上の評価を提供するために利用されている。これらの NAM は、毒物学的プロセスの根底にある生物学的メカニズムの調査、新製品および既存製品の評価支援、健康および環境の危険有害性の分類とリスク評価を目的とした危険有害性の特定と用量反応関係情報の生成に使用されている。本報告書は、開発者とエンドユーザーによる NAM の信頼性の確立を促進するために ICCVAM によって策定された。この信頼性は、NAM の意図した用途を考慮する柔軟で目的に適合したバリデーション戦略を実施することによって達成することができる。本報告書は、NAM の使用方法と目的、生物学的妥当性、技術的特性評価などの概念を説明することにより、信頼性の構築を促進する。本報告書では、米国連邦機関の規制当局による NAM の承認と国際的な調和の可能性について考察し、NAM の品質ツールと技術評価に関するベストプラクティスをレビューする。また、開発者、エンドユーザー、規制当局間のコミュニケーションの必要性も強調する。NAM のような先進的な分野では、妥当性検証に時代遅れの万能戦略は通用しない。ICCVAM は、妥当性検証作業の調整を通じて、また研究協力の機会を確立することにより、諮問委員会としてより現代的なアプローチへの移行を支援することができる組織である。そこで、ICCVAM は、17 の連邦機関および研究所と協力して本報告書を作成し、開発者、エンドユーザー、規制当局が NAM の信頼性を高めることで、規制および非規制の目的で動物実験を補完または置換するために NAM を実施できるようにすることを目指す。

1.0 はじめに

米国の代替法に関する省庁間連絡会議 (ICCVAM) は、毒性および安全性試験に関する情報を要求、使用、生成、または配布する 17 の米国連邦規制および研究機関の代表者から構成される。ICCVAM は、新規、改訂、そして代替試験法の安全性試験法、および規制上の適用性を備えた統合試験戦略の技術的評価を実施する。ICCVAM はまた、動物を用いない方法への置き換え、動物の使用数の削減、動物使用に伴う苦痛の低減（動物の福祉を向上し、痛みや苦痛を軽減または回避する）のための取り組みとして、関連製品の化学物質の安全性と危険有害性を正確に評価する試験法の科学的検証と規制当局による承認または適格性確認を推進する。

1997 年に常任委員会として設立されて間もなく、ICCVAM は「毒性試験法の妥当性検証と規制当局による承認」(ICCVAM、1997) という報告書を発行した。この報告書には、新しい試験法および代替試験法の妥当性検証と規制当局による承認の基準が概説されている。その後、2003 年の出版物「新規、改訂、および代替試験法の推薦と提出に関する ICCVAM ガイドライン」(ICCVAM、2003) において追加のガイダンスが提供された。これらの文書に示された原則は、経済協力開発機構 (OECD) ガイダンス文書 (GD) 34「危険有害性評価のための新規または更新された試験法の妥当性検証と国際的承認に関するガイダンス文書」(OECD、2005) に記載されている国際的基準の策定に引き継がれた。これらのリソースでは、原則的には柔軟性のある検証モデルについて説明したが、実際には時間がかかり、リソースを大量に消費するなど、さまざまな限界が示されている。使用目的によっては、この検証プロセスのすべてのステップを経なくても連邦政府機関にとって貴重なデータが得られる場合がある。さらに、これらの文書は、毒性試験に対する多くの最新アプローチと必ずしも互換性があるとは限らない。これらのアプローチでは、*in vivo* 試験を単一の代替法に置き換えることを重視せず、複数の *in vitro* および *in chemico* 試験と *in silico* アプローチ（計算モデルなど）の結果を統合することを重視している。2018 年の ICCVAM 出版物「米国における化学物質および医薬品の安全性を評価するための新しいアプローチを確立するための戦略的ロードマップ」(ICCVAM、2018) は、特定の用途または使用方法と目的 (COU) について検証された連邦政府機関や規制対象業界による新しい方法の採用を確実にするために、機関と試験法開発者間のより良いコミュニケーションを促進し、信頼の確立における柔軟性を高める概念的枠組みを提供している。以下の文章は、2018 ICCVAM ロードマップに示された原則に基づき、新しいアプローチや方法論 (NAM) の信頼性確立に関する、より具体的な知見を提供する。

本報告書では、NAM という用語は、化学物質の危険有害性とリスク評価に関する情報を提供するために使用できるあらゆる技術、方法論、アプローチ、またはそれらの組み合わせを指し、動物を用いない方法への置き換え、動物の使用数の削減、動物使用に伴う苦痛の低減 (3Rs) を支援する。この報告書は、米国および国際的に承認されている既存のガイダンスおよび文書（表 1）に明記されている原則に基づき、NAM への信頼性

を築くためのより柔軟なアプローチを提唱する。本報告書では、生体システムでの試験（*in vitro*、特定の化学物質、小規模モデル生物など）での試験を含む NAM に適用される妥当性検証、適格性確認、および規制当局による承認の主要な概念を紹介するが、計算モデル予測など、他のタイプの NAM については、さらなる考慮事項が必要になる可能性がある（OECD、2007）。

表 1. NAM の妥当性検証、適格性、規制上の使用に関する既存の米国および国際文書の例

文書タイトル	参考文献
毒性試験法の妥当性検証と規制承認（本報告書の発行時点で廃止）	ICCVAM、1997
新規、改訂、および代替試験法の推薦と提出に関する ICCVAM ガイドライン	ICCVAM、2003
試験および評価に関する OECD シリーズ第 34 号: 危険有害性評価のための新規または更新された試験法の妥当性検証および国際承認に関するガイダンス文書	OECD、2005
CPSC の動物実験に関する方針に関する推奨手順	CPSC、2012
FDA 予測毒性学ロードマップ	FDA、2017a
医療機器開発ツールの適格性: 業界、ツール開発者、および食品医薬品局のスタッフ向けのガイダンス	FDA、2017b
TSCA プログラム内の代替試験法の開発と実施を促進するための EPA 戦略計画	EPA、2018
米国の化学薬品と医薬品の安全性評価に対する新しいアプローチを確立するための ICCVAM 戦略的ロードマップ	ICCVAM、2018
優良インビトロ試験基準に関するガイダンス (GIVIMP)	OECD、2018
業界および試験法開発者向けガイダンス: 代替試験法および統合試験アプローチ並びに FHSA ラベル表示要件をサポートするための方法から生成されたデータに関する CPSC スタッフ評価	CPSC、2020
業界および FDA スタッフ向け医薬品開発ツールのガイダンスの適格性確認プロセス	FDA、2020

文書タイトル	参考文献
EPA の新しいアプローチ方法作業計画	EPA、2021a
化学試験における脊椎動物の使用を削減するための EPA 戦略計画	EPA、2021b
FDA における新しい代替法の推進	FDA、2021a

2.0 柔軟で目的に適合した NAM の妥当性検証における主要概念

OECD GD 34 で説明されている妥当性検証の基本原則が重要であることに変わりはないが、妥当性検証に使用されるプロセスは、目的適合性と信頼性に優れ、対象生物種に関連する情報を提供する NAM を効率的かつタイムリーに開発できるようにすべきである。具体的には、OECD GD 34 は「新しい試験法は、健全な科学を採用し、規制上のニーズを満たすことを保証するために妥当性検証を受ける」と述べている。つまり、その方法は目的適合性があり、規制上の決定に役立つ情報を提供することに加えて、ユースケースのスクリーニングや優先順位付けも含まれる可能性がある。ガイダンスには「妥当性検証プロセスは柔軟で適応可能であるべき」であり、性能は「一連の参照化学物質を使用して実証される」、「既存の関連毒性データに関連して評価される」必要があることも記載されている（OECD、2005）。そのため、NAM への信頼性は、特定のプロセスの完了後に与えられる普遍的なステータスと見なすべきではない（たとえば、リング試験（試験施設間比較試験）が正常に完了した後に NAM を「検証」することを検討する場合）。代わりに、NAM への信頼性を確立することは、試験法の開発者、規制上の意思決定者、および妥当性検証機関の間のコミュニケーションを必要とする発展的かつ反復的なプロセスと見なすべきである（ICCVAM、2018）。

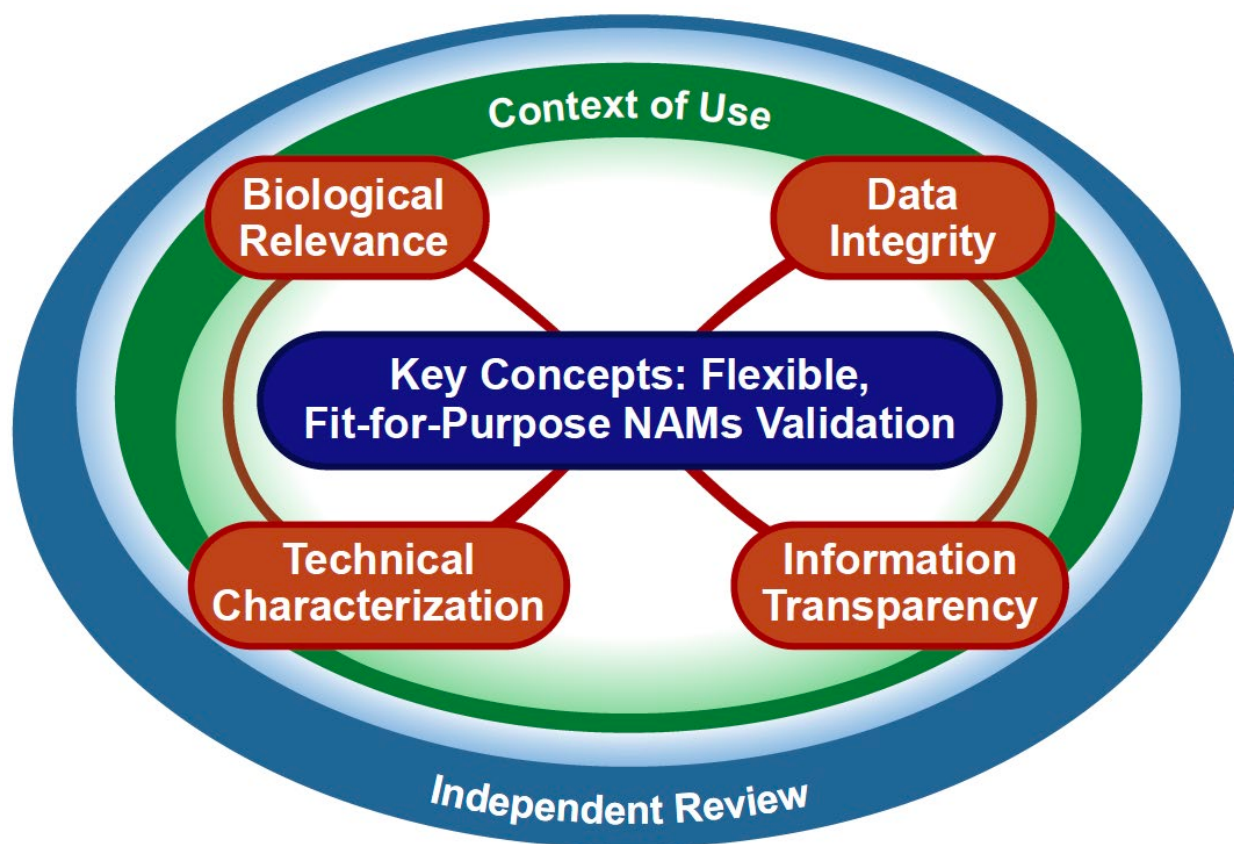


図 1. 柔軟性と目的適合性に優れた NAM のバリデーション戦略の開発および実施中に考慮すべき重要な概念 出典: van der Zalm ら (2022).

図 1 に示すように、柔軟性と目的適合性に優れたバリデーション戦略を設計および実施する際に考慮すべき重要な概念がいくつかある。これらの概念は、さまざまな目的で NAM を使用する際に広く適用できるが、特定のシナリオや用途に合わせて必要に応じて調整すべきである。考慮すべき包括的な概念として、COU、つまり NAM が意図する目的（スクリーニング、危険有害性の特定、強度評価、定量的リスク評価の出発点など）である。COU は、妥当性検証プロセスの柔軟性に関する情報を提供し、特定の COU に対する NAM の適合性が確立される。柔軟で目的に適合した検証プロセスに関わる重要な概念は、生物学的妥当性、技術的特性評価、データ完全性、および情報の透明性である。図に示されるように、これらの側面は関連し、相互に依存する。最後に、妥当性検証プロセスのすべての部分は独立した審査を受ける必要がある。これらの主要な概念は、van der Zalm ら (2022) が NAM の科学的信頼性を確立するためのフレームワークで提案した重要な要素と類似しているが、本書ではより幅広い NAM アプリケーションで有用であるように適用している。これらの主要な概念については、セクション 3.0 で詳しく説明する。

3.0 NAM の信頼性を構築するための主要概念の適用

妥当性検証は、特定の COU に対する NAM の適合性を決定することによって科学的信頼性が確立される、頑健かつ柔軟なプロセスにすべきである。可能かつ適切な場合は、代替法を使用が既存の方法と同等かそれ以上の情報を提供し、規制当局による審査がヒトの健康を守る決定につながることを裏付ける証拠があるべきである。COU は、詳細な基準と主要な概念の実装を決定する。NAM が特定の目的に適しているかどうかを判断するには、COU と NAM が関連する生物学の両方を明確に定義する必要がある。NAM は十分に記述され、対象となるエンドポイントまたはプロセスに生物学的に関連し、および/または健康を保護する、技術的に信頼できる情報を提供する必要がある。NAM が提供する情報の構成は、完全性と信頼性を確保するために独立した審査が可能になるように、十分に透明でなければならない。適切かつ可能な場合、NAM の信頼性構築には、従来の動物実験法によって提供される情報と比較して、規制上の意思決定の対象となる種に関連する情報を、量的にも質的にも同等かそれ以上の品質で関連性のある情報を提供することを実証することが含まれる場合がある。現状に制約されないようにするために、NAM が従来の動物実験法よりも質が高く関連性の高い情報を規制上の意思決定に提供する可能性が確認されなければならない。そのためには、この可能性を考慮し、従来の動物実験法によるデータとの比較が不可能な状況を考慮する妥当性検証フレームワークが必要である。

3.1 使用方法と目的 (COU)

COU、すなわち NAM の意図された利用を確立することには、NAM が意図する使用方法とその規制目的（該当する場合）を完全かつ明確に説明する文章を作成することが含まれる。米国連邦機関は、各機関独自の法令や規制（例えば、Shaffer、2021 を参照）に基づいて運営されているため、NAM が承認され、各機関の個別の要件に適用できるかどうかの基準は異なる。場合によっては、1 つの機関内（例えば、米国食品医薬品局 (FDA) の様々な試験施設、米国環境保護庁 (EPA) の部局内）であっても、規制の分野に応じてニーズが異なる場合がある。これは国際的にさらに複雑になり、他国の規制が米国の規制と異なる場合があり、特定の COU に対してある国では承認されている方法が他の国では承認されない可能性がある。従って、NAM は特定のガイダンス/規制の下で 1 つ（または複数）の COU の特定のエンドポイントに対応するために検証されたとしても、様々な機関、規制管轄区域、および国の間で異なるガイダンス/規制の下では承認されない可能性がある。このような多様な規制状況の理解を深めるために、ICCVAM は国内外のさまざまな機関から情報を収集し、特定の毒性エンドポイントに特有の機関ニーズを特徴づけた（NIEHS、2023a、例は表 2 を参照）。

表 2. 特定のトピックに対する政府機関の試験ニーズの詳細について ICCVAM 作業グループが作成した文書

対象の試験ニーズ	参考文献
急性毒性試験	Strickland ら、2018
生態毒性試験	Ceger ら、2022
<i>In vitro</i> - <i>in vivo</i> 外挿	Chang ら、2022
NAM を用いるナノ材料試験	Petersen ら、2022a
皮膚感作性	Daniel ら、2018; Strickland ら、2019
皮膚刺激性試験および眼刺激性試験	Choksi ら、2019

スクリーニング/優先順位付けや危険有害性の特定など、さまざまな COU を持つ NAM は、毒物学的メカニズムに関する特定の質問に答えたり、定量的リスク評価をサポートしたりするために設計されたものとは異なる場合がある。そのため、それぞれの目的で NAM を評価・妥当性検証するために異なる基準を開発し、適用することができる。学術の現場では、特異的または探索的な性質の場合がある基礎研究の課題に対処するために、多くの新しい方法が開発されている。時には、これらの方法の 1 つが、規制上の意思決定など、基礎研究以外のより応用的な分野で役割を果たす可能性があることが認識される。新しい方法の開発者は、NAM の妥当性を検証する前に、連邦機関および/または対象ユーザー若しくは関係者の両方と協議して、信頼性を確立するために設計された妥当性検証試験が目的とする COU に適切に調整されることを確認することが重要である。

製品の発見と開発には、他の NAM の方が有用な場合がある。NAM は、新しい分子の望ましい活性（例えば、薬理学的活性）または潜在的に望ましくない活性（毒性）をスクリーニングするために採用される場合がある。これらのデータは、今後の開発に取り入れるべき分子を決定するために使用することができる。試験法を標準化は、方法の開発または試験ではなく、新しい技術的または治療法の発見に労力とリソースを集中させることで、イノベーションを支援できる。このような NAM の使用は、ほとんどの連邦機関の規制当局以外で行われる傾向がある。ただし、質の高いデータを確実に取得するには、NAM における科学的信頼性を確立するための重要な概念を適用することが推奨される。このような NAM の利用は、3Rs を直接支援できると同時に、NAM のさらなる開発を支援できる予備的なデータや経験を提供することができ、規制目的での利用を実現する可能性が認識されている。

3.2 生物学的妥当性

NAM の関連性とは、試験と標的生物種の効果との関係を説明するとともに、試験法が規定された目的に対して有意義で有用であるか否かを表すもので、制限事項も特定されている。NAM の関連性を十分に示すことは、NAM の信頼性を高める重要な要因である。生物学的妥当性は、以下に詳述するように、入手可能な情報に応じてさまざまな方法で実証できる。

以下のサブセクションに詳述されている生物学的妥当性に関する考慮事項には、以下が含まれる。

- NAM はどのような種類の情報を提供するか？結果/エンドポイントをもたらす生物学とメカニズムについての理解があるか？
- NAM が照会しようとしている結果をベンチマークするためにどのような参照データが利用可能か？
- 確立された実験方法に対して NAM のアウトプットをベンチマークするかどうか、またどのようにベンチマークするかについてどのような考慮事項があるか？

バリデーションアプリケーションの特定の状況によっては、生物学的妥当性に関する追加の考慮事項も重要になる場合がある。

3.2.1 メカニズムの理解

NAM の関連性を評価する際は、対象となる種（通常は人間であるが、他の種であることも多い）の生物学を考慮することが重要である。NAM により提供された情報と *in vivo* 生物学により提供された情報の比較は、可能な限り、既存のメカニズムに関する知識（例えば、有害性発現経路 [AOP] または毒物学的に関連する生物学的プロセス）によって裏付けられるべきである。分子開始イベントまたは 1 つ以上の重要なイベントを介して確立された AOP によって NAM を支えることは、NAM の生物学的妥当性を実証するのに役立つ。一方、NAM によって予測されているアウトカムについて確立された AOP がいないからといって、NAM が有用である可能性が必ずしも排除されるわけではなく、そのような裏付けとなる情報は、*in vivo* データによって提供されるメカニズム上の洞察と標的生物種の理解から得られる可能性がある。どのような生物的事象が測定されているのか、またそれが有害転帰または懸念される危険有害性とどのように関連するのかを NAM が包括的に特徴づけ、明確に説明することが重要である。

