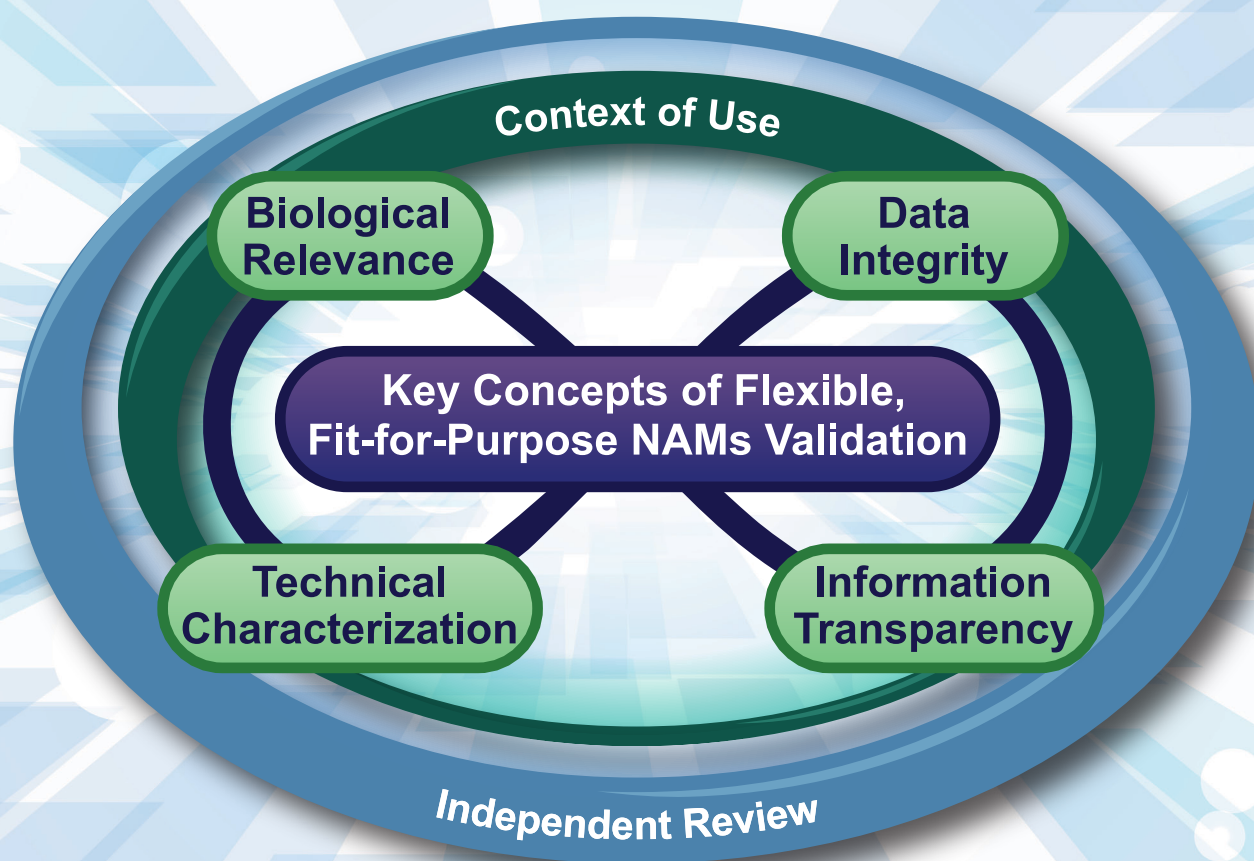


新测试方法的验证、 确认和监管认可

2024 年 3 月



目录

表格列表..... v

图列表..... vi

缩写和首字母缩略词..... vii

ICCVAM 验证工作组..... ix

ICCVAM 机构代表..... xi

免责声明..... 1

执行摘要..... 1

1.0 简介..... 1

2.0 灵活、目的取向的 NAM 验证中的关键概念 3

3.0 应用关键概念建立对 NAM 的信心 5

 3.1 使用情景..... 5

 3.2 生物相关性..... 6

 3.2.1 机制理解..... 6

 3.2.2 参比化合物..... 9

 3.2.3 与现有实验动物方法的比较..... 11

 3.3 技术特征..... 13

 3.3.1 纳入选定的质量工具..... 13

 3.3.2 质量控制最佳实践规范..... 15

 3.3.3 文档..... 20

 3.4 数据完整性..... 22

 3.5 信息透明度..... 22

3.6 独立审查..... 23

4.0 美国联邦机构对 NAM 的认可 24

4.1 了解监管需求和决策背景..... 24

4.2 使用环境考虑因素..... 24

4.3 基于获得经验的信心演变..... 24

5.0 美国和国际协调..... 25

5.1 美国协调：ICCVAM 和 NICEATM 的作用 25

5.2 美国协调：为推进 3R 而开展的额外联邦合作 26

5.3 国际协调..... 26

6.0 鼓励使用 NAM 的宣传和培训 27

7.0 结论与实施..... 28

参考文献..... 29

附录 A：术语表..... 40

附录 B：质量工具..... 45

1.0 流程图..... 45

2.0 因果分析..... 45

3.0 对照图表..... 46

4.0 检查表..... 48

5.0 散点图..... 48

附录 C：分析方法评估..... 50

1.0 检测限和定量限..... 50

2.0 干扰鉴别..... 50

3.0	评估分析精确度	50
4.0	材料稳定性	50
5.0	稳健性测试	51
6.0	回收分析	52
7.0	应用域的技术分析	52
8.0	阳性对照	52
9.0	仪器校准参考标准品	52
10.0	设置质量标准	53

表格列表

表 1. 与 NAM 的验证、确认和监管使用相关的现有美国和国际文件示例	2
表 2. ICCVAM 工作组制作的手稿，提供有关选定主题的机构测试需求的详细信息.....	5
表 3. 已证明 NAM 的生物学和机理相关性的终点示例	7
表 4. 参比化合物的选定用途	10
表 5. 通常用于 NAM 技术表征的质量工具	15
表 6. 分析方法评估	17

图列表

图 1. 在制定和实施灵活、目的取向的 NAM 验证策略时需要考虑的关键概念。	4
图 2. 开发稳健性 NAM 的框架	14
补充图 1. 描述修改后的 MTS 方案的流程图	45
补充图 2. 因果图	46
补充图 3. 对照图数据	47
补充图 4. 硫酸镉 EC ₅₀ 值的相关性	49

缩写和首字母缩略词

3R: 代替、减少或优化对动物的使用

AOP: 有害结局路径

C&E: 因果关系

COU: 使用情景

CPSC: 美国消费品安全委员会

DA: 限定方法

EPA: 美国环境保护署

EURL ECVAM: 欧盟替代方法验证参考实验室

FDA: 美国食品药品监督管理局

GD: 指南文件

GHS: 全球化学品统一分类和标签制度

GIVIMP: 关于良好体外方法规范的指南文件

GLP: 良好实验室规范

IATA: 整合测试评估方法

ICATM: 在替代测试方法上的国际合作

ICCVAM: 替代方法验证跨部门协调委员会

ICH: 国际人用药品技术要求协调理事会

IQ/OQ/PQ: 安装质量/运行质量/性能质量

JaCVAM: 日本替代方法验证中心

MTS: (3-(4,5-二甲基噻唑-2-基)-5-(3-羧甲氧基苯基)-2-(4-磺基苯基)-2H-四氮唑)

NAM: 新测试方法

NCATS: 美国国家转化科学促进中心

NICEATM: 国家毒理学计划替代毒理学方法评估跨机构中心

NIEHS: 美国国家环境卫生科学研究院

OD: 光学密度

OECD: 经济合作与发展组织

SOP: 标准操作规程

TG: 测试指南

Tox21: 21 世纪毒理学

TSAR: 监管认可替代方法跟踪系统

ICCVAM 验证工作组

美国毒物与疾病登记署

Moiz Mumtaz, 博士

消费品安全委员会

John Gordon, 博士(联席主席)

美国国防部

Donald Cronce, 博士

Natalia Garcia-Reyero Vinas, 博士

Matthew Johnson, DVM, DACLAM(至 2022 年 6 月)

Emily N. Reinke, 博士(至 2022 年 4 月)

退伍军人事务部

George Lathrop, Jr., DVM, MS, DACLAM

环境保护署

农药计划办公室

Anna Lowit, 博士

Scott Lynn, 博士

Monique Perron, 博士

研究与发展办公室

Kelly Carstens, 博士

Alison Harrill, 博士

Nisha Sipes, 博士

美国食品药品监督管理局

设备与放射健康中心

Jennifer Goode, 学士

药物评估与研究中心

Paul C. Brown, 博士

食品安全与应用营养中心

Suzanne Fitzpatrick, 博士(联席主席)

Anneliese Striz, 博士

烟草制品中心

Jueichuan (Connie) Kang, 博士
首席科学家办公室
Tracy Chen, 博士

美国国家环境卫生科学研究院

Warren Casey, 博士
国家毒理学计划替代毒理学方法评估跨机构中心(NICEATM)
Helena Hogberg, 博士
Nicole Kleinstreuer, 博士

美国国立卫生研究院

美国国家转化科学促进中心
Menghang Xia, 博士

美国国家标准与技术研究院

Elijah Petersen, 博士(联席主席)

职业安全与健康管理局

Janet Carter, 硕士

NICEATM 支持合同人员 (Inotiv)

David Allen, 博士(至 2023 年 12 月)
Michaela Blaylock
Amber Daniel, MTTox
Agnes Karmaus, 博士(至 2023 年 4 月)

ICCVAM 机构代表

本报告发布时的 ICCVAM 委员会成员

*主要机构代表；+候补主要机构代表

美国毒物与疾病登记署

- * Moiz Mumtaz, 博士
- Patricia Ruiz, 博士

美国消费品安全委员会

- * John Gordon, 博士
- + Kristina Hatlelid, 博士, 公共卫生硕士
- + Eric Hooker, 硕士
- + Joanna Matheson, 博士

美国农业部

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

美国国防部

- * Shannon Marko, DVM, DACLAM
- + Natalia Garcia-Reyero Vinas, 博士(联席主席)
- + Saber Hussain, 博士, ATS 研究员, AFRL 研究员
- + Elaine Merrill, 博士
- Matthew Grogg, 博士
- Donald Cronce, 博士

美国能源部

- R. Todd Anderson, 博士

美国内政部

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

美国交通部

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

美国退伍军人事务部研究与开发办公室

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

美国环境保护署

化学品安全与污染预防办公室

Charles Kovatch

农药计划办公室

Monique Perron, 博士

William (Bill) Eckel, 博士

Cecilia Tan, 博士

污染预防和毒物办公室

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

研究与发展办公室

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

美国食品药品监督管理局

生物制剂评估与研究中心

Leslie Wagner

Allen Wensky, 博士

设备与放射健康中心

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

药物评估与研究中心

Paul C. Brown, 博士

Nakissa Sadrieh, 博士

食品安全与应用营养中心

- * Suzanne Fitzpatrick, 博士(联席主席)
- Omari J. Bandele, 博士
- Patrick Crittenden, 博士
- Shruti Kabadi, 博士

烟草制品中心

- Jueichuan (Connie) Kang, 博士

兽医学中心

- M. Cecilia Aguila, DVM
- Li You, 博士

国家毒理学研究中心

- Mugimane (Manju) Manjanatha, 博士

首席科学家办公室

- Tracy Chen, 博士
- Chad P. Nelson, 博士, MSPH

美国国家癌症研究所

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

美国国家职业安全卫生研究所

- * Stephen Leonard, 博士

美国国家环境卫生科学研究院

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

美国国家标准与技术研究院

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

美国国立卫生研究院

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

美国国家医学图书馆

- * Dina N. Paltoo, 博士, 公共卫生硕士, CPI

职业安全与健康管理局

- * Deana Holmes, 硕士
- + Janet Carter, 硕士

免责声明

本报告是为美国联邦机构和利益相关者提供的一份参考资源，帮助这些机构和利益相关者在寻求新方法代替、减少或优化测试中对动物的使用方面建立信心。本报告中描述的原则根据 17 个联邦机构、多个跨机构工作组、公众和替代毒理学方法科学咨询委员会工作人员的意见制定。因此，本报告不一定体现任何机构或工作组的意见或政策。其不为任何人或任何一方创造权利，任何联邦机构都不应将其视为承诺。本报告未规定任何具有法律约束力的责任。相反，本报告介绍的是替代方法验证跨机构协调委员会(ICCVAM)当前对某个主题的想法，除非引用了具体的监管或法定要求，否则本报告应仅被视为建议。使用“应当”一词意味着建议或推荐做某事，但不是强制要求。

执行摘要

新测试方法(NAM)的开发日益频繁，并被用于对化学品和产品对人类健康和环境的潜在毒性影响进行监管和非监管评估。这些新测试方法被用来研究毒理学过程的生物机制，协助评估新药品和现有药品，并生成危害识别和剂量反应关系信息，用于健康和环境危害分类以及风险评估。本报告由 ICCVAM 制定，旨在帮助开发者和终端用户建立对新测试方法的信心。这种信心可以通过实施灵活的、目的取向的验证策略来实现，而这些策略会考虑到新测试方法的预期应用。本报告介绍了新测试方法的使用背景、生物学相关性和技术特性等概念，有助于建立信心。本报告着眼于美国联邦机构对新测试方法的监管认可以及实现国际协调的可能性，并审查了新测试方法质量工具和技术评估的最佳做法。本报告还强调了开发者、终端用户和监管机构之间进行沟通的必要性。过时的“一刀切”验证策略不适用于像新测试方法这样的先进领域。ICCVAM 是一个可以通过其咨询身份、协调验证工作和创造研究合作机会来支持这种向更现代方法过渡的组织。因此，ICCVAM 与 17 个联邦监管机构和研究实验室合作制定了这份报告，以帮助开发者、终端用户和监管机构建立对新测试方法的信心，从而可以将其用于补充或取代用于监管和非监管目的的动物测试。

1.0 简介

替代方法验证跨机构协调委员会(ICCVAM)由来自 17 个美国联邦监管和研究机构的代表组成，这些机构需要、使用、生成或传播毒理学和安全测试信息。ICCVAM 对新的、修订的和替代的安全性测试方法以及具有监管适用性的综合测试策略进行技术评估。ICCVAM 还促进对准确评估相关产品的化学安全性和危害的测试方法进行科学验证和监管认可或确认，以代替、减少或优化(提高动物福利，减轻或避免给动物造成疼痛和痛苦)对动物的使用。