NAM の説明は、*in vivo* のアウトカムを予測するためのモデルの生物学的妥当性に対処し、生物学的プロセス、メカニズム、または AOP とのメカニズム的関連を提供すべきである。たとえば、ヒト用医薬品が胎児奇形または胎児致死性を誘発する可能性を NAM が予測する場合、胎児発育のメカニズム（細胞移動、分化、血管形成、神経形成、原腸形成など）と、モデルで研究されたその後の発達上の副作用との関係に関する

理解が存在すべきである（FDA、2021b）。予測されている *in vivo* の影響との関係は本質的に相関するかもしれないが、相関関係または経験的關係のみから NAM 予測の信頼を構築することはもっと困難である。したがって、評価対象のプロセスへの生物学的妥当性が明確な試験が好ましい。NAM の生物学的およびメカニズム上の妥当性が理解されていないと、NAM の妥当性検証に使用されるデータによって厳密に定義される境界への適用性が制限され、NAM の確立と妥当性検証に使用される以外の化学クラスに NAM を拡張することが困難になる可能性がある。

AOP は、分子や細胞障害を健康上の有害な転帰と結びつけるために有用な体系化されたフレームワークであり、重要な毒物学的プロセスを表す NAM の開発と定着に使用できる。皮膚感作性など、AOP が十分に確立されているエンドポイントの例がある（Kleinstreuer ら、2018）。AOP が十分に確立されていないエンドポイントでは、代わりに生物学的経路や過程などの要因に基づいて NAM のメカニズム上の妥当性を検討することができる（表 3）。

表 3. NAM の生物学的妥当性とメカニズム上の妥当性が実証されているエンドポイントの例*

エンドポイント	要約	参考文献
皮膚感作性	このエンドポイントには、人間に関連する AOP が十分に開発されており、複数の NAM を組み合わせたディファインドアプローチがマッピングされ、OECD ガイドライン 497 に記載されている。	Kleinstreuer ら、2018; OECD、2021a
内分泌かく乱作用	エストロゲンとアンドロゲン受容体の活性については、統合戦略の一環として補完的 NAM を使用する確立された経路モデルが利用可能である。EPA は、これらの NAM を内分泌かく乱物質スクリーニングプログラムの Tier 1 スクリーニングに認めている。	Judson ら、2015; Kleinstreuer ら、2017; EPA、2023

エンドポイント	要約	参考文献
発達神経毒性	この複雑なエンドポイントの AOP は限られている。代わりに、ヒトの神経発達の重要なプロセスを対象とする一連の NAM が開発された。これらの NAM に関する OECD GD には、試験の実施と評価のための統合的アプローチ (IATA) のケーススタディが含まれている。	Crofton および Mundy、2021; OECD、2022a; OECD、2023
吸入毒性	<i>in vitro</i> ヒト細胞ベースの試験と計算モデリングを使用する別のアプローチが使用され、EPA のヒトの健康リスク評価で使用するための出発点となった。また、このアプローチは、OECD IATA のケーススタディとしても発表された。	Corley ら、2021; EPA、2021c; OECD、2022b
眼刺激性	農薬および農薬製品の眼刺激性を評価するための代替試験フレームワークを説明するガイダンス文書。ヒトの眼の解剖学的構造および毒性メカニズムとの関連性に関して、 <i>in vivo</i> 、 <i>in vitro</i> 、および <i>ex vivo</i> で利用可能な試験法がレビューされた。	EPA、2015; Clippinger ら、2021
皮膚刺激性	既存の情報、物理化学的性質、およびその他の試験以外の方法を用いて、皮膚の腐食性および刺激性に関する IATA を提案するガイダンス文書。	OECD、2014

* さまざまな NAM の承認については、機関独自のガイダンスを参照すること。

理想的には、対象の生物学的効果に対する NAM の関係性の説明は、対象種におけるエンドポイントに関連する生物学または作用機序に関して入手可能な情報に基づくべきである。場合によっては、これが不可能な場合があり、対象種とは異なる種のデータが比較の唯一の有効な根拠となる場合がある。ただし、対象種の生理学は既存の代理試験種とは異なる場合があり、NAM 評価にメカニズムの理解を組み込む必要性がさらに強調される。たとえば、ウサギの眼の解剖学的および生理学的側面は人間とは異なり

(Clippinger ら、2021)、脳の発達中の感受性の範囲は種によって異なる (Smirnova ら、2014; Tsuji および Crofton、2012)。したがって、NAM と既存の参照試験法の両方を評価する際は、対象種との生物学的妥当性および曝露に関する重要な考慮事項を確認

すべきである。生物学的妥当性とメカニズム上の妥当性の評価では、例えば、使用する細胞タイプの関連性、調査対象の関連臓器/組織の生理学的特性、または試験物質に関連する種関連代謝産物の存在などが考慮される可能性がある（ある種に見られる代謝物は、別の種で検出された代謝産物や NAM で生成された代謝産物とは異なる場合がある）。さらに、方法が特定の種固有の作用機序またはメカニズムを評価する能力を調べることもできる（Hartung、2010; Madia ら、2021; Parish ら、2020）。

種の変換だけでなく、個体間のばらつき（*in vitro* 試験で使用する細胞や組織のドナーなど）は、測定された反応に影響を与え得る生物学的変動のさらなる根拠となるだけでなく、潜在的に集団全体の個体間変動に関するさらなる洞察を追加する可能性もある（Harrill、2020）。本書では詳しく説明しないが、NAM のユーザーと開発者は、効果の観察（観察された効果の大きさ、場合によっては観察された効果の有無に関する）、並びに効果が発生する出発点または用量は、細胞株パネル全体に存在する遺伝子配列変異によって調節される可能性があることを認識すべきである（Chiu ら、2017）。細胞ベースの試験に遺伝的に多様な細胞株パネルを使用することは、毒性と表現型反応の集団動態を理解し（Harrill および McAllister、2017）、感受性をもたらす遺伝子配列変異を明らかにし（Frick ら、2015）、特異的に感受性の高い亜集団の同定を可能にし（Church ら、2015）、毒物動態学および毒物動力学の変動の測定を促進する（Rusyn ら、2022）ために利点を提供することができる。腫瘍由来の細胞株が *in vitro* モデルで使用されることがあるが、その遺伝子組成は集団における遺伝的変異を十分に表していない可能性がある。NAM に多様性を組み込むための具体的な推奨事項は、状況によって異なる。したがって、標的生物種の多様性を十分に網羅できるように遺伝学者と協力して、目的に適合した試験デザインを開発すべきである。また、ヒトの遺伝的変異を調べる場合は、生命倫理学者や関連する関係者のグループと協力して開発すべきである（たとえば、ICCVAM、2022 を参照）。個人間の変動を評価するための具体的な推奨事項は、本報告書の範囲外であり、この分野における新しい科学の発展とともに継続的に変化する可能性が高い。

複雑な AOP または複数の AOP の結果であるアウトカムの場合は、アウトカムの適切な予測を得るために有害性発現のメカニズムの複数の側面を評価することが必要な場合がある。一般に、NAM には限界があり、すべての化学クラスを予測したり、すべての化学クラスに技術的に応用したりすることはできない。NAM は、代表できる生物学の複雑さという点でも限界が存在する場合がある。性能、適用性、生物学的適用範囲が異なる複数の NAM をディファインド アプローチ（定義済み総合判定方式、DA）に組み合わせることで、アウトカムを予測する能力を高めることができる（OECD、2017）。これは、適切な NAM を AOP に沿った主要イベントにマッピングして、対象種に関連する有害性発現を予測するためにメカニズム上の十分な代表例を確保することで達成できる。NAM を AOP に沿った主要イベントにマッピングして、規制エンドポイントに対処する DA が得られた例は、OECD の「ガイドライン第 497 号: 皮膚感作性に関するディファインド アプローチ」（OECD、2021a）に記載されている。

3.2.2 参照物質

参照物質には、表 4 に示すように、試験の実施から技術的品質や生物学的妥当性の評価に至るまで、幅広い用途が考えられる。参照物質は、ある目的には有効でも、別の目的には適さない場合がある。例えば、参照物質は、標的生物種に意図された有害作用（例えば、血液脳関門を通過できないため）をもたらさない場合でも、陽性対照として特定の方法（例えば、発達神経毒性試験）では十分に作用する場合があります、したがって生物学的エンドポイントの参照物質としては適さない場合がある。COU によっては、参照物質には単一成分の化学物質、混合物、または複合抽出物（例えば、医療機器に由来）が含まれる場合がある。

表 4. 参照物質の主な用途

個別の化合物セット*	目的	化合物の選択基準
陽性対照	方法が予想どおりに実行されているかどうかを検証する。	その方法の条件下における特定の試験法の測定値において、信頼性が高く、測定可能で、統計的に有意な変化を引き起こすことが確認された化合物。
性能測定用化合物	新しい方法または修正された方法が、確立済みの同様な「ミートゥー（追隨）」方法と比較してどの程度であるかを評価する。	すでに確立済みの方法で確実に反応を誘発する（または反応しない）化合物のセット。
技能測定用化合物	確立済みの方法を使用して、新しい試験施設の性能を評価する。	確立済みの方法で確実に反応を誘発する（または反応しない）化合物。性能測定用化合物のサブセットの場合がある。

個別の化合物セット*	目的	化合物の選択基準
生物学的エンドポイントの参照物質	1) 方法の生物学的妥当性を評価する。 2) 同じ <i>in vivo</i> エンドポイントに関連するアウトカムを測定するために設計された方法であるが、異なるコンテキスト（異なる種、方法、測定方法など）を使用して、その一致度を比較する。	1) 関心対象のエンドポイントで <i>in vivo</i> 効果がある（陽性）化合物または効果がない（陰性）証拠がある化合物（理想的には標的生物または適切な比較種由来）。 2) 関心対象のエンドポイントについて、少なくとも 1 つの信頼できる方法で試験された一般的な化合物セット。

* これらのセット間では該当する化合物が重複している場合がある。

NAM の科学的妥当性を実証するうえで重要な側面は、その性能を現在使用されている既存の試験法と比較して評価することであり、その際、多くの場合、生物学的活性が十分に評価され、理解されている生物学的エンドポイントの参照物質を試験する。生物学的エンドポイントの参照物質には、一般に、理想的にはさまざまな強度の有害作用が生じること、またはそのような作用が生じないことが実証されている化合物が含まれる（Browne ら、2019）。

特定の最終アウトカムに対する作用が知られている生物学的エンドポイント参照物質の同定は困難な場合がある。これは、ヒトの場合には特にそうである。毒物学の分野は、潜在的に危険な化学物質へのヒトの曝露を防ぐために応用されてきたため、データが不足していることが多いからである。この取り組みの成功は、ヒトに関して入手可能なデータが限られていたか、そのような物質がヒト集団で十分に特性評価または測定されていなかったことを意味する。このようなヒトに関するデータが入手できる場合、偶発的な曝露または低レベルの曝露によるものである可能性があり、どちらも曝露レベルまたはアウトカムに関する限られた定量的データしか得られないことがよくある。ケーススタディは危険有害性の兆候を示すことができるが、明確な因果関係の確立には不向きなことが多い。

これらの課題はあるものの、開発中の NAM によって予測されている結果またはメカニズムに関連する人体への影響を持つ化合物を同定するために努力が払われるべきである。このようなデータでは方法の定量的評価を裏付ける能力が限られている場合でも、NAM の関連性を定性的に確認できる場合がある。また、質の高いヒト疫学データを手に入れたり、毒物学データと一致させたりすることも困難な場合があるが、疫学データは、ヒトの健康への悪影響と強い関連がある環境化学物質を特定するのに役立つ。膨大な量の情報をまとめるために、エビデンスベースのフレームワークが開発され、その

後、より狭く関連性の高いデータセットを透明かつ客観的に選択できるようになり（Wikoff ら、2020）、これによってデータのアクセシビリティが向上した。これらのフレームワークには、エビデンスベースの広範な概要を提供する系統的エビデンスマップと、その後の系統的レビューが含まれる。これにより、特定の研究課題をより狭く包括的に評価できるため、人間の疫学データを他の既存の情報源とより迅速に統合できる（Wolffe ら、2019）。これらのデータは、メカニズム上のエビデンスや従来の動物データと組み合わせた場合、生物学的エンドポイントの参照物質を規定し、NAM に対する科学的信頼性を確立するのに役立つ可能性がある（例については、Krishna ら、2021 を参照）。

NAM の頑健性と妥当性を評価する際には、可能な限り、対象種に関連する影響を引き起こすことが示されている化合物を、陽性対照として使用すべきである。Petersen (2021) は、陽性対照（例えば、試験を実施するたびに工程内の対照測定を使用して試験される化学物質）を選択する際に考慮すべき特性を説明しており、NAM を評価するための参照物質についても同じ考慮事項の多くが妥当である。陽性と陰性両方の生物学的エンドポイント参照物質、性能評価用化合物、技能評価用化合物を含めるべきであり、同等の注意を払って選択すべきである。文献には陽性結果を強調する歴史的偏りがあるため（文献で報告されている陽性反応のすべてが真陽性であるとは限らないことを認識して）、標的エンドポイントにとって陰性であり信頼できる参照物質の同定はより困難な場合がある。陰性結果に焦点を当てたデータ厳選の取り組みは、公表することを強く検討すべきであり、幅広い標的にわたって比較的高濃度までの化合物を試験する取り組みは、この問題の解決に役立つ貴重な情報を提供する可能性がある。

すべての COU が 1 つの明確なリストで網羅されているわけではないため、参照物質の選択では、その COU の特定の規制ニーズを考慮すべきである場合によっては、一連の NAM の各々について、評価対象のメカニズム（最終エンドポイントではなく）に関して生物学的エンドポイントの参照物質のサブセットを規定することが適切な場合もある。たとえば、発生神経毒性の参照物質として適格な化合物はほとんどなく、代わりに一連のプロセスで評価される各々のプロセス（増殖や神経突起伸長など）には、独自の試験固有の生物学的エンドポイント参照物質がある。その結果、生物学的エンドポイントの参照物質は、ある試験では陽性であるが、異なるメカニズムまたはプロセスを評価する別の試験では陰性となる場合がある。

OECD が示す正式な妥当性検証プロセスでは、参照物質のリスト（技術移転性の評価などのため）が必要である。ただし、信頼性の高い参照物質の編集には時間がかかり、リソースも大量に消費される。厳密な系統的レビュー、および可能な場合は自動化プロセスを通じて、厳選された参照物質リストの迅速な開発を促進する取り組みにより、NAM の感度と特異性のより確実な評価が可能になっている（Judson ら、2019; Thomas ら、2019）。NAM 適格性確認のための生物学的エンドポイント参照物質リストの例は、2021 年 FDA 改訂版の医薬品規制調和国際会議 (ICH)、S5 (R3) のヒト用医薬品の生

殖毒性および発達毒性の検出に関する業界向けガイダンス（FDA、2021b）に記載されている。このリストには、陽性と陰性のアウトカムを示す化合物が含まれている。ガイダンスには、これらの化合物は他の化合物と同様に、特定の COU の代替試験または一連の試験の適格性を裏付けるために使用できることが記載されている。

3.2.3 既存の実験動物法との比較

NAM に対する科学的信頼性を確立し、規制当局の承認を得るための基準には、一般に、NAM が規制上の意思決定に使用される既存の試験法と同等またはそれ以上の有用性、科学的品質、および/または関連性に関する情報を（各機関の規制の枠組み内で適切に）提供できるかどうかの検討が含まれている。たとえば、OECD GD 34 では、新しい試験法を検証する際の基準として、「その方法は、少なくとも既存の方法を使用して得られたものと同様か、できればそれよりも有用なデータをリスク評価目的で生成する。これにより、ヒトの健康や環境に対する保護レベルが同等かそれ以上になる」

（OECD、2005）としている。改正された米国有害物質規制法（セクション 4 (h) (1) (B)) には、NAM に関する具体的な考慮事項が含まれており、EPA は「規制上の決定に役立つ同等またはそれ以上の科学的品質と関連性のある情報を提供しながら、脊椎動物の使用を減らすか、それ以外に置き換える科学的に有効な試験法および戦略の使用」を奨励および促進することを義務付けている（15 USC §2601、2016）。

これまで、「同等かそれ以上」という概念は、従来の動物実験データとの直接的な比較に依存してきた。しかしながら、「それ以上」の情報という目標を達成するには、NAM は従来の動物試験法で生成されたのと同じ情報を提供しない可能性があり、NAM の結果が従来の動物試験の結果と直接一致しない可能性があることを認識しておく必要がある（たとえば、Clippinger ら、2021; Hoffmann ら、2018、2008; ICCVAM、2018; Kolle ら、2017; Petersen ら、2022b ; Piersma ら、2018 ; Prior ら、2019 ; Sewell ら、2017 を参照）。場合によっては、NAM は、規制上の意思決定プロセスに適した生物学的妥当性の情報、メカニズムの洞察、または十分に感度の高いエンドポイントを提供する場合があります。従来の動物試験法のデータとの比較は不要な場合があります。たとえば、一部の薬理学的または毒物学的標的はヒト以外の種には存在しない場合があるため、動物実験は、そのような標的を介して仲介される潜在的なヒトへの影響を評価することには適さない場合があります。このような場合、ヒトベースの NAM が生物学的に関連する情報を提供する可能性がある。さらに、NAM は、動物実験法で測定された最終エンドポイントのデータではなく、メカニズム上の情報を提供することが多い（例えば、動物で観察された体重の減少は毒性の根底にあるメカニズムを解明できないかもしれないが、NAM はこれらのメカニズム上の洞察を提供できる可能性がある）。また、動物モデルが、COU に関連する複雑な生物学的エンドポイントを測定している場合でも、問題の NAM では十分に網羅されていない状況もある。このような場合、特にメカニズムが完全に理解されていない場合には、従来の動物実験方法との比較が最も目的に適う選択肢となる

可能性がある。NAM の評価の目的は、NAM が既存の方法に基づいて下されるのと同様の規制上の決定につながる情報を提供することを実証することである。

NAM と既存の実験動物法との比較では、可能であれば、両方の種類の方法で観察された差異の理由を理解することを含め、参照動物実験方法の信頼性と反復性を適切に考慮すべきである。利用可能な場合は、対象種の参照データを使用することで、種に関連する反応に対する NAM の評価が可能になる。*in vivo* の時間経過データが利用可能な場合は、生体模倣システムなどの NAM を使用して毒物動態学をシミュレートし、生体模倣システムのコンピュータによる外挿を動物またはヒトのデータと比較することができる。生理学に基づいた薬物動態モデリングにより、基準用量から内部濃度を推定し、*in vitro* の活性濃度と比較することができる。これらのモデルを逆向き線量測定法で適用して、*in vitro* - *in vivo* 外挿 (IVIVE) を行い、NAM で生物活性が観察される血漿または標的組織の濃度が得られる同等の投与量を予測することもできる。ヒトが対象種である場合、効果に関する高品質の疫学的、臨床的、または観察的エビデンスは、NAM の科学的信頼性の構築に関する洞察を提供するのに役立つ可能性がある。ただし、このようなヒト参照データは、ほとんどのエンドポイントや化学物質について入手できることはまれであり、比較は異なる生物種のデータに頼らざるを得ないことがよくある。動物実験で得られたデータを厳選して比較することで、反復性、頑健性に優れた関連性の高い結果が得られる参照標準リストが得られる。生態学的应用など一部の状況では、分子標的の配列相同性や生物学的メカニズムの保存が、異なる種に由来するデータとの比較に役立つ場合がある (Farmahin ら、2013; LaLone ら、2016)。

いくつかの公表論文は、さまざまなエンドポイントにおける動物ベースの参照試験法の結果を評価し、これらの試験の結果がさまざまなレベルの反復性を実証することを示している (Browne ら、2018; Dumont ら、2016; Karmaus ら、2022; Kleinstreuer ら、2018; Luechtefeld ら、2016; Pham ら、2020; Rooney ら、2021)。観察された変動は、たとえば、固有の生物学的要因、報告の誤り、プロトコルの違い、または試験物質の純度の変動を反映している可能性がある。参照動物試験法と NAM データとの間で相反する結果は、対象種の生物学への参照を含め、可能な限り説明すべきである。これは、実験動物データの能力に関する現実的な背景を提供するために必要であり、したがって、参照試験法と比較される NAM の最大性能について適切な期待値を設定するために必要である (Browne ら、2019)。

NAM を従来の動物実験と比較することの相対的価値と、NAM の対象種との生物学およびメカニズム上の妥当性は、入手可能なデータの質に基づいて検討すべきである。現実的には、多くの場合、NAM の信頼性を確立しようとする人々は、これまでの動物実験の利用および既存の方法との比較の必要性を重要な証拠の 1 つとして考慮する必要がある。多くの法的要件や規制要件の下で、NAM から得られた情報は、既存の方法よりも保護性が劣ってはならない。理想的には、この方法はより予測的になり、データが不足していたり、存在しない可能性がある多くの化学物質について、関連データをより

迅速かつ包括的に生成できる可能性があり、規制当局や関係者のコミュニティ間の信頼が高まることが期待される。この規則の例外を認める特定の COU が存在する場合もあるが、そのような NAM が適格な方法として承認されるには、関連機関または規制当局の特定の要件に依存する。このような比較検討の範囲は、NAM と参照方法の両方の入手可能なデータの量と質、並びに対象種との関連性の理解の深さによって異なる場合がある。参照動物試験法に基づいて NAM の性能を評価する場合、意図した COU に基づいて、NAM の許容差の量を、*in vivo* データで観察された変動と相対的に考慮する必要がある。

3.3 技術的特性評価

技術的特性評価は、広く使用される NAM を開発する上で重要な要素である。これには、NAM の変動要因の評価、関連する対照測定を含む試験の設計、試験を使用できる試験物質の範囲（つまり、適用範囲）の評価、および適切な統計データ分析アプローチの開発が含まれる。NAM の科学的妥当性を実証する際の重要な側面は、その試験が技術的観点から特性が十分に評価されて頑健性、信頼性、反復性を保証することである。NAM には、試験システム自体とエンドポイントの定量化方法の両方が含まれる。技術的特性評価には、生物学的妥当性など、これまでの重要な概念と重複する側面があり、さまざまな目的のための参照物質を選択する際に役立つ情報を提供する（表 4）。関連する生物活性を持つ明確に定義された一連の生物学的エンドポイント参照物質に対する NAM の性能の検証はすでに紹介されているため（セクション 3.2.2）、このセクションでは、品質ツール（Ishikawa、1985）、方法開発、文書化、標準などの技術的側面に焦点を当てる。

3.3.1 選択された品質ツールの組み込み

全体として、NAM の技術的特性評価は、1) NAM の科学的関連性の初期審査、2) 概念評価、3) 試験施設内評価、4) 統計データの分析と報告、5) 試験施設間評価（必要な場合¹⁾）という包括的なステップを含むフレームワークに適合する（図 2）。ステップ 2～5 の結果は相互に関連し、各ステップの結果が他のすべてのステップに影響を与える可能性がある。たとえば、試験施設間試験の結果から、追加の試験施設内試験を実施する必要があることや、新しい対照測定を含むようにプロトコルを改訂する必要があることが明らかになる場合がある。

¹大規模なロボット化されたハイスループットスクリーニング試験や機械学習ベースの *in silico* アプローチなど、試験施設間の評価が不要または実行不可能な方法例が存在する可能性がある。

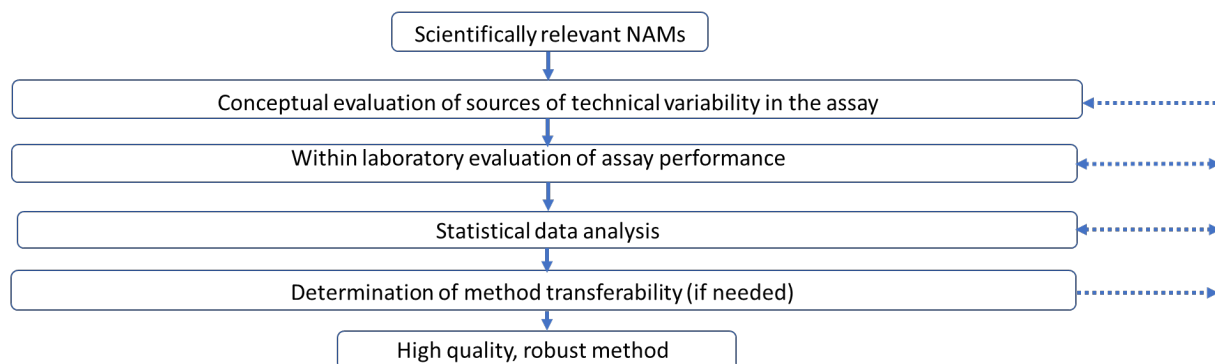


図 2. 頑健な NAM を開発するためのフレームワーク。実線は、推奨順序で実行すべき手順を示す。点線は、必要に応じて前のステップを再検討すべき方向を示す。出典: Petersen ら (2022b) の許可を得て転載。

最初の審査段階には、試験の生物学的妥当性、その潜在的 COU、および十分なレベルの技術的品質に達する可能性の評価が含まれる。NAM に明確な COU や生物学的妥当性がない場合や、解決できない可能性のある重大な技術的品質上の問題がある場合、NAM は規制用途に適さない可能性がある。概念評価段階では、NAM をレビューして予想される変動要因と設計管理測定値を評価する。試験施設内評価では、実験を行って試験の頑健性を評価したり、対照測定値の標準範囲（陰性、溶媒、陽性対照など）を設定したり、対照測定値と試験結果との間に相互作用があるかどうかを確認したりすることができる。概念評価と試験施設内評価から得られる情報により、平均の試験結果を評価し、ばらつきを特徴づけ、陽性または陰性反応の基準と、この決定に対する統計的信頼性を確立するための統計モデルの設計が可能になる。

最後に、一部の NAM では、試験施設間の試験を通じて技術移転性の評価を受けることが必要になる場合がある。試験施設間試験により、試験施設間で解釈が異なるプロトコルの手順や、NAM の技術的品質を向上させるために行うことができる改訂が明らかになる可能性がある。技術移転性の試験では、プロトコルの技術的実装や実行方法の違いから生じる問題（たとえば、ある試験施設のピペッティングと別の試験施設の自動化など）から生じる変更を必要とする問題を特定することもできる。

表 5 は、NAM の技術的特性評価に役立ついくつかの品質ツールを示す。このトピックのより詳細な扱いは、Petersen ら (2022b) に掲載されている。これらのツールから生成された情報の中には、試験法開発者により内部に残るものもあれば、機関の審査を支援するために必要になる場合がある。

表 5. NAM の技術的特性評価によく使用される適格性評価ツール*

適格性評価ツール	説明	利点/有用性
フローチャート	プロトコルの各ステップを図式化する。	対照測定値によって監視できるプロトコルステップの特定と対象範囲を最適化する。
因果関係分析	文献と試験の背景をレビューし、予想される変動の原因をすべて図式化する。	変動の主な原因と、標準化が難しい方法の側面を特定する。頑健性試験と対照測定を選択をガイドする。
管理図	技術的性能を評価するための対照測定には、1 回限りの予備実験、所定の頻度での定期的な測定、および試験を実施するたびに行われる工程内対照測定が含まれる。管理図は、経時的に対照測定値を監視する。	潜在的なバイアスを検定し、機器の性能と校正を評価し、時間の経過や実験間の変動の主な原因を測定する。
チェックシート	主要データ、メタデータ、および対照測定値を記録する。	工程内の対照測定値を監視し、問題が発生した場合のトラブルシューティングをサポートし、データ分析と反復性をサポートする。
散布図	対照測定値と試験物質結果のすべてのデータをプロットする。	異なる工程内対照測定値の間、またはそれらの対照測定値と試験物質の結果との間に相互作用があるかどうかを評価する。

*詳細と例は付録 B に記載されている。

3.3.2 品質管理のベストプラクティス

このセクションでは、DA や試験の実施と評価のための統合的アプローチ (IATA) など、連邦政府機関が NAM の評価のベストプラクティスとして検討する可能性のある技術的要因について説明する。ここに記載されている要素は、すべての NAM に適用される場合と適用されない場合がある。さらに、ガイダンスに記載されていないその他の要

素（追加情報や試験など）は、機関が NAM を評価する際に必要となる場合がある。開発者やスポンサーは、審査に必要となる可能性のあるその他の要素を決定するために、方法またはデータを提出する予定の連邦機関と直接連絡を取ることが推奨される。

規制評価のための NAM の提出には、試験施設内または試験施設間試験を実施する場合、その説明を含めるべきである。NAM が参照方法（別の NAM、*in vivo* 動物データ、またはヒトデータ）と比較されたかどうかを記載すべきである。比較に使用されている参照方法に関する情報（つまり、参照データセット）も評価中に必要になる場合がある。NAM の評価は複数の様々なプロセスを経て行うことができ、特定の試験施設の調査を用いて、方法の性能特性が目的の用途に適しており信頼できることの文書化が含まれる。データの許容可能性は、その方法の評価に使用された基準に直接関係する。

NAM の性能を文書化し、適用性を明らかにすることに加えて、科学的、技術的、品質的に優れたプラクティスにより、評価工程全体が効率的かつ効果的になり、提案された方法への信頼性が高まる。開発者は、機器とソフトウェアのマニュアル、品質と安全性の確認証明書と保証、ソフトウェアバージョンの文書、モデリングアルゴリズム、厳選されたデータベース、トレーニングセットなど、あらゆる機器の校正、操作、および保守に必要なすべての情報を保持すべきである。開発者は、材料、セル、試薬の供給業者の文書も保管すべきである（詳細についてはセクション 3.3.3 を参照）。品質プラクティス、機器手順、および保持すべき文書の詳細については、優良インビトロ試験基準に関する OECD ガイダンス文書（GIVIMP、OECD、2018）を参照されたい。使用する各機器の設備据付時適格性評価設置適格性/運転時適格性評価操作適格性/性能適格性評価性能適格性（IQ/OQ/PQ）報告書が推奨され、試験が優良試験所基準（GLP）に準拠することを目的とする場合は必須である。

3.3.2.1 細胞/組織法の関連情報

どのような試験法でも使用する細胞や組織の起源の記録を維持すべきである。レビュー用に保持する情報には、理想的には以下を含めるべきである（ただし、これらに限定されない）。

- 倫理的、法的、安全上の考慮事項
- 種/系統/性別
- 人口統計情報（該当する場合）
- 情報源/供給元
- ドナー数
- 臓器/組織の起源
- 分離された細胞タイプ
- 人工分子標的の配列（該当する場合）
- 分離技法と日付
- バイオセーフティ分類

- 細胞株の同定と認証
- 感染性病原体の血清学的試験（例: マイコプラズマ試験）
- 細胞継代数/集団倍加数
- 倍加時間
- 遺伝的安定性
- 遺伝およびタンパク質発現情報（該当する場合）
- 前処理

より詳しい情報については、Good Cell and Tissue Culture Practice 2.0 に関する GD（Pamies ら、2022）および OECD GIVIMP 文書（OECD、2018）を参照されたい。

試験に使用する前に、試験施設での適切な保管条件で新しい細胞や組織を隔離することに関する文書を提供すべきである。生細胞または組織を利用する試験には、細胞生存率試験を含めるべきである。

すべての実験装置（プレートリーダー、インキュベーター、冷蔵庫/冷凍庫など）の性能に関する記録と文書は、参加しているすべての試験施設が維持すべきである。

セルバンクに細胞を登録すると、トレーサビリティと長期的な可用性に役立つ可能性がある。セルバンクには、細胞密度と細胞継代数の文書を含むべきである。試薬調製の記録を維持すべきである（例: 使用した試薬や消耗品を確認するためにチェックシートを使用するなど）。これは、細胞/組織を使用しない *in vitro* 法にも当てはまる。例として、以下が挙げられる（ただし、これらに限定されない）。

- サプライヤー
- カタログ番号
- バッチ/ロット番号
- 調製日
- 有効期限
- 分析者名

試験の依頼主または開発者は、開発中の方法または方法論に関連するすべての安全情報、およびすべての危険物の使用、輸送、および廃棄に関するすべての関連規制を提供すべきである。

3.3.2.2 NAM で使用される分析方法の評価

開発者は、検出または定量された物質が目的の化学物質または分析対象物であることを明確に示すデータを提供すべきである。表 6 に、分析方法に関連する最良の品質管理プラクティスとその説明を示す。

表 6. 分析方法評価*

分析方法評価	説明	利点/有用性
検出と定量化の限界	試薬ブランクを超える、または標準曲線内で確実に検出または定量できる検体の最低量または濃度。標準曲線上の最高濃度が定量の上限を決定する。	分析法の範囲を設定する。
干渉の識別	方法のコンポーネントによって検出された信号が誤って変更された場合を識別する。	方法のコンポーネント間に相互作用があるかどうかを判断し、アーティファクトの結果が報告されないようにする。
分析精度の評価	使用された分析方法の精度、およびその他の精度試験（たとえば、異なる担当者が提案された方法を使用する場合や、その方法に異なる機器を使用した場合の性能の変動を評価する試験など）の特性を評価する。これは試験施設間の比較研究によって評価できる。	分析方法の信頼性に対する自信を高め、試験施設のばらつきの原因を評価する。
NAM に使用される材料の安定性	NAM で使用される材料（試験物質、試験装置、曝露システム、試薬、分析物など）が、特定の環境で一定期間にわたって同様に許容できる結果を生成する能力。	方法の材料の特性を評価し、特定の方法について信頼性の高いデータを一貫して取得する。
頑健性試験	得られた結果に予期しない違いが生じることなく、さまざまな条件または状況下で方法を再現できるかどうかを試験する。	試験が適切に機能するパラメータの範囲を決定する。

分析方法評価	説明	利点/有用性
回収の分析	抽出されたサンプルの結果を、類似のマトリックスのスパイクサンプルやスパイクブランクの結果と比較することにより、抽出方法を試験する。	抽出方法の効率と反復性を検証する。
適用範囲の技術分析	試験の対象となる特定の COU に関連する化学物質および/または製品の適切な試験法データを取得し、適用範囲の物理化学的特性を明確に記述する。	さまざまな化学物質や製品を使用する場合の試験性能に関する不確実性を減らし、化学物質がいつ適用範囲内にあるかを判断するための方法と基準を提供する。
陽性対照	エンドポイントに関連する陽性対照化合物を試験の検出ウィンドウ内で同定する。	試験物質の結果を比較するための一貫性のある信頼できる基礎を提供する。
機器校正の参照基準	校正標準や品質管理サンプルを使用して機器を校正するために使用する。	信頼性の高い測定を保証し、潜在的な不確実性の原因を特定する。
規格の設定	統計的アプローチを使用して、試験施設内および/または試験施設間の試験結果に基づいて、工程内対照測定の規格を設定する。	十分に厳しい対照測定基準を保証することで、確実な試験物質結果が得られ、偏りのない結果が得られる。

*詳細と例は付録 C に記載されている。

3.3.2.3 NAM の正確性および/または性能標準との一致度の評価

正確性と一致度は定義上非常に似ており、特定の文書や文章の文脈によっては同義であると思われることがよくある。多くの場合、一致度は、得られた結果に基づいて 2 つの方法または試験を比較することと定義され、正確性は、多くの場合、方法または試験を参照方法または試験結果と比較することと定義される。正確性や一致度の評価は、試験

法や関連データの統計的評価によく使用される。この目的では、正確性とは、結果の総数に占める正しい予測の割合として定義される。方法の正確性または一致度を議論する際に使用されるその他の統計的パラメータには、感度、特異度、陽性と陰性の予測度、偽陽性率と偽陰性率などがある。

可能な限り、明確に定義された性能標準（たとえば、陽性と陰性の結果をもたらすことが知られているバランスの取れた参照物質のセット）を使用して、反応と分析法の妥当性を確認することができる。一致の定量的指標（すなわち、感度、特異度、陽性および陰性の予測度、偽陽性率、偽陰性率）を報告すべきである。提案された試験法を、機能的にもメカニズム的にも類似したデータを生成する性能標準が確立された方法（OECD 試験法ガイドライン (TG) の「ミートウー（追随）」方法など）と比較する場合、両方の方法の一致（不一致データを含む）を互いに、また参照方法と比較して評価すべきである。

3.3.2.4 標準操作手順書と方法の詳細

提案される試験法には、試験システムの取り扱いや装置の校正など方法や関連する試験施設の活動の一貫した実施をサポートするために、十分に文書化された標準操作手順書 (SOPs) を組み込むことが推奨される。SOPs は、テストと分析のあらゆる側面を網羅すべきである。SOPs には以下を含めるべきである

- 試験品の完全性を保証するアカウントビリティシステム（記録管理、セキュリティ、サンプルチェーン管理など）。
- 方法、試薬（該当する場合、メーカー、カタログ番号、ロット番号などを含む）、機器、機器などのサンプル調製および分析ツール。
- 品質管理と結果の検証の手順。
- 製品の詳細な説明と製剤、曝露システム、試験物質の量/重量/溶解度/投与プロトコル、および適切な曝露/用量範囲を含む方法の詳細。

提出者は、試験の性能と結果を判断するための動作特性と運用基準のリストも提供すべきである。NAM を構成する技術システムの運用情報と基準は異なる場合があるが、基準には、すべての管理、基準、曝露（投与量と期間）、および実験グループの品質管理チャートまたはその他の性能標準が含まれる場合がある。データを評価するために使用される統計的手法の説明を含めるべきである。さらに、開発者は、実験の不確実性、統計的不確実性、干渉、および背景がどのように評価されたかを説明すべきである。

3.3.3 文書化

このセクションでは、NAM（DA と IATA を含む）の文書化のベストプラクティスについて説明する。ここに記載されている要素は、すべての NAM に適用される場合と適用されない場合がある。さらに、機関が NAM を評価するために追加の書類が必要になる

場合がある。開発者やスポンサーは、審査に必要となる可能性のあるその他の要素を確実にするために、方法またはデータを提出する予定の連邦機関と直接連絡を取ることが推奨される。一部の機関では、機関による審査の前に、NAM と関連データの独立した査読が必要になる場合がある。

新しい方法については、提案された試験法の説明と、それが規制上の目的にどのように関連するか、または特定の COU にどのように適合するかについての説明を文書に含めるべきである。関連情報には、試験法のメカニズム上の情報および生物学的妥当性、および提案されている COU（リスク評価への適用など）が含まれる（ただし、これらに限定されない）。人体または適切な分類群の適用範囲を文書に含めるべきである。文書化に関するその他の重要な点については、以下で詳しく説明する。

文書化に関するガイダンスを提供するその他のリソースもある。*in vitro* 法については、開発者は OECD GIVIMP 文書（OECD、2018）を参照できる。