ICCVAM 在 1997 年作为常设委员会成立后不久发布了一份报告《毒理学测试方法的验证和监管认可》(ICCVAM, 1997), 概述了新的和替代测试方法的验证和监管认可标准。随后, 2003 年的出版物《ICCVAM 关于提名和提交新的、修订的和替代的测试方法的指南》(ICCVAM, 2003)提供了更多指导。经济合作与发展组织(经合组织, OECD)指南文件(GD)34 “关于新的或更新的危害评估测试方法的验证和国际认可的指南文件”(OECD, 2005)中所述的国际标准沿用了这些文件中概述的原则。这些资源介绍了一种原则上灵活的验证模型, 但在实践中却表现出多种局限性, 例如过于冗长以及资源密集。在某些使用环境中, 方法可能不需经过此验证过程的每一个步骤即可为联邦机构提供有价值的数据。此外, 这些文件并不总是与许多现代毒性测试方法兼容, 这些方法较少强调用单一替代方法取代体内测试, 而更多地强调整合来自多种体外和化学分析以及计算机仿真方法(例如计算模型)的结果。2018 年 ICCVAM 出版物《建立评估美国化学品和医疗产品安全性新方法的战略路线图》(ICCVAM, 2018)提供了一个概念框架, 促进机构和测试方法开发者之间更好地沟通, 并在如何建立信心方面提供更多灵活性, 以帮助确保联邦机构和监管行业在针对特定应用或使用情景(COU)进行验证后采用新方法。下文基于 2018 年 ICCVAM 路线图中概述的原则, 提供了关于建立对新测试方法(NAM)的信心的更多具体见解。

在本报告中, 新测试方法(NAM)一词是指任何可用于提供化学危害和风险评估信息并支持代替、减少或优化动物使用(3R)的技术、方法论、方法或其组合。本报告以美国和国际上接受的现有指南和文件(表 1)中阐述的原则为基础, 倡导以更灵活的方法建立对新测试方法的信心。在这里, 我们介绍适用于 NAM 的验证、确认和监管认可的关键概念, 包括在生物系统中进行测试(例如, 体外、某些化学试验、小模型生物), 同时认识到其他类型的 NAM 可能需要额外的考虑, 例如计算模型预测(OECD, 2007)。

表 1. 与 NAM 的验证、确认和监管使用相关的现有美国和国际文件示例

文件标题	参考文献
毒理学测试方法的验证和监管认可(自本报告发布之日起停用)	ICCVAM, 1997
关于新、修订和替代测试方法提名和提交的 ICCVAM 指南	ICCVAM, 2003
经合组织测试和评估丛书第 34 号: 关于对危害评估的新或更新测试方法验证和国际认可的指南文件	经合组织, 2005
关于消费品安全委员会动物测试政策的建议程序	消费品安全委员会, 2012
FDA 预测性毒理学路线图	FDA, 2017a

文件标题	参考文献
医疗器械开发工具确认：行业、工具开发者和食品药品监督管理局工作人员指南	FDA, 2017b
EPA 促进在 TSCA 计划内开发和实施替代测试方法的战略计划	EPA, 2018
为评估美国化学品及医疗产品安全性建立新方法的 ICCVAM 战略路线图	ICCVAM, 2018
关于良好体外方法规范的指南文件(GIVIMP)	经合组织, 2018
行业和测试方法开发者指南：消费品安全委员会工作人员对替代测试方法和综合测试方法的评估，以及由此类方法生成的数据，以支持 FHSA 的标签要求	消费品安全委员会, 2020
《行业和 FDA 工作人员药品开发工具指南》的确认流程	FDA, 2020
EPA 新方法工作计划	EPA, 2021a
EPA 减少在化学测试中使用脊椎动物的战略计划	EPA, 2021b
在 FDA 推进新的替代方法	FDA, 2021a

2.0 灵活、目的取向的 NAM 验证中的关键概念

经合组织 GD 34 中所述的验证基本原则仍然至关重要，但验证所使用的程序应有助于高效、及时地开发能提供目标物种信息的目的取向的、可靠的 NAM。具体而言，经合组织 GD 34 指出，“新的测试方法经过验证，以确保其具有可靠的科学依据并满足监管需求”，即这些方法是目的取向的，这除了为监管决策提供依据外，还可能包括对使用情况进行筛选和优先排序。该指南还指出，“验证过程应具有灵活性和适应性”，必须“使用一系列参比化学品来证明其性能”，并且“根据现有的相关毒性数据进行评估”(经合组织，2005)。因此，不应将对 NAM 的信心视为特定程序完成后获得的不变状态(例如，在成功完成环比试验研究后认为 NAM 得到了“验证”)；相反，应对 NAM 的信心建立视为一个不断演变和迭代的过程，这个过程需要方法开发者、监管决策者和验证机构之间进行沟通 (ICCVAM, 2018)。

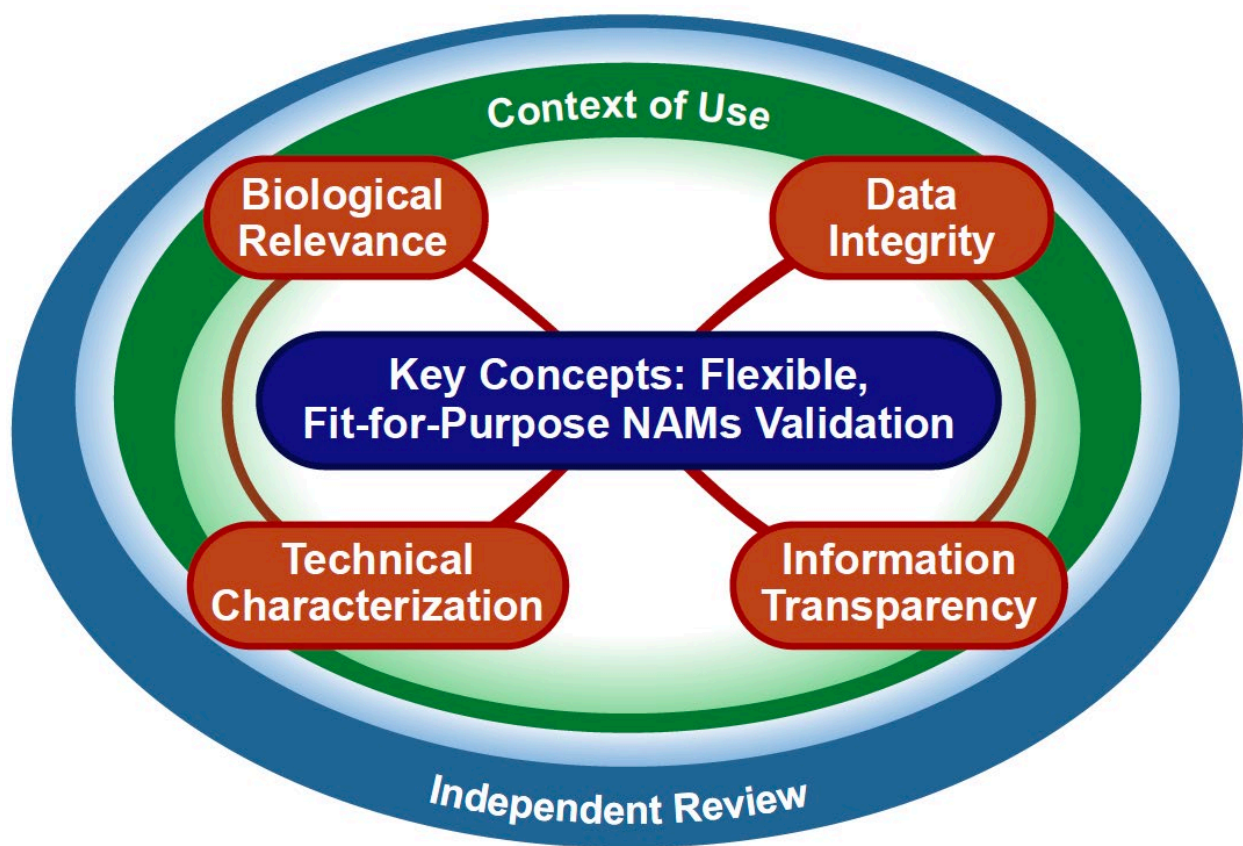


图 1. 在制定和实施灵活、目的取向的 NAM 验证策略时需要考虑的关键概念。改编自 van der Zalm 等人(2022)。

在设计和实施灵活、目的取向的验证策略时，需要考虑几个关键概念，如图 1 所示。这些概念广泛适用于为多种目的使用 NAM，但应根据具体情景和应用的需要进行调整。需要考虑的一个总体概念是 COU，或 NAM 的目的(例如，筛查、危害鉴别、效力评估、定量风险评估的出发点等)。COU 将告知验证过程的灵活性以及如何确定 NAM 对特定 COU 的适用性。这种灵活、目的取向的验证过程所涉及的关键概念是生物相关性、技术特征、数据完整性和信息透明度。如图所示，这些方面是相互关联和相互依存的。最后，验证过程的所有部分都必须接受独立审查。这些关键概念类似于 van der Zalm 等人(2022)在建立 NAM 科学信心的框架中提出的基本要素，但我们正在对其进行调整，使其适用于更广泛的 NAM 应用。第 3.0 节进一步详细介绍了这些关键概念。

3.0 应用关键概念建立对 NAM 的信心

验证应该是一个稳健而灵活的过程，通过确定 NAM 对特定 COU 的适用性来建立科学信心。在可能和适当的情况下，应有证据支持使用替代方法将提供与现有方法相同或更好的信息，并且确保其在监管审查中得出的决策对人类健康有相同的保护作用。COU 确定了关键概念的详细标准和实施。为确定 NAM 是否适合特定目的，应明确界定 COU 和 NAM 意在涵盖的相关生物学。必须对 NAM 进行详细描述，并提供技术上可靠的信息，这些信息与所关注的终点或过程具有生物相关性和/或对健康有保护作用。NAM 提供信息的组成必须足够透明，以便进行独立审查，确保完整性和可信度。在适当和可能的情况下，对 NAM 的信心建立可能包括证明 NAM 与传统动物测试方法提供的信息相比，NAM 在定量或定性上为监管决策提供的信息具有同等或更好的质量和目标物种相关性。为了避免受到现状的限制，必须承认 NAM 有可能为监管决策提供比传统动物测试方法更高质量和更相关的信息。这需要一个允许这种可能性的验证框架，并考虑到可能无法与传统动物测试方法的数据进行比较的情况。

3.1 使用情景

建立 COU 或 NAM 的预期用途包括起草一份声明，全面明确地描述 NAM 的使用方式及其监管目的(如果适用)。美国联邦机构根据每个机构特有的法律法规运作(参见，例如 Shaffer, 2021)，因此，NAM 的认可和适用于每个机构单独要求的标准不同。在某些情况下，即使在一个机构内部(例如，美国食品药品监督管理局 [FDA] 的不同中心、美国环境保护署 [EPA] 的不同办公室)，也可能根据监管领域的不同有不同的需求。这种情况在国际上更加复杂，因为其他国家的法规可能与美国的法规不同，因此一个国家认可的特定 COU 的方法在其他地方可能不被认可。因此，尽管根据特定指南/法规，NAM 可能被验证为解决一个(或多个)COU 的特定终点，但在不同机构、监管司法管辖区和国家的不同指南/法规下，其可能不被认可。为了更好地了解这种多样化的监管格局，ICCVAM 从国内和国际的各种机构收集了信息，以描述该机构对特定毒性终点的特定需求(NIEHS, 2023a; 示例如表 2 所示)。

表 2. ICCVAM 工作组制作的手稿，提供有关选定主题的机构测试需求的详细信息

需要关注的测试	参考文献
急性毒性测试	Strickland 等人，2018
生态毒性测试	Ceger 等人，2022
体外到体内外推	Chang 等人，2022
使用 NAM 进行纳米材料测试	Petersen 等人，2022a

需要关注的测试	参考文献
皮肤过敏性	Daniel 等人，2018；Strickland 等人，2019
皮肤和眼睛刺激测试	Choksi 等人，2019

具有不同 COU(例如筛选/优先顺序和危害鉴别)的 NAM 可能与旨在回答有关毒理学机制的特定问题和/或支持定量风险评估的 NAM 不同，因此，可以制定不同的标准并将其应用于评估和验证每种目的的 NAM。学术环境中开发了许多新方法，以解决本质上可能具有特定或探索性的基础研究问题。有时，人们认识到，其中一种方法可能在监管决策等更具应用性的领域发挥基础研究以外的作用。对于新方法的开发者来说，在验证 NAM 之前，与联邦机构和/或目标用户或利益相关者会面很重要，以确保旨在建立信心的验证研究能够根据预期的 COU 进行量身定制。

其他 NAM 可能对药品发现和开发更有用。可以采用 NAM 来筛选新分子的理想活性(例如药理学活性)或潜在的不良活性(毒性)。这些数据可用于决定应采用哪些分子进行进一步开发。测试方法的标准化可以将精力和资源集中在新的技术或治疗发现上，而不是方法开发或测试上，从而支持创新。NAM 的这些用途通常不在大多数联邦机构的监管范围内。但是，仍然建议应用建立 NAM 科学信心的关键概念，以确保获得高质量的数据。人们认识到，NAM 的这些用途既可以直接支持 3R，又可以提供初步数据和经验，支持 NAM 的进一步发展，有可能使其用于监管目的。

3.2 生物相关性

NAM 的相关性描述了测试与目标物种效应之间的关系，以及该测试方法对既定目的是否有意义和有用，并确定了其局限性。充分证明 NAM 的相关性是增强对 NAM 信心的重要因素。基于现有信息，可以通过多种方式证明生物相关性，详情如下。

以下小节中详述的有关生物相关性的注意事项包括：

- NAM 提供什么类型的信息？对导致结果/终点的生物学和机制有了解吗？
- 有哪些参考数据可用来衡量 NAM 计划质疑的结果？
- 关于是否以及如何将 NAM 的产出与既定实验室方法进行衡量的考虑因素是什么？

根据验证申请的具体情况，围绕生物相关性的其他考虑因素也可能很重要。

3.2.1 机制理解

在评估 NAM 的相关性时，考虑目标物种(一般是人类，但通常是其他物种)的生物学非常重要。在可能的情况下，应以现有的机制知识(例如有害结局路径 [AOP] 或毒理学相关的生物过程)来支持将 NAM 提供的信息与体内生物学进行比较。通过分子启动事件或一个

或多个关键事件将 NAM 锚定到既定的 AOP 可以帮助证明 NAM 的生物相关性。另一方面，对于 NAM 预测的结果，缺乏既定的 AOP 并不一定会排除 NAM 的潜在用处，此类支持信息可能来自体内数据提供的机制见解以及对目标物种生物学的理解。必须全面描述 NAM 的特点，明确描述正在测量的生物事件及其与不良结果或令人担忧的危害有何关系。