3.3.3.1 試験物質の同一性と純度

NAM で試験されたすべての物質（対照、参照物質、その他の試験物質など）について、少なくとも物質の同一性、理想的には固有の識別子（CASRN、SMILES、InchiKey など）、およびサプライヤーから提供された物質の純度に関する情報を報告すべきである。試験施設に必要な能力があれば、分析方法を使用して物質の純度と同一性を評価する追加の品質管理措置を実施することには、さらなるメリットがある。一部の試験物質（組成が不明または変動する複雑な混合物を表す環境サンプルなど）は十分に特性評価されていない場合があるため、この情報を収集できるかどうかは、COU によっても異なる。

3.3.3.2 方法の開発

方法から得られたデータに基づいて、その方法に関する具体的で詳細な書面による説明を作成すべきである。これは、プロトコル、研究計画、レポート、および/または SOPs の形式にする可能性がある。方法の各ステップは、環境、マトリックス、または手順の変数が分析対象物の検出および/または定量にどの程度影響するかを判断するために調査すべきである。

開発中は、次のような要因に細心の注意を払う必要がある。

- 試薬の選択（適切な生物学的妥当性、特異性、および安定性を備えること）。
- 検出方法または検出機器（つまり、性能および校正手順がすぐに利用可能）。

- 消耗品（マイクロタイタープレートやその他のプラスチックなど）と試験測定値および試験物質との適合性。
- 分析方法/統計的手法。
- 試験のばらつきを引き起こす可能性のある手順または工程。

特に開発中に使用された化合物、マトリックス、または機器が方法の技術的特性評価に使用されたものと異なる場合は、方法のアプリケーション全体にわたって外部またはマトリックス効果を最小限に抑える（または少なくともそれらの効果の特性を評価する）ための適切な措置を講じるべきである。

3.3.3.3 エンドポイントとパラメータの測定

各エンドポイントまたは分析対象物の測定値は、十分に試験して文書化すべきである。新しい NAM の方法開発には、その方法がすべての関連パラメータを正常に測定できるという実証を含めるべきである。定量的データを実験条件や試験化学物質や濃度などの定性情報にリンクし、試験のばらつきやバッチ効果に寄与する可能性のある外部要因（実験の実施日や実施技師など）を追跡するには、各実験に関連するメタデータをキャプチャして記録すべきである。これらのデータとメタデータは、妥当性検証の独立したレビュー時に参照できるように、アクセス可能な形式でエクスポートして保存すべきである。このトピックについては、付録 B のコントロールチャートとチェックシートに関するセクションでも検討される。

3.3.3.4 使用の制限

試験法の具体的な長所と限界を明確に特定し、説明すべきである。潜在的な干渉源をリストアップし、試験を妨害する可能性のある化学物質または化学物質のクラスを特定すべきであるまた、文書には、NAM を使用して試験できる材料に関する既知の制限も明記すべきである。

3.3.3.5 明確に定義されたエンドポイント

試験法によって生成されたデータは、対象となるエンドポイントを適切に測定または予測すべきであり、そのエンドポイントは、セクション 3.2 で説明するように、生物学的妥当性の説明とともに明確に定義すべきである。その一例が、AOP の特定の主要事象に関する情報を提供する NAM である。データには、新しい試験法と既存の試験法との関係、または新しい試験法と標的生物種への影響との関係についても記載すべきである。システムの安定性を確保するために、NAM の陽性、陰性、または不確定な結果の基準を明確に定義し、時間をかけて評価すべきである。

3.3.3.6 統計モデルの構築

統計モデルは、試験施設内試験のデータを使用して構築できる。これらのモデルは、ベイズ統計学または頻度主義統計学いずれかのアプローチを使用して構築できる。ヒストグラムは、工程内の対照測定で得られたデータの分布を評価し、どのタイプの分布（正規分布など）がデータに適合するかを評価するために使用できる（詳細については、Petersen ら、2022b を参照）。試験物質の変動に関するデータだけを使用するのではなく、試験物質の測定値と工程内対照測定値の両方から NAM の累積変動性を計算するモデルを開発すると便利である。この情報を使用して、決定（たとえば、試験物質が陽性か陰性か）とその決定の統計的信頼性を導き出すことができる統計モデルを構築できる。試験物質評価の平均値を閾値と単純に比較しただけでは、試験結果のばらつきは考慮されず、決定に対する統計的信頼性は得られない。

NAM の統計モデルを開発する際の主な懸念事項の 1 つは、「陰性」と「弱陽性」の結果をどのように区別するかである。統計モデルの閾値は、入手可能な場合は *in vivo* データで判断できる（Friedman ら、2023; Karmaus ら、2022; Pham ら、2020）。これには、NAM の反復性を評価するために「境界領域にある」化合物を反復試験することが必要な場合がある。たとえば、ガイドライン 497（OECD、2021a）を参照されたい。たとえば、用量反応関係性を評価する統計モデルを使用して、特定の効果を引き起こす出発点または濃度（たとえば、 EC_{50} の値など）、およびそれらの値に関する関連するデータ主導の信頼区間を評価できる。T 検定、Z 係数、またはその他の適切な統計基準などの統計的アプローチを使用して NAM の品質を評価することが適切な場合もある（Zhang ら、1999; Zhang、2011）。

3.3.3.7 試験結果の反復性

該当する場合、提出される情報には、技術的反復性に関する文書を含めるべきである。方法の反復性は、品質管理やサンプルを含む反復測定によって評価できる。この評価には、反復性の評価に用いられる物質の選択の理論的根拠（可能性として実施される試験施設内および試験施設間試験における）と、それらが考えられる試験結果の範囲をどの程度表すかについての議論を含めるべきである。外れ値を特定し、検討すべきである。試験施設内および／または試験施設間のばらつきの程度に関する定量的統計分析または変動係数分析を含めるべきである。過去の対照データ（陰性、陽性、および該当する場合は溶媒）について、中心傾向と変動の測定値を要約すべきである同じ化合物を複数回試験する場合、比較は定量的（たとえば、得られた EC_{50} の値など）または定性的（危険有害性の分類など）であり得る。提案された試験法が、既存の性能標準（OECD TG など）で確立された試験法と機構的にも機能的にも類似している場合は、2 つの試験法の信頼性を比較し、相違点による潜在的な影響について検討する必要がある。

3.3.3.8 データ解釈手順

陽性反応と陰性反応の基準を含むデータ解釈手順は、NAM それぞれで明確に説明すべきである。NAM を DA にまとめるには、客観的で専門家の判断を含まない固定データ解釈手順が必要であり、異なるグループに適用しても同じ結果が得られることを確実にする（OECD、2017）。結論の反復性を確保するために、機械学習モデルなどの計算アルゴリズムやソフトウェア（バージョン番号を含む）の使用を十分に文書化すべきである。

3.4 データの完全性

データの完全性は、NAM から得られる情報の信頼性と信頼性を確保する上で重要な側面である。方法の開発者は、データの完全性と結果の信頼性を確保するために、生データの取得、転送、処理に使用される工程の内部評価を実施してから、それらのデータを外部の独立した関係者に提出して評価と査読を行うことを推奨する。試験は、必要に応じて、GLP の原則（21 CFR § 58、40 CFR § 160、40 CFR § 792、OECD、1998）に従って可能な限り実施されるべきである。さらに、代替法評価に関する毒性学プログラム省庁間センター（NICEATM）などの評価機関は、NAM の開発工程の質と完全性の評価を促進することができる（NIEHS、2023b）。データの完全性を最大限に高めるガイダンスを提供するその他のリソースも入手可能である。*in vitro* 法については、開発者は OECD GIVIMP 文書（OECD、2018）を参照できる。デジタルツールとデジタルデータの管理については、開発者は 2016 年に公開された「科学データの管理とスチュワードシップに関する FAIR 指導原則」（Wilkinson ら、2016）に従うことができる。

3.5 情報の透明性

透明性があると、NAM の使用に対する信頼性が高まるため、政府機関の規制上の意思決定工程や規制当局の承認または適格性確認のペースが速まる。NAM と種、COU、および技術的特性評価との関連性は、査読者、科学界、および一般市民に透明に伝えるべきである。必要に応じて、NAM の COU、生物学的妥当性、および技術的特性評価を説明する査読済みの記事や情報は、オープンアクセスジャーナルに掲載すべきであり、および/または一般に公開される規制文書に要約されるべきである。理想的には、NAM の原則、プロトコル、データの分析とグラフ化に使用される生データファイルとスクリプト、および報告基準は一般に公開されるべきである。知的財産を含む NAM に対して、OECD は、ライセンス契約に関する合理的かつ差別のない条件など、透明性を維持するためのツールを提供している（OECD、2021b）。独自技術または特許取得済みの技術または機器をある方法で使用することで、規制試験のニーズを満たすことができる可能性がある。一部の機関では、規制当局の承認や認定を支援するために、NAM 開発者が専有情報または特許情報を伝える必要性が存在する場合がある。機関固有のガイダンスにより、政府機関の情報ニーズについて試験法開発者に指示することができる。

代替試験法国際協力のパートナー（ICATM；NIEHS、2023c）は、規制当局の承認に向けた規制当局の承認に向けた代替法の追跡システム（TSAR）（EURL ECVAM、2021）を通じて、NAM 評価と査読に関する情報を公開している。TSAR は、規制の目的で使用する標準として認められた基準としての NAM の承認段階を、概要の説明と承認されたプロトコルまたは SOPs とともに示す。TSAR には、提出、妥当性検証、査読、推奨事項、規制当局の承認または適格性確認といったプロセス全体のさまざまな段階とつながりのある NAM に関連する関連記録や文書も含まれている。エンドユーザーがプロセスを理解し、実際の環境で応用できるように、NAM が生成するデータをどのように解釈するか、および関連する承認基準を明確に伝えるべきである。

3.6 独立審査

NAM の COU、生物学的妥当性、および技術的特性評価を裏付ける情報およびデータは、独立した第三者（そのメンバーには利益相反がない）によって科学的にレビューされる場合がある。ただし、必要なレビューレベルは各機関の規制と方針によって異なる。

NAM の評価は、いくつかの異なるプロセスを経て行うことができ、方法の性能特性を文書化して目的のアプリケーションへの適性と信頼性を判断することが含まれる。方法の信頼性には、多くの場合、再現性、反復性、および頑健性が含まれるが、これらに限定されない。ただし、一部の方法のレビューには、本報告書に記載されていない追加情報が必要な場合がある。NAM の性能と適用性に加えて、科学的、技術的、品質的に優れたプラクティスにより、独立審査プロセスが効率的かつ効果的になり、提案された方法に対する信頼性が高まる。試験施設では、NAM の実施に使用される機器の運用と保守に必要な情報（機器およびソフトウェアのマニュアル、品質および安全性の確認証明書など）や、材料、細胞、試薬のサプライヤーの文書など、NAM の評価に関連するすべての情報を保持すべきである。GLP に準拠することを意図した試験では、使用した各機器の IQ/OQ/PQ レポートが関連する場合もある。非 GLP 試験では、機器が意図したとおりに機能することを示すための適切な据付と試験実施に関する文書を保管すべきである。

NAM の生データおよび NAM を説明する情報は、独立した第三者および/または規制機関の意思決定者が確認できるようにアクセス可能にすべきである。NAM の評価と独立した査読は、NICEATM、欧州動物実験代替法評価センター（EURL ECVAM）とその科学諮問委員会、および日本動物実験代替法評価センター（JaCVAM）などの検証機関によって組織される場合がある。NAM を独自に審査できる他の機関や国際機関には、米国連邦殺虫剤、殺菌剤、殺鼠剤法科学諮問委員会、欧州食品安全機関科学委員会、OECD などがある。あるいは、開発者が方法の独立したレビューに資金を提供することもできる（ただし、直接その管理を主導することはできない）。査読済み文献は、試験情報を

科学界と共有するために有用である。また、独立した第三者より正式なレビューを補完して、規制目的の方法の承認と使用を裏付けることもできる。

独立審査の範囲は、COU、規制の枠組み、および評価される具体的な方法によって異なる。開発者が独立した審査のために送る情報には、NAM を他の NAM、*in vivo* 動物、またはヒトのデータと比較したかどうかなど、試験施設内または試験施設間の試験の記録が含まれるものがある。これらの評価は、NAM の試験施設間の技術移転性に役立つ可能性がある。

4.0 米国連邦機関による NAM の承認

4.1 規制のニーズと意思決定の背景の理解

連邦機関には、毒物データを要求、取得、および使用するためのさまざまな権限がある。これらは、法令、規制、および製品カテゴリによって異なる。NAM は、一部の法的または規制の状況では有用で適切な場合もあれば、そうでない場合もある。したがって、NAM の開発者とエンドユーザーは、それが使用される目的を考慮する必要がある。考えられる目的には、DA、IATA、またはスタンドアロン方法が含まれるが、これらに限定されない。

NAM が規制目的またはその他の目的で承認されるかどうかは、下される決定の性質、NAM の使用目的に対する NAM の妥当性、および規制当局への提出が安全性、有効性、および/またはリスクの判断を裏付ける NAM の結果にどの程度依存するかに応じて異なる。懸念されるシグナルを特定する高感度で特異度が低い NAM は、多数の化合物をスクリーニングしてさらなる評価を優先させるアプリケーションで役立つ可能性がある。このような NAM は、単一化合物のどの用量が曝露しても安全かを判断する場合にはあまり役立たない場合がある。場合によっては、政府機関がデータ不足の状況で懸念の兆候（危険有害性の特定）を探すことがある。また、政府機関が化合物についてより定量的な決定（リスク評価）を行う必要があり、その決定に利用される NAM からのアウトプットに高いレベルの不確実性があることを認めたくない場合もある。性能特性の実装は、NAM の使用方法を反映すべきであり、各 NAM とその COU に合わせて個別化すべきである。

このプロセスを円滑に進めるために、一部の機関では規制上の意思決定プロセスで NAM を使用する前に評価することを希望する場合がある。評価では、COU が明確に定義されている NAM に焦点を当てる可能性が高い。評価の目的の 1 つは、規制当局が NAM の基礎となるすべての裏付けデータを再検討することなく、NAM を使用して得られた結果を適用できるようにすることである（表 1 に引用されている文書を参照）。

4.2 使用方法と目的 (COU) に関する考慮事項

NAM の目的（危険有害性の特定、強度評価、定量的リスク評価の出発点など）を明確に伝えるべきであり、その目的に基づいて NAM を評価すべきである。新しく開発された NAM のアプリケーションは 1 つの COU に集中することが適切である。ただし、適切な裏付けデータがあれば、後から追加の COU を追加できる。COU は通常、特定の規制ニーズに焦点を当てる必要がある。NAM の適格性確認に向けた最初の取り組みは、COU が狭い場合に最も成功する場合がある。COU は、必要に応じて追加のデータで拡張できる規制ニーズは機関によって異なるため、特定の NAM の COU も異なる場合がある。全面的な適格性確認作業を行う前に、提案された NAM の適切な COU を確立することが不可欠である。

適切な COU の決定は、通常、NAM 開発者と COU が関係する機関との間で検討すべきである。許容できる COU を定義するには、COU を何回か繰り返すことが必要な場合がある。NAM の適用可能性と制限が明確になるにつれて、COU はデータ収集中に変化することもある。

4.3 獲得した経験に基づく信頼性の変化

NAM を規制用途に組み込むには、規制当局と規制対象業界による方法に対する十分な信頼が必要である。妥当性検証と適格性確認はこの信頼性を裏付けるが、実装を確実にするには不十分な場合がある。NAM がその目的のために広く受け入れられるようになるには、通常、NAM に関する教育と経験が必要である。妥当性検証や適格性確認のデータが入手できたとしても、コスト、複雑さ、試薬の入手可能性、NAM の実施と解釈に関するトレーニングを受けたスタッフの存在などが、NAM の採用を制限する要因となる。

既存のアプローチを補完または置き換えることを目的とした NAM の性能は、通常、既存のアプローチと比較する必要がある。多くの場合、豊富な経験がある既存のアプローチには高い信頼性がある。これらの既存のアプローチは正式な妥当性検証を受けていない可能性があるが、既存のアプローチを反復して使用し、動物実験に固有の妥当性を想定することで、アプローチにかなりの信頼性が生まれることがよくある。エンドユーザーは、その結果を使用して安全性に関する意思決定を行う際に、NAM が既存のアプローチと同等かそれ以上になることを知る必要がある。たとえば、優先順位付けやスクリーニング、危険有害性の特性評価、リスク評価へと進むにつれて、信頼性を高める必要があるため、さまざまなレベルの不確実性が許容される場合がある。

NAM の信頼性を高めるためのメカニズムの 1 つは、NAM が置き換えたり補完したりしようとしている既存の方法のデータと並行して、NAM からのデータをエンドユーザーが提供することである。業界内および業界間で NAM に関する情報を自発的に共有することは、方法への信頼を裏付けるのに十分な量のデータを確立するのに役立つ。やが

て、政府機関や業界は、安全基準を損なうことなく、NAM が既存の評価パラダイムにどのように適合するかを理解することができる。

5.0 米国と国際調和

米国連邦機関間、さらに広くは各国の規制当局との調整を行うことで、NAM を検証し、その適用と実施を支援するためのアプローチの調和を確保するのに役立つ。ICCVAM と NICEATM は、国内外のコミュニケーションとコラボレーションを促進する上で重要な役割を果たしている。

5.1 米国の調和: ICCVAM と NICEATM の役割

ICCVAM 承認法は、ICCVAM の以下の目的を示している（42 U.S.C 2851-3、2000; NIEHS、2023d）。

- 米国連邦機関の試験法レビューの効率と有効性を高める。
- 不必要な重複作業を排除し、米国連邦規制機関間で経験を共有する。
 - これは、月例会議、ワークグループ、公開会議（公開フォーラム）、科学諮問グループ（代替毒性試験法に関する科学諮問委員会）など、いくつかの手段によって達成される。
- 米国連邦機関外の科学的専門知識の活用を最適化する。
 - これは多くの場合、ICATM パートナーやカンファレンスで多数の科学者との会議やワークショップを開催することで達成される。
- 新しい試験法や改訂された試験法が、米国連邦機関のニーズを満たすように検証されていることを保証する。
 - 連邦機関の科学者と規制当局は開発者と協力して、開発中の方法が規制上のニーズに対応することを保証する。
- 可能であれば、試験に使用する動物の数を減らしたり、試験に伴う苦痛を削減したり、動物を使用しない方法に置き換えたりする。

ICCVAM は、動物実験の使用動物数の削減、動物使用に伴う苦痛の低減、動物を使用しない方法への置き換えを促進する代替法の開発、規制当局の承認または適格性確認、並びに使用を促進する省庁間および国際的な協力を促進する。ICCVAM は、試験法開発者にガイダンスを提供し、代替毒物試験法の専門家による査読からの推奨事項を評価し、レビューされた試験法の使用に関する推奨を適切な連邦機関に提供する。

ICCVAM は、NICEATM が管理する特別な技術作業グループを通じてその機能を果たし、動物実験の代替案の開発または妥当性検証に重要な特定のタスクを実施する。そのような継続的な例の 1 つは、化学スクリーニングのための *in vitro* ヒト甲状腺微小組織

試験に基づいて EPA が開発した NAM の試験施設間の事前検証試験の調整における ICCVAM の支援である（Deisenroth ら、2020）。

NICEATM は、国立環境衛生学研究所 (NIEHS) 内のトランスレーショナル毒性学部内の部門であり、ICCVAM および ICCVAM 作業グループの活動、査読パネル、専門家パネル、ワークショップ、および妥当性検証作業に関する技術的、科学的、運用上のサポートを提供する。ICCVAM のサポートに加えて、NICEATM は以下を実施する。

- 試験法の分析と評価を実施し、新規かつ優先度の高い代替試験アプローチに関する独立した妥当性検証研究を調整する。
- NICEATM Web サイト、統合化学環境、および関心のあるトピックに関するワークショップを通じて、試験法開発者、規制当局、規制対象業界に情報を提供する。
- NIEHS トランスレーショナル毒性学部門の活動、特に米国連邦機関間の 21 世紀の毒性試験 (Tox21) コンソーシアムに貢献する活動をサポートする。

ICCVAM により提供される機関間コミュニケーションのおよび NICEATM により提供されるサポートに関するフォーラムは、規制適用対象となる NAM の妥当性検証と適格性確認のための米国連邦機関の取り組みを調整するために限られた資源が効果的に活用されていることを確実にする役目を果たす。

5.2 米国の調和: 3Rs を推進する機関間連携の強化

米国連邦機関は、ICCVAM への参加に加えて、さまざまな方法で NAM に関して連携している。たとえば、Tox21 (Tox21, n.d.) は、EPA、NIEHS トランスレーショナル毒性学部、国立衛生研究所内の国立先進トランスレーショナル科学センター (NCATS)、および FDA による連邦機関の共同研究である。Tox21 の目標は、特定の化合物が人体の生物学的プロセスを妨害し、健康に悪影響を及ぼす可能性があるかどうかを効率的に試験するために、より優れた毒性評価方法を開発することである。Tox21 は、EPA の内分泌かく乱物質スクリーニングプログラム (EPA、2023) など、規制上の用途に活用されている独創的な出版物や分析を数多く作成している。また、NICEATM と EPA は、規制試験における使用動物数を削減や、動物を用いない試験法に置き換えのための複数の研究でも緊密に協力している。これには、1) 農薬および農薬製品に対する皮膚急性毒性 (EPA、2020)、2) 農薬のヒトの健康リスク評価における皮膚吸収因子に対して *in vitro* データのみを使用するだけで十分かどうかの判断 (Allen ら、2021) 3) 非標的水生脊椎動物と同じレベルの保護が、3 種未満の魚種に関する *in vivo* 急性毒性試験で達成できるかどうかの判断 (Ceger ら、2023) について、動物の使用を排除するための後ろ向き分析が含まれる。FDA もまた、NCATS と協力して生体模倣システム技術をさらに開発し、その進歩を促進し、トランスレーショナルの利用を加速させている (FDA、2023)。

5.3 国際調和

ICCVAM を通じて促進される国際的な共同作業に加えて、米国の機関はそれぞれ国際的に協力して NAM の承認を進めている。例として、ICATM、化学品の分類および表示に関する世界調和システム (GHS) に関する国連専門家小委員会との米国の関与、NAM の使用を説明するガイダンスの作成のための ICH との関与、OECD 試験ガイドラインプログラムおよび危険有害性評価作業部会への参加などがある。

ICATM は、米国 (ICCVAM)、日本 (JaCVAM)、欧州連合 (EURL ECVAM)、およびカナダ（カナダ保健省内の環境健康科学研究局）の妥当性検証機関間のパートナーシップとして設立された。その他の参加組織には、韓国代替法検証センター、ブラジル代替法検証センター、中国食品医薬品局および広東省疾病予防管理センターなどがある。このグループの包括的な目標は次のとおりである。

- 代替法/戦略が世界中でより容易に承認されるように、妥当性検証研究、独立した査読、調和のとれた勧告の開発という重要な分野で国際協力を確立すること。
- 規制上の用途に採用される新しい代替試験法/戦略が、科学的に可能な場合はいつでも、動物の使用に取って代わり、使用動物数の減少、または動物福祉（痛みや苦痛の低減）を図りながら、ヒト、動物、環境に対する保護と同等または改善されることを保証するために必要な国際協力を確立すること。

国連内では、GHS 小委員会が作業グループを設立し、GHS のさまざまな章を更新し、さまざまな危険有害性クラス（皮膚刺激性/腐食性、眼刺激性/重篤な眼損傷、皮膚感作性など）における NAM の使用に関する具体的な基準を確立している。これらの取り組みは、国際レベルで NAM を推進する上で実りある成果をもたらし、危険有害性判定のための承認方法として DA を確立している。このような国際的な取り組みは、NAM を推進するだけでなく、この分野で十分な資源を持たない可能性のある国際地域に専門知識を提供する上でも重要である。

NAM の使用に対する調和のとれたアプローチの価値は、ICH の活動にも表れている。ICH が作成したいくつかのガイダンスには、世界中の複数の規制当局や業界団体が許容できるアプローチと見なしている代替法の使用方法が記載されている。これらの方法は、ICH プロセスの専門家作業グループによって評価され、必要に応じてガイダンスに組み込まれた。例として、光毒性の評価のための *in chemico* 法および *in vitro* 法の使用が挙げられる（ICH、2013）。さらに、生殖毒性および発生毒性に関する ICH ガイダンスには、代替試験の使用目的や、そのような試験の適格性を確認するためのアプローチに関する推奨事項、および生物学的エンドポイント参照物質のリストが含まれている（FDA、2021b）。

OECD は、規制試験ガイドラインとガイダンス文書を調和させるための国際フォーラムであり、NAM の妥当性検証と使用にますます重点を置いている。ICCVAM は、OECD の健康影響テストガイドラインプログラムにおける米国の立場を調整し、貢献する上で重要な役割を果たしている。米国ナショナルコーディネーターは、ナショナルコーディネーターの作業グループの年次総会やその他のテストガイドライン作成活動に米国を代表して参加する。その役割において、米国ナショナルコーディネーターは、3Rs のあらゆる側面を含む OECD TG 活動について、関連する ICCVAM 機関からの意見を求める。ICCVAM 機関の対象分野の専門家は、複数の OECD 専門家グループに所属し、試験ガイドライン、DA、ガイダンス文書など、OECD 製品の開発に関する科学的ガイダンスを提供している。OECD のテストガイドラインは、化学物質の安全性を評価するために 38 の OECD 加盟国の関係者が使用している。OECD のデータ相互承認条項により、OECD の試験ガイドラインを使用して生成された安全性データがすべての加盟国で承認され、試験の重複が回避される。ICCVAM 機関は、OECD IATA ケーススタディプロジェクトにも貢献している。このプロジェクトでは、規制の使用目的における化学物質の安全性を評価するために IATA で新しい方法の使用を各国が共有し、協力することができる。

6.0 NAM の使用を促進するためのコミュニケーションとトレーニング

特定の NAM の受け入れ可能性に関する機関からのコミュニケーションや NAM に関するトレーニングは、その利用を促進することができる適切かつ実現可能な場合、政府機関固有の規則や方針にもよるが、NAM がいつ、どのように受け入れられるかを公に伝えることもできる。たとえば、NAM は GD または一般に公開されている Web サイトに記載されている場合がある。規制機関は、既存の研修プログラムを使用するか、職員向けの新しいプログラムを実施し、可能であれば、新しい方法に関する研修を一般に公開することができる。科学界の教育を通じて、NAM の妥当性が検証または適格性が確認される前から、新しいアプローチへの信頼を築くことが可能である。こうした早期教育の取り組みは、新しいアプローチの基礎科学を中心とすることができる。アプローチが成熟し、アプローチの妥当性を裏付けるデータが蓄積されるにつれて、教育活動はこれらのデータに関する知識をコミュニティに提供するように移行することができる。

NAM の特定の使用法と解釈に関するトレーニングは、特定の COU について NAM を評価した後に実施できる。OECD、ICCVAM、その他のバリデーション機関や科学団体などの組織も、NAM に関するトレーニングや情報へのアクセスを提供することで、NAM の使用に対する信頼性を高めることができる。

NAM 開発者、業界ユーザー、規制当局間の交流により、方法の開発と採用が容易になる。このような交流は、そのような方法が議論される科学会議にすべての関係者が参加

することや、機関固有のプロセスに従ったより正式な規制上の交流を通じて発生する可能性がある。前述のように、適切な COU と適格性確認データセットの開発は反復的なプロセスになることがある。データが蓄積されるにつれて、NAM の開発および調査中に、COU や適用範囲が変化する可能性がある。この工程中の全関係者間の継続的なコミュニケーションは、NAM の承認への適切な経路を確実にするために役立つ。

NAM がエンドポイントとして利用可能で、規制当局によって承認された場合でも、申請のスポンサーは従来の動物実験などの他のアプローチを使用することを選択する可能性があるため、NAM は必ずしも規制申請で提出されるとは限らない。このような規制当局の制御が及ばない理由はいくつか考えられる。規制機関は、従来の検査の代わりに 1 つまたは複数の NAM を使用することを提案または推奨できるが、化合物のスポンサーが必ずしもその提案に従う義務を負うとは限らない。すべての関係者の間で NAM の採用に対する十分な信頼を築くには、NAM に関する教育と知識が必要である。

7.0 結論と実装

本報告書の目的は、試験法開発者、規制対象業界の関係者、および連邦政府機関が、科学的に関連する NAM の開発、妥当性検証、適格性確認、承認を支援することである。本書では、目的適合性と信頼性が高く、対象種に関連する情報を提供できる NAM を効率的かつタイムリーに開発するために考慮すべき重要な概念について説明した。すべての情報は、特定の方法、DA、または IATA に適用される場合と適用されない場合がある。方法、DA、または IATA に該当する概念が他にもあるかもしれないが、本書では検討していない。開発者は、特に規制審査におけるリスク評価アプリケーションにおいて、NAM の使用目的を見据えて、連邦機関やエンドユーザーと緊密に連携することが重要である。NAM に対する科学的な信頼を確立し、特定の目的や COU のための方法の妥当性検証または適格性確認を実施することは、関係者間の多方向のコミュニケーションを通じて発展する反復的な工程にすべきである。

NAM の分野は進化を続けており、本報告書の執筆時点では予想されていなかった NAM の妥当性検証と適格性評価に関する新たな考慮事項が浮かび上がる可能性がある。したがって、NAM の開発、妥当性検証、適格性確認に携わる関係者は、本報告しで書で説明されていない考慮事項を柔軟かつ積極的に取り入れる必要があるかもしれない。本報告書は定期的に、また必要に応じて更新される。

参考文献

- 15 USC §2601, 2016. 15 USC §2601: Findings, policy, and intent [WWW Document]. URL [https://uscode.house.gov/view.xhtml?req=\(title:15%20section:2601%20edition:prelim\)](https://uscode.house.gov/view.xhtml?req=(title:15%20section:2601%20edition:prelim)) (accessed 11.2.22).
- 21 CFR § 58, n.d. CFR - Code of Federal Regulations Title 21 [WWW Document]. URL <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=58> (accessed 11.3.22).
- 40 CFR § 160, n.d. CFR - Code of Federal Regulations Title 40 [WWW Document]. URL <https://www.ecfr.gov/current/title-40/chapter-I/subchapter-E/part-160> (accessed 5.9.23).
- 40 CFR § 792, n.d. CFR - Code of Federal Regulations Title 40 [WWW Document]. URL <https://www.ecfr.gov/current/title-40/chapter-I/subchapter-R/part-792> (accessed 5.9.23).
- 42 U.S.C 2851-3, 2000. ICCVAM Authorization Act of 2000 [WWW Document]. 42 U.S.C. 2851-3 Public Law 106-545. URL <https://www.congress.gov/bill/106th-congress/house-bill/4281/text> (accessed 5.9.23).
- Allen, D.G., Rooney, J., Kleinstreuer, N., Lowit, A., Perron, M., 2021. Retrospective analysis of dermal absorption triple pack data. *ALTEX - Alternatives to animal experimentation* 38, 463–476. <https://doi.org/10.14573/altex.2101121>
- Browne, P., Delrue, N., Gourmelon, A., 2019. Regulatory use and acceptance of alternative methods for chemical hazard identification. *Current Opinion in Toxicology* 15, 18–25.
- Browne, P., Kleinstreuer, N.C., Ceger, P., Deisenroth, C., Baker, N., Markey, K., Thomas, R.S., Judson, R.J., Casey, W., 2018. Development of a curated Hershberger database. *Reproductive Toxicology* 81, 259–271. <https://doi.org/10.1016/j.reprotox.2018.08.016>
- Ceger, P., Allen, D., Blankinship, A., Choksi, N., Daniel, A., Eckel, W.P., Hamm, J., Harwood, D.E., Johnson, T., Kleinstreuer, N., Sprankle, C.S., Truax, J., Lowit, M., 2023. Evaluation of the fish acute toxicity test for pesticide registration. *Regulatory Toxicology and Pharmacology* 139, 105340. <https://doi.org/10.1016/j.yrtph.2023.105340>
- Ceger, P., Garcia-Reyero Vinas, N., Allen, D., Arnold, E., Bloom, R., Brennan, J.C., Clarke, C., Eisenreich, K., Fay, K., Hamm, J., Henry, P.F.P., Horak, K., Hunter, W., Judkins, D., Klein, P., Kleinstreuer, N., Koehn, K., LaLone, C.A., Laurenson, J.P., Leet, J.K., Lowit, A., Lynn, S.G., Norberg-King, T., Perkins, E.J., Petersen, E.J., Rattner, B.A., Sprankle, C.S., Steeger, T., Warren, J.E., Winfield, S., Odenkirchen, E., 2022. Current ecotoxicity testing needs among selected U.S. federal agencies. *Regulatory Toxicology and Pharmacology* 133, 105195. <https://doi.org/10.1016/j.yrtph.2022.105195>

- Chang, X., Tan, Y.-M., Allen, D.G., Bell, S., Brown, P.C., Browning, L., Ceger, P., Gearhart, J., Hakkinen, P.J., Kabadi, S.V., Kleinstreuer, N.C., Lumen, A., Matheson, J., Paini, A., Pangburn, H.A., Petersen, E.J., Reinke, E.N., Ribeiro, A.J.S., Sipes, N., Sweeney, L.M., Wambaugh, J.F., Wange, R., Wetmore, B.A., Mumtaz, M., 2022. IVIVE: Facilitating the use of in vitro toxicity data in risk assessment and decision making. *Toxics* 10, 232. <https://doi.org/10.3390/toxics10050232>
- Chiu, W.A., Wright, F.A., Rusyn, I., 2017. A tiered, Bayesian approach to estimating of population variability for regulatory decision-making. *ALTEX* 34, 377–388. <https://doi.org/10.14573/altex.1608251>
- Choksi, N.Y., Truax, J., Layton, A., Matheson, J., Mattie, D., Varney, T., Tao, J., Yozzo, K., McDougal, A.J., Merrill, J., Lowther, D., Barroso, J., Linke, B., Casey, W., Allen, D., 2019. United States regulatory requirements for skin and eye irritation testing. *Cutan Ocul Toxicol* 38, 141–155. <https://doi.org/10.1080/15569527.2018.1540494>
- Church, R.J., Gatti, D.M., Urban, T.J., Long, N., Yang, X., Shi, Q., Eaddy, J.S., Mosedale, M., Ballard, S., Churchill, G.A., Navarro, V., Watkins, P.B., Threadgill, D.W., Harrill, A.H., 2015. Sensitivity to hepatotoxicity due to epigallocatechin gallate is affected by genetic background in diversity outbred mice. *Food Chem Toxicol* 76, 19–26. <https://doi.org/10.1016/j.fct.2014.11.008>
- Clippinger, A.J., Raabe, H.A., Allen, D.G., Choksi, N.Y., van der Zalm, A.J., Kleinstreuer, N.C., Barroso, J., Lowit, A.B., 2021. Human-relevant approaches to assess eye corrosion/irritation potential of agrochemical formulations. *Cutan Ocul Toxicol* 40, 145–167. <https://doi.org/10.1080/15569527.2021.1910291>
- Corley, R.A., Kuprat, A.P., Suffield, S.R., Kabilan, S., Hinderliter, P.M., Yugulis, K., Ramanarayanan, T.S., 2021. New approach methodology for assessing inhalation risks of a contact respiratory cytotoxicant: computational fluid dynamics-based aerosol dosimetry modeling for cross-species and in vitro comparisons. *Toxicological Sciences* 182, 243–259. <https://doi.org/10.1093/toxsci/kfab062>
- CPSC, 2020. Proposed Guidance for Industry and Test Method Developers: CPSC Staff Evaluation of Alternative Test Methods and Integrated Testing Approaches and Data Generated from Such Methods to Support FHSA Labeling Requirements [WWW Document]. URL <https://www.regulations.gov/document/CPSC-2021-0006-0001> (accessed 5.4.22).
- CPSC, 2012. Recommended Procedures Regarding the CPSC’s Policy on Animal Testing [WWW Document]. URL <https://www.cpsc.gov/Business--Manufacturing/Testing-Certification/Recommended-Procedures-Regarding-the-CPSCs-Policy-on-Animal-Testing> (accessed 5.4.22).
- Crofton, K.M., Mundy, W.R., 2021. External Scientific Report on the Interpretation of Data from the Developmental Neurotoxicity In Vitro Testing Assays for Use in Integrated

- Approaches for Testing and Assessment. EFSA Supporting Publications 18, 6924E.
<https://doi.org/10.2903/sp.efsa.2021.EN-6924>
- Daniel, A.B., Strickland, J., Allen, D., Casati, S., Zuang, V., Barroso, J., Whelan, M., Régimbald-Krnel, M.J., Kojima, H., Nishikawa, A., Park, H.-K., Lee, J.K., Kim, T.S., Delgado, I., Rios, L., Yang, Y., Wang, G., Kleinstreuer, N., 2018. International regulatory requirements for skin sensitization testing. *Regul Toxicol Pharmacol* 95, 52–65. <https://doi.org/10.1016/j.yrtph.2018.03.003>
- Deisenroth, C., Soldatow, V.Y., Ford, J., Stewart, W., Brinkman, C., LeCluyse, E.L., MacMillan, D.K., Thomas, R.S., 2020. Development of an *In Vitro* Human Thyroid Microtissue Model for Chemical Screening. *Toxicological Sciences* 174, 63–78.
<https://doi.org/10.1093/toxsci/kfz238>
- Dumont, J., Euwart, D., Mei, B., Estes, S., Kshirsagar, R., 2016. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit Rev Biotechnol* 36, 1110–1122. <https://doi.org/10.3109/07388551.2015.1084266>
- Elliott, J.T., Rösslein, M., Song, N.W., Toman, B., Ovaskainen, A.K.-, Maniratanachote, R., Salit, M.L., Petersen, E.J., Sequeira, F., Romsos, E., Kim, S.J., Lee, J., Moos, N.R. von, Rossi, F., Hirsch, C., Krug, H.F., Suchaoin, W., Wick, P., 2017. Toward achieving harmonization in a nanocytotoxicity assay measurement through an interlaboratory comparison study. *ALTEX - Alternatives to animal experimentation* 34, 201–218.
<https://doi.org/10.14573/altex.1605021>
- EPA, 2023. Availability of New Approach Methodologies (NAMs) in the Endocrine Disruptor Screening Program (EDSP) [WWW Document]. URL <https://www.regulations.gov/document/EPA-HQ-OPP-2021-0756-0002> (accessed 2.20.23).
- EPA, 2021a. EPA New Approach Methods Work Plan [WWW Document]. URL https://www.epa.gov/system/files/documents/2021-11/nams-work-plan_11_15_21_508-tagged.pdf (accessed 5.5.22).
- EPA, 2021b. EPA Strategic Plan to Reduce the Use of Vertebrate Animals in Chemical Testing [WWW Document]. URL <https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/strategic-plan-reduce-use-vertebrate-animals-chemical> (accessed 5.5.22).
- EPA, 2021c. Chlorothalonil: Revised Human Health Draft Risk Assessment for Registration Review [WWW Document]. URL <https://www.regulations.gov/document/EPA-HQ-OPP-2011-0840-0080> (accessed 3.13.23).
- EPA, 2020. Guidance for Waiving Acute Dermal Toxicity Tests for Pesticide Technical Chemicals & Supporting Retrospective Analysis [WWW Document]. URL <https://www.regulations.gov/document/EPA-HQ-OPP-2016-0093-0181> (accessed 3.7.23).