对 NAM 的描述应涉及用于预测体内结果的模型的生物学合理性和/或提供与生物过程、机制或 AOP 的机制联系。例如，如果 NAM 要预测人用药物诱发胎儿畸形或胚胎-胎儿死亡的可能性，则应了解胚胎-胎儿发育的机制(例如细胞迁移、分化、血管生成、神经生成、原肠胚形成)，以及与使用该模型研究的后续发育不良反应的关系(FDA，2021b)。尽管与所预测的体内效应的关系本质上可能是相关的，但仅凭相关或经验关系来建立对 NAM 预测的信心更具挑战性；因此，首选与所评估过程具有明显生物相关性的测试。缺乏对 NAM 的生物学和机理相关性的理解可能会限制其适用性，使其仅限于由验证 NAM 所使用的数据所严格定义的范围内，并使得将 NAM 扩展到建立和验证 NAM 所用化学品之外的化学品类别变得困难。

AOP 是将分子和细胞小变异与不良健康后果联系起来的有用组织框架，可用于开发和锚定代表重要毒理学过程的 NAM。有些终点确实具有完善的 AOP，例如皮肤过敏性(Kleinstreuer 等人，2018)。对于缺乏完善 AOP 的终点，可以改为根据生物路径或过程等因素来考虑 NAM 的机理相关性(表 3)。

表 3. 已证明 NAM 的生物学和机理相关性的终点示例*

终点	摘要	参考文献
皮肤过敏性	该终点具有完善的人类相关 AOP，经合组织指南 497 中描绘和描述了结合多个 NAM 的既定方法。	Kleinstreuer 等人，2018；经合组织，2021a
内分泌失调	使用互补 NAM 作为综合策略一部分的既定途径模型可用于雌激素和雄激素受体活性。EPA 在内分泌干扰物筛查计划中接受这些 NAM 用于一级筛查。	Judson 等人，2015；Kleinstreuer 等人，2017；EPA，2023
发育神经毒性	这个复杂终点的 AOP 有限。取而代之的是，已经开发出了一系列涵盖人类神经发育关键过程的 NAM。经合组织关于一系列 NAM 的 GD 已经出炉，其中包括整合测试评估方法(IATA)案例研究。	Crofton 和 Mundy，2021；经合组织，2022a；经合组织，2023

终点	摘要	参考文献
吸入毒性	采用基于体外人体细胞的测定和计算建模的替代方法来推导 EPA 人类健康风险评估的出发点。这种方法也作为经合组织 IATA 案例研究发表。	Corley 等人, 2021; EPA, 2021c; 经合组织, 2022b
眼睛刺激	一份指南文件, 描述了评估农药和农药产品可能对眼睛造成刺激的替代测试框架。对现有的体内、体外和活体外测试方法进行了审查, 以了解其与人体眼部解剖结构和毒性机制的相关性。	EPA, 2015; Clippinger 等人, 2021
皮肤刺激	一份指南文件, 建议使用现有信息、物理化学特性和其他非测试方法对皮肤腐蚀和刺激进行 IATA。	经合组织, 2014

* 有关不同 NAM 认可的信息, 请参阅各机构的具体指南。

理想情况下, 描述 NAM 与感兴趣的目标生物效应的关系应基于目标物种的关注终点的相关生物学或作用机制的现有信息。在某些情况下, 这或许是不可能的, 来自不同于目标物种的其他物种的数据可能是唯一可用的比较依据。但是, 目标物种的生理学可能与现有的替代测试物种不同, 这进一步强调了将机制学的理解纳入 NAM 评估的必要性。例如, 兔眼的解剖学和生理学方面与人类不同(Clippinger 等人, 2021), 大脑发育过程中的易感窗口因物种而异(Smirnova 等人, 2014; Tsuji 和 Crofton, 2012)。因此, 在评估 NAM 和现有参比测试方法时, 应确认与目标物种的生物相关性和关键暴露考虑因素。例如, 对生物学和机理相关性的评估可以考虑所用细胞类型的相关性、所研究的相关器官/组织的生理特征, 或与测试物质相关的物种相关代谢物的存在(注意在一个物种中发现的代谢物可能不同于在另一种物种中发现的代谢物或在 NAM 中产生的代谢物)。此外, 人们可以研究一种方法评估特定物种的作用模式或机制的能力(Hartung, 2010; Madia 等人, 2021; Parish 等人, 2020)。

除了物种翻译之外, 个体之间的变异(例如体外分析中使用的细胞或组织的捐赠者)还可能成为生物变异的基础, 这些变异既会影响测得的反应, 也有可能增加对人群中个体间变异的更多见解(Harrill, 2020)。尽管本文件没有广泛涉及, 但 NAM 的用户和开发者应意识到, 对效应(量, 在某些情况下还包括观测效应的存在)的观察, 以及效应发生的起点或剂量, 可以通过细胞系中存在的基因序列变异来调节(Chiu 等人, 2017)。在基于细胞的测试中使用遗传多样化的细胞系试剂盒可以带来优势, 有助于了解毒性和表型反应的种群动态(Harrill 和 McAllister, 2017)、发现赋予易感性的基因序列变异(Frick 等人, 2015)、识别独特的易感亚群(Church 等人, 2015)以及促进毒物动力学和毒物动力学变异的测量(Rusyn 等人, 2022)。虽然肿瘤衍生的细胞系有时用于体外模型, 但其遗传组成可能无法充分代

表人群中遗传变异。将多样性纳入 NAM 的具体建议将视具体情况而定。因此，应与遗传学家合作开发目的取向的研究设计，以确保充分覆盖目标物种变异，如果要质疑人类遗传变异，则应与生物伦理学家和利益相关群体合作开发(例如参见 ICCVAM, 2022)。评估个体间变异的具体建议超出了本报告的范围，并且可能会随着该领域的新兴科学的发展而继续发展。

在某些情况下，复杂的 AOP 或由多个 AOP 导致的结果可能需要评估不良结果机制的多个方面，以获得对结果的充分预测。NAM 通常具有局限性，无法预测或在技术上适用于所有化学类别。就 NAM 所代表的生物学的复杂性而言，NAM 也可能受到限制。将多个具有不同性能、适用性和生物覆盖范围的 NAM 组合成限定方法(DA)，可以增强其预测结果的能力(OECD, 2017)。这可以通过将适当的 NAM 映射到 AOP 的关键事件来实现，以确保有足够的机制表现来预测目标物种的相关不良后果。经合组织“第 497 号指南：皮肤过敏性的定义方法”(OECD, 2021a)中提供了将 NAM 映射到 AOP 的关键事件的示例，从而得到解决监管终点的限定方法。

3.2.2 参比化合物

参比化合物可能具有广泛的潜在用途，从进行测定到评估其技术质量和生物相关性，如表 4 所示。参比化合物可能适用于一种用途，但不足以满足另一种用途。例如，一种参比化合物即使不会对目标物种产生相关的不利影响(例如，由于无法穿过血脑屏障)，也可能适合作为阳性对照(例如发育神经毒性测定)，但它可能不适合作为生物终点参比化合物。根据 COU 的不同，参比化合物可能包括单组分化学品、混合物或复合提取物(例如，来自医疗器械)。

表 4. 参比化合物的选定用途

离散化合物集*	目的	化合物选择标准
阳性对照	验证该方法是否按预期运行。	经证实可在方法条件下使特定测定读数发生可靠、可测量且具有统计学意义变化的化合物。
高性能化合物	评估新方法或修改后的方法与类似的既定方法相比的程度(即“模仿”)。	一组在既定方法中可靠地引起反应(或没有反应)的化合物。
能力验证化合物	使用既定方法评估新实验室的性能。	在既定方法中可靠地引起反应(或没有反应)的化合物；可以是高性能化合物的子集。
生物终点参比化合物	1) 评估方法的生物相关性。 2) 比较不同方法之间的一致性，这些方法旨在在不同的背景下（例如，不同的物种、方法或测量模式）衡量与相同体内终点相关的结果。	1) 有证据表明对目标终点有体内效应(阳性)或无影响(阴性)的化合物(最好来自目标生物或合适的比较物种)。 2) 一组常见的化合物，这些化合物已通过至少一种可靠的方法进行了目标终点的测试。

* 这些组合中符合条件的化合物可能有重叠之处。

证明 NAM 科学有效性的一个关键方面是根据正在使用的现有测试方法评估其性能，通常是通过测试生物学活性已得到充分表征和理解的生物终点参比化合物。生物终点参比化合物通常包括已被证明会产生不良影响(理想情况下具有一系列效力)或不产生此类作用的化合物(Browne 等人，2019)。

鉴定对特定顶端结果具有已知影响的生物终点参比化合物可能具有挑战性。对于人类来说尤其如此，因为毒理学领域已被应用于防止人类接触潜在的危险化学品，因此数据往往缺乏。成功防止人类接触危险化学品的工作意味着可用的人类数据有限，或者此类物质在人群中没有得到充分表征或测量。当此类人类数据出现时，它们可能来自意外暴露或低水平暴露，这两种暴露水平或结果的定量数据通常都有限。案例研究可以提供危害指示，但通常不适合建立明确的因果关系。

尽管存在这些挑战，但仍应努力确定与正在开发的 NAM 预测的结果或机制相关的对人类造成影响的化合物。即使此类数据支持对该方法进行定量评估的能力有限，其也可以对 NAM 的相关性进行定性检查。高质量的人类流行病学数据也可能难以获取，也难以与毒理学数据保持一致；但是，流行病学数据可用于识别与人类健康不良影响密切相关的环境化学物质。已经开发了基于证据的框架，允许汇编大量信息，然后透明和客观地选择范围更窄的相关数据集(Wikoff 等人，2020)，从而改善了数据的可访问性。这些框架包括系统证据图，可广泛概述证据基础，以及随后的系统评价，后者对特定的研究问题进行更狭义和全面的评估，从而可以更快地将人类流行病学数据与其他来源的现有信息整合(Wolffe 等人，2019)。结合机制证据和传统动物数据，这些数据可用于定义生物学终点参比化合物，以帮助建立对 NAM 的科学信心(例如参见 Krishna 等人，2021)。

在评估 NAM 稳健性和相关性期间，应尽可能使用已被证明对目标物种产生相关影响的化合物作为阳性对照。Petersen (2021)描述了选择阳性对照时需要考虑的特征(例如，每次进行测定时都使用过程控制测量进行测试的化学物质)，许多相同的考虑因素将与评估 NAM 的参比化合物有关。阳性、阴性生物终点参比化合物、高性能化合物和能力验证化合物都应包括在内，并应同样谨慎地进行选择。由于文献中强调阳性结果的历史偏见(认识到并非文献中报告的所有阳性结果都是真正的阳性)，因此对目标终点呈阴性的可靠参比化合物可能更难识别。应充分考虑公布以阴性结果为重点的数据整理工作，而在广泛的靶标上测试浓度相对较高的化合物的工作可能会提供有价值的信息来帮助解决这个问题。

一份明确的清单可能无法涵盖所有的 COU，参比化合物的选择应考虑该 COU 的特定监管需求。在某些情况下，还应针对一系列 NAM 中每个的评估机制(而不是顶端终点)定义生物终点参比化合物的子集。例如，很少有化合物符合发育神经毒性的参比化合物，相反，一系列中评估的每个过程(例如增殖或神经突生长)都有自己的特定测定生物终点参比化合物。因此，一个生物终点参比化合物在一项测定中可能为阳性，而在另一项评估不同机制或过程的测定中可能为阴性。

在经合组织概述的正式验证过程中，需要参比化合物清单(例如，用于评估可转移性)。但是，编制可靠的参比化合物既耗时又耗费资源。在可行的情况下，通过严格的系统评价和自动化流程，努力促进参比化合物精选清单的快速制定，能够对 NAM 灵敏度和特异度进行更有力的评估(Judson 等人，2019；Thomas 等人，2019)。用于 NAM 确认的生物终点参比化合物清单的示例，可在 FDA 2021 年修订的《国际人用药品技术要求协调理事会(ICH)，S5(R3)《人用药物生殖和发育毒性检测行业指南》(FDA，2021b)中找到。该清单包括具有阳性和阴性结果的化合物。该指南指出，这些化合物以及其他化合物可用于支持特定 COU 的替代测定或一系列检测的确认。

3.2.3 与现有实验动物方法的比较

建立对 NAM 的科学信心并获得监管认可的标准通常包括考虑 NAM 能否提供与用于监管决策的现有测试方法(视每个机构的监管框架情况而定)相同或更好的有用性、科学质量和/或相关性的信息。例如，OECD GD 34 中验证任何新测试方法的标准是“该方法生成的数

据用于风险评估目的，这些数据至少与使用现有方法获得的数据一样有用，最好更好。这将为人类健康或环境提供同等或更好的保护水平”(经合组织，2005)。修订后的《美国有毒物质控制法》(第 4(h)(1)(B)条)包括 NAM 的具体注意事项，并要求 EPA 鼓励和促进“使用科学上有效的测试方法和策略，减少或取代脊椎动物的使用，同时提供具有同等或更高科学质量和相关性的信息，以支持监管决策”(15 USC §2601，2016)。

从历史上看，“同等或更好”的概念依赖于与传统动物测试数据的直接比较。但是，为了实现“更好”信息的目标，必须承认，NAM 可能无法提供与传统动物测试方法生成的相同信息，NAM 的结果可能与传统动物测试的结果不直接一致(例如，见 Clippinger 等人，2021；Hoffmann 等人，2018，2008；ICCVAM，2018；Kolle 等人，2017；Petersen 等人，2022b；Piersma 等人，2018；Prior 等人，2019；Sewell 等人，2017)。在某些情况下，NAM 可能提供足够监管决策过程的生物学相关信息、机制见解或足够敏感的终点，可能没有必要与传统动物测试方法的数据进行比较。例如，某些药理学或毒理学靶标可能不存在于非人类物种中，因此动物研究可能与评估通过此类靶标介导的潜在人类影响无关。在这些情况下，基于人类的 NAM 可能会提供与生物学相关的信息。此外，NAM 通常提供机制信息，而不是动物测试方法测得的顶端终点数据(例如，虽然观察到的动物体重下降可能无法阐明潜在的毒性机制，但 NAM 可能能够提供这些机制见解)。在某些情况下，动物模型可能正在测量复杂的生物学终点，该终点与 COU 有关，但未被有关 NAM 充分覆盖。在这些情况下，与传统动物测试方法进行比较可能是最权宜的选择，尤其是在机制尚不完全了解的情况下。对 NAM 进行评估的目的是证明 NAM 提供的信息可以促成与根据现有方法作出类似的监管决定。