- EPA, 2018. Strategic Plan to Promote the Development and Implementation of Alternative Test Methods Within the TSCA Program.
- EPA, 2015. Use of an Alternate testing framework for classification of eye irritation [WWW Document]. URL https://www.epa.gov/sites/default/files/2015-05/documents/eye_policy2015update.pdf (accessed 11.13.23).
- EURL ECVAM, 2021. EU Reference Laboratory for alternatives to animal testing (EURL ECVAM) [WWW Document]. URL https://joint-research-centre.ec.europa.eu/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam_en (accessed 11.3.22).
- Farmahin, R., Manning, G.E., Crump, D., Wu, D., Mundy, L.J., Jones, S.P., Hahn, M.E., Karchner, S.I., Giesy, J.P., Bursian, S.J., Zwiernik, M.J., Fredricks, T.B., Kennedy, S.W., 2013. Amino Acid Sequence of the Ligand-Binding Domain of the Aryl Hydrocarbon Receptor 1 Predicts Sensitivity of Wild Birds to Effects of Dioxin-Like Compounds. *Toxicological Sciences* 131, 139–152. <https://doi.org/10.1093/toxsci/kfs259>
- FDA, 2023. MOU 225-23-003 Memorandum of understanding between the National Institutes of Health (NCATS) and the Food and Drug Administration (FDA) for the Microphysiological Systems Program [WWW Document]. FDA. URL <https://www.fda.gov/about-fda/domestic-mous/mou-225-23-003> (accessed 5.9.23).
- FDA, 2021a. Advancing New Alternative Methodologies at FDA [WWW Document]. URL <https://www.fda.gov/media/144891/download> (accessed 7.19.23).
- FDA, 2021b. S5(R3) Detection of Reproductive and Developmental Toxicity for Human Pharmaceuticals Guidance for Industry [WWW Document]. URL <https://www.fda.gov/media/148475/download>
- FDA, 2020. Qualification Process for Drug Development Tools Guidance for Industry and FDA Staff [WWW Document]. URL <https://www.fda.gov/media/133511/download> (accessed 7.19.23).
- FDA, 2018. Bioanalytical Method Validation Guidance for Industry [WWW Document]. URL <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/bioanalytical-method-validation-guidance-industry> (accessed 7.31.23).
- FDA, 2017a. FDA’s Predictive Toxicology Roadmap [WWW Document]. URL <https://www.fda.gov/media/109634/download> (accessed 5.5.22).
- FDA, 2017b. Qualification of Medical Device Development Tools: Guidance for Industry, Tool Developers, and Food and Drug Administration Staff [WWW Document]. FDA. URL <https://www.fda.gov/medical-devices/science-and-research-medical-devices/medical-device-development-tools-mddt> (accessed 5.5.22).

- Frick, A., Suzuki, O.T., Benton, C., Parks, B., Fedoriw, Y., Richards, K.L., Thomas, R.S., Wiltshire, T., 2015. Identifying genes that mediate anthracycline toxicity in immune cells. *Front Pharmacol* 6, 62. <https://doi.org/10.3389/fphar.2015.00062>
- Friedman, K.P., Foster, M.J., Pham, L.L., Feshuk, M., Watford, S.M., Wambaugh, J.F., Judson, R.S., Setzer, R.W., Thomas, R.S., 2023. Reproducibility of organ-level effects in repeat dose animal studies. *Comput Toxicol* 28, 1–17. <https://doi.org/10.1016/j.comtox.2023.100287>
- Harrill, A.H., 2020. ToxPoint: In the era of precision medicine, diversity should not be neglected in chemical safety assessment. *Toxicol Sci* 173, 3–4. <https://doi.org/10.1093/toxsci/kfz232>
- Harrill, A.H., McAllister, K.A., 2017. New rodent population models may inform human health risk assessment and identification of genetic susceptibility to environmental exposures. *Environ Health Perspect* 125, 086002. <https://doi.org/10.1289/EHP1274>
- Hartung, T., 2010. Evidence-based toxicology - the toolbox of validation for the 21st century? *ALTEX* 27, 253–263. <https://doi.org/10.14573/altex.2010.4.253>
- Hoffmann, S., Kleinstreuer, N., Alépée, N., Allen, D., Api, A.M., Ashikaga, T., Clouet, E., Cluzel, M., Desprez, B., Gellatly, N., Goebel, C., Kern, P.S., Klaric, M., Kühnl, J., Lalko, J.F., Martinozzi-Teissier, S., Mewes, K., Miyazawa, M., Parakhia, R., van Vliet, E., Zang, Q., Petersohn, D., 2018. Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database. *Crit Rev Toxicol* 48, 344–358. <https://doi.org/10.1080/10408444.2018.1429385>
- Hoffmann, S., Saliner, A.G., Patlewicz, G., Eskes, C., Zuang, V., Worth, A.P., 2008. A feasibility study developing an integrated testing strategy assessing skin irritation potential of chemicals. *Toxicol Lett* 180, 9–20. <https://doi.org/10.1016/j.toxlet.2008.05.004>
- ICCVAM, 2022. Symposium Webinar: Using New Approach Methodologies to Address Variability and Susceptibility Across Populations [WWW Document]. URL <https://ntp.niehs.nih.gov/go/popvar> (accessed 3.13.23).
- ICCVAM, 2018. A Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States [WWW Document]. URL https://ntp.niehs.nih.gov/iccvam/docs/roadmap/iccvam_strategicroadmap_january2018_document_508.pdf (accessed 5.3.22).
- ICCVAM, 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods [WWW Document]. URL https://ntp.niehs.nih.gov/iccvam/suppdocs/subguidelines/sd_subg034508.pdf (accessed 5.3.22).

- ICCVAM, 1997. Validation and Regulatory Acceptance of Toxicological Test Methods [WWW Document]. URL https://ntp.niehs.nih.gov/iccvm/docs/about_docs/validate.pdf (accessed 3.7.23).
- ICH, 2013. ICH Harmonized Tripartite Guideline Photosafety Evaluation of Pharmaceuticals S10 [WWW Document]. URL https://database.ich.org/sites/default/files/S10_Guideline.pdf
- Ishikawa, K., 1985. What is Total Quality Control? The Japanese Way. Translated by Lu, David J., 1st ed. Prentice-Hall.
- Judson, R.S., Magpantay, F.M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., Xia, M., Huang, R., Rotroff, D.M., Filer, D.L., Houck, K.A., Martin, M.T., Sipes, N., Richard, A.M., Mansouri, K., Setzer, R.W., Knudsen, T.B., Crofton, K.M., Thomas, R.S., 2015. Integrated model of chemical perturbations of a biological pathway using 18 in vitro high-throughput screening assays for the estrogen receptor. *Toxicol Sci* 148, 137–154. <https://doi.org/10.1093/toxsci/kfv168>
- Judson, R.S., Thomas, R.S., Baker, N., Simha, A., Howey, X.M., Marable, C., Kleinstreuer, N.C., Houck, K.A., 2019. Workflow for defining reference chemicals for assessing performance of in vitro assays. *ALTEX* 36, 261. <https://doi.org/10.14573/altex.1809281>
- Karmaus, A.L., Mansouri, K., To, K.T., Blake, B., Fitzpatrick, J., Strickland, J., Patlewicz, G., Allen, D., Casey, W., Kleinstreuer, N., 2022. Evaluation of variability across rat acute oral systemic toxicity studies. *Toxicological Sciences* 188, 34–47. <https://doi.org/10.1093/toxsci/kfac042>
- Kleinstreuer, N.C., Ceger, P., Watt, E.D., Martin, M., Houck, K., Browne, P., Thomas, R.S., Casey, W.M., Dix, D.J., Allen, D., Sakamuru, S., Xia, M., Huang, R., Judson, R., 2017. Development and validation of a computational model for androgen receptor activity. *Chem. Res. Toxicol.* 30, 946–964. <https://doi.org/10.1021/acs.chemrestox.6b00347>
- Kleinstreuer, N.C., Hoffmann, S., Alépée, N., Allen, D., Ashikaga, T., Casey, W., Clouet, E., Cluzel, M., Desprez, B., Gellatly, N., Göbel, C., Kern, P.S., Klaric, M., Kühnl, J., Martinozzi-Teissier, S., Mewes, K., Miyazawa, M., Strickland, J., van Vliet, E., Zang, Q., Petersohn, D., 2018. Non-animal methods to predict skin sensitization (II): an assessment of defined approaches *. *Crit Rev Toxicol* 48, 359–374. <https://doi.org/10.1080/10408444.2018.1429386>
- Kolle, S.N., Van Cott, A., van Ravenzwaay, B., Landsiedel, R., 2017. Lacking applicability of in vitro eye irritation methods to identify seriously eye irritating agrochemical formulations: Results of bovine cornea opacity and permeability assay, isolated chicken eye test and the EpiOcular; ET-50 method to classify according to UN GHS. *Regulatory Toxicology and Pharmacology* 85, 33–47. <https://doi.org/10.1016/j.yrtph.2017.01.013>

- Krishna, S., Berridge, B., Kleinstreuer, N., 2021. High-throughput screening to identify chemical cardiotoxic potential. *Chem Res Toxicol* 34, 566–583.
<https://doi.org/10.1021/acs.chemrestox.0c00382>
- LaLone, C.A., Villeneuve, D.L., Lyons, D., Helgen, H.W., Robinson, S.L., Swintek, J.A., Saari, T.W., Ankley, G.T., 2016. Editor's Highlight: Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS): A Web-Based Tool for Addressing the Challenges of Cross-Species Extrapolation of Chemical Toxicity. *Toxicol. Sci.* 153, 228–245.
<https://doi.org/10.1093/toxsci/kfw119>
- Luechtefeld, T., Maertens, A., Russo, D.P., Rovida, C., Zhu, H., Hartung, T., 2016. Analysis of public oral toxicity data from REACH registrations 2008-2014. *ALTEX* 33, 111–122.
<https://doi.org/10.14573/altex.1510054>
- Madia, F., Pillo, G., Worth, A., Corvi, R., Prieto, P., 2021. Integration of data across toxicity endpoints for improved safety assessment of chemicals: the example of carcinogenicity assessment. *Arch Toxicol* 95, 1971–1993. <https://doi.org/10.1007/s00204-021-03035-x>
- NIEHS, 2023a. Testing Regulations and Guidelines [WWW Document]. National Toxicology Program. URL <https://ntp.niehs.nih.gov/go/837330> (accessed 7.19.23).
- NIEHS, 2023b. Funding Opportunities for Test Method Developers [WWW Document]. Funding Opportunities for Test Method Developers. URL <https://ntp.niehs.nih.gov/go/alt-funding> (accessed 5.9.23).
- NIEHS, 2023c. International Cooperation on Alternative Test Methods [WWW Document]. International Cooperation on Alternative Test Methods. URL <https://ntp.niehs.nih.gov/go/icatm> (accessed 5.9.23).
- NIEHS, 2023d. About ICCVAM [WWW Document]. About ICCVAM. URL <https://ntp.niehs.nih.gov/go/iccvam> (accessed 5.9.23).
- OECD, 2023. No. 377: Initial Recommendations on Evaluation of Data from the Developmental Neurotoxicity (DNT) In-Vitro Testing Battery, OECD Series on Testing and Assessment. OECD Publishing, Paris.
- OECD, 2022a. No. 364: Case study on the use of Integrated Approaches for Testing and Assessment for DNT to prioritize a class of Organophosphorus flame retardants, Series on Testing and Assessment. OECD Publishing, Paris.
- OECD, 2022b. No. 367: Case Study on the use of an Integrated Approach for Testing and Assessment (IATA) for New Approach Methodology (NAM) for Refining Inhalation Risk Assessment from Point of Contact Toxicity of the Pesticide, Chlorothalonil, OECD Series on Testing and Assessment. OECD Publishing, Paris.

- OECD, 2021a. Guideline No. 497: Defined Approaches on Skin Sensitisation, OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris.
- OECD, 2021b. No. 298: Guiding Principles on Good Practices for the Availability/Distribution of Protected Elements in OECD Test Guidelines, Series on Testing and Assessment. OECD Publishing.
- OECD, 2018. No. 286: Guidance Document on Good In Vitro Method Practices (GIVIMP), OECD Series on Testing and Assessment. OECD Publishing.
<https://doi.org/10.1787/9789264304796-en>
- OECD, 2017. No. 255: Guidance Document on the Reporting of Defined Approaches to be Used Within Integrated Approaches to Testing and Assessment, OECD Series on Testing and Assessment. OECD Publishing, Paris. <https://doi.org/10.1787/9789264274822-en>
- OECD, 2014. No. 203: New Guidance Document on an Integrated Approach on Testing and Assessment (IATA) for Skin Corrosion and Irritation, OECD Series on Testing and Assessment. OECD Publishing, Paris.
- OECD, 2007. No. 69: Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models, OECD Series on Testing and Assessment. OECD Publishing, Paris. <https://doi.org/10.1787/9789264085442-en>.
- OECD, 2005. No. 34: Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, OECD Environment, Health and Safety Publications Series on Testing and Assessment. OECD Publishing, Paris.
- OECD, 1998. No. 1: OECD Principles of Good Laboratory Practice, Series on Principles of Good Laboratory Practice and Compliance Monitoring. OECD Publishing, Paris.
- Pamies, D., Leist, M., Coecke, S., Bowe, G., Allen, D.G., Gstraunthaler, G., Bal-Price, A., Pistollato, F., Vries, R.B.M. de, Hogberg, H.T., Hartung, T., Stacey, G., 2022. Guidance document on Good Cell and Tissue Culture Practice 2.0 (GCCP 2.0). ALTEX - Alternatives to animal experimentation 39, 30–70.
<https://doi.org/10.14573/altex.2111011>
- Parish, S.T., Aschner, M., Casey, W., Corvaro, M., Embry, M.R., Fitzpatrick, S., Kidd, D., Kleinstreuer, N.C., Lima, B.S., Settivari, R.S., Wolf, D.C., Yamazaki, D., Boobis, A., 2020. An evaluation framework for new approach methodologies (NAMs) for human health safety assessment. Regulatory Toxicology and Pharmacology 112, 104592.
<https://doi.org/10.1016/j.yrtph.2020.104592>
- Petersen, E., 2021. Characteristics to consider when selecting a positive control material for an in vitro assay. ALTEX. <https://doi.org/10.14573/altex.2102111>

- Petersen, E.J., Ceger, P., Allen, D.G., Coyle, J., Derk, R., Garcia-Reyero, N., Gordon, J., Kleinstreuer, N.C., Matheson, J., McShan, D., Nelson, B.C., Patri, A.K., Rice, P., Rojanasakul, L., Sasidharan, A., Scarano, L., Chang, X., 2022a. U.S. federal agency interests and key considerations for new approach methodologies for nanomaterials. *ALTEX - Alternatives to animal experimentation* 39, 183–206. <https://doi.org/10.14573/altex.2105041>
- Petersen, E.J., Elliott, J.T., Gordon, J., Kleinstreuer, N.C., Reinke, E., Roesslein, M., Toman, B., 2022b. Technical framework for enabling high-quality measurements in new approach methodologies (NAMs). *ALTEX - Alternatives to animal experimentation*. <https://doi.org/10.14573/altex.2205081>
- Petersen, E.J., Uhl, R., Toman, B., Elliott, J.T., Strickland, J., Truax, J., Gordon, J., 2022c. Development of a 96-well electrophilic allergen screening assay for skin sensitization using a measurement science approach. *Toxics* 10, 257. <https://doi.org/10.3390/toxics10050257>
- Pham, L.L., Watford, S., Pradeep, P., Martin, M.T., Thomas, R., Judson, R., Setzer, R.W., Paul Friedman, K., 2020. Variability in in vivo studies: Defining the upper limit of performance for predictions of systemic effect levels. *Comput Toxicol* 15, 1–100126. <https://doi.org/10.1016/j.comtox.2020.100126>
- Piersma, A.H., van Benthem, J., Ezendam, J., Kienhuis, A.S., 2018. Validation redefined. *Toxicol In Vitro* 46, 163–165. <https://doi.org/10.1016/j.tiv.2017.10.013>
- Prior, H., Casey, W., Kimber, I., Whelan, M., Sewell, F., 2019. Reflections on the progress towards non-animal methods for acute toxicity testing of chemicals. *Regul Toxicol Pharmacol* 102, 30–33. <https://doi.org/10.1016/j.yrtph.2018.12.008>
- Rooney, J.P., Choksi, N.Y., Ceger, P., Daniel, A.B., Truax, J., Allen, D., Kleinstreuer, N., 2021. Analysis of variability in the rabbit skin irritation assay. *Regulatory Toxicology and Pharmacology* 122, 104920. <https://doi.org/10.1016/j.yrtph.2021.104920>
- Rösslein, M., Elliott, J.T., Salit, M., Petersen, E.J., Hirsch, C., Krug, H.F., Wick, P., 2015. Use of Cause-and-Effect Analysis to Design a High-Quality Nanocytotoxicology Assay. *Chem. Res. Toxicol.* 28, 21–30. <https://doi.org/10.1021/tx500327y>
- Rusyn, I., Chiu, W.A., Wright, F.A., 2022. Model systems and organisms for addressing inter- and intra-species variability in risk assessment. *Regul Toxicol Pharmacol* 132, 105197. <https://doi.org/10.1016/j.yrtph.2022.105197>
- Sewell, F., Doe, J., Gellatly, N., Ragan, I., Burden, N., 2017. Steps towards the international regulatory acceptance of non-animal methodology in safety assessment. *Regul Toxicol Pharmacol* 89, 50–56. <https://doi.org/10.1016/j.yrtph.2017.07.001>