NAM 与现有实验动物方法之间的比较应酌情考虑参比动物测试方法的可靠性和重现性，包括尽可能了解两种方法中观测到的差异的原因。如果可用，使用目标物种的参考数据，可以对照物种相关反应对 NAM 进行评估。如果有体内时间历程数据，则可以使用微生理系统等 NAM 来模拟毒代动力学，然后将微生理系统的计算推断与动物或人类数据进行比较。基于生理学的药代动力学建模允许根据参考剂量估算内部浓度，并与体外活性浓度进行比较。这些模型还可以应用于反向剂量测定方法，进行体外到体内外推(IVIVE)，并预测在 NAM 中观察到生物活性时，能够在血浆或靶组织中产生相应浓度的等效给药剂量。当人类是目标物种时，高质量的流行病学、临床或观察效应证据可能有助于为 NAM 建立科学信心提供见解。但是，对于大多数终点和化学品，很少有这种人类参考数据，而且比较通常必须依赖于来自不同物种的数据。可以整理动物研究的数据，并将其与产生可重复、稳健和相关结果的参考标准清单进行比较。在某些情况下，例如生态学应用、分子靶标的序列同源性或生物机制的保护，可能有助于支持与来自不同物种数据的比较 (Farmahin 等人，2013；LaLone 等人，2016)。

一些出版物评估了一系列终点的动物参比测试方法的结果，并表明这些测试的结果显示出不同的重现性水平(Browne 等人，2018；Dumont 等人，2016；Karmaus 等人，2022；Kleinstreuer 等人，2018；Luechtefeld 等人，2016；Pham 等人，2020；Rooney 等人，2021)。例如，观测到的变异性可反映固有的生物学因素、报告错误、方案差异或测试物

质纯度的变化。应尽可能解释参比动物测试方法和 NAM 数据之间的冲突结果，包括参考目标物种的生物学。这对于提供有关实验动物数据能力的真实背景是必要的，因此，必须围绕与参比测试方法进行比较的 NAM 的最大性能容量设定适当的期望(Browne 等人，2019)。

应根据现有数据的质量来考虑将 NAM 与传统动物测试进行比较的相对价值以及与 NAM 目标物种的生物学和机理相关性。实际上，在许多情况下，那些寻求建立对 NAM 信心的人需要将动物研究的历史用途以及与现有方法进行比较的要求作为一项重要的证据。根据许多法律和监管要求，来自 NAM 的信息的保护性不得低于现有方法。理想情况下，该方法将更具预测性，可以更快、更全面地生成许多化学品的相关数据(这些化学品可能很少或没有数据)，从而在监管机构和利益相关者群体中建立信心。可能有特定的 COU 允许对此规则有例外情况；但是，认可此类 NAM 作为合格方法将取决于相关机构或监管机构的具体要求。此类比较的范围可能会有所不同，具体取决于现有数据的数量和质量以及对 NAM 和参考方法中有关目标物种的理解深度。当根据参比动物测试方法评估 NAM 的性能时，必须根据在体内数据中观察到的变化，基于预期的 COU，考虑 NAM 中可接受的变化。

3.3 技术特征

技术特性是开发可广泛使用的 NAM 的关键组成部分。其包括评估 NAM 的变异来源，设计包含相关对照测量结果的测定方法，评估该测定可以使用的测试物质范围(即应用域)，以及开发合适的统计数据分析方法。证明 NAM 科学有效性的一个关键方面是，从技术角度来看，该测定方法具有足够良好的表征确保其稳健、可靠和可重现。NAM 既包括测试系统本身，也包括终点的量化方法。技术表征的某些方面与先前的关键概念重叠，例如生物相关性，其为不同用途的参比化合物的选择提供了依据(表 4)。已经通过一组具有相关生物活性的明确定义生物终点的参比化合物来检查 NAM 的性能(第 3.2.2 节)，因此本节将重点关注质量工具(Ishikawa, 1985 年)、方法开发、文档和标准等技术方面。

3.3.1 纳入选定的质量工具

总体而言，NAM 的技术特征符合包含总体步骤的框架(图 2)：1)对 NAM 的科学相关性进行初步审查，2)概念评估，3)实验室内部评估，4)统计数据分析和报告，以及 5)实验室间评估(如果需要¹)。步骤 2 到 5 的结果相互关联，每个步骤的结果可能会影响所有其他步骤。例如，实验室间测试的结果可能表明，应进行额外的实验室内部测试，或者需要修订方案以纳入新的对照测量。

¹可能有一些不需要或不可行的实验室间评估方法的例子，例如大规模机器人化高通量筛选测定和基于机器学习的计算机仿真方法。

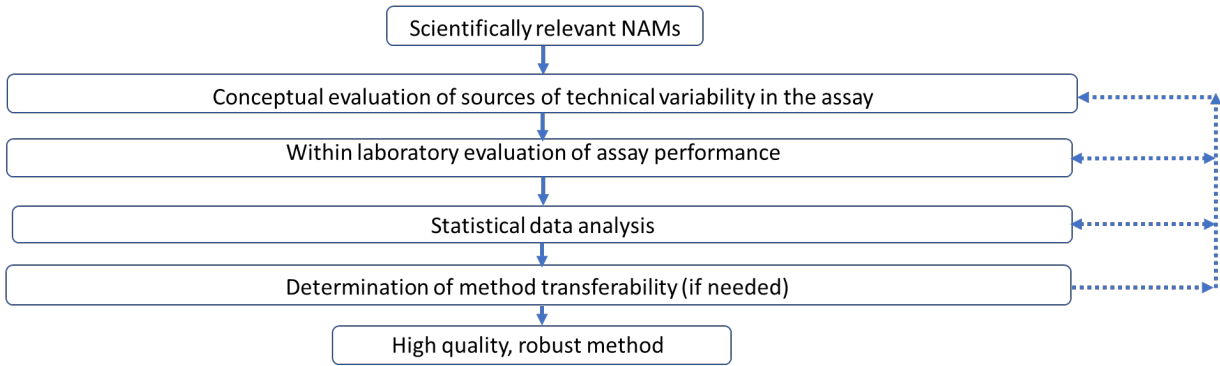


图 2. 开发稳健性 NAM 的框架。实线表示应按建议顺序采取的步骤。虚线表示必要时可以采取的重新评估任何先前步骤的方向。经 Petersen 等人许可，改编和转载(2022b)。

最初的审查步骤包括评估测定的生物相关性、其潜在的 COU 以及达到足够技术质量水平的可能性。如果 NAM 缺乏明确的 COU 或生物相关性，或者存在可能无法解决的重大技术质量问题，则 NAM 可能不适用于监管用途。在概念评估阶段，对 NAM 进行审查以评估预期的可变性来源和设计对照测量。在实验室内部评估中，可以进行实验来评估测定的稳健性，确定对照测量的典型范围(例如阴性、载体和阳性对照)，并确定对照测量与测试结果之间是否存在相互作用。通过概念评估和实验室内部评估获得的信息，可以设计一种统计模型，以评估平均测定结果、表征变异性，并确定阳性或阴性反应的标准以及该测定的统计置信度。

最后，可能需要通过实验室间测试对某些 NAM 进行可转移性评估。实验室间测试可能会揭示方案中的步骤，这些步骤在不同实验室之间有不同的解释，可以进行修订以提高 NAM 的技术质量。可转移性研究还可以识别因技术实施或方案执行方面的差异(例如，一个实验室的移液与另一个实验室的自动化)而产生的问题，这些问题需要修正。

表 5. 列出了几种有助于对 NAM 进行技术表征的质量工具。Petersen 等人(2022b)对该主题进行了更详细的介绍。这些工具生成的某些信息可能仍保留在测试方法开发者内部，而其他信息可能会被要求用于支持机构审查。

表 5. 通常用于 NAM 技术表征的质量工具*

质量工具	描述	效益/效用
流程图	绘制方案中每个步骤的图表。	优化方案步骤的识别和覆盖范围，其可通过对照测量进行监控。
因果分析	审查文献和测定背景；绘制所有预期的变异来源图表。	确定变异性的关键来源以及可能难以标准化的方法的各个方面。指导稳健性测试和对照测量的选择。
对照图表	用于评估技术性能的对照测量，包括一次性初步实验、按预定频率进行的定期测量以及每次进行测定时进行的过程控制测量。对照图表监控一段时间内的对照测量。	潜在偏差测试，评估仪器性能和校准，并测量随时间和不同实验之间变异的关键来源。
检查表	记录关键数据、元数据和对照测量。	监控过程控制测量并在出现问题时支持故障排除，支持数据分析和重现性。
散点图	绘制来自对照测量和测试物质结果的所有数据。	评估不同的过程控制测量之间，或这些对照测量与测试物质结果之间是否存在相互作用。

*详情和示例可在附录 B 中找到。

3.3.2 质量控制最佳实践规范

本节描述了联邦机构可能考虑作为评估 NAM 最佳实践规范的技术因素，包括 DA 和整合测试评估方法(IATA)。此处列出的因素可能适用于，也可能不适用于所有 NAM。此外，机构在评估 NAM 时可能需要指南中未列出的其他因素，例如额外的信息或测试。鼓励开发者和/或申办方直接与他们打算向其提交方法或数据的联邦机构进行沟通，以确定可能需要的其他审查因素。

提交的用于监管评估的 NAM 应包括对任何实验室内或实验室间研究(如果已进行)的描述。其应说明是否将 NAM 与参考方法进行了比较：另一种 NAM、体内动物数据或人类

数据。评估期间可能还需要有关用于比较的参考方法(即参考数据集)的信息。NAM 的评估可以通过几种不同的过程来完成,包括使用特定的实验室研究,记录一种方法的性能特征是否适用于预期的应用,并且是可靠的。数据的可接受性与用于评估该方法的标准直接相关。

除了记录性能和表征 NAM 的适用性,良好的科学、技术和质量规范还能确保整个评估过程高效有效,从而增强人们对拟议方法的信心。开发者应保留校准、操作和维护任何设备所需的所有信息,例如设备和软件手册、质量和安全确认证书和保修书、软件版本文档、建模算法、精选数据库和培训集。如果材料、细胞和试剂供应商的文件与 NAM 的评估有关,开发者还应保留这些信息(详见第 3.3.3 节)。有关质量规范、设备程序和要保留文件的更多详细信息,请参阅经合组织关于良好体外方法规范的指南文件(GIVIMP; OECD, 2018)。建议为所使用的每台仪器提供安装质量/运行质量/性能质量 (IQ/OQ/PQ) 报告,如果研究要符合药物非临床研究质量管理规范 (GLP),则必须提交安装质量/运行质量/性能质量 (IQ/OQ/PQ) 报告。

3.3.2.1 细胞/组织方法的相关信息

应保留任何测试方法中使用的细胞和组织起源的文件。理想情况下,保留供审查的信息应包括(但不限于):

- 伦理、法律和安全方面的考虑
- 物种/菌株/性别
- 人口统计信息(如果相关)
- 来源/供应商
- 捐赠者人数
- 起源器官/组织
- 分离出来的细胞类型
- 工程分子靶标序列(如果适用)
- 分离技术和日期
- 生物安全分类
- 细胞系识别和认证
- 传染病原体的血清学检测(例如支原体检测)
- 细胞培养传代/群体倍增数
- 倍增时间
- 遗传稳定性
- 遗传和蛋白质表达信息(如果相关)
- 预处理

更多详细信息,请参阅《良好细胞和组织培养规范 2.0》(Pamies 等人, 2022)和经合组织 GIVIMP 文件(经合组织, 2018)。

应提供关于在实验室中将新细胞和组织在适当储存条件下进行检疫的文件，然后再用于测试。利用活细胞或组织的检测应包括细胞活力测定。

所有参与的实验室都应保存所有实验室设备(读板机、孵化器、冰箱/冰柜等)的性能记录 and 文件。

在细胞库中注册细胞有助于提高可追溯性和长期可用性。细胞库应包括细胞密度和细胞传代数的记录。应保留试剂制备记录(例如，使用检查表来追踪所使用的任何试剂或消耗品)。这也与不使用细胞 /组织的 体外方法有关。一些例子是(但不限于):

- 供应商
- 目录号
- 批次/批号
- 制备日期
- 到期日期
- 分析师姓名

研究的申办方或开发者应提供与正在制定的方法或方法论有关的所有安全信息，以及有关使用、运输和处置所有危险材料的所有相关法规。

3.3.2.2 评估 NAM 中使用的分析方法

开发者应提供数据，清楚地证明检测到或量化的物质是预期的化学物质或目标分析物。表 6 列出并描述了与分析方法有关的最佳质量控制规范。

表 6. 分析方法评估*

分析方法评估	描述	效益/效用
检测和定量限	在试剂空白之上或标准曲线内可以可靠地检测或定量的分析物的最低量或浓度。标准曲线上的最高浓度决定了定量上限。	确定分析方法的范围。
识别干扰	识别该方法的组成部分何时错误地改变了检测到的信号。	确定方法组成部分之间是否存在相互作用并妨碍报告可疑结果。

分析方法评估	描述	效益/效用
评估分析精确度	描述所用分析方法和任何其他精确度测试的精确度，例如在不同人员使用拟议方法时或该方法使用不同仪器时评估性能变异性的测试。这可以通过实验室间比较研究进行评估。	建立对分析方法可靠性的信心，并评估实验室变异性的来源。
NAM 中使用的材料的稳定性	NAM 中使用的材料(例如测试物质、测试设备、暴露系统、试剂和分析物)在一段时间内在给定环境中产生相似且可接受的结果的能力。	表征方法材料并确保特定方法持续获得可靠数据。
稳健性测试	测试方法在不同条件或环境下重现的能力，而所得结果不会出现意外差异。	确定测定可接受的参数范围。
回收分析	通过比较提取样本的结果与类似基质和/或加标空白的加标样本结果来测试提取方法。	验证提取方法的效率和重现性。
应用域的技术分析	获得足够的化学品和/或产品测试方法数据，这些数据代表了与拟议测试的特定 COU 相关的化学品和/或产品，并清楚地描述应用域的物理化学特性。	降低了用于不同化学品和/或产品的测定性能的不确定性，并提供了确定化学品何时处于应用域的方法和标准。
阳性对照	在测定的检测窗口内识别与终点相关的阳性对照化合物。	为测试物质结果的比较提供一致且值得信赖的基础。
仪器校准参考标准品	使用校准标准品和/或质量控制样本校准仪器。	确保可靠的测量并确定潜在的不确定性来源。

分析方法评估	描述	效益/效用
设置质量标准	使用统计方法，根据实验室内部和/或实验室间的测试结果，设定过程控制测量的质量标准。	确保对照测量的标准足够严格，从而获得可靠的测试物质结果且不存在偏差。

*详细信息和示例见附录 C。

3.3.2.3 评估 NAM 与性能标准的准确度和/或一致性

准确度和一致性在定义上非常相似，通常被认为是可以互换的，具体取决于给定文件或段落的上下文。一致性通常定义为根据获得的结果对两种方法或测试进行比较，而准确度通常定义为一种方法或测试与参考方法或测试结果比较。对准确度和/或一致性的评估通常用于测试方法和相关数据的统计评估。在这种情况下，准确度被定义为正确预测在结果总数中所占的比例。讨论方法的准确度或一致性时使用的其他统计参数包括灵敏度、特异度、阳性和阴性预测度以及假阳性和假阴性率。

只要有可能，就可以使用明确的性能标准(例如，一组已知会产生阳性和阴性结果的平衡参考物质)来检查反应和方法的有效性。应报告一致性的定量检测标准(即灵敏度、特异度、阳性和阴性预测度、假阳性和阴性率)。当将拟议的测试方法与产生功能和机械上相似数据的具有既定性能标准的方法(例如，经合组织测试指南 [TG] 的“模仿”方法)进行比较时，应评估两种方法的一致性(包括不一致的数据)，并根据参考方法进行一致性评估。

3.3.2.4 标准操作程序和方法详情

建议拟议的测试方法采用有据可查的标准操作规程(SOP)，以支持该方法和相关实验室活动(例如测试系统操作和设备校准)的稳定性能。标准操作程序应涵盖测试和分析的各个方面。标准操作程序应包括：

- 确保测试物完整性的问责制度(例如，记录保存、安全和样本链保管)。
- 样品制备和分析工具，例如方法、试剂(如果适用，包括制造商、目录号、批号等)、设备和仪器。
- 质量控制和结果验证程序。
- 方法详情，包括完整的产品描述和配方、暴露系统、测试物质体积/重量/溶解度/剂量方案以及适当的暴露/剂量范围。

提交方还应提供用以判断测试性能和结果的操作特点清单和操作标准。组成 NAM 的技术系统的运行信息和标准可能有所不同，但标准可能包括所有控制措施、标准、暴露(剂量和持续时间)和实验组的质量控制图表或其他性能标准。应包括对用于评估数据的统计方法的描述。此外，开发者应描述如何评估实验不确定性、统计不确定性、干扰和背景。

3.3.3 文档

本节介绍了 NAM(包括 DA 和 IATA)文档的最佳实践规范。此处列出的因素可能适用于,也可能不适用于所有 NAM。此外,机构可能需要其他文件来评估 NAM。鼓励开发者和/或申办方直接与他们打算向其提交方法或数据的联邦机构沟通,以查明可能需要审查的其他因素。对于某些机构而言,在机构审查之前,可能需要对 NAM 和相关数据进行独立的同行审查。

对于新方法,文档应包括对拟议测试方法的描述,以及该方法与监管目的的相关性或如何适合特定的 COU。相关信息将包括(但不限于)测试方法和任何拟议的 COU(例如,应用于风险评估)的任何机制信息和生物相关性。文档中应包含人类或相应分类群的应用域。文档的其他重要方面详述如下。

还有其他资源可以为文档提供指导。对于体外方法,开发者可以查阅经合组织 GIVIMP 文件(经合组织,2018)。

3.3.3.1 测试物质特性和纯度

对于在 NAM 中测试的所有物质(例如对照品、参比化合物、其他测试物质),至少应报告物质的特性,最好是唯一的标识符(例如 CASRN、SMILES、InChIKey),以及供应商提供的有关该物质纯度的信息。如果实验室具备必要的能力,则使用分析方法执行额外的质量控制措施来评估物质的纯度和特性有额外的好处。收集这些信息的能力也可能取决于 COU,因为某些测试物质(例如,代表成分未知或可变的复杂混合物的环境样品)可能无法很好地表征。

3.3.3.2 方法开发

应根据该方法生成的数据对该方法进行具体、详细的书面描述。这可以采用方案、研究计划、报告和/或 SOP 的形式。应研究该方法中的每个步骤,以确定环境、基质或程序变量可能在多大程度上影响分析物的检测和/或定量。

在开发过程中,应密切注意以下因素:

- 试剂选择(具有适当的生物相关性、特异度和稳定性)。
- 测试方法或仪器(即随时可用的性能和校准程序)。
- 一次性用品(即微量滴定板和其他塑料用品)与测定测量值和测试物质的兼容性。
- 分析方法/统计方法。
- 可能引入测定变异性的步骤或过程。

应采取适当步骤,在整个方法应用过程中最大限度地减少外部或基质效应(或至少表征这些影响),特别是在开发过程中使用的化合物、基质或设备与该方法技术表征期间使用的化合物、基质或设备不同的情况下。

3.3.3.3 终点和参数测量

每个终点或分析物的测量都应经过充分测试和记录。新型 NAM 的方法开发应包含该方法可以成功测量所有相关参数的证明。应采集和记录每项实验的相关元数据，将定量数据与定性信息联系起来，例如实验条件或测试化学物质和浓度，并跟踪可能导致测定变异或批次效应的外部因素(例如进行实验的日期或技术人员)。这些数据和元数据应以可访问的格式导出和保存，以便在进行独立审查验证时能够进行参考。附录 B 中有关对照图表和检查表的部分也讨论了该主题。

3.3.3.4 使用限制

应明确确定和描述测试方法的具体优势和局限性。应列出任何潜在的干扰源，并应确定任何可能干扰测试的化学物质或化学品类别。该文件还应确定 NAM 可对哪些材料进行测试的任何已知限制。

3.3.3.5 定义明确的终点

测试方法生成的数据应充分测量或预测目标终点，如第 3.2 节所述，应明确定义该终点，并解释生物相关性。这方面的一个例子是 NAM 提供 AOP 中特定关键事件的信息。数据还应描述新的测试方法与现有测试方法之间的任何联系，或者新的测试方法与目标物种的效果之间的任何联系。应明确界定 NAM 取得阳性、阴性或无定论结果的标准，并逐步进行评估，以确保该系统的稳定性。

3.3.3.6 建立统计模型

可以使用实验室内部测试的数据来建立统计模型。这些模型可以使用贝叶斯统计方法或频率统计方法建立。直方图可用于评估通过过程控制测量所获得的数据的分布，并评估哪种类型的分布(例如正态分布)与数据拟合(有关详细信息，请参阅 Petersen 等人，2022b)。开发模型，通过测试物质的测量值和过程控制测量值来计算 NAM 的累积变异性，而不仅仅是使用测试物质的变异性数据，这很有帮助。这些信息可用于建立统计模型，该模型可以得出决策(例如，测试物质是阳性还是阴性)以及该决策的统计置信度。测试物质评估的平均值与阈值的简单比较没有考虑测试结果的可变性，也无法为决策提供统计置信度。

开发 NAM 统计模型时的一个关键问题是如何区分“阴性”和“弱阳性”结果。统计模型的阈值可以通过体内数据(可用时)来确定(Friedman 等人，2023; Karmaus 等人，2022; Pham 等人，2020)。这可能需要重复测试“临界”化合物，以评估 NAM 的重现性；例如参见指南 497(经合组织，2021a)。例如，评估剂量反应关系的统计模型可用于评估导致明确效应(例如 EC₅₀ 值)的起始点或浓度，以及围绕这些值的相关数据驱动的置信区间。使用诸如 T 检验、Z 因子或其他适当的统计标准之类的统计方法来评估 NAM 的质量也可能很重要(Zhang 等人，1999; Zhang，2011)。

3.3.3.7 测定结果的重现性

如果适用，应在提交的信息中包括技术重现性文档。该方法的重现性可以通过重复测量来评估，包括质量控制和样品。该评估应包括讨论选择用于评估重现性的物质的理由(可能在进行的任何实验室内和实验室间研究中)，以及其在多大程度上代表可能的测试结果范围。应识别和讨论异常值。应包括对任何实验室内和/或实验室间变异或变异系数分析程度的定量统计分析。应汇总历史对照数据(阴性、阳性和载体[如果适用])的中心趋势和变异测量。当多次测试相同的化合物时，比较可以是定量的(例如，获得的 EC₅₀ 值)，也可以是定性的(例如危害分类)。如果拟议的测试方法在机制和功能上与具有现有性能标准的既定测试方法(例如来自经合组织 TG)相似，则应比较两种测试方法的可靠性，并讨论任何差异的潜在影响。

3.3.3.8 数据解释程序

应明确说明每个 NAM 的数据解释程序，包括阳性和阴性反应的标准。将 NAM 合并成 DA 需要固定的数据解释程序，这些程序必须客观且不包括专家判断，从而确保在不同团体应用时得出相同的结果(经合组织，2017)。应详细记录计算算法(例如机器学习模型)和软件(包括版本号)的使用，以确保结论的重现性。

3.4 数据完整性

数据完整性是确保来自 NAM 的信息可信和可靠的关键方面。建议方法开发者对用于获取、转移和处理原始数据的流程进行内部评估，然后再将这些数据提交给外部独立方进行评估和同行审查，以确保数据的完整性和结果的可信度。必要时，应尽可能根据 GLP 的原则进行研究(21 CFR § 58; 40 CFR § 160; 40 CFR § 792; 经合组织，1998)。此外，评估机构，例如国家毒理学计划替代毒理学方法评估跨机构中心(NICEATM)，可以促进对 NAM 开发过程的质量和完整性的评估(NIEHS，2023b)。还有其他资源可以为最大限度地提高数据完整性提供指导。对于体外方法，开发者可以查阅经合组织 GIVIMP 文件(经合组织，2018)。对于数字工具和数字数据管理，开发者可以遵循 2016 年发布的“科学数据管理和组织工作的 FAIR 指导原则”(Wilkinson 等人，2016)。

3.5 信息透明度

透明度促进了人们对使用 NAM 的信任，从而加快了机构的监管决策过程和潜在的监管认可或确认。NAM 与物种、COU 和技术特征的相关性应透明地传达给同行评审者、科学界和公众。在适当的情况下，应在可以开放获取的期刊上发表描述 NAM 的 COU、生物相关性和技术特征的同行评审文章和/或在面向公众的监管文件中进行总结。理想情况下，NAM 的原则、方案、用于分析和绘制数据的原始数据文件和脚本以及报告标准应公开传播。对于包含知识产权的 NAM，经合组织提供了保持透明度的工具，包括合理和非歧视性的许可承诺条款(经合组织，2021b)。在一种方法中使用的专有或专利技术或设备可能用于满足监管测试需求。对于某些机构而言，NAM 开发者可能需要传达专有或专利信

息，以支持监管认可或确认；特定机构的指南可以指导测试方法开发者了解该机构的信息需求。

替代测试方法国际合作组织(ICATM; NIEHS, 2023c)的合作伙伴通过监管认可替代方法跟踪系统(TSAR)(EURL ECVAM, 2021)发布有关 NAM 评估和同行评审的信息。TSAR 指出了 NAM 在被接受为监管背景下的公认标准方面已达到的阶段，并提供了摘要说明和公认的方案或标准操作程序。如有可能，TSAR 还包括与 NAM 相关的记录和文件，这些记录和文件与整个流程的不同步骤有关：提交、验证、同行评审、建议以及监管认可或确认。应清楚地传达如何解释 NAM 生成的数据以及相关的认可标准，以便终端用户了解该过程并将其应用到实际环境中。

3.6 独立审查

支持 NAM 的 COU、生物相关性和技术特征的信息和数据可以由独立第三方(其成员没有利益冲突)进行科学审查；但是，必要的审查级别将取决于每个机构的法规和政策。

NAM 的评估可以通过几个不同的过程来完成，包括记录方法的性能特征，以确定其是否适合和可靠地用于预期的应用。方法的可靠性通常包括(但不限于)重现性、重复性和稳健性。但是，可能还有本报告中未列出的其他信息，这些信息可能是审查某些方法所必需的。除了 NAM 的性能和适用性外，良好的科学、技术和质量实践还确保了独立审查过程的效率和效力，从而增强了人们对拟议方法的信心。实验室应保留与评估 NAM 有关的所有信息，例如操作和维护用于 NAM 的设备所需的信息(例如设备和软件手册以及质量和安全合格证书)，以及材料、细胞和试剂供应商的文件。对于旨在符合 GLP 的研究，还可能需要为所使用的每种仪器提供 IQ/OQ/PQ 报告。对于非 GLP 研究，应保留正确安装和测试的文档，以显示设备性能符合预期。

来自 NAM 的原始数据和描述信息应可供独立第三方和/或监管机构决策者审查。NAM 的评估和独立同行评审可以由验证机构组织，例如 NICEATM、欧盟动物测试替代参考实验室(EURL ECVAM)及其科学咨询委员会以及日本替代方法验证中心(JaCVAM)。其他可以独立审查 NAM 的机构或国际组织包括美国联邦杀虫剂、杀菌剂和灭鼠剂法案科学咨询小组、欧洲食品安全局科学委员会和经合组织。或者，开发者可以资助(但不能直接领导管理)对该方法的独立审查。经过同行评审的出版物有助于与科学界共享测定信息，并且可以补充独立第三方更正式的审查，以支持在监管背景下认可和使用该方法。

独立审查的范围将根据 COU、监管框架和评估的具体方法而有所不同。开发者发送给独立审查的部分信息将包括任何实验室内或实验室间研究的记录，包括该 NAM 是否与另一种 NAM、动物体内或人类数据进行了比较。这些评估有可能有助于提高 NAM 的实验室间可转移性。

4.0 美国联邦机构对 NAM 的认可

4.1 了解监管需求和决策背景

联邦机构有不同的权限来请求、获取和使用毒理学数据。这因法律、法规和产品类别而异。NAM 在某些法律或监管背景下可能有用和充分，但在其他情况下则不然。因此，NAM 的开发者和用户必须考虑其使用背景。一些潜在的背景将包括但不限于 DA、IATA 或独立方法。

NAM 是否可以用于监管或其他目的取决于所作决定的性质、NAM 是否足以满足其预期用途，以及监管机构提交的文件在多大程度上取决于 NAM 结果以支持安全性、有效性和/或风险决策。对于需要筛选大量化合物以确定进一步评估优先顺序的应用，识别出一个值得关注的信号、具有高灵敏度和低特异度的 NAM 可能很有用。在决定哪种剂量的单一化合物可以安全暴露时，这种 NAM 可能起到的作用有限。在某些情况下，各机构可能会在数据匮乏的情况下寻找关注信号(危害鉴别)。在其他情况下，各机构可能需要就化合物做出更多的定量决策(风险评估)，并且可能不愿接受用于为这些决策提供信息的 NAM 产出的高度不确定性。性能特征的实施应反映 NAM 的使用方式，并根据每个 NAM 及其 COU 进行个性化设置。

为了促进这一进程，一些机构可能希望在将 NAM 用于监管决策过程之前对其进行评估。评估可能会侧重于拥有明确 COU 的 NAM。评估的目的之一是使监管机构能够应用使用 NAM 生成的结果，而无需重新审查 NAM 的所有基础支持数据(见表 1 中引用的文件)。