- Shaffer, R.M., 2021. Environmental Health Risk Assessment in the Federal Government: A Visual Overview and a Renewed Call for Coordination. *Environ. Sci. Technol.* 55, 10923–10927. <https://doi.org/10.1021/acs.est.1c01955>
- Slezák, P., Waczulíková, I., 2011. Letter to the Editor: Reproducibility and Repeatability. *Physiol Res* 60, 203–205.
- Smirnova, L., Hogberg, H.T., Leist, M., Hartung, T., 2014. Developmental neurotoxicity - challenges in the 21st century and in vitro opportunities. *ALTEX* 31, 129–156. <https://doi.org/10.14573/altex.1403271>
- Strickland, J., Clippinger, A.J., Brown, J., Allen, D., Jacobs, A., Matheson, J., Lowit, A., Reinke, E.N., Johnson, M.S., Quinn, M.J., Mattie, D., Fitzpatrick, S.C., Ahir, S., Kleinstreuer, N., Casey, W., 2018. Status of acute systemic toxicity testing requirements and data uses by U.S. regulatory agencies. *Regul Toxicol Pharmacol* 94, 183–196. <https://doi.org/10.1016/j.yrtph.2018.01.022>
- Strickland, J., Daniel, A.B., Allen, D., Aguila, C., Ahir, S., Bancos, S., Craig, E., Germolec, D., Ghosh, C., Hudson, N.L., Jacobs, A., Lehmann, D.M., Matheson, J., Reinke, E.N., Sadrieh, N., Vukmanovic, S., Kleinstreuer, N., 2019. Skin sensitization testing needs and data uses by US regulatory and research agencies. *Arch Toxicol* 93, 273–291. <https://doi.org/10.1007/s00204-018-2341-6>
- Thomas, R.S., Bahadori, T., Buckley, T.J., Cowden, J., Deisenroth, C., Dionisio, K.L., Frithsen, J.B., Grulke, C.M., Gwinn, M.R., Harrill, J.A., Higuchi, M., Houck, K.A., Hughes, M.F., Hunter, E.S., Isaacs, K.K., Judson, R.S., Knudsen, T.B., Lambert, J.C., Linnenbrink, M., Martin, T.M., Newton, S.R., Padilla, S., Patlewicz, G., Paul-Friedman, K., Phillips, K.A., Richard, A.M., Sams, R., Shafer, T.J., Setzer, R.W., Shah, I., Simmons, J.E., Simmons, S.O., Singh, A., Sobus, J.R., Strynar, M., Swank, A., Tornero-Valez, R., Ulrich, E.M., Villeneuve, D.L., Wambaugh, J.F., Wetmore, B.A., Williams, A.J., 2019. The next generation blueprint of computational toxicology at the U.S. Environmental Protection Agency. *Toxicol Sci* 169, 317–332. <https://doi.org/10.1093/toxsci/kfz058>
- Tox21, n.d. Tox21 - Toxicology in the 21st Century [WWW Document]. Tox21. URL <https://tox21.gov/> (accessed 5.9.23).
- Tsuji, R., Crofton, K.M., 2012. Developmental neurotoxicity guideline study: issues with methodology, evaluation and regulation. *Congenit Anom (Kyoto)* 52, 122–128. <https://doi.org/10.1111/j.1741-4520.2012.00374.x>
- van der Zalm, A.J., Barroso, J., Browne, P., Casey, W., Gordon, J., Henry, T.R., Kleinstreuer, N.C., Lowit, A.B., Perron, M., Clippinger, A.J., 2022. A framework for establishing scientific confidence in new approach methodologies. *Arch Toxicol* 96, 2865–2879. <https://doi.org/10.1007/s00204-022-03365-4>

- Wikoff, D., Lewis, R.J., Erraguntla, N., Franzen, A., Foreman, J., 2020. Facilitation of risk assessment with evidence-based methods – A framework for use of systematic mapping and systematic reviews in determining hazard, developing toxicity values, and characterizing uncertainty. *Regulatory Toxicology and Pharmacology* 118, 104790. <https://doi.org/10.1016/j.yrtph.2020.104790>
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wolffe, T.A.M., Whaley, P., Halsall, C., Rooney, A.A., Walker, V.R., 2019. Systematic evidence maps as a novel tool to support evidence-based decision-making in chemicals policy and risk management. *Environ Int* 130, 104871. <https://doi.org/10.1016/j.envint.2019.05.065>
- Zhang, J.H., Chung, T.D., Oldenburg, K.R., 1999. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen* 4, 67–73. <https://doi.org/10.1177/108705719900400206>
- Zhang, X.D., 2011. Illustration of SSMD, z score, SSMD*, z* score, and t statistic for hit selection in RNAi high-throughput screens. *J Biomol Screen* 16, 775–785. <https://doi.org/10.1177/1087057111405851>

付録 A: 用語集

免責事項:本用語集は、本報告書内の用語の使用を支援することを目的とする。特定の用語の意味または使用法は、規制機関によって異なる場合がある。

3Rs: 人道的な実験手法の原則。具体的には、研究や化学物質の安全性試験における動物の使用の削減（使用する動物の数を最小限に抑える）、置き換え（系統発生的下位動物または非動物試験の使用）、または苦痛の低減（特定の技術により実験動物が受ける可能性のある痛みや苦痛を最小限に抑える）に関するもの。

正確性、精度: 試験法の結果と許容される基準値との近さ。

有害性発現経路 (AOP): 生物がある物質に曝露されたときに有害影響をもたらす、生物学的組織の様々なレベルで発生する一連の事象を構造化して表したものの。

適用範囲: ある方法を使用して試験できる化学物質の種類、またはその方法で得られた結果が許容範囲内と見なされる化学物質の種類。

アッセイ: 使用された実験システム。多くの場合、「試験」および「試験法」と同じ意味で使用される。

バランス精度: 分類モデルの性能評価の一つの指標。一方の「クラス」が他方の「クラス」よりもはるかに多く出現する（例えば、陽性数よりも陰性数のほうが多い）不均衡なデータセットを説明するために使用される統計的指標。バランス精度は、感度と特異度の算術平均として計算される。

生物学的妥当性: 対象分類群における化学物質の影響を評価するための適切性の指標。

一致度: 2 つの変数間の一致または一貫性の近さ。この用語は、NAM の生物学的妥当性を、参照動物試験法および/または（入手可能な場合）ヒトの参考データと比較して定性的に説明するために使用できる。また、試験されたすべての化学物質のうち、陽性または陰性として正しく分類される割合を定量的に表すためにも使用できる。この場合、この用語はしばしば「正確性」と同じ意味で使用される。

使用方法と目的 (COU): 特定の方法、アプローチ、またはアプリケーションの使用方法和目的の明確な説明。

キュレーション データベース/リスト: 注意深く編集され維持管理されている、よく特徴付けられた信頼できる情報から構成された情報一式。

ディファインドアプローチ（定義済み総合判定方式、DA）：定義された一組の情報源から得られた入力データとあらかじめ規定されたデータ解釈手順によって判定結果を導き出すアプローチ。特定の規制ニーズを満たすために、単独で使用することも、IATA 内の他の情報源と組み合わせることもできる。試験と評価におけるディファインドアプローチは、化学物質の危険有害性の特定、危険有害性の特性評価、および/または安全性評価を支援するために使用できる（OECD、2017）。

不一致度：試験されたすべての化学物質のうち、陽性または陰性として誤って分類された化学物質の割合。

エンドポイント：試験法によって評価された生物学的または化学的なプロセス、反応、または効果。

偽陰性：特定の参照データと比較して、ある試験法によってある物質が誤って陰性と識別されること。

偽陽性：特定の参照データと比較して、ある試験法によってある物質が誤って陽性と識別されること。

目的適合性：NAM（または一組の NAMs）によって提供される情報の種類と確実性が、ある判断に必要な情報の種類と確実性に一致すること（EPA、2021a）。

優良細胞培養基準 (GCCP)：*in vitro*（細胞ベース）作業の再現性を確保し、科学的データの質を高めるために、試験施設での実用化に向けて開発された一連の原則（Pamies ら、2022）。

優良インビトロ試験基準 (GIVIMP)：当該項目に関する OECD ガイダンス文書（OECD、2018）に記載されている包括的なフレームワークで、*in vitro* 試験法の開発、妥当性検証、規制当局による承認または適格性確認、および使用に関する推奨事項を提供するもの。

優良試験所基準 (GLP)：EPA、FDA、OECD などの当局によって制定された規制で、各国の規制機関へのデータ提出の基礎となる試験施設の検査記録の記録管理および品質保証手順が規定されているもの。

危険有害性、ハザード：健康または生態系に有害な影響を及ぼす可能性。

危険有害性の分類：特定の毒性エンドポイントに対する標準試験法の結果に基づいて、化学物質または製品の危険有害性をカテゴリーに区分すること。主にラベル表示の目的で使用される。

危険有害性の特定: 特定の物質への曝露が健康または生態系への有害影響と関連しているかどうか、または関連している可能性があるかどうかを判断することに関連するリスク評価の一部。

試験の実施と評価のための統合的アプローチ (IATA): 化学物質の危険有害性の特定、危険有害性の特性評価、および/または安全性評価に使用される複数の情報源に基づくアプローチ (OECD、2017)。

施設間再現性: 同じプロトコルと試験化学物質を使用する、適格性を満たした異なる試験施設において、質的にも量的にも同様の結果が得られるかどうかを示す指標。施設間再現性は、事前検証と妥当性検証のプロセス中に評価され、試験法の試験施設間の技術移転性を示す。

メカニズム上の妥当性: 化学物質が効果を発揮する生化学的プロセスまたは経路を評価するための適切性の指標。

陰性予測度: ある試験方法で定義された参照データと比較して、陰性と判定された物質のうち、正しく陰性を示した割合。これは試験法の正確性の 1 つの指標である。陰性予測度は、試験方法の特異度と、試験対象物質における陰性物質の割合の関数である。

新しいアプローチや方法論 (NAM): 化学物質の危険有害性とリスク評価に関する情報提供に使用することができ、動物を用いない方法への置き換え、動物の使用数の削減、動物使用に伴う苦痛の低減 (3Rs) を支援する、あらゆる技術、方法、アプローチ、またはこれらの組み合わせに対する広義の参照語。

性能: 試験法の正確性および信頼性の特性 (「正確性」、「信頼性」を参照)。

性能標準: 検証済み試験法に基づいて、作用機構も機能も類似している提案された試験法の比較可能性を評価するための基礎を提供する。これには、(1) 試験法の必須構成要素、(2) 検証済み試験法の許容可能な性能を実証するために使用された化学物質の中から選択された参照化学物質のリスト、(3) 最小限の参照化学物質リストを使用して評価した場合に、提案された試験法が実証すべき、検証済み試験法で得られたものと同等な正確性と信頼性の水準が含まれる。

陽性予測度: ある試験法で定義された参照データと比較して、陽性と判定された物質のうち、正しく陽性を示した割合。これは試験法の正確性の 1 つの指標である。陽性予測度は、試験法の感度と、試験対象物質における陽性物質の割合の関数である。

強度: 物質の相対的な生物学的または化学的活性を示す指標。1 つの物質の強度は、様々な生物学的または生化学的作用ごとに異なる場合がある。

精度: 単一サンプルを複数回分析した後の、分析対象物の個々の測定値の近さ。精度は、しばしば変動係数として表される。

プロトコル: 必要なすべての試薬のリストと、試験データを生成および評価するためのすべての基準と手順のリストを含む、試験法の正確な記述。

適格性: 検証済みのモデルまたは試験を用いた評価の結果が、製品開発や規制上の意思決定において特定の解釈及び適用を行うために信頼できるという結論。

品質管理: 製品または方法の品質が意図したとおりに維持されていることを確認するために使用される一連の活動またはサンプル。

参照物質: 参照試験法または対象種での反応がわかっているため、提案された試験法の研究、開発、または評価に使用するために選択される化学物質（「参照試験法」を参照）。

参照試験法: 試験物質が対象種に有害である可能性を評価するために規制上の目的で使用するものが認められている試験法。

信頼性: ある試験法が長期にわたって試験施設内および試験施設間でどの程度再現可能に実施できるかを示す指標。これは、試験施設内および試験施設間の再現性と試験施設内の反復性を計算することによって評価される。

反復性: 一定期間内に、同一条件下で同じ試験プロトコルと試験サンプルを使用して得られた個々の試験結果の一貫性。「... 同じ被験者（または被験物質、または試験材料）について、同一条件下で同一の方法により得られた独立した結果間の近さ」（Slezák および Waczulíková、2011 年）。

再現性: 同じ試験プロトコルと試験サンプルを使用して得られた個々の試験結果の一貫性。「同じ被験者（または被験物質、試験材料）について、異なる条件下（異なる観察者、試験施設など）で、同じ方法により得られた独立した結果間の近さ」（Slezák および Waczulíková、2011 年）。

回収: 抽出された対象試料の結果とその類似マトリックス試料および/またはブランク試料に添加した結果を比較し、抽出法の抽出効率および再現性を定量的に評価する方法。

リスク評価: 外因性物質の曝露による標的生物種における健康および環境の潜在的な有害影響の評価。

頑健性: ある方法が、得られた結果に予期しない違いが生じることなく、さまざまな条件または状況下で再現されうる能力。

感度: すべての陽性化学物質のうち、試験法で陽性と正しく分類された化学物質の割合。感度は、検出限界の観点から、許容できる正確性および/または精度で測定できる最も低い検体の濃度として定義することもできる。

特異度: すべての陰性化学物質のうち、試験法で陰性として正しく分類された割合。試験法の正確性の指標である。この用語は、特定の分析対象物を検出する分析方法の能力を表すためにも使用される。

安定性: 試験材料（試験物質、試験装置、試薬、分析対象物など）について、特定の環境で一定期間にわたって同様かつ許容可能な結果が得られる能力。

標準曲線: 試験データを定量的にプロットする方法。未知のサンプルを既知の濃度の標準物質（多くの場合、陽性対照物質を使用）と比較することにより、未知のサンプルに含まれる物質の濃度を決定するために用いられる。

標準操作手順書 (SOPs): 試験施設での特定の作業や業務をどのように実施するかを説明する正式な手順書。これは GLP ガイドラインで義務付けられている。

標的臓器: 化学物質の潜在的な毒性に関する情報を求める対象となる臓器。

標的生物種: 化学物質の潜在的な毒性に関する情報を求める対象となる生物種。

試験: 使用する実験系。「試験法」および「アッセイ」と同じ意味で使用される。

試験法: 物質または製品の特性に関する情報を取得するために使用される工程または手順。毒性試験法は、特定の条件下で特定の生物学的効果を発現する物質または薬品の能力に関する情報を得ることができる。「試験」および「アッセイ」と同じ意味で使用される。「バリデーションされた試験法」および「参照試験法」も参照。

試験法開発者: 最初に試験法を考案し、その再現性と意図された使用目的への適合性を確保する組織または個人。

試験法提案者:（行政的）検討のために試験法を提案する組織または個人。試験法開発者と同じ組織または個人の場合もある。

試験法提出: 規制当局またはその他の特定の用途向けに提案されている試験法の補足文書の概要。試験法提出には、通常、特定の規制試験要件または用途に対する試験法の有用性と限界を示すため、完了したバリデーション試験の記録と、試験法提出ガイドラインに従って作成された科学的妥当性に関するその他の適切な文書が含まれる。

技術移転性: 試験法または手順をさまざまな適格試験施設で正確かつ確実に実施できる能力。「施設間再現性」も参照。

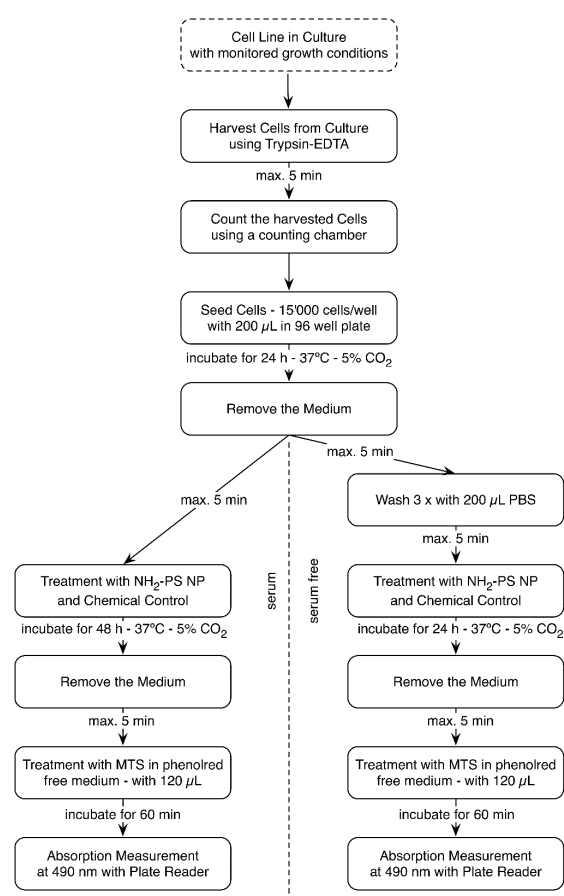
バリデーションされた試験法: 特定の提案された用途に対するこの方法の正確性、信頼性および妥当性を判断するためのバリデーション試験が完了して、承認された試験法。

バリデーション、妥当性検証: 特定の目的に合った手順の正確性、信頼性、妥当性を確立する工程。ある特定の目的のための検証は、他の特定の目的のための検証を意味するものではない。特定の使用方法と目的では、さらに適格性確認が必要になる場合がある。

付録 B: 品質ツール

1.0 フローチャート

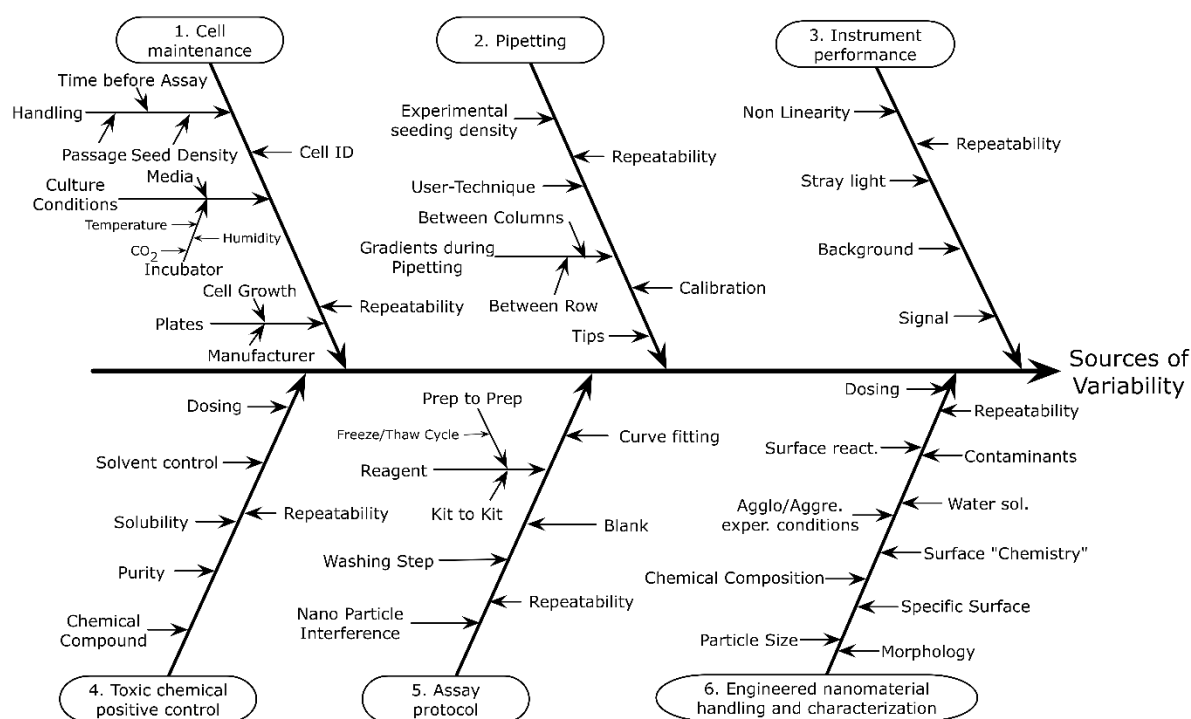
フローチャートでは、プロトコルの各ステップが図式化される（例については図 S1 を参照）。これは、各ステップを網羅し、実行可能な場合に追跡できるようにすることで、実験の対照測定値を設計するのに役立つ。1つのステップを複数の対照測定で網羅することも可能で、逆に1つの対照測定が複数のステップを網羅することもある。さらに、NAM間の対照測定値を比較すると、異なるNAM間で同様のステップが明らかになる場合がある。これらのステップでは、変動の原因も似ている可能性がある。



補足図 1. 試験施設間試験用の MTS (3- (4,5-ジメチルチアゾール-2-イル) -5- (3-カルボキシメトキシフェニル) -2- (4-スルホフェニル) -2H-テトラゾリウム) 修正プロトコルを説明するフローチャート。この図は、Elliott ら (2017) の許可を得て転載。

2.0 因果関係分析

因果関係 (C&E) 分析は、考えられる主な変動要因を特定し、C&E 図式（フィッシュボーン図式とも呼ばれる）を使用してそれらを表示するための概念的なツールである（例については図 S2 を参照）。この図式を作成する工程には、ブレインストーミングや試験に関する文献のレビューが含まれる場合がある。C&E 図式の各分岐は、予想される変動の主な原因を示す。C&E 図式は新しい NAM の開発にも役立つ。C&E 図式の共有分岐（例えば、同じタイプの細胞毒性試験を使用して測定を行う）はより迅速に作成でき、同様の対照測定が必要になり、同様の変動緩和戦略が採用される可能性が高いためである。C&E 図式の分析は、標準化が難しい方法の側面（校正が難しい機器や不安定な試験試薬など）を特定するのに役立つ。C&E 図式は、頑健性試験と対照測定の選択の指針となるため、理想的には、C&E 図式の各分岐と下位分岐における変動の原因を分析できる。



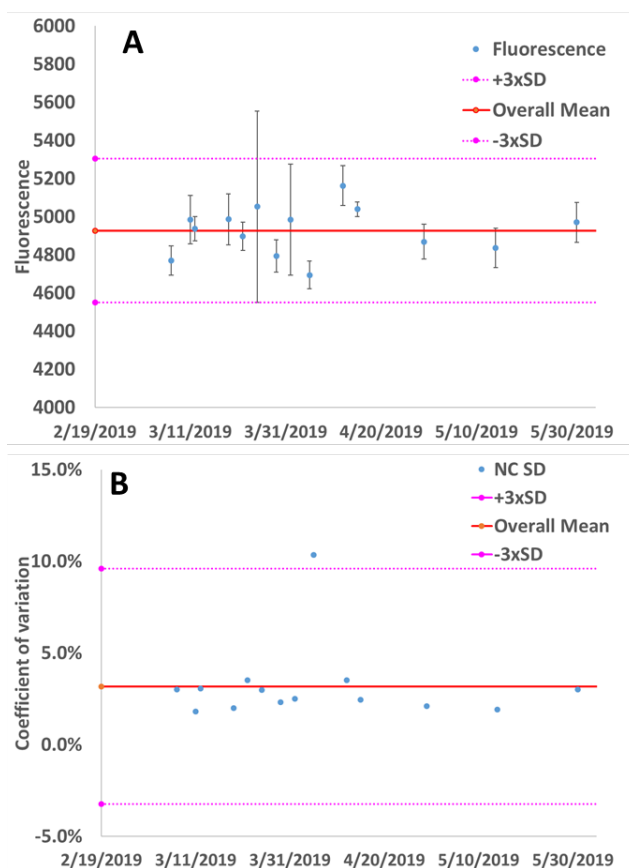
補足図 2. 人工ナノ材料用に設計された MTS 細胞毒性試験の因果関係図式。Rösslein ら (2015) の許可を得て転載。

3.0 管理図

対照測定には、1 回限りの予備実験（潜在的なバイアスに対する試験など）、所定の頻度で実行される定期的な測定（機器性能の評価や機器の校正など）、および試験を実行するたびに行われる工程中の対照測定がある。限られた数のサンプルを同時に分析でき

る試験（例えば、フロースルー曝露システムを使用してエアロゾル化された化学物質に曝露する吸入試験）では、定期的な測定が役立つ場合がある。例えば、ナノ粒子が重要な試験試薬を吸着する可能性があるかどうかなど、被験物質がバイアスを引き起こす可能性があるかどうかを評価するには、1 回限りの予備実験が役立つ。工程内対照測定は、試験を実施するたびに主要な変動要因を測定するために使用できる。たとえば、一般的な工程内管理測定の 1 つが陽性対照である。この対照測定により、最大反応（100% の効果）に達したかどうかが明らかになる。また、より低く、より中程度の反応が得られる化学物質濃度に対する試験反応の感度を示すためにも使用できる。陽性対照を選択する際の重要な考慮事項は、Petersen ら (2021) によって説明されている。その他の一般的な工程内対照測定として、細胞や追加の試薬がない対照、細胞と追加の試薬は含まれるが試験物質を含まない対照がある。考えられるすべての工程内対照測定を含めることはできない場合があり、どの工程内対照測定を含めるかという点でトレードオフが生じる場合がある。

変動の主な原因は、工程内対照測定の平均値と変動を表示する管理図を使用して、時間の経過や実験間で監視できる（例については図 S3 を参照）。これは、時間の経過とともに平均値または変動値に系統的な変化があり、試験内の何か（試薬の不安定性など）が変化した可能性があるかどうかを評価するのに役立つ。チェックシートを確認すると、変更が発生した理由を特定するのに役立つ。



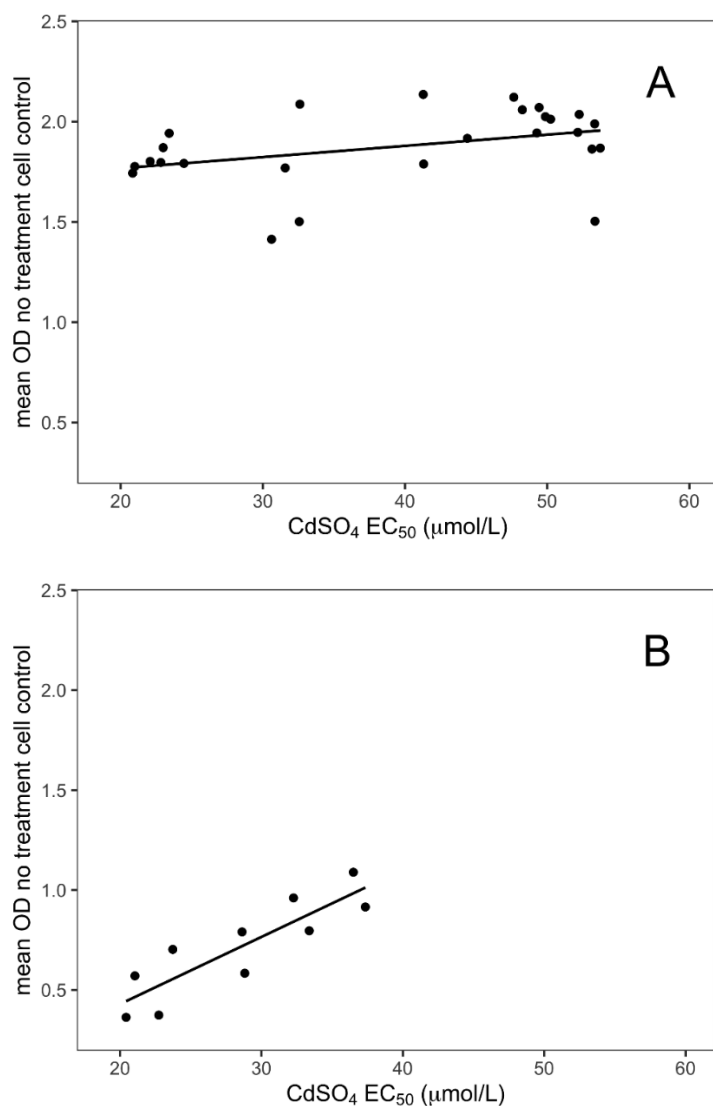
補足図 3. 陰性インキュベーター対照の求電子アレルゲンスクリーニングアッセイ蛍光法の対照表データ: (A) 実施日に応じたすべての実験の (A) 平均と (B) 変動係数。グラフ B では、1 つの値が変動係数の外れ値であり、この研究の規格外である（全体平均は平均標準偏差値の ± 3 倍）。また、時間の経過に伴う平均変動値または変動係数の値を示す系統的な傾向はない。この図は Petersen ら (2022c) の許可を得て修正および転載。図のキャプションは Petersen ら (2022b) から修正および転載した。

4.0 チェックシート

重要なメタデータを記録し、GLP の要件を満たすには、チェックシートを使用できる。方法の開発時に GLP を使用しない場合、その NAM が規制当局への提出を目的とし、GLP 遵守が必要な場合には、チェックシートを使用して重要なデータを記録するなどの GLP に定められた原則が重要になる。試験開発中の代替関連リソースとしては、GCCP 2.0 (Pamies ら、2022) と GIVIMP (OECD、2018) に関するガイダンス文書がある。これらは細胞ベースの試験に特化したもので、反復性と高品質の科学データを確保するための品質管理リストが追加されている。チェックシートは、工程内対照測定値を監視したり、問題が発生したときのトラブルシューティングを支援したりするのにも役立つ。たとえば、試薬や消耗品（ピペットやマイクロウェルプレートなど）のロット番号の変更を記録するために使用できる。これは、ある規格が他の規格よりも頻繁に満たされない場合のトラブルシューティングに役立つ。データ計算ツールを含むチェックシートを使用して、実施した試験のデータがすべての規格を満たしているかどうかを評価し、試験結果の統計的評価を行うこともできる。書面によるチェックシートの代わりに、電子ノートや実験情報管理システムを使用できる（保管、試験施設間の技術移転性、検索が容易になるなどの利点がある）。

5.0 散布図

散布図を使用して、さまざまな工程内対照測定値間、またはそれらの対照測定値と試験物質の結果との間に相互作用があるかどうかを評価できる（例については図 S4 を参照）。試験物質の結果は、規格で定められた範囲内の工程内対照測定の結果とは無関係とすべきである。



補足図 4. MTS 試験を用いて測定した硫酸カドミウム (CdSO_4) EC_{50} 値と、被験物質または陽性対照に曝露されていない A549 細胞の平均吸光度 (OD) 値との相関関係。これらのデータは、平均 OD 値の範囲に応じて、 EC_{50} 値間の相互作用の欠如 (パート A) または相互作用 (パート B) のいずれかを示す。実線は線形回帰近似である。パート B の傾きは統計的に 0 とは異なる。これは、 EC_{50} の値が OD 値と相関していることを示す。Petersen ら (2022b) の許可を得て修正および転載。図のキャプションは Petersen ら (2022b) の許可を得て転載。

付録 C: 分析方法評価

付録 C の情報は、「毒性試験法の妥当性検証と規制承認」（ICCVAM、1997）、「業界向けバイオ分析法検証ガイダンス」（FDA、2018）、OECD GD 34（OECD、2005）、および OECD GIVIMP 文書（OECD、2018）など、いくつかの参考文献からのものである。付録 C の情報は、方法評価の一般的なフレームワークを提供することを目的とする。一部の側面は、すべての方法に適用できる場合と適用できない場合がある。

1.0 検出と定量化の限界

検出と定量化の限界によって、分析法の範囲が決まる。分析方法是、多くの場合、定量化の下限または検出限界のいずれかで、システムにより適した方で試験される。この工程では、試薬ブランクの上または標準曲線内で、確実に検出（検出限界）または定量化（定量下限）できる検体の最低量または濃度を設定することにより、分析方法の範囲の下限を決定する。標準曲線上の最高濃度が定量の上限を決定する。

2.0 干渉物質の同定

試験法開発者は、試験法の検出シグナル（蛍光/吸光度、ルシフェラーゼ、酵素など）への干渉を含め、試験法の重要なコンポーネントと重要でないコンポーネントから生じる可能性のある干渉物質を記録すべきである。干渉は、内分泌攪乱物質試験法に含まれる特定のプラスチックや、その方法に含まれる他の試薬などの消耗品からも発生する可能性がある。潜在的な干渉物質には、内因性マトリックス成分、代謝産物、分解生成物、およびその他の生体異物が含まれるが、これらに限定されない。試験法が複数の分析対象物の定量を目的としている場合は、対象となるすべての分析対象物を試験して、干渉がないことを確認すべきである。ブランクサンプルは干渉のテストによく使用される。

3.0 分析精度の評価

使用した分析方法の精度の説明や、その他の精度試験（たとえば、異なる担当者が提案された方法を使用する場合や、その方法に異なる機器を使用した場合の性能の変動を評価する試験など）を含めると、可能であれば試験施設間の比較研究に参加することが役立つ。

4.0 材料の安定性

NAM で使用される材料（試験物質、試験装置、試薬、分析対象物など）に関連する安定性とは、特定の環境で一定期間にわたって同様に許容できる結果が得られる材料の能力を指す。NAM 自体の安定性については、前述のとおり（付録 B、セクション 3.0）、

経時的な管理図の評価とともに説明した。サンプルの収集、取り扱い、保管条件の影響を評価すべきである。安定性の問題の影響を受ける可能性のある材料には、適切な有効期限を設定すべきである。分解生成物による干渉や実際の投与量の変化を避けるため、試験物質、試薬、および試験装置（プラスチックマイクロプレートなど）の安定性を確保すべきである。いずれかのコンポーネントについて実施された安定性試験を報告書に含めるべきである。開発、妥当性検証、または適格性確認の試験に使用する前に、試験当日までに調製した化学混合物または調製サンプルの安定性を評価すべきである。

特定の時間間隔における特定の条件下での特定の混合物またはマトリックスの化学的安定性は、いくつかの方法で評価される。試験前の安定性評価は、試験実施中に予想されるサンプルの取り扱いおよび保管条件（試験場所、出荷中、その他すべての二次施設での条件を含む）を対象とすべきである。特定の混合物、マトリックス、およびコンテナシステムにおける分析対象物の安定性は、その混合物、マトリックス、およびコンテナシステムにのみ関係し、他のシステムに当てはめるべきではない。安定性試験では、長期間（目的の保管温度で凍結）および短期（ベンチトップ、室温）保管時、および凍結/解凍サイクル後および分析プロセス後の分析対象物の安定性を評価すべきである。安定性試験で使用する条件は、実際のサンプルの取り扱いと分析中に発生する可能性のある状況を反映すべきである。試験用サンプルの分析中に、保管条件が試験法の技術的特性評価中に評価されたサンプル保管条件を変更したり、超えたりした場合、これらの新しい条件下での安定性を確立すべきである。重要な試薬が遭遇する可能性のあるあらゆる条件を対象として、独立した安定性試験を実施すべきである。試薬や材料の安定性試験は、凍結融解サイクル中やベンチトップでの使用、ストック溶液、加工サンプル、長期間の研究など、さまざまな状況で検討すべきである。

5.0 頑健性試験

頑健性は、得られた結果において予想外の差が発生しない様々な条件または状況下で反復できる方法の安定性である。NAM の頑健性の 1 つの側面は、付録 B のセクション 3.0 で説明されるように、時間の経過に伴う管理図の結果の一貫性である。頑健性試験は、実験用試薬やプロトコルに意図しないばらつきがあることによる結果の変化を検出するために使用されることが多い。試験法のあらゆる面で頑健性試験が推奨され、すべてのパラメータと測定範囲を可能な限りいつでもどこでも確立すべきである。

たとえば、ある試験では 5 分のインキュベーション時間が最適と定められていたが、頑健性試験の結果、データは 5 分プラスマイナス 30 秒ですべての品質管理要件を満たした。したがって、頑健性試験で許容されるインキュベーション時間は 5 分±30 秒である。

許容範囲を確立すべき試験パラメータには、以下が含まれる。

- インキュベーション時間

- インキュベーション温度
- pH
- 試薬の供給源
- 細胞密度 (該当する場合)
- 実験条件
- 分析ソフトウェア

モデルの設計によって、許容範囲が必要な実験条件が決まる。たとえば、流れる媒体を含むモデルでは、流れに関連するパラメータの許容範囲が必要になる。より複雑なモデルほど、対照を必要とする試験パラメータがある可能性が高くなる。

該当する場合は、重要な試薬と重要でない試薬の頑健性試験の結果も報告すべきである。試薬を 1 つのサプライヤーから購入すべきか、それとも複数のサプライヤーを使用できるかを判断するために、（可能な限り）さまざまなサプライヤーを試験すべきである。これは、サプライチェーンの混乱を回避したり、メーカーが特定の製品の製造を中止したりする場合にも役立つ。

必要に応じて、機器の頑健性試験も実施すべきである。試験施設間の妥当性検証では、多くの場合、性能パラメータや性能が異なるさまざまなブランドの機器が使用されるため、データのばらつきが増したり、実施される方法のパラメータが変化したりする可能性がある。

6.0 回収の分析

特定のマトリックスからの抽出について試験を実施する研究では、抽出法の効率と反復性を検証するために回収研究を実施すべきである。回収研究は、多くの場合、抽出されたサンプルの結果を、同様のマトリックスのスパイクサンプルやスパイクブランクの結果と比較することによって実施される。

7.0 適用範囲の技術分析

試験の対象となる特定の COU に関連する化学物質および/または製品の適切な試験法データが存在すべきである。NAM、DA、または IATA の文書には、適用範囲の物理化学的特性を明確に記述すべきである。これには、特定の分子量範囲の化学物質、揮発性、溶解性、安定性など、方法の制限も含まれる。

適用範囲は試験の初期開発時に分析的に評価すべきであるが、この方法を幅広く使用することで追加のデータが得られる可能性がある。これらのデータは、技術的にさらなる対処が必要な可能性のある干渉を明らかにする可能性がある（Petersen ら、2022a）。

特定の物理化学的特性を持つ化学物質について試験が評価されないと、分析的（例えば、試験の性能に影響するバイアスの有無）と潜在的な一致性（例えば、*in vivo* 試験と同様の結果が得られるかどうか）の両方から、これらの化学物質の試験性能に関する不確実性が高まる。したがって、物理化学的特性が異なる化学物質間の試験の一致性を評価すると、より広範囲の化学物質への試験の適用性に対する信頼性を高めることにも役立つ場合がある。

8.0 陽性対照

試験開発には、少なくとも 1 つの陽性対照を含めるべきである。これらの対照は、標準曲線と組み合わせて使用することも、陽性対照の単一濃度として使用することもできる。陽性対照は、試験の検出エンドポイントに妥当な物にすべきである。陽性対照の濃度は、試験の検出ウィンドウ内に存在すべきである。多くの場合、1 つの濃度のみを試験し、最大用量の分解能に関するデータが得られるならば、最大検出範囲の 50%~80% の反応が得られるように陽性対照濃度を選択する。標準曲線は多くの場合、システムの陽性対照を使用して生成され、標準曲線は、試験反応と既知の分析対象物濃度との関係として定義される。標準曲線は、経時的に再現可能であるべきである。標準曲線は、試験の対象となる試験サンプルと同じ賦形剤または溶媒で作成すべきである。同じ混合物またはマトリックスが得られない場合は、適切な文書化と正当化を行って代理混合物またはマトリックスを使用できる。

9.0 機器校正の参照標準物質

該当する場合は、校正標準や品質管理サンプルを使用して機器を校正すべきである。純度、安定性、供給源、ロット番号、分析証明書、有効期限など、標準サンプルまたは品質管理サンプルの出所を文書化すべきである。標準には、米国薬局方大要基準や、純度が証明された信頼できる商業的供給元から入手した市販の資料など、いくつかの種類がある。

理想的には、一部の機器（質量分析計など）の対照標準物質は、試験の対象となる分析対象物と同一にすべきである。多くの場合、これは不可能であり、純度がわかっている確立された化学形態を使用できる。

10.0 規格の設定

試験施設内または試験施設間の試験結果に基づいて、工程内対照測定値の規格を設定できる。使用できる方法は複数存在する。たとえば、一般的に使用される方法では、工程内対照測定値に 95% の信頼区間を適用したり、標準偏差値の 2~3 倍の $\pm$ 平均値を使用したりするが、他の統計的アプローチもよく使用される。規格を設定する際は、さまざまな目的のバランスを取るようにする。試験が期待どおりに機能していないことを示すデータや、工程内対照測定値データが被験物質の結果に偏りをもたらす可能性のある範

囲内にあることを示すデータを除外するために、十分に厳密な規格にすることが重要である。また、試験データが標準範囲と見なされる試験で、規格を満たせない試験の実行回数が多くなり過ぎるほど規格を厳密に設定しないことも重要である。