4.2 使用环境考虑因素

应明确说明 NAM 的目的(例如，危害鉴别、效力评估、定量风险评估的起始点等)，并根据该目的对 NAM 进行评估。将新开发的 NAM 的应用集中在单个 COU 上是恰当的。但是，之后可以使用合适的支持数据添加其他 COU。COU 通常需要专注于特定的监管需求。如果 COU 范围很窄，最初为 NAM 确认所做的努力可能是最成功的。可以根据需要使用其他数据来扩展 COU。各机构的监管需求各不相同，因此特定 NAM 的 COU 也可能有所不同。在进行全面确认之前，为拟议的 NAM 建立适当的 COU 至关重要。

通常应由 NAM 开发者和与 COU 相关的机构讨论确定合适的 COU。在定义可接受的 COU 之前，可能需要对 COU 进行多次迭代。随着 NAM 的适用性和局限性的进一步定义，COU 甚至可能会在数据收集过程中发生变化。

4.3 基于获得经验的信心演变

将 NAM 纳入监管用途需要监管机构和受监管行业对该方法有足够的信心。验证和确认支持了这种信心，但可能不足以确保实施。在 NAM 的用途被广泛接受之前，通常需要有关 NAM 的教育和经验。即使有验证和确认数据，试剂的成本、复杂性、可用性以及在实施和解释 NAM 方面接受过培训的人员也可能会限制 NAM 的使用。

旨在补充或取代现有方法的 NAM 的性能通常需要与现有方法进行比较。人们往往对有丰富经验的现有方法充满信心。这些现有方法可能尚未经过正式验证，但是反复成功使用现有方法以及假设动物测试的固有有效性通常会增强人们对该方法的信心。用户需要知道，在使用结果做出安全性决策时，NAM 将与现有方法一样好或更好。对于不同的 COU，不同程度的不确定性可能是可以接受的，例如，在确定优先顺序、筛选、危害特征描述再到风险评估的过程中，需要增加信心。

建立对 NAM 信心的一种机制是，用户提供来自 NAM 的数据，同时提供来自 NAM 计划取代或补充的现有方法的数据。在各行业内部和跨行业自愿共享关于 NAM 的信息有助于建立足够大的数据体系，以支持人们对这些方法的信心。随着时间的推移，各机构和行业可能会看到 NAM 如何在不影响安全性标准的情况下融入现有的评估模式。

5.0 美国和国际协调

美国联邦机构之间以及更广泛地与国际监管机构的协调将有助于确保 NAM 验证方法的协调统一，并支持其应用和实施。ICCVAM 和 NICEATM 在促进国内和全球的沟通与协作方面发挥着重要作用。

5.1 美国协调：ICCVAM 和 NICEATM 的作用

ICCVAM 授权法概述了 ICCVAM 的以下目的(42 U.S.C 2851-3, 2000; NIEHS, 2023d):

- 提高美国联邦机构测试方法审查的效率和有效性。
- 消除不必要的重复工作，并在美国联邦监管机构之间分享经验。
 - 这是通过多种方式实现的，例如月度会议、工作组、公开会议(公共论坛)和科学咨询小组(替代毒理学方法科学咨询委员会)。
- 优化美国联邦政府以外的科学专业知识的利用。
 - 这通常是通过与 ICATM 合作伙伴以及在会议上与众多科学家举行会议和研讨会来实现的。
- 确保新的和修订的测试方法经过验证，以满足美国联邦机构的需求。
 - 联邦机构的科学家和监管机构人员与开发者合作，确保正在开发的方法能够满足监管需求。
- 在可行的情况下，减少、优化或代替在测试中对动物的使用。

ICCVAM 促进跨机构和国际合作，促进替代测试的开发、监管认可或确认和使用，以鼓励减少、改进或替代动物测试方法。ICCVAM 为测试方法开发者提供指南，评估专家对替代毒理学测试方法的同行评审提出的建议，并就使用经过审查的测试方法向合适的联邦机构提出建议。ICCVAM 通过由 NICEATM 管理的特设技术工作组履行其职能，执行对开发或验证动物测试替代至关重要的特定任务。其中一个持续的例子是 ICCVAM 支持协

调一项针对 NAM 的实验室间预验证研究，该研究由环保局根据用于化学筛查的体外人体甲状腺微组织测定而开发(Deisenroth 等人，2020)。

NICEATM 是美国国家环境健康科学研究所(NIEHS)转化毒理学部下属的一个办事处，为 ICCVAM 和 ICCVAM 工作组活动、同行评审小组、专家小组、研讨会和验证工作提供技术、科学和运营支持。除了支持 ICCVAM 之外，NICEATM 还会：

- 进行测试方法分析和评估，并协调有关新颖和高优先级替代测试方法的独立验证研究。
- 通过 NICEATM 网站、综合化学环境和有关关注主题的研讨会，向测试方法开发者、监管机构和受监管行业提供信息。
- 支持 NIEHS 转化毒理学部的活动，尤其是那些为美国政府跨机构 21 世纪毒理学 (Tox21)联盟做出贡献的活动。

ICCVAM 提供的跨机构沟通论坛和 NICEATM 提供的支持旨在确保有效利用有限的资源来协调美国联邦机构对 NAM 进行监管申请的验证和确认工作。

5.2 美国协调：为推进 3R 而开展的额外联邦合作

除了参与 ICCVAM 外，美国联邦机构还通过多种方式就 NAM 进行合作。例如，Tox21(Tox21, n.d.)是环保局、NIEHS 转化毒理学部、美国国立卫生研究院美国国家转化科学促进中心(NCATS)和 FDA 之间的联邦合作项目。Tox21 的目标是开发更好的毒性评估方法，以有效测试某些化合物是否有可能破坏人体的生物过程并对健康造成负面影响。Tox21 已经发布了许多开创性的出版物和分析，这些出版物和分析已投入监管机构使用，例如用于环保局的内分泌干扰物筛查计划(EPA, 2023)。NICEATM 和 EPA 还密切合作开展了多项研究，以减少或取代在监管测试中对动物的使用。这包括回顾性分析，以消除将动物用于：1)对农药和农药制剂的皮肤急性毒性(EPA, 2020)；2)确定仅使用体外数据是否足以推导皮肤吸收因子进行农药的人体健康风险评估(Allen 等人，2021)；以及 3)确定是否可以通过不到三种鱼类体内急性毒性测试对非靶向水生脊椎动物达到相同水平的保护(Ceger 等人，2023)。FDA 还与 NCATS 合作，进一步开发微生理系统技术，以促进其进步并加速转化使用(FDA, 2023)。

5.3 国际协调

除了 ICCVAM 促进的国际合作外，美国各机构还独立开展国际合作，以推动 NAM 的认可。例如，包括 ICATM、美国参与联合国全球化学品统一分类和标签制度(GHS)专家小组委员会、与 ICH 合作制定描述 NAM 使用情况的指南，以及参与经合组织测试指南计划和危害评估工作组。

ICATM 是美国(ICCVAM)、日本(JaCVAM)、欧盟(EURL ECVAM)和加拿大(加拿大卫生部环境健康科学与研究局)的验证组织合作建立的。其他参与组织包括韩国替代方法验证

中心、巴西替代方法验证中心以及中国食品药品监督管理局和广东疾病预防控制中心。该小组的总体目标是：

- 在验证研究、独立同行审查和制定统一建议等关键领域建立国际合作，以确保替代方法/策略更容易在全世界得到认可。
- 建立必要的国际合作，确保监管部门采用的新替代测试方法/策略能够为人、动物和环境提供同等或更好的保护，同时在科学可行的情况下取代、减少或改进(减少痛苦和痛苦)动物使用。

在联合国内部，全球统一制度(GHS)小组委员会成立了一个工作组，负责更新全球统一制度的各个章节，为各种危害类别(例如皮肤刺激/腐蚀、眼睛刺激/严重眼损伤、皮肤过敏)中使用 NAM 制定具体标准。这些努力在国际层面推进 NAM 方面取得了丰硕成果，并将 DA 确立为认可的危害确定方法。这些国际努力不仅对推进 NAM 很重要，而且对向在这一领域可能没有足够资源的国际区域提供专业知识也很重要。

ICH 的活动也说明了统一 NAM 使用方法的價值。ICH 制定的几份指南描述了替代方法的使用，这些方法被世界各地的多个监管机构和行业团体视为认可的方法。这些方法由 ICH 流程中的专家工作组进行了评估，并酌情纳入了指南。例如，包括使用化学和体外方法评估光毒性(ICH, 2013)。此外，ICH 关于生殖和发育毒性的指南包括替代检测的一些使用背景，以及对此类检测的确认方法的建议和生物终点参比化合物清单(FDA, 2021b)。

经合组织是协调监管测试指南和指导文件的国际论坛，越来越关注 NAM 的验证和使用。ICCVAM 在协调和促进美国对经合组织健康影响测试指南计划的立场方面发挥着重要作用。美国国家协调员代表美国参加国家协调员工作组的年度会议和其他测试指南制定活动。在这一角色中，美国国家协调员就涉及 3R 任何方面的经合组织测试指南(TG)活动征求相关 ICCVAM 机构的意见。来自 ICCVAM 机构的主题专家在多个经合组织专家组中任职，为经合组织产品的开发提供科学指导，例如测试指南、DA 和指南文件。38 个经合组织成员国的利益相关者使用经合组织测试指南来评估化学品安全。经合组织相互认可数据条款确保使用经合组织测试指南生成的安全性数据将被所有成员国接受，从而避免重复测试。ICCVAM 机构还为经合组织 IATA 案例研究项目做出了贡献，该项目允许各国分享和合作使用 IATA 的新方法，在监管背景下评估化学品安全。

6.0 鼓励使用 NAM 的宣传和培训

各机构就特定 NAM 的可接受性进行沟通，而关于 NAM 的培训可以促进其使用。在适当和可行的情况下，各机构可以公开沟通何时以及如何可以接受 NAM，具体取决于特定机构的规则和政策。例如，可以在 GD 或公开网站上描述 NAM。监管机构可以使用现有的培训计划或为员工实施新的计划，并在可能的情况下就新方法提供公开的培训。甚至在 NAM 通过科学界的教育得到验证或确认之前，就可以开始对新方法建立信心。这些早期

教育工作可以侧重于新方法的基础科学。随着方法的成熟和支持方法有效性的数据积累，教育工作可以转向让相关社群熟悉这些数据。当针对特定 COU 对 NAM 进行评估时，可以进行有关 NAM 的具体用法和解释的培训。经合组织、ICCVAM 等实体机构以及其他验证组织和科学协会也可以提供有关 NAM 的培训和获取信息的机会，这将增强人们对使用 NAM 的信心。

NAM 开发者、行业用户和监管机构之间的互动可以促进方法的开发和采用。这种互动的实现可以通过所有各方参与科学会议来讨论此类方法，以及根据机构特定程序进行更正式的监管互动。如上所述，开发适当的 COU 和确认数据集可能是一个迭代的过程。随着数据的积累，在 NAM 的开发和探索过程中，COU 和应用域可能会发生变化。在此过程中，各方继续进行沟通有助于确保采取适当途径认可 NAM。

即使 NAM 可用于某个终点并被监管机构认可，申请的申办方也可以选择使用其他方法，例如传统的动物测试，因此不一定会在监管申请中提交 NAM。这可能有几个监管机构无法控制的原因。尽管监管机构可以建议或推荐用一种或多种 NAM 代替传统测试，但化合物的申办方可能并不总是有义务遵循该建议。要为采用 NAM 建立足够的信心，所有利益相关方都必须接受教育并熟悉 NAM。

7.0 结论与实施

本报告旨在帮助方法开发者、受监管的行业利益相关者和联邦机构开发、验证、确认和认可具有科学意义的 NAM。在这里，我们描述了应考虑的关键概念，以便高效、及时地开发目的取向、可靠并提供与目标物种相关信息的 NAM。所有信息可能适用于，也可能不适用于任何特定的方法、DA 或 IATA。可能还有其他适用于方法、DA 或 IATA 的概念未在本报告中讨论。开发者必须与联邦机构和终端用户密切合作，着眼于 NAM 的预期用途，尤其是监管审查中的风险评估应用。建立对 NAM 的科学信心，验证或确认用于特定目的和 COU 的方法，应是通过利益相关者之间的多向沟通逐步演变的迭代过程。

NAM 领域是一个不断演变的领域，在 NAM 验证和确认方面可能会出现一些在撰写本报告时未曾预料到的新考虑因素。因此，参与 NAM 开发、验证和确认的利益相关方可能需要对此处未述及的考虑因素保持灵活性和持开放态度。该报告将根据需要和定期更新。

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>

附录 A：术语表

免责声明：本术语表旨在支持本报告中术语的使用。特定术语的含义或用法可能因监管机构而异。

3R：人道实验技术的原则，特别涉及到减少(即尽量减少使用的动物数量)、代替(即使用低阶物种或非动物进行测试)或优化(即最大限度地减少研究动物可能因特定技术而遭受的疼痛和痛苦)进行研究和化学品安全测试中的动物使用。

准确度：测试方法结果与公认的参考值之间的接近程度。

有害结局路径 (AOP)：以结构化形式表示当生物体暴露于某种物质时，在生物组织的不同层面上发生的、导致不良影响的连续事件。

应用域：可以使用某种方法测试的化学品类型，或由该方法得出的相关结果被认为是可以接受的化学品类型。

测定：使用的实验系统。通常与“测试”和“测试方法”互换使用。

平衡准确度：一种统计指标，用于解释不平衡的数据集，其中一个“类别”的出现次数比另一个“类别”多得多(例如，阴性数多于阳性数)。平衡准确度以灵敏度和特异度的算术平均值计算。

生物相关性：一种评估目标分类群中化学品影响的适当性的衡量标准。

一致性：两个变量之间的接近程度或一致性。一致性可用于定性描述 NAM 与参比动物测试方法和/或人类参考数据(如果有)相比的生物相关性。一致性还可用于定量描述所有测试的化学品中被正确归类为阳性或阴性的比例，在这种情况下，该术语通常与“准确度”互换使用。

使用情景(COU)：清晰阐述特定模式、方法或应用的使用方式和目的的描述。

精选数据库/列表：一组经过精心编译和维护的结构化、具有良好特征且可靠的信息。

限定方法(DA)：由一组定义的信息源和固定的数据解释程序生成的输入数据组成，以得出可单独使用或与 IATA 内的其他信息源一起使用的结果，以满足特定的监管需求。测试和评估的限定方法可用于支持化学品的危害鉴别、危害描述和/或安全评估(OECD, 2017)。

不一致：所有经过测试的化学品中被错误地归类为阳性或阴性的比例。

终点：测试方法评估的生物或化学过程、反应或影响。

假阴性：相对于指定的参考数据，通过测试方法错误地鉴别为阴性的物质。

假阳性：相对于指定的参考数据，通过测试方法错误地鉴别为阳性的物质。

目的取向：将一个 NAM(或一组 NAM)提供的信息的类型和确定性与给定决策所需的信息类型和确定性相匹配(EPA, 2021a)。

良好细胞培养规范(GCCP)：为实验室实际使用而开发的一组原则，旨在确保体外(基于细胞)工作的重现性并提高科学数据的质量(Pamies 等人, 2022)。

良好体外方法规范(GIVIMP)：经合组织良好体外方法规范指南文件(经合组织, 2018)中描述的综合框架，为体外方法的开发、验证、监管认可或确认以及使用提供了建议。

良好实验室规范(GLP)：美国环保局、美国食品药品监督管理局和经合组织等机构颁布的法规，描述了实验室记录的记录保存和质量保证程序，这些程序将作为向国家监管机构提交数据的依据。

危害：可能对健康或生态造成不利或有害影响。

危害分类：根据特定毒性终点的标准测试方法的结果，将化学品或产品危害划分为某个类别；最常用于标签目的。

危害鉴别：风险评估的一部分，与确定暴露于某特定物质是否会或可能与不利的健康或生态影响有关。

整合测试评估方法(IATA)：一种基于多种信息源的方法，用于化学品的危害鉴别、危害特征描述和/或安全评估(OECD, 2017)。

实验室间重现性：衡量使用相同方案和测试化学品的不同的合格实验室能否产生定性和定量的相似结果。实验室间重现性是在预验证和验证过程中确定的，其表明测试方法可以在多大程度上在实验室之间成功转移。

机理相关性：评估化学品可能发挥作用的生化过程或途径的适当性度量。

阴性预测度：在经测试方法检测为阴性的物质中，正确的阴性反应相对于定义的参考数据的比例。它是测试方法准确度的一个指标。阴性预测度是测试方法的特异度和测试物质中阴性比例的函数。

新测试方法(NAM)：广泛指代能够提供化学危害和风险评估信息并支持替代、减少或优化动物使用(3R)的任何技术、方法论、方法或其组合。

性能：测试方法的准确度和可靠性特征(参见“准确度”、“可靠性”)。

性能标准：基于经验证测试方法的标准，为评估在机制和功能上相似的拟议测试方法的可比性提供了基础。包括(1)基本测试方法组成部分；(2)参比化学品列表(从一些化学品中选出，这些化学品用于展示经验证的测试方法的公认性能)；以及(3)在使用最少量的参比化学品进行评估时，拟议的测试方法应展示出与经验证的测试方法相当的准确度和可靠性水平。

阳性预测率：在经测试方法检测为阳性的物质中，正确的阳性反应相对于定义的参考数据的比例。它是测试方法准确度的一个指标。阳性预测率是测试方法的灵敏度以及所测试物质中阳性比例的函数。

效力：衡量物质的相对生物或化学活性的指标。由于不同的生物或生化效应，单一物质的效力可能会有所不同。

精确度：某一个分析物(通常是单个样品)在多次分析后，单个测量值的接近程度。精确度通常用变异系数表示。

方案：对测试方法的精确度的逐步描述，包括所有必要试剂的列表以及生成和评估测试数据的所有标准和程序。

确认：得出结论，即使用经过验证的模型或测定的评估结果可以用来对产品开发 and 监管决策进行具体的解释和应用。

质量控制：一组活动或样本，用于验证产品或方法的质量是否保持在预期水平。

参比化合物：选择在研究、开发或评估拟议测试方法期间使用的化学品，因为其在参比测试方法中的反应或目标物种是已知的(见“参比测试方法”)。

参比测试方法：用于监管目的的公认测试方法，用于评估测试物质对目标物种的潜在危害。

可靠性：衡量实验室内部和实验室之间在一段时间内可以重复执行测试方法的程度。它是通过计算实验室内和实验室间重现性以及实验室内重复性来评估的。

重复性：在给定时间段内、在相同条件下对同一物质进行检测时获得的测试结果的一致性；“... 在相同条件下，使用相同方法对相同对象(或物体或测试材料)获得的独立结果之间的接近程度”(Slezák 和 Waczulíková, 2011)。

重现性：使用相同测试方案和测试样本获得的单个测试结果的一致性；“...使用相同方法对相同对象(或物体或测试材料)但在不同的条件(不同的观察者、实验室等)下获得的独立结果之间的接近程度”(Slezák 和 Waczulíková, 2011)。

回收：一种定量方法，通过将提取样品的结果与类似基质和/或加标空白的加标样本的结果进行比较，来验证提取方法的效率和重现性。

风险评估：评估暴露于外源物对目标物种的潜在不利健康和环境影响。

稳健性：一种方法在不同条件或情况下重现的能力，而不会在获得的结果中出现意外的差异。

灵敏度：在测试方法中被正确归类为阳性的所有阳性化学物质的比例。在检出限的背景下，灵敏度也可以定义为可以以可接受的准确度和/或精确度测量的最低分析物浓度。

特异度：在测试方法中被正确归类为阴性的所有阴性化学物质的比例。它是测试方法准确度的衡量标准。这个词还用于描述分析方法检测特定分析物的能力。

稳定性：测试材料(例如测试物质、测试设备、试剂或分析物)在一段时间内、在给定环境中产生相似且可接受的结果的能力。

标准曲线：一种绘制测定数据的定量方法，通过比较未知样本与已知浓度的标准样本(通常使用阳性对照材料)来确定未知样本中物质的浓度。

标准操作规程(SOP)：描述如何进行具体实验室操作的正式书面程序。这些是 GLP 指南所要求的。

靶器官：寻求有关化学品潜在毒性信息的器官。

目标物种：寻求有关化学品潜在毒性信息的物种。

测试：使用的实验系统；可与“测试方法”和“测定”互换使用。

测试方法：一种用于获取物质或制剂特性信息的过程或程序。毒理学测试方法生成有关物质或制剂在特定条件下产生特定生物效应的能力的信息。可与“测试”和“测定”互换使用。另请参阅“经验证测试方法”和“参比测试方法”。

测试方法开发者：最初设计测试方法并确保其重现性和适用于预期用途的组织或个人。

测试方法赞助者：提交测试方法以供审议的组织或个人；也可以是与测试方法开发者相同的组织或个人。

测试方法提交：为监管或其他特定应用提出的测试方法的支持文件汇编。测试方法提交通常包括已完成的验证研究记录，这些记录旨在描述该测试方法对特定拟议监管测试要求或应用的用处和局限性，以及根据测试方法提交指南编写的其他充分的科学有效性文件。

可转移性：测试方法或程序在不同的合格实验室中准确、可靠地执行的能力。另见“实验室间重现性”。

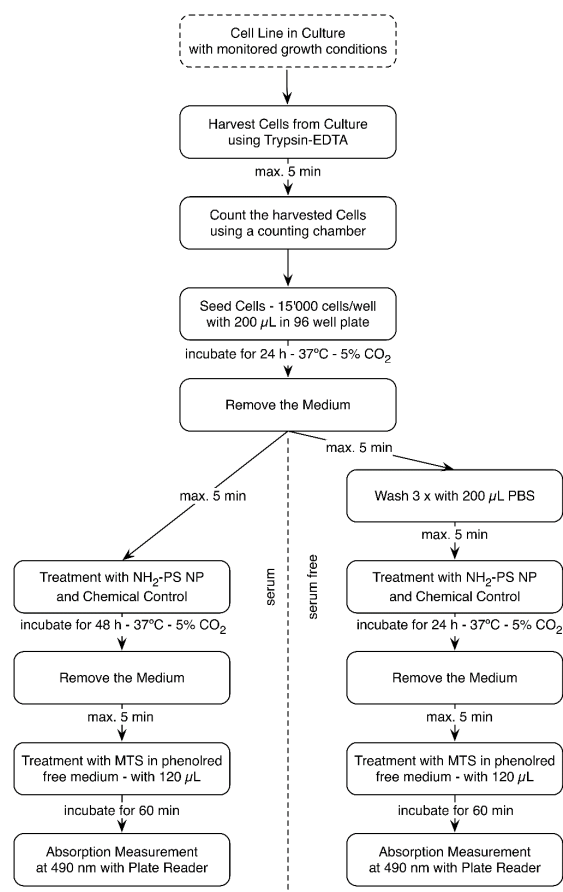
经验证测试方法：一种公认的测试方法，已完成验证研究，以确定该方法与特定拟议用途的准确度、可靠性和相关性。

验证：为特定目的确定程序的准确度、可靠性和相关性的过程。为一个特定目的进行验证并不意味着为其他特定目的进行验证。对于特定的使用环境，可能需要进一步的确认。

附录 B：质量工具

1.0 流程图

在流程图中，方案中的每个步骤都予以图示(有关示例，请参见图 S1)。通过确保在可行的情况下涵盖和跟踪每个步骤，这有助于设计实验的对照测量。多个对照测量可能涵盖一个步骤，反之，一个对照测量可能涵盖多个步骤。此外，比较 NAM 之间的对照测量结果可能会发现不同 NAM 之间的相似步骤。对于这些步骤，变异性的来源也可能相似。

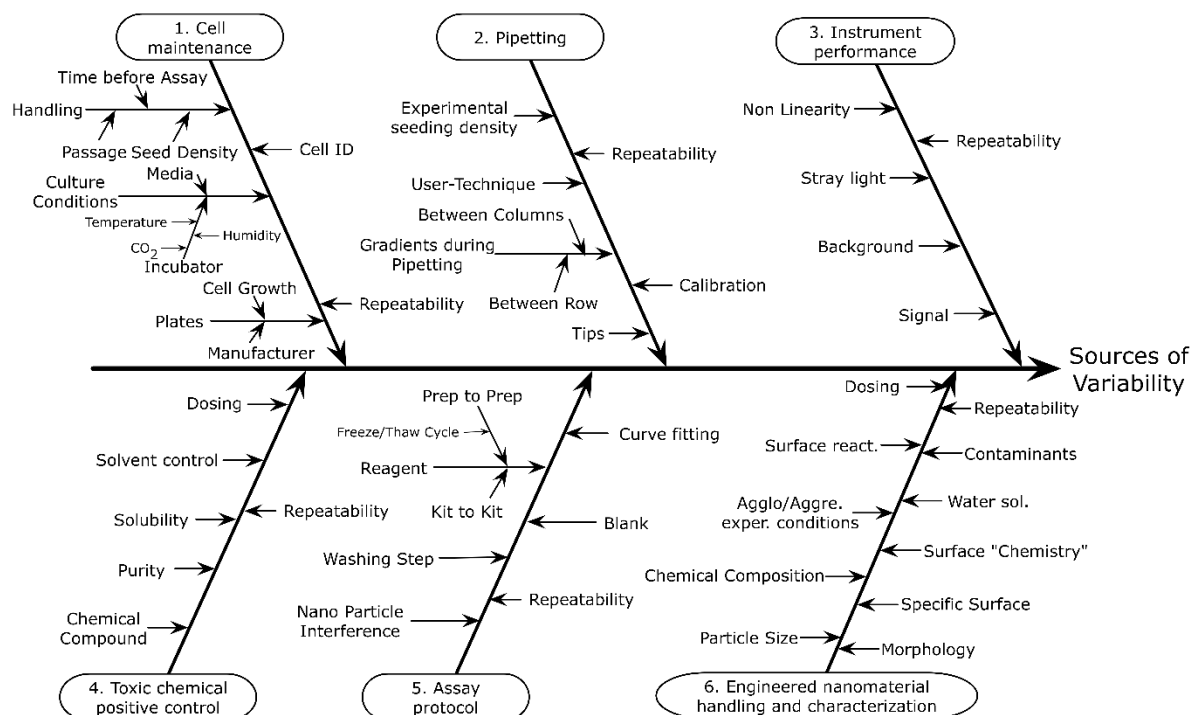


补充图 1. 描述实验室间研究修改后的 MTS(3-(4,5-二甲基噻唑-2-基)-5-(3-羧甲氧基苯基)-2-(4-磺基苯基)-2H-四氮唑)方案的流程图。本图经 Elliott 等人(2017)许可转载。

2.0 因果分析

因果分析 (C&E) 是一种概念工具，可用于鉴别可能的关键变异性来源，并使用 C&E 图 (有时称为鱼骨图) 进行展示(有关示例，参见图 S2)。绘制这些图表的过程可能包括集思广

益和审查有关测定的文献。C&E 图的每个分支都指示预期变异性的关键来源。C&E 图还可以支持新 NAM 的开发，因为 C&E 图的共享分支(例如，使用相同类型的细胞毒性测定进行测量)可以更快地制定，可能需要类似的对照测量，并且可能具有类似的减少变异性策略。对 C&E 图的分析可以帮助确定可能难以标准化的方法的各个方面(例如，难以校准的仪器或不稳定的测定试剂)。C&E 图可以帮助指导稳健性测试和对照测量的选择，因此，理想情况下，可以分析 C&E 图的每个分支和子分支的变异来源。



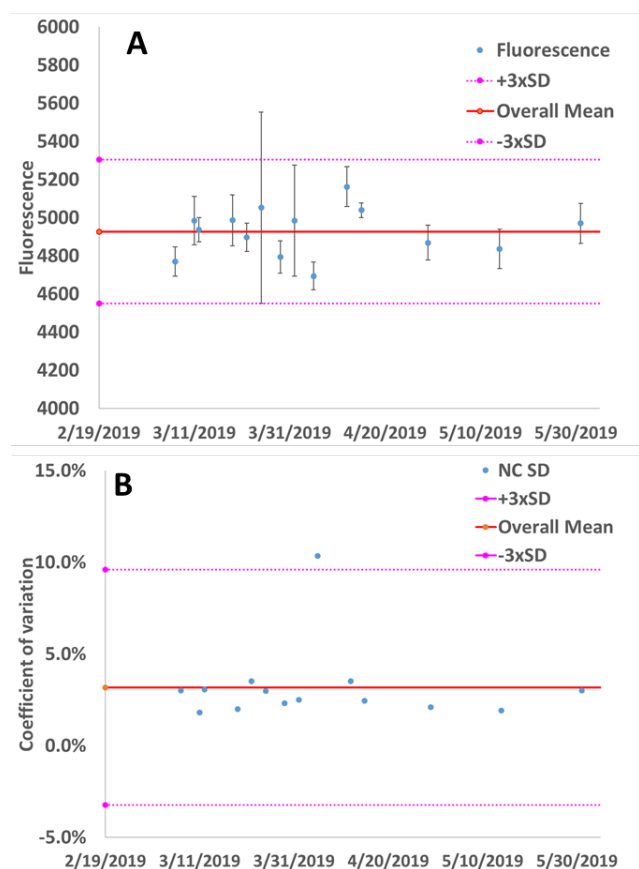
补充图 2. 设计用于工程纳米材料的 MTS 细胞毒性测定的因果图。经 Rösslein 等人(2015) 许可转载。

3.0 对照图表

对照测量可以是一次性的初步实验(例如，测试潜在偏差)、以预定频率进行的定期测量(例如，仪器性能和仪器校准评估)，以及每次进行测定时进行的过程控制测量。对于允许同时分析有限数量样本的某些检测(例如，使用流通式暴露系统进行吸入测定，暴露于气溶胶化学物质)，定期测量可能会有所帮助。一次性初步实验有助于评估测试物质是否可能导致偏差，例如，纳米颗粒是否会吸附关键测定试剂。每次进行测定时，过程控制测量可用于测量关键的变异性来源。例如，一种常见的过程控制测量是阳性对照。这种对照测量可以揭示是否达到最大响应(100% 效应)，还可用于证明测定对产生更低、更温和反应的化学浓度的灵敏度。Petersen 等人(2021)描述了选择阳性对照的关键考虑因素。其他常见的过程控制测量包括没有细胞和没有额外测定试剂的对照组以及添加了细胞和其他测定试

剂但没有测试物质的对照组。它可能无法包括所有潜在的过程控制测量，并且在包括哪些过程控制测量方面可能需要权衡取舍。

使用展示过程控制测量值的平均值和变异性的对照图，可以在不同时间段和实验之间监控关键变异性来源(示例参见图 S3)。这有助于评估平均值或变异值是否随时间推移发生系统性变化，这表明测定中的某些内容可能发生了变化(例如，试剂的不稳定性)。审查检查表有助于确定发生变更的原因。



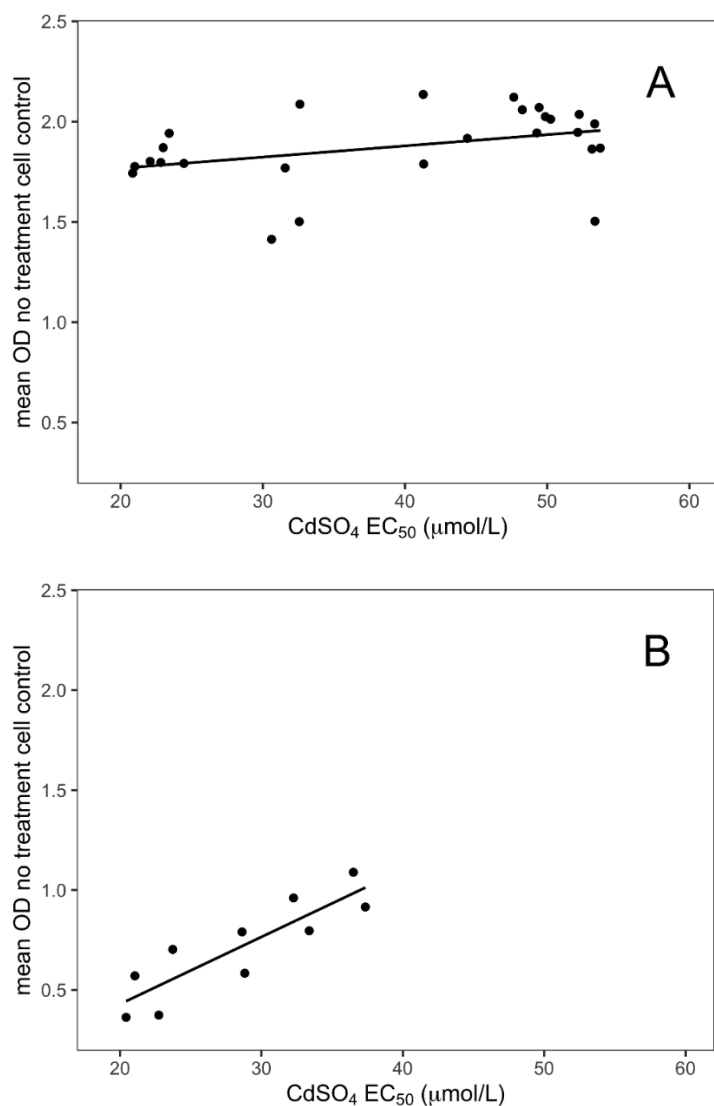
补充图 3. 阴性培养箱对照的亲电过敏原筛选测定荧光方法的对照图数据：所有实验的(A)平均值和(B)变异系数，具体取决于测定日期。在图 B 中，一个值是变异系数的异常值，不在本研究的质量标准范围内（总体平均值 ± 3 倍平均标准差值）。此外，不存在随时间推移的均值或变异系数值的系统性趋势。经 Petersen 等人(2022c)许可，对该图进行了修改和重印，而图形标题则由 Petersen 等人(2022b)修改和转载。

4.0 检查表

要记录关键元数据并满足 GLP 要求，可以使用检查表。如果在开发方法时不使用 GLP，则 GLP 中规定的原则(例如使用检查表记录关键数据)对于 NAM 用于监管提交(可能需要遵守 GLP)至关重要。测定开发期间的替代相关资源是 GCCP 2.0(Pamies 等人，2022)和 GIVIMP(经合组织，2018)的指南文件，这些文件专门针对基于细胞的工作，以及额外的质量管理列表，以确保重现性和高质量的科学数据。检查表还可用于监控过程控制测量并在出现问题时支持故障排除。例如，其可用于记录试剂和消耗品(例如移液器和微孔板)批号的变化。如果一项质量标准未达标的频率高于其他质量标准，这对于故障排除很有价值。包含数据计算器的检查表也可以用来评估测定运行的数据是否符合所有质量标准，并对测定结果进行统计评估。可以使用电子笔记本和实验室信息管理系统代替书面检查单(具有易于存储、实验室间可转移性和可搜索等优点)。

5.0 散点图

散点图可用于评估不同的过程控制测量之间或这些控制测量值与测试物质结果之间是否存在相互作用(有关示例，请参见图 S4)。在质量标准设定的范围内，测试物质结果应独立于过程控制测量结果。



补充图 4. 使用 MTS 测定法测定的硫酸镉(CdSO_4) EC_{50} 值与未暴露于测试物质的 A549 细胞的平均光密度(OD)值或阳性对照之间的相关性。这些数据要么显示 EC_{50} 值(A 部分)之间缺乏相互作用, 要么显示相互作用(B 部分), 具体取决于平均 OD 值的范围。实线是线性回归拟合。从统计学上讲, B 部分中的斜率与 0 不同, 这表明 EC_{50} 值与 OD 值相关。经 Petersen 等人(2022b)许可修改和重印, 图形标题经 Petersen 等人(2022b)许可转载。

附录 C：分析方法评估

附录 C 中的信息来自多个参考文献，包括“毒理学测试方法的验证和监管认可”(ICCVAM, 1997)、“行业生物分析方法验证指南”(FDA, 2018)、经合组织 GD 34(经合组织, 2005)和经合组织 GIVIMP 文件(经合组织, 2018)。附录 C 中的信息旨在为方法评估提供一个总体框架。某些方面可能适用于，也可能不适用于所有方法。

1.0 检测限和定量限

检测限和定量限定义了分析方法的范围。分析方法通常在定量下限或检测下限进行测试，以更适合该系统者为准。该过程通过确定分析物的最低数量或浓度来确定分析方法的范围的下限，该分析物在试剂空白以上或标准曲线内可以可靠地进行检测(检测限)或定量(定量下限)。标准曲线上的最高浓度决定了定量上限。

2.0 干扰鉴别

方法开发者应记录干扰物质，其可能来自方法的关键和非关键成分，包括对该方法检测到的信号(例如荧光/吸收、荧光素酶、酶)的任何干扰。干扰也可能来自消耗品，例如内分泌干扰物测试方法中的某些塑料以及该方法中的其他试剂。潜在的干扰物质包括但不限于内源性基质成分、代谢物、分解产物和其他外源性物质。如果该方法旨在量化多个分析物，则应测试所有预期分析物，以确保没有干扰。空白样本通常用于测试干扰。

3.0 评估分析精确度

描述所用分析方法的精确度以及任何其他精确度测试会很有帮助，例如评估不同人员使用拟议方法时的性能差异性，或者该方法在使用不同仪器进行测试时的性能差异性，以及在可能的情况下参与实验室间的比较研究。

4.0 材料稳定性

与 NAM 中使用的材料(例如测试物质、测试设备、试剂和分析物)相关的稳定性是指材料在一段时间内、在给定环境中产生相似且可接受结果的能力。前面讨论了 NAM 本身的稳定性(附录 B, 第 3.0 节)，并评估了随时间推移的对照图表。应评估样本收集、处理和储存条件的影响。应适当为可能受到稳定性问题影响的材料指定有效期。应确保测试物质、试剂和测试设备(例如塑料微孔板)的稳定性，以避免降解产物的干扰和实际应用剂量的变化。对任何成分进行的稳定性研究都应包含在报告中。在研究当天之前制备的任何化学混合物或制备样品的稳定性在用于开发、验证或确认研究之前，应评估其稳定性。

通过多种方式评估给定混合物或基质在给定时间间隔的特定条件下的化学稳定性。研究前的稳定性评估应涵盖研究期间的预期样本处理和储存条件，包括测试场地、装运期间和所有其他二级场所的条件。分析物在特定混合物、基质和容器系统中的稳定性仅与该混合

物、基质和容器系统有关，不应推断到其他系统。稳定性测试应评估分析物长期(冷冻在预期储存温度)和短期(台面、室温)储存，以及冷冻、解冻循环和分析过程后的稳定性。稳定性实验中使用的条件应反映在实际样品处理和分析过程中可能遇到的情况。如果在研究样品分析期间，储存条件发生变化和/或超出了该方法技术表征期间评估的样品储存条件，则应在这些新条件下建立稳定性。应进行独立的稳定性研究，并涵盖关键试剂可能遇到的任何情况。应考虑在各种情况下对试剂或材料进行稳定性测试，包括在冻融循环期间以及在台面、库存溶液、处理过的样品中以及长期研究中使用。

5.0 稳健性测试

稳健性是一种方法在不同条件或情况下重现的能力，而不会在获得的结果中出现意外的差异。NAM 稳健性的一个方面是对照图表的结果随时间推移的一致性，如附录 B 第 3.0 节所述。稳健性测试通常用于检测实验试剂或实验方案中意外变化导致的结果变化。建议对测试方法的各个方面进行稳健性测试，并应尽力在任何可能的情况下为所有参数和测量值建立范围。

例如，研究的最佳孵育时间为 5 分钟，但经过稳健性测试，数据在 5 分钟 $\pm$ 30 秒时通过了所有质量控制要求。因此，经稳健性测试的可接受孵育时间为 5 分 $\pm$ 30 秒。

一些应具有既定接受范围的研究参数包括：

- 孵化时间
- 孵化温度
- pH 值
- 试剂来源
- 细胞密度(如果适用)
- 实验条件
- 分析软件

该模型的设计将决定哪些实验条件需要可接受范围。例如，具有流动介质的模型需要与流量相关的参数的可接受范围。更复杂的模型更有可能具有需要控制的研究参数。

如果适用，还应报告关键和非关键试剂的稳健性测试结果。应对不同的供应商进行测试(只要可行)，以确定是应从一个供应商那里购买试剂或是可以使用多个供应商。这也有助于避免供应链中断或生产商停止生产特定产品。

如果适用，还应进行仪器稳健性测试。跨实验室验证通常涉及具有不同性能参数和能力的不同品牌仪器，这可能会增加数据的变异性或改变所用方法的参数。

6.0 回收分析

对给定基质提取物进行测试的研究应进行回收研究，以验证提取方法的效率和重现性。回收研究通常通过将提取样本的结果与类似基质和/或加标空白的加标样本结果进行比较来进行。

7.0 应用域的技术分析

对于与拟议测试特定 COU 相关的化学品和/或产品，应有足够的测试方法数据。NAM、DA 或 IATA 的文件应明确描述应用域的物理化学特性。这将包括该方法的任何限制，例如特定分子量范围、挥发性、溶解度或稳定性的化学品。

虽然应在测定方法的初始开发过程中对应用域进行分析评估，但通过更广泛地使用该方法，可能会获得更多数据。这些数据可揭示可能需要在技术上进一步解决的干扰问题 (Petersen 等人, 2022a)。如果不对具有某些物理化学特性的化学物质进行测定，则无论是从分析(例如，是否存在影响检测性能的偏差)还是从潜在的一致性(例如，该测定是否产生与体内测定相似的结果)的考虑，这些化学品的测定性能的不确定性就会更大。因此，评估具有不同物理化学特性的化学物质的测定一致性也可能很有价值，这样可以提高人们对该测定适用于更广泛化学物质的信心。

8.0 阳性对照

测定开发中应至少有一个阳性对照。这些对照可以与标准曲线结合使用，也可以作为阳性对照的单一浓度使用。阳性对照应与检测的终点相关。阳性对照的浓度应在测定的检测窗内。通常，如果仅测试单一浓度并且有关于最大剂量分离度的数据，则选择产生最大检测范围的 50%至 80%响应的阳性对照浓度。标准曲线通常使用系统的阳性对照生成，标准曲线定义为测定反应与已知分析物浓度之间的关系。标准曲线应可随着时间的推移而重现。标准曲线应在与研究的预期测试样本相同的载体或溶剂中制备。当无法获得相同的混合物或基质时，可以在有适当记录和理由的情况下使用替代混合物或基质。

9.0 仪器校准参考标准品

如果适用，应使用校准标准品和/或质量控制样本对仪器进行校准。应记录标准品或质量控制样品的来源，包括纯度、稳定性、来源、批号、分析证书和有效日期。有几种类型的标准品，例如美国药典汇编标准品和商业供应的材料，这些材料应从信誉良好的商业来源获得的，纯度有据可查。

理想情况下，某些仪器(例如质谱仪)的对照标准品应与测定中的目标分析物相同。这通常是不可能的，可以使用已知纯度的既定化学剂型。

10.0 设置质量标准

可以根据实验室内部或实验室间的测试结果设置过程控制测量的质量标准。有多种方法可以使用。例如，常用方法将 95% 的置信区间应用于过程控制测量，或者使用平均值 $\pm$ 两到三倍的标准差范围，但也经常使用其他统计方法。设定质量标准涉及平衡不同的目标。重要的是要使质量标准足够严格，以排除表明测定未按预期进行的数据，或过程控制测量的数据处于可能对测试物质结果产生偏差的范围内的数据。同样重要的是，不要将质量标准设置得过紧，以免测试数据在正常范围内时，测试中出现过多不符合质量标准的分析运行。